IV

PLV Complete Genomes

Contents

IV-1	Introduction	**223**
IV-2	Sequences	**226**
IV-3	Alignments	**229**

IV-1 Introduction

Complete genomes of HIV-1, HIV-2, and SIV sequences are represented in the alignment. For the selection, one of each of the major subtypes and a few CRFs of the HIV-1 M group were chosen, plus one N group, 2 O group, and one P group sequence. Most of the SIV-CPZs were included, and this year 3 SIV-gorilla sequences have been added. Three HIV-2 A group and 2 HIV-2 B group sequences were included along with the single genomes sequenced to date for HIV-2 G and HIV-2 U. 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. Also, many shorter sequences from PLVs are included in the PLV protein sequence alignments later in this compendium. This year we include two phylogenetic trees, built using PhyML on the *pol* and *env* gene regions. As discussed in the 2003[1] and 2005[2] Compendia, there is some evidence of recombination events and/or differential rates of evolution in the history of the primate lentiviruses such that phylogenies of different regions of the genomes are not completely congruent.

The alignment was based on the previous version published last year, produced 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 some empty columns have been inserted. The HIV-2/SIVsmm *vpx* gene is postulated to be a duplication of the *vpr* gene 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, but 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 MAC239. The translations are provided as a visual aid for finding landmarks in the genomes.

[1] `http://www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2003/7.pdf`

[2] `http://www.hiv.lanl.gov/content/sequence/HIV/COMPENDIUM/2005/7.pdf`

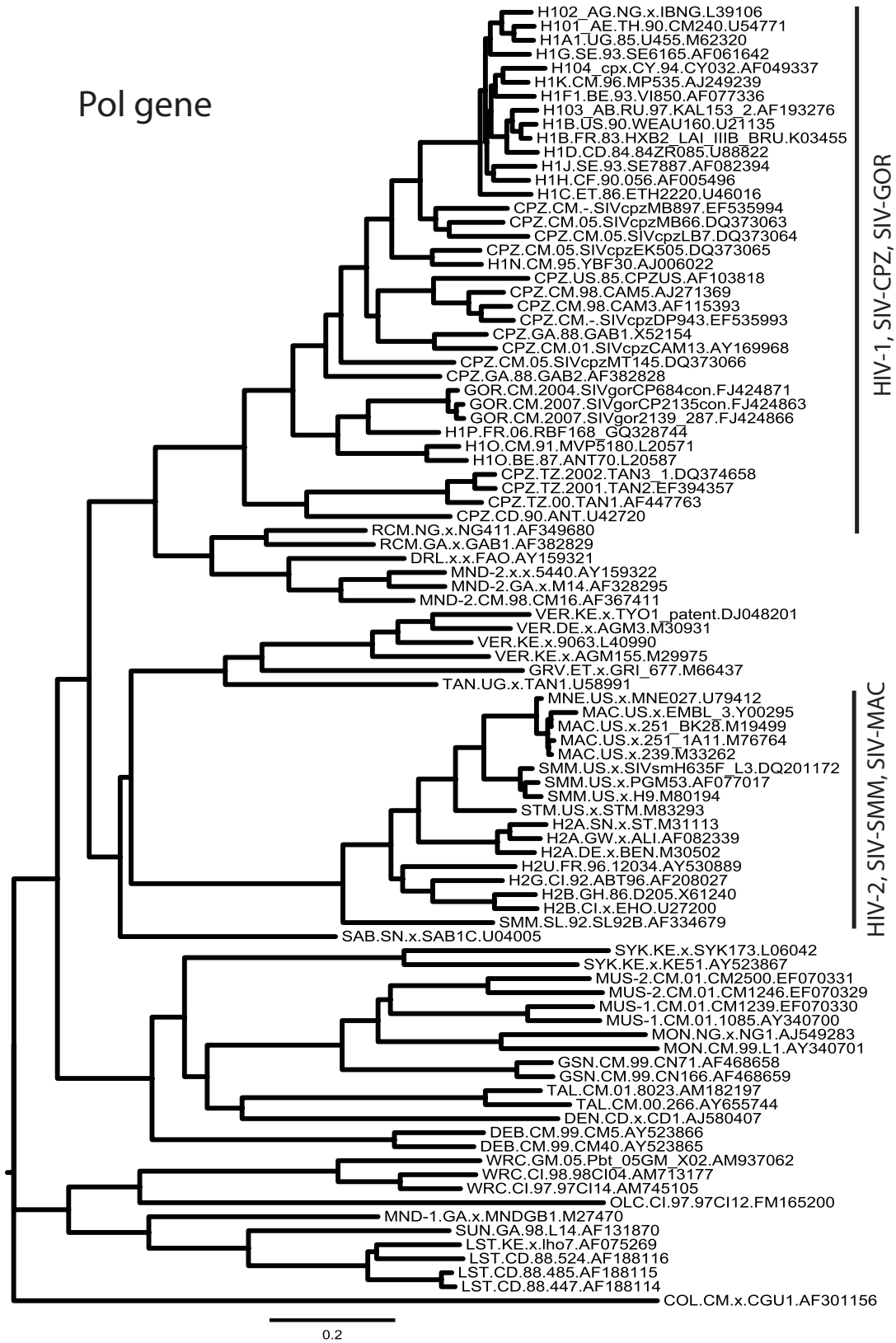


Figure IV-1: Genetic relationships among primate lentiviruses (PLVs), based on the *pol* gene. This tree shows the inferred phylogenetic relationships in the following PLV complete genome alignment, based on the *pol* gene, and using PhyML to calculate the tree.

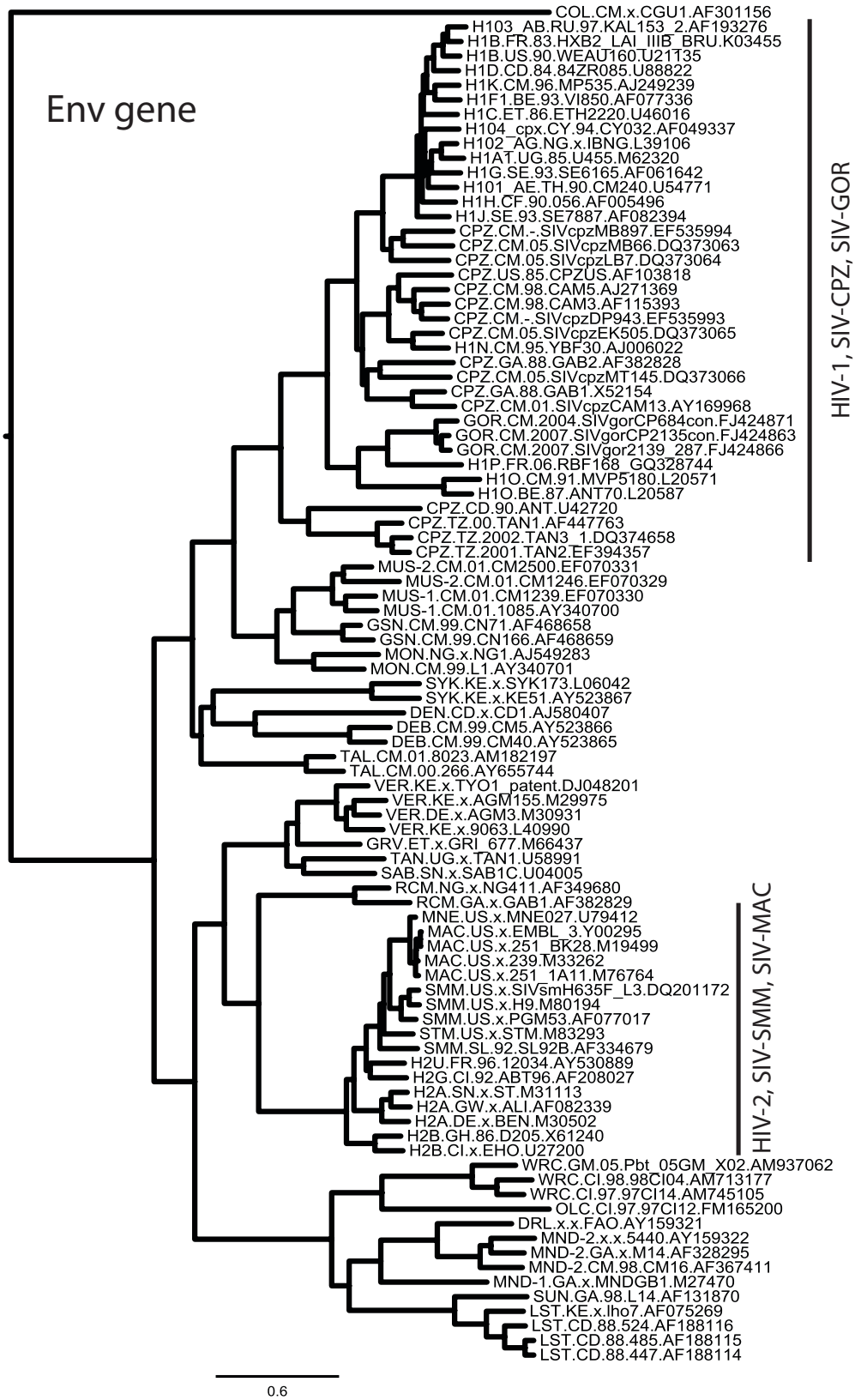


Figure IV-2: Genetic relationships among primate lentiviruses (PLVs), based on the *env* gene. This tree shows the inferred phylogenetic relationships in the following PLV complete genome alignment, based on the *env* gene, and using PhyML to calculate the tree. Note that the scale bar is different for *env* than for *pol*, reflecting the greater distances between *env* gene sequences than *pol* gene sequences.

IV-2 Sequences

Primate lentivirus (PLV) Sequences in the complete genome alignments.

Name	Accession	Country	Author	Reference
H1B.FR.83.HXB2	K03455	France	van Beveren, C.P.	*(in) Weiss,RL, Teich,N, Varmus,H and Coffin,J(Eds); RNA TUMOR VIRUSES, SECOND EDITION,* **2**, Vol 2: 1102-1123; Cold Spring Harbor Laboratory, Cold Spring Harbor (1985)
H1A1.UG.85.U455	M62320	Uganda	Oram, J.D.	*ARHR* **6**(9):1073-1078 (1990)
H1B.US.90.WEAU160	U21135	United States	Wei, X.	*Nature* **422**(6929):307-312 (2003)
H1C.ET.86.ETH2220	U46016	Ethiopia	Salminen, M.O.	*ARHR* **12**(14):1329-1339 (1996)
H1D.CD.84.84ZR085	U88822	D.R.C.	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H1F1.BE.93.VI850	AF077336	Belgium	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
H1G.SE.93.SE6165	AF061642	Sweden	Carr, J.K.	*Virology* **247**(1):22-31 (1998)
H1H.CF.90.056	AF005496	C.A.R.	Gao, F.	*J Virol* **72**(7):5680-5698 (1998)
H1J.SE.93.SE7887	AF082394	Sweden	Laukkanen, T.	*ARHR* **15**(3):293-297 (1999)
H1K.CM.96.MP535	AJ249239	Cameroon	Triques, K.	*ARHR* **16**(2):139-151 (2000)
H101_AE.TH.90.CM240	U54771	Thailand	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)
H102_AG.NG.x.IBNG	L39106	Nigeria	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
H103_AB.RU.97.KAL153_2	AF193276	Russia	Liitsola, K.	*ARHR* **16**(11):1047-1053 (2000)
H104_cpx.CY.94.CY032	AF049337	Cyprus	Gao, F.	*J Virol* **72**(12):10234-10241 (1998)
H1N.CM.95.YBF30	AJ006022	Cameroon	Simon, F.	*Nat Med* **4**(9):1032-1037 (1998)
H1O.BE.87.ANT70	L20587	Belgium	Vanden Haesevelde, M.	*J Virol* **68**(3):1586-1596 (1994)
H1O.CM.91.MVP5180	L20571	Cameroon	Gurtler, L.G.	*J Virol* **68**(3):1581-1585 (1994)
H1P.FR.06.RBF168	GQ328744	France	Plantier, J.-C.	*Nat Med* **15**(8); 871-2 (2009)
CPZ.CM.01.SIVcpzCAM13	AY169968	Cameroon	Nerrienet, E.	*J Virol* **79**(2):1312-9 (2005)
CPZ.CM.05.SIVcpzEK505	DQ373065	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzLB7	DQ373064	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzMB66	DQ373063	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.05.SIVcpzMT145	DQ373066	Cameroon	Keele, B.F.	*Science* **313**(5786):523-526 (2006)
CPZ.CM.-.SIVcpzMB897	EF535994	Cameroon	Heuverswyn, F.V.	*Virology* **368**(1):155-171 (2007)
CPZ.CM.-.SIVcpzDP943	EF535993	Cameroon	Heuverswyn, F.V.	*Virology* **368**(1):155-171 (2007)
CPZ.CM.98.CAM3	AF115393	Cameroon	Corbet, S.	*J Virol* **74**(1):529-534 (2000)
CPZ.CM.98.CAM5	AJ271369	Cameroon	Muller-Trutwin, M.C.	*J Med Primatol* **29**(3-4); 166-72 (2000)
CPZ.GA.88.GAB1	X52154	Gabon	Huet, T.	*Nature* **345**(6273):356-359 (1990)
CPZ.GA.88.GAB2	AF382828	Gabon	Bibollet-Ruche, F.	*ARHR* **20**(12):1377-1381 (2004)
CPZ.US.85.CPZUS	AF103818	United States	Gao, F.	*Nature* **397**(6718):436-441 (1999)
CPZ.CD.90.ANT	U42720	D.R.C.	Vanden Haesevelde, M.M.	*Virology* **221**(2):346-350 (1996)
CPZ.TZ.00.TAN1	AF447763	Tanzania	Santiago, M.L.	*J Virol* **77**(3):2233-2242 (2003)
CPZ.TZ.2001.TAN2	EF394357	Tanzania	Takehisa, J.	*J Virol* **81**(14):7463-7475 (2007)
CPZ.TZ.2002.TAN3_1	DQ374658	Tanzania	Takehisa, J.	*J Virol* **81**(14):7463-7475 (2007)
GOR.CM.2007.SIVgor2139_287	FJ424866	Cameroon	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)
GOR.CM.2007.SIVgorCP2135con	FJ424863	Cameroon	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)
GOR.CM.2004.SIVgorCP684con	FJ424871	Cameroon	Takehisa, J.	*J Virol* **83**(4):1635-1648 (2009)
H2A.DE.x.BEN	M30502	Germany	Kirchhoff, F.	*Virology* **177**(1):305-311 (1990)
H2A.GW.x.ALI	AF082339	Guinea-Bissau	Azevedo-Pereira, J.M.	Unpublished
H2A.SN.x.ST	M31113	Senegal	Kumar, P.	*J Virol* **64**(2):890-901 (1990)

Name	Accession	Country	Author	Reference
H2B.CI.x.EHO	U27200	Cote D'Ivoire	Rey-Cuille, M.A.	*Virology* **202**(1):471-476 (1994)
H2B.GH.86.D205	X61240	Ghana	Dietrich, U.	*Nature* **342**(6252):948-950 (1989)
H2G.CI.92.ABT96	AF208027	Cote D'Ivoire	Brennan, C.A.	*ARHR* **13**(5):401-404 (1997)
H2U.FR.96.12034	AY530889	France	Damond, F.	*ARHR* **20**(6):666-672 (2004)
SMM.SL.92.SL92B	AF334679	Sierra Leone	Chen, Z.	*J Virol* **70**(6):3617-3627 (1996)
SMM.US.x.H9	M80194	United States	Courgnaud, V.	*J Virol* **66**(1):414-419 (1992)
SMM.US.x.PGM53	AF077017	United States	Novembre, F.J.	*J Virol* **72**(11):8841-8851 (1998)
SMM.US.x.SIVsmH635F_L3	DQ201172	United States	Kuwata, T.	*J Virol* **80**(3):1463-75 (2006)
MAC.US.x.239	M33262	United States	Kestler, H.	*Science* **248**(4959):1109-1112 (1990)
MAC.US.x.251_1A11	M76764	United States	Marthas, M.L.	*J Med Primatol* **18**(3-4):311-9 (1989)
MAC.US.x.251_BK28	M19499	United States	Hirsch, V.	*Cell* **49**(3):307-319 (1987)
MAC.US.x.EMBL_3	Y00295	United States	Franchini, G.	*Nature* **328**(6130):539-543 (1987)
STM.US.x.STM	M83293	United States	Novembre, F.J.	*Virology* **186**(2):783-787 (1992)
MNE.US.x.MNE027	U79412	United States	Kimata, J.T.	*J Virol* **72**(1):245-256 (1998)
DEB.CM.99.CM40	AY523865	Cameroon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEB.CM.99.CM5	AY523866	Cameroon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
DEN.CD.x.CD1	AJ580407	D.R.C.	Dazza, M.C.	*J Virol* **79**(13):8560-8571 (2005)
LST.CD.88.447	AF188114	D.R.C.	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.485	AF188115	D.R.C.	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.CD.88.524	AF188116	D.R.C.	Beer, B.E.	*J Virol* **74**(8):3892-3898 (2000)
LST.KE.x.lho7	AF075269	Kenya	Hirsch, V.M.	*J Virol* **73**(2):1036-1045 (1999)
SUN.GA.98.L14	AF131870	Gabon	Beer, B.E.	*J Virol* **73**(9):7734-7744 (1999)
MND-1.GA.x.MNDGB1	M27470	Gabon	Tsujimoto, H.	*Nature* **341**(6242):539-541 (1989)
OLC.CI.97.97CI12	FM165200	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.CI.97.97CI14	AM745105	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.CI.98.98CI04	AM713177	Cote D'Ivoire	Liegeois, F.	*J Virol* **83**(1):428-439 (2009)
WRC.GM.05.Pbt_05GM_X02	AM937062	Gambia	Locatelli, S.	*Virology* **376**(1):90-100 (2008)
MND-2.CM.98.CM16	AF367411	Cameroon	Takehisa, J.	*ARHR* **17**(12):1143-1154 (2001)
MND-2.GA.x.M14	AF328295	Gabon	Souquiere, S.	*J Virol* **75**(15):7086-7096 (2001)
MND-2.x.x.5440	AY159322		Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
DRL.x.x.FAO	AY159321		Hu, J.	*J Virol* **77**(8):4867-4880 (2003)
MON.CM.99.L1	AY340701	Cameroon	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
MON.NG.x.NG1	AJ549283	Nigeria	Barlow, K.L.	*J Virol* **77**(12):6879-6888 (2003)
GSN.CM.99.CN166	AF468659	Cameroon	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
GSN.CM.99.CN71	AF468658	Cameroon	Courgnaud, V.	*J Virol* **76**(16):8298-8309 (2002)
TAL.CM.00.266	AY655744	Cameroon	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
TAL.CM.01.8023	AM182197	Cameroon	Liegeois, F.	*Virology* **349**(1):55-65 (2006)
MUS-1.CM.01.1085	AY340700	Cameroon	Courgnaud, V.	*J Virol* **77**(23):12523-12534 (2003)
MUS-1.CM.01.CM1239	EF070330	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS-2.CM.01.CM1246	EF070329	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
MUS-2.CM.01.CM2500	EF070331	Cameroon	Aghokeng, A.F.	*Virology* **360**(2):407-418 (2007)
SYK.KE.x.KE51	AY523867	Kenya	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
SYK.KE.x.SYK173	L06042	Kenya	Hirsch, V.M.	*J Virol* **67**(3):1517-1528 (1993)
RCM.GA.x.GAB1	AF382829	Gabon	Bibollet-Ruche, F.	*J Virol* **78**(14):7748-7762 (2004)
RCM.NG.x.NG411	AF349680	Nigeria	Beer, B.E.	*J Virol* **75**(24):12014-12027 (2001)
GRV.ET.x.GRI_677	M66437	Ethiopia	Fomsgaard, A.	*Virology* **182**(1):397-402 (1991)

Name	Accession	Country	Author	Reference
SAB.SN.x.SAB1C	U04005	Senegal	Jin, M.J.	*EMBO J* **13**(12):2935-2947 (1994)
TAN.UG.x.TAN1	U58991	Uganda	Soares, M.A.	*Virology* **228**(2):394-399 (1997)
VER.DE.x.AGM3	M30931	Germany	Baier, M.	*Virology* **176**(1):216-221 (1990)
VER.KE.x.9063	L40990	Kenya	Hirsch, V.M.	*J Virol* **69**(2):955-967 (1995)
VER.KE.x.AGM155	M29975	Kenya	Johnson, P.R.	*J Virol* **64**(3):1086-1092 (1990)
VER.KE.x.TYO1_patent	DJ048201	Kenya	Omori, T.	Patent: WO 2007049749-A 13 03-MAY-2007; DNAVEC CORPORATION
COL.CM.x.CGU1	AF301156	Cameroon	Courgnaud, V.	*J Virol* **75**(2):857-866 (2001)

```
                                 5' LTR U3 start
H1B.FR.83.HXB2                 .TGGAAGGGCTAATTCACTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCA...GGGCCAGGG...ATCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGA 163
H1A1.UG.85.U455                ..................................................................................................................................................................... 0
H1B.US.90.WEAU160              .-------------T--------AA------------------------G-----------------T----------G---------------..---------...-CT-----C------TG-------------T---------------G--A----- 163
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
H101_AE.TH.90.CM240            ..................................................................................................................................................................... 0
H102_AG.NG.x.IBNG              ..................................................................................................................................................................... 0
H103_AB.RU.97.KAL153_2         ..................................................................................................................................................................... 0
H104_cpx.CY.94.CY032           ..................................................................................................................................................................... 0
H1N.CM.95.YBF30                ..................................................................................................................................................................... 0
H1O.BE.87.ANT70                C--------T-----T-------TAA----GC---A-----G-----T---G-G--TA----T--G--A-T-----------G----------------..--A-----A...-C---G-TC--------A----------TG-TT--A-----------GTCAGA--- 164
H1O.CM.91.MVP5180              C----T---T-----T-------TAAG---GC---A-----G-----C-----A--T-----T--G--A-T-----------G----TGT--------G..--A-----A...CCT----TC--------A----------TG-TT--A--G--------GTCAG---- 164
H1P.FR.06.RBF168               ..................................................................................................................................................................... 0
CPZ.CM.01.SIVcpzCAM13          ..................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzEK505          ..................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzLB7            ..................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzMB66           ..................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzMT145          ..................................................................................................................................................................... 0
CPZ.CM.-.SIVcpzMB897           ..................................................................................................................................................................... 0
CPZ.CM.-.SIVcpzDP943           ..................................................................................................................................................................... 0
CPZ.CM.98.CAM3                 ..................................................................................................................................................................... 0
CPZ.CM.98.CAM5                 ..................................................................................................................................................................... 0
CPZ.GA.88.GAB1                 .--------T--G--T-----AGGA----------G--------C--C---G----T------------T---------C-G------------A--..--A-----A...-CA----TC-----CTGT-------------T---A-----G--CC-GACAGAG-- 163
CPZ.GA.88.GAB2                 ..................................................................................................................................................................... 0
CPZ.US.85.CPZUS                .--------T--G--T-T---AG-A----------G-----A-----C---G-T-----T---------AAT---------C-G------T-----C---..-----T--A...G-A--------C--A--A-A---GC------TT---T----C--TC-CACAGA--- 163
CPZ.CD.90.ANT                  ..................................................................................................................................................................... 0
CPZ.TZ.00.TAN1                 ..................................................................................................................................................................... 0
CPZ.TZ.2001.TAN2               ..................................................................................................................................................................... 0
CPZ.TZ.2002.TAN3_1             ..................................................................................................................................................................... 0
GOR.CM.2007.SIVgor2139_287     ..................................................................................................................................................................... 0
GOR.CM.2007.SIVgorCP2135con    ..................................................................................................................................................................... 0
GOR.CM.2004.SIVgorCP684con     ..................................................................................................................................................................... 0
H2A.DE.x.BEN                   .--C-----A-GT--T--AGTAGGA-G-----TAGA-----A--CA-A-ACC-AG-AA-AGAGG----GATAA-A--A----G------T--T--T-AT..---------A...G-A--G--C---A--TA---C--G---CTG-GG---------T----A--A.... 158
H2A.GW.x.ALI                   .--------A-GT--T--AGTG-GA-------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G----GATAA-T--A----G---------T--T-AT..---------A...--A--G--C--GA--TT------G---CTG-GG-------------A--T..... 158
H2A.SN.x.ST                    ..................................................................................................................................................................... 0
H2B.CI.x.EHO                   .--------A-TT-CT-TAGTG-GA----G--TA-A--A--A--CACA-ACT-AG-AA-TGA-G-----ATTG-GT---GA-G-------------AT..---------A...G-A--------CAA-TT------C---CTA-GG--------G---A-AA-T.... 158
H2B.GH.86.D205                 .--------A-GTA-T-TAGTG-GA-------CAGA--AT-A--CACA-ATT-TG-GA-TGA-G-----ATTG-GT---GC-G---A-----T--T-AT..----------..--A--GC----CAAATA------T---CTG-GG------G---------A---...T 160
H2G.CI.92.ABT96                ..................................................................................................................................................................... 0
H2U.FR.96.12034                ..................................................................................................................................................................... 0
SMM.SL.92.SL92B                ..................................................................................................................................................................... 0
SMM.US.x.H9                    ..................................................................................................................................................................... 0
SMM.US.x.PGM53                 ..........................................................................-A----G---A--T------T---..---------A..---------C--TA--TT------C---CTA-GG--AT-G--C-----AAGT..... 84
SMM.US.x.SIVsmH635F_L3         .--------A-TTA-T--AATG-GAA------TAGA--AT-A---A---AC--GG-AA-GGA-G----AATAA-A--A----G---A--T------TT-..---------A...-CT-----C--TA--TA------G---CT--GG--AT----C-----A--T..... 158
                                 5' LTR U3 start
MAC.US.x.239                   .--------A-TTA-T--AGTGC-A-------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A----G----G-T-----CT--..--A-----A..--T-----C---AA---A-----C---CTA-GG--AT----C--T--AA-T..... 158
MAC.US.x.251_1A11              .--------A-TTA-T-TAGTGC-A-------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A----G----G-T-----CT--..--A-----A..--T-----C---AA---A-----C---CTA-GG--AT----C--T--AA-T..... 158
MAC.US.x.251_BK28              .--------A-TTA-T--AGTGC-A-------TAGA---T-A--CA---ACT-AG-AA-GGA-G-----AT-A-A--A----G----G-T-----CT--..--A-----A..--T-----C---AA---A-----C---CTA-GG--AT----C--T--AA-T..... 158
MAC.US.x.EMBL_3                ..................................................................................................................................................................... 0
STM.US.x.STM                   .--------A-TTA-T--AGTG-.............................................................................................................................................. 22
MNE.US.x.MNE027                ..................................................................................................................................................................... 0
DEB.CM.99.CM40                 ..................................................................................................................................................................... 0
DEB.CM.99.CM5                  ..................................................................................................................................................................... 0
DEN.CD.x.CD1                   ..................................................................................................................................................................... 0
LST.CD.88.447                  ..................................................................................................................................................................... 0
LST.CD.88.485                  ..................................................................................................................................................................... 0
LST.CD.88.524                  ..................................................................................................................................................................... 0
LST.KE.x.lho7                  .-----------TA-TGG-----GA-------G--C---T-AAT---C-ACTGTG-GA-TGAGTGG--TCT-A-TGGA--C-TCAT----------GAT..--A-----A...-CA-----C--T-----A----------TG-GGC--T---A------G-CT...TG 160
SUN.GA.98.L14                  .----------CTGGT-----AG-ACC---G----A---T-G---T---ATGCAG-AA-TGAGTGG----T-A-AA---GC-G---AG-T-----CAA-..-----T--A...G-G---------AA-G-------C---CTA-GG-------C------CACC...AT 160
MND-1.GA.x.MNDGB1              ..................................................................................................................................................................... 0
OLC.CI.97.97CI12               .........................................................................AA-AGG-A-C-TC---GCAACAGAGGAGGTA-AAGAG--CACGT-A--GAGA--TAA---GCCA--G----TA-GGG-T--TAGGAG---G--A..... 94
WRC.CI.97.97CI14               .----G-ATA--T--TGG--AGCTA-----G----A--AG-GA-C--C-ATGCAC--A-TGA-TGG--GCTGA-AGA--GC-G---AGCT------GA-..--A------..--A----GG---AA---A-----G-------GG--AT-G--T-----AC-C..... 158
WRC.CI.98.98CI04               .----GTATA--T--TGG--T-CTA-----G----A--AG-GA-C--T-ATGC--T-A-TGA-TGG--ACTGA-AGAA-G--G----GCA------GAG..--A-----C..--AC-C-GG---AA---A-----C-----T-GG--AT----G--T--G--A..... 158
WRC.GM.05.Pbt_05GM_X02         ..................................................................................................................................................................... 0
MND-2.CM.98.CM16               ..................................................................................................................................................................... 0
MND-2.GA.x.M14                 ..................................................................................................................................................................... 0
MND-2.x.x.5440                 ..................................................................................................................................................................... 0
DRL.x.x.FAO                    ..................................................................................................................................................................... 0
MON.CM.99.L1                   ..................................................................................................................................................................... 0
MON.NG.x.NG1                   ..................................................................................................................................................................... 0
GSN.CM.99.CN166                ..................................................................................................................................................................... 0
GSN.CM.99.CN71                 ..................................................................................................................................................................... 0
TAL.CM.00.266                  ..................................................................................................................................................................... 0
TAL.CM.01.8023                 ..................................................................................................................................................................... 0
MUS-1.CM.01.1085               ..................................................................................................................................................................... 0
MUS-1.CM.01.CM1239             ..................................................................................................................................................................... 0
MUS-2.CM.01.CM1246             ..................................................................................................................................................................... 0
MUS-2.CM.01.CM2500             ..................................................................................................................................................................... 0
SYK.KE.x.KE51                  ..................................................................................................................................................................... 0
SYK.KE.x.SYK173                ..............................................................................................................---A...-C-C-G--C--GT-A-T-----------TA-GGG-A-----T---AA----..... 55
RCM.GA.x.GAB1                  ..................................................................................................................................................................... 0
RCM.NG.x.NG411                 ..................................................................................................................................................................... 0
GRV.ET.x.GRI_677               .----T---A--TA-T----TG--A-----G--A-G-----GA--T---ATGC--TGA--GAGTGG--AATAA-AGA-----G---AGCT---T------..--C--G---..--A--G--C--GAGAGT------C-TC----TT--------C-----G--C..... 158
SAB.SN.x.SAB1C                 .--------A-TTA-T--AGTG--A-G---A-GA-G---T-G-----C-ATGCTCTTA-TGAGTGG---ATAG-AGA--GA-G------T--T---GAT..---------T...-C---G--C---AA-TG--------------TT--A--G-----T--G--C..... 158
TAN.UG.x.TAN1                  .----G--CA--TA-TGG--A-C-AAG---G-GC-A-----GA----C-ATGC-CTAA-TGA-TGG---ATTA-AGA-----G---AGCA--TT-C--T..---------C...-C---GA-G--C-GATG---C--C-TT----TTG--T----G--G--A--T..... 158
VER.DE.x.AGM3                  ..................................................................................................................................................................... 0
VER.KE.x.9063                  .----T---A-TTA-T-----G--A-G--GA-TAGA-----CA-C----ATGCTCTTA-TGA-TGG--AAT-A-TGA-----G-A-CGC--GGT--A--..--A-----C...G-A---------C-GCTG------C-TT----T------------G--A---..... 158
VER.KE.x.AGM155                .----T---A-TTA-T-----G-TA-G---A-TA-G------A------ATGC-CT-A-TGA-TGG--AAT-A-TGA-----G-A-CGCA-GGT--AA-..--A--T---..--A-----C--GAG-TG------C-TC----T------------G----CC..... 158
VER.KE.x.TYO1_patent           .----T---A-TTA-T-----G-TA-G---A-TA-G-----GA------ATGCTCTTA-TGA-TGG--GATAA-TGA-----G-A-TGC--GGT-GAAG..--A-----A..--A----TC--TAAATG------G-TC----TT--------G-----G--C..... 158
COL.CM.x.CGU1                  ..................................................................................................................................................................... 0
```

```
H1B.FR.83.HXB2             GAAGTTAGAAGAA............GCCAACAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGC............CTGCATGGAATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACATGGCCCGAGAGCTGCATC  309
H1A1.UG.85.U455            ..............................................................................................................................  0
H1B.US.90.WEAU160          ----G-------------..--T--TG------------A----------------A-----------------------G----------AT----A------A---AT-----A----------AAA---------------G-A----------------  309
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
H101_AE.TH.90.CM240        ..............................................................................................................................  0
H102_AG.NG.x.IBNG          ..............................................................................................................................  0
H103_AB.RU.97.KAL153_2     ..............................................................................................................................  0
H104_cpx.CY.94.CY032       ..............................................................................................................................  0
H1N.CM.95.YBF30            ..............................................................................................................................  0
H1O.BE.87.ANT70            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-AC-GAT-----A------T--ATCT----GCAACACC--TG-T--TATGATAACT--C-  313
H1O.CM.91.MVP5180          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-AC--A-A---CA------T--ATCAT---GC--AACA--T--A----TGC-AAA---C-  313
H1P.FR.06.RBF168           ..............................................................................................................................  0
CPZ.CM.01.SIVcpzCAM13      ..............................................................................................................................  0
CPZ.CM.05.SIVcpzEK505      ..............................................................................................................................  0
CPZ.CM.05.SIVcpzLB7        ..............................................................................................................................  0
CPZ.CM.05.SIVcpzMB66       ..............................................................................................................................  0
CPZ.CM.05.SIVcpzMT145      ..............................................................................................................................  0
CPZ.CM..SIVcpzMB897        ..............................................................................................................................  0
CPZ.CM..SIVcpzDP943        ..............................................................................................................................  0
CPZ.CM.98.CAM3             ..............................................................................................................................  0
CPZ.CM.98.CAM5             ..............................................................................................................................  0
CPZ.GA.88.GAB1             -C--G-----C-------------------TG------T----A-T--C----G--T--CA-AT-TCAGCATGGGATGGAAG---AGGAC............A---G--C--G-TC---C-C---------A-G--G--C--AAGA--T--T---A----A-AA----  309
CPZ.GA.88.GAB2             ..............................................................................................................................  0
CPZ.US.85.CPZUS            -G-AG----GCGG-------------T----------C--A--TAC--C-G-----CA--T--CAGCATGGAATGGAAG---A-CAT............G-C-----C---ATC---CA--------CTGAA--G--TCGGAGA----GA--TAA-----------  309
CPZ.CD.90.ANT              ..............................................................................................................................  0
CPZ.TZ.00.TAN1             ..............................................................................................................................  0
CPZ.TZ.2001.TAN2           ..............................................................................................................................  0
CPZ.TZ.2002.TAN3_1         ..............................................................................................................................  0
GOR.CM.2007.SIVgor2139_287 ..............................................................................................................................  0
GOR.CM.2007.SIVgorCP2135con ............................................................................................................................  0
GOR.CM.2004.SIVgorCP684con ............................................................................................................................  0
H2A.DE.x.BEN               .CTC-C-C----G.............-AG-GG---AT---GC--A-T----AG-------A-CACAAACAAGCAGACATGATG---AGCAT............G-G--GACA----T----CA------TC-ATG--G--C-A-A-CT---A----TTCACT--T-C-  303
H2A.GW.x.ALI               .GTCCC-C-----------------.GGG-GG-CACT----CTCA-T--C--C------A--ACAAACAAGCAGGCATGATG-CACGCAT............G-G--GACA----TT-----A------CCTAAG--G--TCA-G--T---AA---TTTATT--A--C-  303
H2A.SN.x.ST                ..............................................................................................................................  0
H2B.CI.x.EHO               -T-A--C-----------------.C-AG-GG-C-AG--A-C-CATT-TC--G-G--T--A-CACAGACCTCCGCATGGGATG-CCCCCAC............GA---GACCC-T-TC---CA------TC--T--------A-G-CT-TG-----TTCAGCAG-TTC-  303
H2B.GH.86.D205             -CCAGC--CGACC............CGAG-GG-G-AG--A-C-CATT--C-AA-G-----G-CACAGATCTCCTCATGGGATG-CATCCAT............G-G--GACTC-TATC---CA------TTC--T---G----A-G--T-TG----TTTCA-TAG-TT--  306
H2G.CI.92.ABT96            ..............................................................................................................................  0
H2U.FR.96.12034            ..............................................................................................................................  0
SMM.SL.92.SL92B            ..............................................................................................................................  0
SMM.US.x.H9                ..............................................................................................................................  0
SMM.US.x.PGM53             .GTC-C---T---------------TC-AG---AC----CACATTA---AG-G--T--A-CTCAGATATCCCAGTGGGATG--CCCTG............G----G--TC-G-CA----A------TTCA-AAT----C-ACAGAT-TGA---TTTTATTAGA--C-  229
SMM.US.x.SIVsmH635F_L3     .GTC-C---T---------------C-GG---AC----CACATT--C--G-G--T--A-CACAGACTCATCAGTGGGATG-CCCCTG............G----G--AC-G-CA----A------TCCAGAAT-G--T-A-AGCT-T-A---ATTTATTAA-T-C-  303

MAC.US.x.239               .GTA-C---T--G.............--AC-GG-G-AT---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCAGTGGGATG-CCCTTG............G----G--TC---CA----A-----TCCAACT--G--C-ACAC-T-TGA---ATAT-TTAGAT-C-  303
MAC.US.x.251_1A11          .GTA-C---T--G.............--AC-GG-G-AT---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCAGTGGGATG-CCCTTG............G----G--TCC--CA----A-------CCAACT--G--C-ACAC-T-TGA---ATAT-TTAGAT-C-  303
MAC.US.x.251_BK28          .GTA-C---T--G.............--AC-GG-G-AT----GGCATTAT--AA-G--G--A-CTCAAACTTCCAAGTGGGATG-CCCTTG............G----G--TC---C-----A------TCCAACT-----C-ACAC-T-TGA---ATAT-CTAGAT-C-  303
MAC.US.x.EMBL_3            ..............................................................................................................................  0
STM.US.x.STM               ..............................................................................................................................  22
MNE.US.x.MNE027            ..............................................................................................................................  0
DEB.CM.99.CM40             ..............................................................................................................................  0
DEB.CM.99.CM5              ..............................................................................................................................  0
DEN.CD.x.CD1               ..............................................................................................................................  0
LST.CD.88.447              ..............................................................................................................................  0
LST.CD.88.485              ..............................................................................................................................  0
LST.CD.88.524              ..............................................................................................................................  0
LST.KE.x.lho7              TG-TGAGT-TA--------------.ATCC-TC--AT----CACAAT--C-T--G--TT-AAGCCAGCTGGGGGTCCTGGAAG-CCCCTG............G----GAG-C-TATC---CAC--CA-TCCAATGT-G---...................-T-GACTT-A  288
SUN.GA.98.L14              TG-TGA---CAGG.............-ATCCT--TCATCCTTGTCAAGC-C-AC-G--TT-CAGCCAGCAAGGGGTTAATGAAG-CCCTTG............G----GAGAC-TATC----CA-----TCC-ACT--G--C..................T-TGATTT-A  288
MND-1.GA.x.MNDGB1          ..............................................................................................................................  0
OLC.CI.97.97CI12           .CCA--GCCT--G.............A--CTTG---AT----GGGATGC------G--T---A-TCAGACCCCTCAAACA-CTG-G-AGT-A-GG...............-C-GG-T----GTC--CA-TCCA-ATAG----...................A-TAC--C-  218
WRC.CI.97.97CI14           .TTCAGT------TCAGATAATACA-A-CCAT-CTAT----GA-AT-AG--AC-G--T--A-CAT-TGCTGAG...CCAGA-G--CCCTG............G-----CATC-G-TC----A-------CCTAAG--G--C......TGTGATTTTGT--CAGGAAGA-  306
WRC.CI.98.98CI04           .TTTAGT-----GTCTAATGACAGCT-TC-GT-CTAC-----G-A--AGC-------T--A-GCT--CAAGAA...TCAGA-G--CCTTG............G-----CAC--G-TC----A-------CCT--A-----C......TGTGACTTTGTG-CAAGTAGA-  306
WRC.GM.05.Pbt_05GM_X02     ..............................................................................................................................  0
MND-2.CM.98.CM16           ..............................................................................................................................  0
MND-2.GA.x.M14             ..............................................................................................................................  0
MND-2.x.x.5440             ..............................................................................................................................  0
DRL.x.x.FAO                ..............................................................................................................................  0
MON.CM.99.L1               ..............................................................................................................................  0
MON.NG.x.NG1               ..............................................................................................................................  0
GSN.CM.99.CN166            ..............................................................................................................................  0
GSN.CM.99.CN71             ..............................................................................................................................  0
TAL.CM.00.266              ..............................................................................................................................  0
TAL.CM.01.8023             ..............................................................................................................................  0
MUS-1.CM.01.1085           ..............................................................................................................................  0
MUS-1.CM.01.CM1239         ..............................................................................................................................  0
MUS-2.CM.01.CM1246         ..............................................................................................................................  0
MUS-2.CM.01.CM2500         ..............................................................................................................................  0
SYK.KE.x.KE51              ..............................................................................................................................  0
SYK.KE.x.SYK173            .-TAGAG-G-TGC............TTGG-GT-T-AG-----T--ACT-C-CC-C--T--A-CCTCT...GGGACAGGGGA-TTC-A-CAT............G-G--GAACC-CAT-----AC--CA--CC--A---G--C..................T-TACA-CAG  179
RCM.GA.x.GAB1              ..............................................................................................................................  0
RCM.NG.x.NG411             ..............................................................................................................................  0
GRV.ET.x.GRI_677           .CT-CAT--G--G.............--ACG---CT-T----GACA-T-TC--A-G--T--A-CACAGATGGGG......GAAG--CCTGAT-GAATAGATC-TG-------C--G-TC----A-------CCGAAGT-G--G...................-T-GA-T-C-  291
SAB.SN.x.SAB1C             .CT--C---G--G.............--T--G--TA-T------CA-T--C--C-G--T----CTCAAGTGGCC......TATG-G-ATGAC-CCTGG......-A---GACCC-G-T-----A-------CCA-TA------...................-T-GACT--G  285
TAN.UG.x.TAN1              .GTC-CCC-G---..............---C-GG-T-AG...-GACA-T-T---C-C-------CCCAGATTGAA......TG-G-AA-CGACCC-TGG......-A---GACA--G-T-----AA---A-TCCA-T-------...................-TT-A-T-CA  282
VER.DE.x.AGM3              ..............................................................................................................................  0
VER.KE.x.9063              .-T-CAT-----G.............--AG-G-CCT-T----GACATTA----GCG--T----CACAAATAAAA.....GA-G-CCCCGAT-GCAT-AGTC-TG---GACC--G-TC----A-------CCTATGT-----...................-T--A-T--G  291
VER.KE.x.AGM155            .CT-CAT--G---..............--AG-A-C-T-T--A-GGCATT-----G-------A-CACAACTGCAT......GAAG-CCCTGAT-G-ATAAATC-TG-----A-A--G-CA----A------TCCAATGT-G--T..................-TT-A-T-CG  291
VER.KE.x.TY01_patent       .TTACAT--G---..............--AC-A-C-T-T--A-GACATT--C-AG-C--T--A-C-CAGATGGGA......GAAG--CC-GAT-G-AT-AGCC-TG---GA-C--G-T-----A------TCCTATGT-G---...................ATA-A-T-CG  291
COL.CM.x.CGU1              ..............................................................................................................................  0
```

```
                                      TCF-1 alpha
H1B.FR.83.HXB2               CGGAGTACTTCAAGAAC..........................................TGCTGACA.................................... 334
H1A1.UG.85.U455              ........................................................................................................ 0
H1B.US.90.WEAU160            ----------T--AG--..........................................--------.................................... 334
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
H101_AE.TH.90.CM240          ........................................................................................................ 0
H102_AG.NG.x.IBNG            ........................................................................................................ 0
H103_AB.RU.97.KAL153_2       ........................................................................................................ 0
H104_cpx.CY.94.CY032         ........................................................................................................ 0
H1N.CM.95.YBF30              ........................................................................................................ 0
H1O.BE.87.ANT70              -A---CT----C----G..........................................GA--A-A-ACTGCTG............................. 345
H1O.CM.91.MVP5180            -A---CT----CCC--G.............................................-A--TGACACTGC............................. 343
H1P.FR.06.RBF168             ........................................................................................................ 0
CPZ.CM.01.SIVcpzCAM13        ........................................................................................................ 0
CPZ.CM.05.SIVcpzEK505        ........................................................................................................ 0
CPZ.CM.05.SIVcpzLB7          ........................................................................................................ 0
CPZ.CM.05.SIVcpzMB66         ........................................................................................................ 0
CPZ.CM.05.SIVcpzMT145        ........................................................................................................ 0
CPZ.CM.-.SIVcpzMB897         ........................................................................................................ 0
CPZ.CM.-.SIVcpzDP943         ........................................................................................................ 0
CPZ.CM.98.CAM3               ........................................................................................................ 0
CPZ.CM.98.CAM5               ........................................................................................................ 0
CPZ.GA.88.GAB1               ---------A---AG--...............................................----T................................... 331
CPZ.GA.88.GAB2               ........................................................................................................ 0
CPZ.US.85.CPZUS              -A---------CG----................................................---TG................................... 331
CPZ.CD.90.ANT                ........................................................................................................ 0
CPZ.TZ.00.TAN1               ........................................................................................................ 0
CPZ.TZ.2001.TAN2             ........................................................................................................ 0
CPZ.TZ.2002.TAN3_1           ........................................................................................................ 0
GOR.CM.2007.SIVgor2139_287   ........................................................................................................ 0
GOR.CM.2007.SIVgorCP2135con  ........................................................................................................ 0
GOR.CM.2004.SIVgorCP684con   ........................................................................................................ 0
H2A.DE.x.BEN                 -A--AG-G--TGG-C--AAG...TCAGGATTGCCAGAGAAAGAA...TGGAAGGCAAAACTGAAAGCAAGAGGGATACCATATAGTGAA....................-A---GGAACAACCATACTTGGTCA.................................... 411
H2A.GW.x.ALI                 -A---G-A--TGG-T--AAG...TCAGGCCTGCCAGAAGATGAG...TGGAAGGCAAGACTGAAAGCAAGAGGGATACCATTTAGTAAGAACAGGAACAGC..........---TTTGGTCA............................................ 409
H2A.SN.x.ST                  ........................................................................................................ 0
H2B.CI.x.EHO                 -A---G-G--TGG-T-TCAG...TCAGGAATGCCAGAGAAAGAG...TGGAAGGCTAAACTGAGAGCAAGAGGAATACCTACAGAG.......................-AGACAGGAAATAACAGCACAAAGAGGAACT.............................. 414
H2B.GH.86.D205               -A--AG-G---GG-T--CAG...TCAGGGCTACCAGAGGAGGAG...TGGAAGGCTAGACTAAAAGCAAGAGGGATACCTACAGAT.......................-AGGCAAGAGACAGCAGCATAAACAGGAACT.............................. 417
H2G.CI.92.ABT96              ........................................................................................................ 0
H2U.FR.96.12034              ........................................................................................................ 0
SMM.SL.92.SL92B              ........................................................................................................ 0
SMM.US.x.H9                  ........................................................................................................ 0
SMM.US.x.PGM53               -A--AG-G--TGGT-GTAAG...TCAGGCTTGTCAGGGGAAGAG...GTAAAGAGAAGGCTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACGAGC...---.................................................... 333
SMM.US.x.SIVsmH635F_L3       -A--AG-G--TGGT-GTAAG...TCAGGCTTGTCAGAGGAAGAG...GTAAAGAGAAGGCTAACCGCAAGAGGCATTTATAAAATGGCTGACAAGAGGGAAACAAGC...---.................................................... 407
                                 TCF-1 alpha
MAC.US.x.239                 -A--AG-G--TGGA-G-AAG...TCAGGCCTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTCGC...---.................................................... 407
MAC.US.x.251_1A11            -A--AG-G--TGGA-G-AAG...TCAGGCCTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTAGA...---.................................................... 407
MAC.US.x.251_BK28            -A--AG-G.-TGGA-G-AAG...TCAGGCCTGTCAGAGGAAGAG...GTTAGAAGAAGGCTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAGGGAAACTCGC...---.................................................... 406
MAC.US.x.EMBL_3              ........................................................................................................ 0
STM.US.x.STM                 ................................................GAGAAGGCTAACCGCAAGAGGCCTTCTAAAGATGGCTGACAAGAAGGAAACAAGC...---.................................................... 80
MNE.US.x.MNE027              ........................................................................................................ 0
DEB.CM.99.CM40               ........................................................................................................ 0
DEB.CM.99.CM5                ........................................................................................................ 0
DEN.CD.x.CD1                 ........................................................................................................ 0
LST.CD.88.447                ........................................................................................................ 0
LST.CD.88.485                ........................................................................................................ 0
LST.CD.88.524                ........................................................................................................ 0
LST.KE.x.lho7                TA-CCCT-AAA---C-GCCTGCGAAGATACAAAATACT......GCTTTCGCTTTTGACTGCAAGAGGAAG.......................................-AG-CAAACCGCAAACCGCAGCACATCCTCTTGCAGCCCGGTTGCTAA............ 401
SUN.GA.98.L14                G--CCATACAG---C--CCTGAAGAGTTCAAGCATGTT.........ACTAGTCTGCAGTGGGAGGTG..........................................-AG-TTAACCGCAAACCGCATCCTCTTGCATCGCCT........................ 383
MND-1.GA.x.MNDGB1            ........................................................................................................ 0
OLC.CI.97.97CI12             TTA----AAGG...-CATGTTCAGACTTGCATAGGATG............TGGTTTGTGTGGTTAAGC..........................................-AGA-GACTTGGTTGCCTGGGAGACGGGGGTTGG....ACATGGGCGGGCT......... 318
WRC.CI.97.97CI14             TCCCTGGACAGTCA-T-GCCACTGGGAAAGAGGCCATTGAG...AGTCTCTATAAGAACAAGAAGCAGTGCAAG...................................-AGAGACATGTTGCAGTAGGAAGTGGCTTTGCGGTTAAGCTACAT.............. 422
WRC.CI.98.98CI04             TCCCTGGACAT-TA-TAGCCCAAGGAGAAGATGCTATTGAGACCCTCCATAAAGAGAAGCTGAAGCGGCACAAG...................................-AGAGACATGTTGCAGTAGGATGTGGTTTTGTGGTTACGCTACAT.............. 425
WRC.GM.05.Pbt_05GM_X02       ........................................................................................................ 0
MND-2.CM.98.CM16             ........................................................................................................ 0
MND-2.GA.x.M14               ........................................................................................................ 0
MND-2.x.x.5440               ........................................................................................................ 0
DRL.x.x.FAO                  ........................................................................................................ 0
MON.CM.99.L1                 ........................................................................................................ 0
MON.NG.x.NG1                 ........................................................................................................ 0
GSN.CM.99.CN166              ........................................................................................................ 0
GSN.CM.99.CN71               ........................................................................................................ 0
TAL.CM.00.266                ........................................................................................................ 0
TAL.CM.01.8023               ........................................................................................................ 0
MUS-1.CM.01.1085             ........................................................................................................ 0
MUS-1.CM.01.CM1239           ........................................................................................................ 0
MUS-2.CM.01.CM1246           ........................................................................................................ 0
MUS-2.CM.01.CM2500           ........................................................................................................ 0
SYK.KE.x.KE51                ........................................................................................................ 0
SYK.KE.x.SYK173              GCTG-G-GA-GGCCCG-TTGCAGTTAGAGAGACAAACA.........GGAAAACCACAAGAGCTGCAGTCAGCACTTAGTAAGAACATCTCC..................-AGG-GACTCCATGGTGACAAGCTCGG....................... 289
RCM.GA.x.GAB1                ........................................................................................................ 0
RCM.NG.x.NG411               ........................................................................................................ 0
GRV.ET.x.GRI_677             GCCC-......G-C-TGTTTAAGGACATGCACGAACAT.........GCAAAGCGC.......................................................-AGTGTCAG................................................ 341
SAB.SN.x.SAB1C               TT-CT-GGAGGCT-C--CCAGAGCAGGTGCCGTCAGCA.........CAAGGC..........................................................-A--CACAAAACCACATCCTATG.................................... 352
TAN.UG.x.TAN1                ACCC-......G-C-G-TTTAAGGACATGCATGGTCTT.........GTTAAGAGGAAG....................................................-AG-CTAACCGCAGGCTTGCGGTTAAGCTAACTAGGCCGTTGCCTAGGAGATGACATAAA 386
VER.DE.x.AGM3                ........................................................................................................ 0
VER.KE.x.9063                ACCCAA--AGAC--T-TTTTGAAGACATGCATGCAATA.........GTGAAGAGGAAG....................................................-AG-TAACCGCAGGCTTGTGGTTAAG............................... 367
VER.KE.x.AGM155              ACCCC-CAAGGG--T--TTTACAGACTTATATTCAACA.........GTTGGTACAGGAAAC.................................................-AG-CGACCACAGGCTTGCGGTTTCC............................... 370
VER.KE.x.TYO1_patent         ACCCCA-TCGGG--T--TTTACTGACATGCATGGGCTG.........GTGAAGAGGAAG....................................................-AG-CAGACCGCAAGCCTGCGGTTAGA............................... 368
COL.CM.x.CGU1                ........................................................................................................ 0
```

```
                                                 NF-κ-B-II                          NF-κ-B-I          Sp1-III         Sp1-II   Sp1-I                        TATA Box
H1B.FR.83.HXB2              ....................TCGAGCTTGCT...ACA...AGGGACTTTCCG....................CTGGGGACT...TTCCAG.............GG.AGGCGTGG.CCTGGGCGGGACTGGGGAGTGGCGA.......GCCCT..CAGATCCTGCATAT 429
H1A1.UG.85.U455             ....................................................................................................................................................................... 0
H1B.US.90.WEAU160           ....................-----T--T--..---..-----------A....................---------..----G-.............--.-------A.T----------A-----------......-----..-----G-------- 429
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
H101_AE.TH.90.CM240         ....................................................................................................................................................................... 0
H102_AG.NG.x.IBNG           ....................................................................................................................................................................... 0
H103_AB.RU.97.KAL153_2      ....................................................................................................................................................................... 0
H104_cpx.CY.94.CY032        ....................................................................................................................................................................... 0
H1N.CM.95.YBF30             ....................................................................................................................................................................... 0
H1O.BE.87.ANT70             ACCTGAAGATTGCTGACACTG-..................--A-------A..GCAAAGACTGCTGACACTGC-------.------..............TG--A-G-ACAGG------TT-G---AGTG-C.......TAA----.-----AG-------- 464
H1O.CM.91.MVP5180           .......................................--------A.......GACTGCTGACACTGC-------.------.............C-TG--A-G-ATAAG------TT-G---AGTG-C.......TAA----.-----G-------- 437
H1P.FR.06.RBF168            ....................................................................................................................................................................... 0
CPZ.CM.01.SIVcpzCAM13       ....................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzEK505       ....................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzLB7         ....................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzMB66        ....................................................................................................................................................................... 0
CPZ.CM.05.SIVcpzMT145       ....................................................................................................................................................................... 0
CPZ.CM.-.SIVcpzMB897        ....................................................................................................................................................................... 0
CPZ.CM.-.SIVcpzDP943        ....................................................................................................................................................................... 0
CPZ.CM.98.CAM3              ....................................................................................................................................................................... 0
CPZ.CM.98.CAM5              ........GCGCCTGCGCAG-A--AAC----G..---A...---------TA.A......................------.-----A..............-GG--A-GCTCA..--------.ACA---C----TTTC......A----.----AG-------- 105
CPZ.GA.88.GAB1              TTAG.ACTGGCGCATGCGC.ACAAGAAC----G..--TCTGC----------AA......................------.----G-..............--AGA-GTG-TCGGGA------GT-----T-----TTT..............----.AG-------- 449
CPZ.GA.88.GAB2              ....................................................................................................................................................................... 0
CPZ.US.85.CPZUS             GCTGTAACCGCGCAGGCGC.AATA-AAC----G..-..CTG.A--------TAA....GGGACTTTCCAA......-----G..----.................AAGG--GTG--T-A.------A.ACA---C----TTTT......A----..--A-GA-------A 456
CPZ.CD.90.ANT               ....................................................................................................................................................................... 0
CPZ.TZ.00.TAN1              ....................................................................................................................................................................... 0
CPZ.TZ.2001.TAN2            ....................................................................................................................................................................... 0
CPZ.TZ.2002.TAN3_1          ....................................................................................................................................................................... 0
GOR.CM.2007.SIVgor2139_287  ....................................................................................................................................................................... 0
GOR.CM.2007.SIVgorCP2135con ....................................................................................................................................................................... 0
GOR.CM.2004.SIVgorCP684con  ....................................................................................................................................................................... 0
H2A.DE.x.BEN                AGGCAGGAAGTAGCTA.CTAAGAA-C.AGCTGA..GGCTGC-----------A.......................GAA--GGC...-GTA-C.............CAAG--A-G-AC.A----A--AG----T-G-GAA-...............-CAT-CT-A---T---A 529
H2A.GW.x.ALI                GGGCAGGAAGTAACTACTG..AAA-C.AGCTGA..GACTGC-----------A.......................GAA--GGC...-GTA-C.............CAGG--AAG-AC.A----A--AG----T-G-GAA-...............-CAT-CTC.---T---A 525
H2A.SN.x.ST                 ....................................................................................................................................................................... 0
H2B.CI.x.EHO                AGCTAACACTGCATAGAGAAGGAA-CTAGCTGA..TACTGC---.-------A.......................GAA--GGC...-GTA-C.............A-TG--A-G-AC.A----A--AG-A---AG-.AA-..............-ATCA---CT---T---A 532
H2B.GH.86.D205              AGCTGACACTGCACAA.GAAGGAA-CTAGCAGA..CACTGC-----------A.......................AAA-AGGC..-GTA-C.............A.TG--A-GA-C.AA---A---GT---TTG-GAA-..............-ATTA-ATCT---T---A 535
H2G.CI.92.ABT96             ....................................................................................................................................................................... 0
H2U.FR.96.12034             .........................................................................................................................GAA-...............-CAA-TGAT---T---A 25
SMM.SL.92.SL92B             ....................................................................................................................................................................... 0
SMM.US.x.H9                 ....................................................................................................................................................................... 0
SMM.US.x.PGM53              ..................................GACAGC-----------A.......................CAAA-GGGA..-GTT-T...............--GGA-GTACTG..G--AG-AACTG-TT-GACC.................-ACTT-TT-T---T---A 418
SMM.US.x.SIVsmH635F_L3      ..................................GACAGC-----------A..CAAGGGACTTTCCA......CAA--GGC...-GT--TGGG..GAGGTACT--GGA-GAGCTG.GCT--.AACG......................---ACTT-TTCT---T---A 505

                                                 NF-κ-B-II                                   NF-κ-B-I                                                              TATA Box
MAC.US.x.239                ..................................-ACAGC-----------A.......................CAA--GGA...-GTT-CGGG..GAGGTACT--GGA-GAGCCG.G-C--GAACG.....................---ACTTTCT-GA--T---A 492
MAC.US.x.251_1A11           ..................................GACAGC-----------A.......................CAA--GG....-GTT-CGGG..GAGGTACT--GGA-GAGCCG.G-C--GAAC-.....................---ACTTTCT-GA--T---A 491
MAC.US.x.251_BK28           ..................................GATAGC-----------A.......................CAA--GGA...-GTT-T..............-GGA-GAGCCG.G-C--GAAC-.....................---ACTTTCT-GA--T---A 480
MAC.US.x.EMBL_3             ....................................................................................................................................................................... 0
STM.US.x.STM                ..................................GACAGC-----------A.......................CAA--GGC...-GTA-.............CA--GGA-G..TAC.TG--AG-A-CTG-T--G-.AA-..............-----TTT-CT-T---T---A 165
MNE.US.x.MNE027             ....................................................................................................................................................................... 0
DEB.CM.99.CM40              ....................................................................................................................................................................... 0
DEB.CM.99.CM5               ....................................................................................................................................................................... 0
DEN.CD.x.CD1                ....................................................................................................................................................................... 0
LST.CD.88.447               ....................................................................................................................................................................... 0
LST.CD.88.485               ....................................................................................................................................................................... 0
LST.CD.88.524               ....................................................................................................................................................................... 0
LST.KE.x.lho7               GGCAACCGGGCTAGCGCATGCG-A-TGGCTTGCTG--GAGC-----------A...............GGACGGGC---GGAGG.--T............................................................GCC--CT-G-------- 490
SUN.GA.98.L14               AGGCAACGGGGCTAGC.GCAT......GC--GC..TAGAGT------------......................................................GACAG--AGGG--A-GC-T..............ACC--CT-A-------- 463
MND-1.GA.x.MNDGB1           ...........................................................................................................................................GGAGTCTCTACTACAGA 17
OLC.CI.97.97CI12            A.................................T..CCTT--TTG-----G-TTTTGCCC...........................................................................................T---A 349
WRC.CI.97.97CI14            CCTGAGACTGTGTACCAACGGA-TTAGCAA-CATGCT-ATG-ACTGGGACATTGGGAGGAGACT.GGGAGGTGC--T-T...........................................................GGTTT-G-CCT---A 515
WRC.CI.98.98CI04            CCTGGACCTGTGTACCAACGAA-TTAGCAA-CATGCT-ATGGACTGGGAC-T-G.GTGGAGACTTGGGAGGAGA--T-T...........................................................GGTTT-G-CCT---A 518
WRC.GM.05.Pbt_05GM_X02      ....................................................................................................................................................................... 0
MND-2.CM.98.CM16            .........................................................................................-GTCGGGTGGGGGAGTG-TC-AC-CAACCTGGCT--ATATAA-CA-CTC-CTTTGCGCTTGAAGTGAGTCTC---CTG-G-G 82
MND-2.GA.x.M14              ................................................................................................................................................................G-G-G 5
MND-2.x.x.5440              ....................................................................................................................................................................... 0
DRL.x.x.FAO                 ....................................................................................................................................................................... 0
MON.CM.99.L1                ....................................................................................................................................................................... 0
MON.NG.x.NG1                ....................................................................................................................................................................... 0
GSN.CM.99.CN166             ....................................................................................................................................................................... 0
GSN.CM.99.CN71              ....................................................................................................................................................................... 0
TAL.CM.00.266               ....................................................................................................................................................................... 0
TAL.CM.01.8023              ....................................................................................................................................................................... 0
MUS-1.CM.01.1085            ....................................................................................................................................................................... 0
MUS-1.CM.01.CM1239          ....................................................................................................................................................................... 0
MUS-2.CM.01.CM1246          ....................................................................................................................................................................... 0
MUS-2.CM.01.CM2500          ....................................................................................................................................................................... 0
SYK.KE.x.KE51               ....................................................................................................................................................................... 0
SYK.KE.x.SYK173             CCCAGGGGA................................-----C-GGGC..................GGGG-AA-G--...CA--CT..............-CTGC......................................................---A 335
RCM.GA.x.GAB1               ....................................................................................................................................................................... 0
RCM.NG.x.NG411              ....................................................................................................................................................................... 0
GRV.ET.x.GRI_677            ...............................-ACTTTGCGGTT-----------.CCAGGGACTTTCCACA.....-T--GTGG................ATCG-AGGC-GTACAGGGGC--TACTGGGA-T-GC-T.............T-CCCTCAGAG-------A 444
SAB.SN.x.SAB1C              GAGTTGTCATGGTGATGACAT-AAGAAC----GAC..TGAG-----------A...........................GGG...-GGAGA.............CTGG-CG--ACTGGGA-T..........-GC-T..............G-CCCGT-C-G--CA---- 454
TAN.UG.x.TAN1               GAAGTGC........................-GA..TCAGC-----------A...........................................G--TG-TGCGA-.GGC--TACTTGGGA-T-GC-T...............-CCCTCAGAG-------A 464
VER.DE.x.AGM3               ....................................................................................................................................................................... 0
VER.KE.x.9063               CCGTTGCCAGGGAGATGACAT-T--AAC----GAC-G.....----------A....AGGGACTTTCCA.....................................G--CG---CAT-.GGC--T.AC-GGGA-T-GT-T.............TACCCTCAGAG-------A 478
VER.KE.x.AGM155             TGGTTGCCTAGGAGATGACAT-AAGAAC----GAC.......----------A.GCACGGGACTTTCCA.....................................A--CG--ACAT-.GGC--T.AC-GGGA-T-GC-T.............TACCCTCAGAG-------A 482
VER.KE.x.TYO1_patent        ACATCACCATGGAGATGACAT-AA-AAC----GAC.......---------AGCGAAGGGACTTTCCA.....................................A--CG--ACAT-.GGC--T.CC-GGGA-T-GC-T.............TACCCTCAGAA-------A 481
COL.CM.x.CGU1               ....................................................................................................................................................................... 0
```

```
                                                                                                 +1 mRNA start site
                                            TAR element start                                                                                         TAR element end
                       TATA Box                                                             5' LTR U3 end 5' LTR R repeat begin
H1B.FR.83.HXB2              AAGCAGCTGCT..TTTTGCCTGTACT........................................................GGGTCTCTCTGGTTAGACCAGA.TCTGAGCCTGGGAG..CTCTCT..GGCTA.ACTA..GGG.AACCCACTGCTTA.....AGCCT.. 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         ..................................................................................----------T----------G.--.-----C-----..------..-----.G-A-..---.-------------A....-----.. 72
H102_AG.NG.x.IBNG           ....................................................................................................----.--------------..------..-----.G-GG..A--.-------------.....-----.. 54
H103_AB.RU.97.KAL153_2      .......................................................................................................................................................................... 0
H104_cpx.CY.94.CY032        .......................................................................................................................................................................... 0
H1N.CM.95.YBF30             .............--C-CG--TGTACT.......................................................----------T-C-G-------.-TA-----------..------..-----.G---..---.-------------.....-----.. 85
H1O.BE.87.ANT70             -------C---..--C---T-----C........................................................-------GG-TAGAG------G.--------C-----..---C--..---CT.CTAG..CT-.-----G-------.....-CG--.. 560
H1O.CM.91.MVP5180           -----------..--CC--T---.ACC.......................................................------TAG-TAGAG------G.--------C-----..---C--..---CT.CTAG..CT-.-----G-------.....-CG--.. 533
H1P.FR.06.RBF168            .......................................................................................................................................................................... 0
CPZ.CM.01.SIVcpzCAM13       ....................................................................................------.-CAC-G-------.--------C-----..------..-C-AGTG.....A--.-------------.....-----.. 67
CPZ.CM.05.SIVcpzEK505       ....................................................................................--------T-----------.-TA-----------..------..----GCA.....A--.-------------.....-----.. 69
CPZ.CM.05.SIVcpzLB7         ....................................................................................--------T-----------T--------------..------..----GCT.-...A--.-------------A....-----.. 71
CPZ.CM.05.SIVcpzMB66        ....................................................................................--------T-----------T--------------..------..----GCT.-...A--.-------------.....-----.. 70
CPZ.CM.05.SIVcpzMT145       ....................................................................................--------T-G---------.--------------..------..----ACA.....---.-------------.....-----.. 70
CPZ.CM.-.SIVcpzMB897        ...........................................................................................................................G.TCTCTCT-GTTAG-CCA-ATCTGAGC---GGAGCTCTCT-G--AA 46
CPZ.CM.-.SIVcpzDP943        ...........................................................................................................................G.TCTCTCT-GCTAG-CCA-ATT-GAGC-C-GGAGCTCTCT-GTCAG 46
CPZ.CM.98.CAM3              .......................................................................................................................................................................... 0
CPZ.CM.98.CAM5              -------C---..--C---T-----G........................................................T---------T---G-------.TA------------..------..---GCC-GC-..---.-------------.....-----.. 200
CPZ.GA.88.GAB1              -----------..--C----------........................................................--------T-CAC-G-------.T-------------..------..A.-C-GTGA-..---.-------------.....-----.. 544
CPZ.GA.88.GAB2              .......................................................................................................................................................................... 0
CPZ.US.85.CPZUS             -------C---..C-C---T------........................................................--------TC-C-AGA.-----.-AA-----C-----..------..---GTG--....---.-------------.....-----.. 550
CPZ.CD.90.ANT               .......................................................................................................................................................................... 0
CPZ.TZ.00.TAN1              ...............................GCTCTTGCCTAATCTG..........................................................-----.--------------..------..--TAGTGGCTGGCTAG-GA--G-------.....-CG--.. 75
CPZ.TZ.2001.TAN2            ...........................................................................................................................GCTCTT-C---A---GCCA-ATCTGAGC---GGAGCTCTCT-GTAGT 47
CPZ.TZ.2002.TAN3_1          ...........................................................................................................................GCTCTT-C---A---GCCA-ATCTGAGC---GGAGCTCTCT-GTAGT 47
GOR.CM.2007.SIVgor2139_287  ...........................................................................................................................G.TCTCT-TG-C.AG-CCA-ATC-GAGC---GGAGCTCTCT-GTCTT 45
GOR.CM.2007.SIVgorCP2135con ...........................................................................................................................G.TCTCT-TG-C.AG-CCA-ATC-GAGC---GGAGCTCTCT-GTCTT 45
GOR.CM.2004.SIVgorCP684con  .......................................................................................................................................................................... 0
H2A.DE.x.BEN                -T-T-C-C---.TC--GCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATCGAGCCCTGAGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CTG-.A-----A..............CCAGT--TTGGCCG..GTACTGG--AGAC 674
H2A.GW.x.ALI                -T-T-C-C---.GC--GCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGTG-TTGGCCG..GCACTGG--AGAC 670
H2A.SN.x.ST                 ..............................AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGTG-TTGGCCG..GCACTGG--AGAC 119
H2B.CI.x.EHO                -T-T-C-C---.T-G-GCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATCGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAC 677
H2B.GH.86.D205               -T-T-C-C---.TC-CGCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TAGGCCA..GTGTTGG--AGAC 680
H2G.CI.92.ABT96             ................................................................................................................---CT--.A-----A..............CCAGCG-TTGGCCG..GTGCTGG--AGAG 41
H2U.FR.96.12034             -T-T-C-C---CT--.GCAT----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-G..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGT--TTGGCCG..GTACTGG--AGAG 170
SMM.SL.92.SL92B             ..............................GTCGCTCTACGGAGAGGCTGCCAGA.TAGAGCCCCGAGAGGTTCTCT.GCAGCAC-TA.GTA---.--G-CTGGGTGTTCT--ACT-A.A-----A.............ACCAGC-A-AGGCCA..TTGCTGG-TAGAC 119
SMM.US.x.H9                 ...........................C..AGTCACTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CTG-.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG 120
SMM.US.x.PGM53              -TA--A----A.T--CGCT-----T-C...AGTCGCTCTGGGCAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A--C.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG 563
SMM.US.x.SIVsmH635F_L3      -TA--A----A.T--CGCT-----G-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG 650
```

```
                                                     +1 mRNA start site
                                               TAR element start
                       TATA Box        5' LTR U3 end 5' LTR R repeat begin

MAC.US.x.239                -TATCA----A.T--CGCT-----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG 637
MAC.US.x.251_1A11           -TATCA----A.T---GCTA----G-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGT-TTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGG--AGAG 636
MAC.US.x.251_BK28           -TATCA----A.T--CGCT-----T-C...AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCA..GTGCTGG--AGAG 625
MAC.US.x.EMBL_3             ...........................C..AGTCGCTCTGGC.AGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCA..GTGCTGG.-AGAG 118
STM.US.x.STM                -TA--A-C---.TA-AGCT-----T-C...AGTCACTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGAGTGTTC----CT--.A-----A..............CCAGC--TTGGCCA..GTGCTGG-TAGAG 310
MNE.US.x.MNE027             ..............................AGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCT.CCAGCAC-A..G-A---.---G-CTGGGTGTTC----CT--.A-----A..............CCAGC--TTGGCCG..GTGCTGT--AGAG 119
DEB.CM.99.CM40              ......................................................................CGGGAGTCGCGCCTTG-GAGTCCTGGT--GT---GCGGCCTGGGA-CCCTTG.---ACT--GATAG-CT.A-CTCC-G-GG-CTGGG-GCCCTGAG--GG 98
DEB.CM.99.CM5               ......................................................................CGGGAGTCGCGCCTAG-GAGTCCTGGTAGGT---GCGGCCTGGGA-CCCTTGC--GCTA-TGACT--CT.A-CTCC-GTGG-CTGGG-GCCCTGAG--GG 99
DEN.CD.x.CD1                ...........................................................................................G-AC-T-CTTGCGAGTCTCT--A-C-GCCTGGGAGCCCT-AG-GG-CTTA-C...-G-GG-CTGGG-GCCCTCTAAGGT 76
LST.CD.88.447               .......................................................................................................................................................................... 0
LST.CD.88.485               .......................................................................................................................................................................... 0
LST.CD.88.524               .......................................................................................................................................................................... 0
LST.KE.x.lho7               -----------..C-CC--T---T-GG...GAG.......................................TCCTT.GCTCT--GGA-CAGCAG-C-CT-CTGGGT.----ACACTG-.T-AGCTA.A-CAGTC-GCCCT---AGG-TACACTTA..GCACCCA---GG 610
SUN.GA.98.L14               -----------.C--CGC.----T-GGG..AGTCTCTGCTCAGAGG....CTAGGCGACGACCCT.......................GAGA-G-CTCA-CT.AGAGCT-G--T-TC-.CTC--GA...............GT-AGC--TTG---A..GCACCTCA--CGA 584
MND-1.GA.x.MNDGB1           GG...CTAAGGGT-G-AT-TCTG-G.CAGATCCCCTTAGAGCAAGGACCAGAGTCCT.............GAGTGACTGGGTCT-AGCAC--CA.CTCGGGGCTGATCACCT-GA--TAGTGGAA--CCT..-GCTTG................................ 135
OLC.CI.97.97CI12            ---TCTG----..C-GC--T---TAGA...GGTCT.TGCTCGGAGA..CCAGCC..................................................T-GA-------A-T-CTT--GGCAGCCTGCTTTG-GC-CCTGGGTGTGCTTGGCAGGTGGTTAAAG 461
WRC.CI.97.97CI14            ---T--A----..C-G-AG--A-CT......................................GGTTGCAGTTTTCTGCCTGA--CTAGGTGCC-GC-TAGT-GTC--AC-TGGCCT--CA----GAGT-C--GAGTGGTG-TCCC--TTGGCT-AG-GACTC--GTAGG 645
WRC.CI.98.98CI04            ---T--A----..C-G-AG--A-C.......................................CTTTGCAGTCTTCTGCCTGA--CTAGGTGCC-GC-TAGT-GTC--AC-TGGCCT--CT----GAGT-C--GAGTGGTG-TCCC--TTGGCT-AG-GACTC--GTAGG 647
WRC.GM.05.Pbt_05GM_X02      .......................................................................................................................................................................... 0
MND-2.CM.98.CM16            GCTACCAGAT-.GAGCCTGGGTGTTCTCTGGTGAGTCTT.GGAAGGAGTGCCTGCTTGTAGCCCTGGGTGGTTCGCAGGCCCCT-G--TGTA-C-CT-GG---CTCG-C--GTGTTCCTGAAGAGGCTCAT-GGGGGCTCCCTTGCTTGGT--TGG--GACCTCTAGCAG 250
MND-2.GA.x.M14              GCTACCAGAT-GAGCC--GG-TGTTCTCTGGTGAGTCCT.TGAAGGAGTGCCTGCTTGTAGCCCTGGGCGGTTCGCAGGCCCCT-G--TGTA-C-CT-GGT--CTCG-C--GTGTTCT-GAAAG.G-CTC----AGTG-GCACTCTTG-TTGG..--TTGGTAGA---CT 171
MND-2.x.x.5440              .......................................................................................................................................................................... 0
DRL.x.x.FAO                 .......................................................................................................................................................................... 0
MON.CM.99.L1                ........................................................................................................AG-G-CT-G-CTCCA.CAG-GGC.CT-GG-GC-CTTG-A-CT-GGT-AA--GGCCTGGG----CTT 64
MON.NG.x.NG1                .......................................................................................................................................................................... 0
GSN.CM.99.CN166             ..........................................................................................................TGCTT-G-CTTTCACAG-GGC................CTGGGAG-C--TG..GAGCTAG-TAAAG 47
GSN.CM.99.CN71              ..........................................................................................................TGCCT--GTCTCCACAG-GGC................CTGGGAG-C--TG..GAGCTAG-TAGAG 47
TAL.CM.00.266               ...................................................................................................................T-TCTG-AGAG-GGTCAC-GCG-CCT---AG---TTGGC-ACCGGTAA---G-CC 55
TAL.CM.01.8023              .......................................................................................................................................................................... 0
MUS-1.CM.01.1085            ..........................................................................................................AGT-CTAGCTCC-.CAG-GGC................CTGGGAG-C--CG..GAACTAG-TA-AG 46
MUS-1.CM.01.CM1239          .........................................................................................................CAGTTCTAGCTCC-.CAG-GGC................CTGGGAG-C--CG..GAACTAG-TAAAG 47
MUS-2.CM.01.CM1246          ..........................................................................................................TCTTCTG-CTCC-.CAG-GGC................CTGGGGG-C--CA..GAACTAG-TAGAG 46
MUS-2.CM.01.CM2500          ......................................................................................................-TC-TCTCCTG-CTCC-.CAG-GGC................CTGGGAG-C--CG..GAGCTG--TAGAG 50
SYK.KE.x.KE51               .........................CGA..AGTTGCAGCTAGCTCTCCTCAGCGGCCTGGGAGCCC.TAGGAGAC..............................................TGG-AG.A-CGGCCTGGGAGCCCTTG--AGT.....-AAAGGCTAG-AA 90
SYK.KE.x.SYK173             ----C-A----..C-G-AG--A-C-GCGG.AGTTGCAACTAGCTCTCCAGCGGCCTGGGAGCCCTGAGAGAC.................................................TGG-AG.A-CGGCCCGGGAGCCCTTG--AGT......AAAGGCTAG-GA 448
RCM.GA.x.GAB1               .......................................................................................................................................................................... 0
RCM.NG.x.NG411              .......................................................................................................................................................................... 0
GRV.ET.x.GRI_677            ------A----.CGC-G--T----ACTC..AGTCTCTTACTAGGAG....ACCAGCTAGAGCC.TGGGT.GTTCGC.....................................---TTA.G-CTAAC.CC-G-.TGGCCA.CC-GGGGTA-GGA--.C............ 551
SAB.SN.x.SAB1C              ------------.--GC-------AAAC..AGTCTCTGCTTGGAG.G.CTGCCAGATTGAGCC.TGGGT.GTTCTCTGGTAA-TC-GAAC-A-C--.--G-CTGGGTGTTCG-----GAAGGCAAGC.CTAAG-GCACTCT--TG-GGT-TTG-TA..GACCCTCG-CTG 612
TAN.UG.x.TAN1               ------A----.CGC-G--T----ACTC..AGTCTCTTACTAGGAGA..........................................C-A-C--.--G-CTGGGAGTTCG----TGAGC--AA-C.T--TC.TGGCCATCCAGGGGTA-GA--C-.C........... 574
VER.DE.x.AGM3               ...........................C..AGTCTCTTACTAGGAGA..........................................C-A-C--.--G-CTGGGTGTTCG----TTA.G-CTAAC.CT-G-.TGGCCA.CCAGGGGTA-GGA--..C........... 81
VER.KE.x.9063               ------A----.CGC-G--T----ACTC..AGTCTCTTACTAGGAGA..........................................C-A-C--.--G-CTGGGTGTTCG----TTA.G-CTAAC.CT-G-.TGGCCA.CCAGGGGTA-GGA--..C........... 585
VER.KE.x.AGM155             ------A----.CGC-G--T----ACTC..AGTCTCTTACTAGGAGA..........................................C-A-C--.--G-CTGGGTGTTCG----TTA.G-CTAAC.CT-G-.TGGCCA.CCAGGGGTA-GGA--..C........... 589
VER.KE.x.TYO1_patent        ------A----CGC-GGC.T----ACTC..AGTCTCTTACTAGGAGA..........................................C-A-C--.--G-CTGGGTGTTCG----TTA.G-CTAAC.CT-G-.TGGCCA.CCAGGGGTA-GGA--..C........... 588
COL.CM.x.CGU1               ...............................TGGAGTTCTGACTCTT............................................G-CT-GCT-CG-CTGAG-CCT-AG---CTTGC-AGCTA.TCAGCCT.TGCCT-A-TCTGAGTAGCT-G....CTTGG-TG 90
```

```
                                                                                                    Extensive secondary structure
                                                                  Poly-A signal      5' LTR R repeat end  5' LTR U5 start
H1B.FR.83.HXB2                 ..........................................................C.AATAAAGC..TTGCCTTG..AGT....GCTTCA.AGTA...GTGTGTGCCCGTCTG.TT.GT.........GT  576
H1A1.UG.85.U455                ..........................................................-.--------.--------.---....----G-.----..----------------.--AT-..........--   46
H1B.US.90.WEAU160              ..........................................................-.--------.--------.---....------.----..----------------.--.--..........--  576
H1C.ET.86.ETH2220              ......................................................................................................................................    0
H1D.CD.84.84ZR085              ..........................................................-.--------.--------.---....----T-.----..----------------.--TT..........--   90
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            ..........................................................-.--------.--------.---....----A-.---G..---------------.-G.T-.........AG  123
H102_AG.NG.x.IBNG              ..........................................................-.--------.--------.---....------.----..---------------.A-T--..........--  106
H103_AB.RU.97.KAL153_2         ......................................................................................................................................    0
H104_cpx.CY.94.CY032           ......................................................................................................................................    0
H1N.CM.95.YBF30                ..........................................................-.--------.--------.---....---AA-..--G..----------A--CA.--C-..........--  135
H1O.BE.87.ANT70                ..........................................................-.--------.--------.---....-AGAA...-C-..---------T-A----.--C...........  606
H1O.CM.91.MVP5180              ..........................................................-.--------.--------.---....-AGAA...-C-..---------T-A----.--C...........  579
H1P.FR.06.RBF168               ..........................................................-.--------.--------.---....-AG-A-A.-C-..---------T-A----.--C.........AG   39
CPZ.CM.01.SIVcpzCAM13          ..........................................................-.--------.--------.---....------.----..----------TA----.--CA..........--  118
CPZ.CM.05.SIVcpzEK505          ..........................................................-.--------.--------.---....------.----..-----------A--CA.--C-..........--  119
CPZ.CM.05.SIVcpzLB7            ..........................................................-.--------.--------.---....---G--.----..-----------A-TCA.C-TT..........--  122
CPZ.CM.05.SIVcpzMB66           ..........................................................-.--------.--------.---....-AC-T-.----..---------T-A--CA.--GT..........--  121
CPZ.CM.05.SIVcpzMT145          ..........................................................-.--------.--------.---....---G--.----..---------------A.--CA..........--  121
CPZ.CM.-.SIVcpzMB897           TAGGA...GAACCCACTGCTTAAGCCT...................................-.--------.------G..---...AAC-TG.--C-..-----------A--CA.A-TT..........--  120
CPZ.CM.-.SIVcpzDP943           TAGA....GAACCCACTGCTTAAGCCT...................................-.--------.--------.---....-AC-TG.--C-..----------TA--CA.--CA..........--  120
CPZ.CM.98.CAM3                 ......................................................................................................................................    0
CPZ.CM.98.CAM5                 ..........................................................-.--------.--------.---....-AC-TG.--C-..----------TA--CA.--CA..........--  251
CPZ.GA.88.GAB1                 ..........................................................-.--------.--------.---...GTA--TG.--C-..----------TA---A.--CA.........-A  596
CPZ.GA.88.GAB2                 ......................................................................................................................................    0
CPZ.US.85.CPZUS                ..........................................................-.--------.--------.---....----T-.----..----------TA--C.--CA..........--  600
CPZ.CD.90.ANT                  ......................................................................................................................................    0
CPZ.TZ.00.TAN1                 ..........................................................-.--------.C----G...--A...-TG-T..-AC-..-----------A-T-CA-ACCGCGT.....C-  129
CPZ.TZ.2001.TAN2               GGTGGCTGGAGACCGCTGCTTAACGCT...................................-.--------.C-----.-AG--.......-T..-AC-..-----------A-T-C.A-AT........C-C  124
CPZ.TZ.2002.TAN3_1             GGTGGCTGGAGACCGCTGCTTAACGCT...................................-.--------.C-----.-AG--.......-T..-AC-..-----------A-T-C.A-AC........C-C  124
GOR.CM.2007.SIVgor2139_287     AGCA....GAACCCACTGCTTAACGCT...................................-.--------.-----A-TG.---....-AA-T.--C-..---------T-A---A.--CA..........  118
GOR.CM.2007.SIVgorCP2135con    AGCA....GAACCCACTGCTTAACGCT...................................-.--------.-----A-TG.---....-AA-T.--C-..---------T-A---A.--CA..........  118
GOR.CM.2004.SIVgorCP684con     ..........................................................-.--------.--------.---....-AA-T.--C-..---------T-A---A.--CA..........   37
H2A.DE.x.BEN                   ...GGCTC.CACGCTTG.....................CTTG..CT.T....AAAG.......ACCTCTT........-.--------.-.---A.-..TTAGAAGCAAGTT.-AGT..------T---A---.C-CCTAGTCGCCGCC-  767
H2A.GW.x.ALI                   ...GGCTC.CACGCTTG.....................CTTG..CT.T....AAAG.......ACCTCTT........-.--------.-.---AAT...TAGAAGCAGGTT.-AAG..------T---A---.C-CCTAGTCGCCGCC-  762
H2A.SN.x.ST                    ...GGCTC.CACGCTTG.....................CTTG..CT.T...AAAAG.......ACCTCTT........-.--------.-.---AGT...TAGAAGCAAGTT.-AGT..-----CT---A---.C-CCTAGTCGCCGCC-  212
H2B.CI.x.EHO                   ...GGCTC.CACGCTTG.....................CTTG..CT.T....AAAA.......CCCTCTT........-.------A-..-.-----.A..TTAGAAGCAAGAC.-AGT..------T---A---.C-CCTAGTCGCCGCC-  769
H2B.GH.86.D205                 ...GGCTC.CACGCTTG.....................CTTG..CT.T....AAAA.......CCCTCTT........-.--------.-.---A.A..TTAGAAGCAAGT..-AGT..------T---A---.C-CCTAGTCGC.GCC-  769
H2G.CI.92.ABT96                Y..GGCTC.CACGCTTG.....................CTTG..CT.T..ACAAAG.......ACCTCTT........-.--------..-GC-AAAT...TAGAAGTAAG--AGTGT..------T---A---.C-CCTAGTCGCCGCC-  138
H2U.FR.96.12034                T..GACTC.CACGCTTG.....................CTTG..CT.T....AAAA.......CCCTCT.........-.--------.-.---AGT...TAGAAGCAAG-TA-.GT..------T---A---.C-CCTAGTCGCCGCC-  263
SMM.SL.92.SL92B                ...GGCTC.CACGCTTG.....................CTTG..CT.T...............GACTCT.........-.-------G.C-.--AACT...TAGAAGCAAGTC.-AG-..------T-T-A---.C-CCTAGTCGCCGCC-  208
SMM.US.x.H9                    T..RGCTC.CACGCTTG.....................CTTG..CT.T...AAAAG.......ACCTCTT........-.--------.C-.---AGT...TAGAAGTAAG-C.-.GT..------T---A---.C-CCTAGMCGCCGCC-  214
SMM.US.x.PGM53                 C..GGCTC.CACGCTTG.....................CTTG..CT.T....AAAG.......ACCTCTT........-.--------.-.---GCT...TAGAAGTAAG.C.-AGT..------T---A---.C-CCTAGTCGCCGCC-  656
SMM.US.x.SIVsmH635F_L3         T..GGCTC.CACGCTTG.....................CTTG..CT.T....AAAG.......ACCTCTT........-.--------.-.---A-T...TAGAAGTAAG--.-.GT..------T---A---.C-CCTAGTCGCCGCC-  743
                       TAR element end                                                  Poly-A signal      5' LTR R repeat end 5' LTR U5 start

MAC.US.x.239                   T..GACTC.CACGCTTG.....................CTTG..CT.T....AAAG.......CCCTCTT........-.--------.-.---A-T..TTAGAAGTAAG-T.-.GT..------T---A---.C-CCTAGCCGCCGCC-  731
MAC.US.x.251_1A11              T..GACTC.CACGCTTG.....................CTTG..CT.T....AAAG.......CCCTCTT........-.--------.-.---A-T..TTAGAAGTAAG-T.-.GT..------T---A---.C-CCTAGTCGCCGCC-  730
MAC.US.x.251_BK28              T..GGCTC.CACGCTTG.....................CTTG..CT.T....AAAG.......ACCTCTT........-.--------.-.---A-T..TTAGAAGTAAG-C.-.GT..------T---A---.C-CCTAGTCGCCGCC-  719
MAC.US.x.EMBL_3                T..GGCTC.CACGCTTG.....................CTTG..CT.T....AAAG.......ACCTCTT........-.--------.C-.---CATTT.TAGAAGTAAG-C.-.GT..----C-T---A---.C-CCTAGTCGCCGCC-  213
STM.US.x.STM                   T..GGCTC.CACGCTTG.....................CTTG..CT.T....ATAT.......ATCTCTT........-.--------.-.---AAT..TTAGAAGTAAG.C.-AGT..------T---A---.C-CCTAGTCGCCGCC-  404
MNE.US.x.MNE027                T..GGCTC.CACGCTTG.....................CTTG..CT.T....AAAG.......ACCTCTT........-.--------.-.---A-T...TAGAAGTAAG-C.-.GC..------T---A---.C-CCTAGTCGCCGCC-  212
DEB.CM.99.CM40                 GG....................TAGGCACTCCTCGGCACGCTTG..CT.T.....GCTACGCTTCTATAGCCCAG........--------..TG--GTAA...C-GCAA...................------CTG-CCTCATCATTGACTCCAGG-.  196
DEB.CM.99.CM5                  GG...................TTAGGTACTCCTCGGCACGCTTG..CT.T.....GCTGCATTTT.ATTGCCCAG........--------..TG--GTAA...C-GCAA...................------CTG-CCTCATCATTGACTCCAGG-.  197
DEN.CD.x.CD1                   ATTG.................CAACAAGTACCCAGACAAGCTTG..CTCG.....TGCTGCTTTTGATGACCATA........--------..TCA-TTAA...C-GGCA..............TCT---CTTT-CTCCTCATTAGTGGGTTCC.  178
LST.CD.88.447                  ......................................................................................................................................    0
LST.CD.88.485                  ......................................................................................................................................    0
LST.CD.88.524                  ......................................................................................................................................    0
LST.KE.x.lho7                  GCCTGATCA.............GCCTAGGGAGC.TAAGGGCTCCTAG.....CTTG..CT......ACTTTAAA...G....CCTT........-.------AA.G----TCAT..T--TAGAAAG-AC.-AGC...TA----CT-ATAG-.AG-CTCTTTCTCCTTC  727
SUN.GA.98.L14                  G..CCTGA.TCAGCTTG.....................GGGAGCTAGAGGCTC...CTTG..CT.T..TGCAGTACA...AG...CCTT........-.--------..CC------A..GT-TTACTGCAAG.CAA..........----T-----G-TC-C.TCTTTCTCCTTC  696
MND-1.GA.x.MNDGB1              ......................................CTTG..CTAT.....................TGTCTT........-.-------TAAC-TAGAA...TTAGAGCAAG-G-.GTA-........----TAT--AT-G--CGCCTCTCTTCT....  211
OLC.CI.97.97CI12               AGTAAGAACCTCAG........................CTTTG.CT.GTGT.GAGAGA......................-.------......................GCTGAG-AA-AGAAGACTC-C---CT-G-GTGTT--AC-CATCCAGGT-  548
WRC.CI.97.97CI14               AGGCTGTCACTCGCTTG.....................CTTG..CT.A............GG.GTTT.GCAGTT........-------AGTC-TTGACTTAGACAAGCAG-GAGTCAGTAGTC-TAT-TGGGCAAA-TAAG-AACCCTGGTTC  753
WRC.CI.98.98CI04               AGTCTGTC.CTCGCTTG.....................CTTG...TTA............GGAGTTTTGCA.TT........-------AGCC-TTTGCTTG-CAAGCA-TGAGTTT-GTAGTT-CAT-TA-A--ACAG-AG-AACCCTGGTTA  755
WRC.GM.05.Pbt_05GM_X02         ......................................................................................................................................    0
MND-2.CM.98.CM16               TCTCGG.......T.GGCCAAGAGGCTGTGGGATTGACTACCG..........CTTG..CT.T...GCTTT.......TAAGTGCT........-.------AG.C--A--CAAGTTAAGAA.AAGCA-.GTG-........----ACT-G-C-A-.T--GTCTTTGGT....  367
MND-2.GA.x.M14                 AGCAATCTC....TAGGCCAGGAGATTGTGGGTTAGACTACCG..........CTTG..CT.T....GCT.......ATTGATGCT........-.-------G.C--A---GA..TTAGAA.TAGCA-.GTA-........----ACT-GTCCA-.T--GTCCTTGAT....  289
MND-2.x.x.5440                 ......................................................................................................................................    0
DRL.x.x.FAO                    ......................................................................................................................................    0
MON.CM.99.L1                   G.....................CCTAGAGAGTTGGTGAAGCTGACCCCGGCTCGCTTG..CT.T........TAGTA...AGGCCTTACAGCC...-.--------.............................................  130
MON.NG.x.NG1                   ......................................................................................................................................    0
GSN.CM.99.CN166                CGGCCTGGGAGCCCTTT....GCTAG..........AAGACGGTAGAATTTCGCTTG..CT.T...GCCTAAGGCT...TATGACT..........-------..CT-CCAAAG-..-TAA.CGAG-CTGCTC-C....AT-CATTGG-.TC-AGGTCGGGCTTG....  168
GSN.CM.99.CN71                 CGGCCTGGGAGCCCTTGC...CTAGAG...........GACGGTAGATTCTCGGCTTG..CT.T...GCCCTAGGGC...TTTAGTGGCT......-.--------.CT-C-AAAG-..TACAGCAAG-CTGCTC-C....AT-CATTGG---G-GA-C-GGCTTGG....  176
TAL.CM.00.266                  TGGGAGACCTTGCCGGTAAACGCTCAGTAGACATCCCGCTGCTTGCTTCGCGAAGCGCTTTAAACATAGCC..........................-.--------..TG--TTCA.AAGAGCAGCAG-GTGTGCAGCCCA-C---TGGA-C-GCGGCG-GTGCTGGTCCC  196
TAL.CM.01.8023                 ......................................................................................................................................    0
MUS-1.CM.01.1085               CGGCCTGGGAGCCCTTG....CCTAG..........GAGTACTGATTCTCGGCTTG..CT.T..TGCTTCGAAGC...TTATTG.AAGCCCA..-.-------.G...........................................  123
MUS-1.CM.01.CM1239             CGGCCTGAGAGCCCTTG....CCTAG..........GAGAACTGGTTCTCGGCTTG..CT.T..TGCTTCGAAGC...TTATTG..AGCCCA..-.-------TAG-A-AACC.AAG-CAAGTCT.................................  144
MUS-2.CM.01.CM1246             CGGCCTGAGAGCCCTTG....CCTAG........GGGTTGAGACCCCGGCTCGCTTG..CT.T...GCAGAAA...GGCTTTTACATAGCC...-.--------TCC---AAAC.AACC-GCGA-TC-A..................----TCA-.C-GT--GGCGTCTAAG  172
MUS-2.CM.01.CM2500             CGGCCTGGGAGCCCTTG....CTCAG..........AAGGAATAGACCCCGGCTTG..CT.T...GCTTACTTCAGGCTTTATC..AGCC....-.--------.CT---AAC-.AACT-GCGA-TC-A..................----TCA-.CA-T-GCGGCGC....  169
SYK.KE.x.KE51                  G..ACAAGGACTG....ATCCACCTAGCCTCTACAATTCTCCGTTA.......CTTA..CT.T...TACTACGCG..CTGATCAGCC........-.--------T.G-GTTAACT..-CAAGTAA-GAGT...GC...TCTCA-CTGTTCTGCCGGC--CTCCCAGAAG  220
SYK.KE.x.SYK173                G.GTCCGGGTTTG....GCCGCCCGGTCTTCGGCAGCTCTCCGTTG.......CTTG..CT.T...TGCTTACGCT..GATCAGCCCA.......-.--------T.G-G-TAAAC..GCAAGTAA-G--T.GC-C...TCA-C--TT-TG---GG-CACCTCCCGAA-G  581
RCM.GA.x.GAB1                  ......................................................................................................................................    0
RCM.NG.x.NG411                 ......................................................................................................................................    0
GRV.ET.x.GRI_677               ......................................CTTGG.CT.TCATATAGCT..........................-.------C-...---TCGCT.TAGTCGCTA-ATTGGAGTCAA---CTCATT...............GCTGC-C  622
SAB.SN.x.SAB1C                 GCGAC.TGGCCATTGCCAGTAGCAGAGACTCCG.....................CTTG..CT.T...GCTT.GA.......TTGCCT...........-.--------T.C-------A..GT-TACAAG-AA-.GCA-..-----------A-T-.A--CC-CAGGACAACC  728
TAN.UG.x.TAN1                  ......................................CTTG..CT.T...AGAA.........GGCTATTG........-T-------..------GC..TTAAGAGT--ATC.T-AGCAA---CC.T-.AT-GC...GC-CCC.TTCCTC.  654
VER.DE.x.AGM3                  ......................................CTTGG.CT.T...AGAAAG..........................-T------T..C--CG--GC..-T-AGAGCT-CTC.T-AG...TCAA------TCA-.TGAC-CCTCACTCTTG  160
VER.KE.x.9063                  ......................................CTTGG.CT.T...GGAAAG..........................-T------A-.T------GC..-T-AGAGCT-ATC.T-AG...TCAA------TCA-.TGAC-CCTCACTC..A  663
VER.KE.x.AGM155                ......................................CTTGG.CT.T...AGAAAG..........................-T-------...C------GC..-T-AGAGCT-ATC.T-AG...TCAA------TCA-.TAAT-CCTCACTC..-  665
VER.KE.x.TYO1_patent           ......................................CTTGG.CT.T...AGAAAG..........................-T------.-..------GC..-T-AGAGCT-ATC.T-AG...TCAA---T--TCA-.TGAC-CCTCACTCTC-  666
COL.CM.x.CGU1                  AAGGGCTTGAGGGACTC.......CAGTCCCCTCGCGCTAAGGGTAGATAGGTGCTTTGCAT.T................AT.AAGAG........-.-------T.GA-TTTGAA..GC-TCAA--AG-G.TT-G...AGTAC-AGTTTCAC-TCCGCCTTCTC.....  214
```

```
                                  Extensive secondary structure                                    Lys tRNA primer binding site
                                                                                    5' LTR U5 end
H1B.FR.83.HXB2                GACTCT.GGTAACTAGAGATCCCTCAGACCCTTTTAGTCA.GTGTGG.........................AAAA.TCTCTAGCAG....TGGCGCCCGAACAGGG..ACCTGA.............................................  657
H1A1.UG.85.U455               ------.-----------------------AC-A---A-T.----A..........................----.----------...----------------..--TC--............................................  126
H1B.US.90.WEAU160             ------.-----------------------A---------.------.........................----.----------...----------------..---C--............................................  657
H1C.ET.86.ETH2220             ........................................................................---.----------...----------------.G------............................................   36
H1D.CD.84.84ZR085             ------.----------------------AGC-------C.-G----.........................----.----------...----------------..------............................................  171
H1F1.BE.93.VI850              .......................................................................................................--..--T---............................................    8
H1G.SE.93.SE6165              ...................................................................GGCTT.----.----------...----------------..--T---............................................   40
H1H.CF.90.056                 ...........................................................................................................T---............................................    4
H1J.SE.93.SE7887              .........................................................................................................................................................    0
H1K.CM.96.MP535               .........................................................................................................................................................    0
H101_AE.TH.90.CM240           ------.-----------------------AC-C---A-T.-A--A..........................----.-------C---...----------------.C--TC--..........................................  204
H102_AG.NG.x.IBNG             ------.-----------------------AC-C---A-T.----A..........................----.----------...--------------A...-T---...........................................  185
H103_AB.RU.97.KAL153_2        .........................................................................................................................................................    0
H104_cpx.CY.94.CY032          ...........................................................................................................T---............................................    4
H1N.CM.95.YBF30               A----.----C-----------------A-C-AGAC.T.-A---A...........................----.----------...----------------..--T---............................................  215
H1O.BE.87.ANT70               A--C--.---GT---------------T-AC--AGAC.T.-AAGCAG.........................----.----------...----------------..--T---............................................  687
H1O.CM.91.MVP5180             A--C--.---GT---------------T-AC--AGAC.T.-AAGCAG.........................----.----------...----------------..-GCG..............................................  659
H1P.FR.06.RBF168              A-----.---GT---------------G-AC--CTAG-C.-A----..........................----.----------...----------------..------............................................  114
CPZ.CM.01.SIVcpzCAM13         T-----.----T----------------TAAAA--TAGTC.CAG...................G.----.----------...----------------..------............................................  198
CPZ.CM.05.SIVcpzEK505         A-----.----T---------------T.AA--CTA-TGAGTA.............................----.----------...----------------..--T---............................................  198
CPZ.CM.05.SIVcpzLB7           ------.----T---------------A-....-GT-GTCTGCGC-G.........................----.----------...----------------..--T---............................................  202
CPZ.CM.05.SIVcpzMB66          ------.----T---------------TA..AC-GTAGTCTGTGC-GCC.......................----.----------...----------------..--T---............................................  204
CPZ.CM.05.SIVcpzMT145         ------.----T---------------A-.A-C--TATTGA-AGA...........................----.----------...----------------..--T---............................................  201
CPZ.CM.-.SIVcpzMB897          ------.----T---------------TAT-AAGT---TGTGCG-...........................----.----------...----------------..--T---............................................  201
CPZ.CM.-.SIVcpzDP943          ------.--------------------A-.AAA-TAAGTA-GT.-...........................----.----------...----------------..--T---............................................  199
CPZ.CM.98.CAM3                ........................................................................---.----------...----------------..--T---............................................   26
CPZ.CM.98.CAM5                ------.--------------------A....CA-T-AT.AA--A-..................GTG.----.----------...----------------..--T---............................................  331
CPZ.GA.88.GAB1                C-----.--------------------T....-A-A-T-.TA--CA..................AGG..---.----------...----------------..--T---............................................  676
CPZ.GA.88.GAB2                ...........................................................................................................T---............................................    4
CPZ.US.85.CPZUS               ------.--------------------A....C-CT-A-.AA--A...................GGTG.----.----------...----------------..--T---............................................  680
CPZ.CD.90.ANT                 ......................................................................................................---..--T---............................................    9
CPZ.TZ.00.TAN1                -C-....CTGGGG--------------T.......................TTGTA..GTGGCTAAGTA.----.-------C---...----------------..--T---............................................  206
CPZ.TZ.2001.TAN2              -TT-GCCCTGGGG--------------TTTG-AGT-G-TAAGTA............................----.-------C---...----------------..--T---............................................  205
CPZ.TZ.2002.TAN3_1            -T--GCCCTGGGG--------------TTTG-AGT-G-TAAGTA............................----.-------C---...----------------..--T---............................................  205
GOR.CM.2007.SIVgor2139_287    A-----.---GT---------------GTAAA----CTGA-CAGA...........................----.----------...----------------..--T---............................................  199
GOR.CM.2007.SIVgorCP2135con   A-----.---GT--------T------GTAAAG---CTGA-CAG............................----.----------...----------------..--T---............................................  198
GOR.CM.2004.SIVgorCP684con    A-----.---GT---------------GTAAAG---CTGA--AGA...........................----.----------...----------------..--T---............................................  118
H2A.DE.x.BEN                  -GTCA-..TCGGTGTTCATCTGAGT-AC.AAGACCCTGGT.C---TAGGACCCTTCTCGCTTTGGGA..ATCCAAGGCAGG----.--C-------G..T----------------..--T---..........................................  880
H2A.GW.x.ALI                  -GTCA-..TCGGTGTTCACCTG.AGTA--AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACGGAGGCAGG----.--C-------G..T----------------..--T---..........................................  875
H2A.SN.x.ST                   -GTCA-..TCGGTGTTCATCTAAAGTA--AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCAAGGCAGG----.--C-------G..T----------------..--T---..........................................  326
H2B.CI.x.EHO                  -GTCA-..TCGGTGCTCC-CTAG.......AAACCCTGGT.C---TAGGACCCCTTCTGCTTTGGGA..AACCAAGGCAGG----.--C-------A..T----------------..------..........................................  876
H2B.GH.86.D205                -GTCAC..TCGGTGCTCC-CTGA......TAGACCCTGGT.C---TAGGACCCTTCTTGCTTTGGGA..AACCGAAGCGGG----.--C-------A..T---------------.-----G...........................................  877
H2G.CI.92.ABT96               -GTCAC.CTCGGTGTTC-CGTAGC.....AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGRG..AACCAGAGCGGG----.--C-------A..T----------------..--T---..........................................  248
H2U.FR.96.12034               -GTCA-..TCGGAG-TTC-C-TTAT-.....AACCCTGGTAC---TAGGACCCTTCTTGCTTTGGGA..AACCAGAGCAGG----.--C-------G..T----------------..--T---..........................................  373
SMM.SL.92.SL92B               -GTCA-..TCGGGGTCCCGGATAA-....GAG.CCCTGGTAC---CAGAACCCTCTCACTCTTGGGGGCAACTCCTGAGTG----A-----G---T..T----------------..--T---..........................................  321
SMM.US.x.H9                   -GTCA-.CTCGGTACTC--A-ATAAG...AAGSCCCTGGT.C---TAGGACYCTTTCTGCTTTGGGAAAC..CGAGGCAGG----.--C-------G..T----------------..--T---..........................................  326
SMM.US.x.PGM53                -GTCA-.CTCGGTACTC--A-ATAA--..AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGC..AACCGAGGCAGG----.--C-------A..T----------------..--T---..........................................  769
SMM.US.x.SIVsmH635F_L3        -GTCA-.CTCGGTACTC--CA-A-A-AG.AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAGGCAGG----.--C-------G..T----------------..--T---..........................................  857
                                                                                      5' LTR U5 end    Lys tRNA primer binding site

MAC.US.x.239                  -GTCAA.CTCGGTACTCA--AATAAG...AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAAGCAGG----.--C-------A..T-------T--------..--T---..........................................  843
MAC.US.x.251_1A11             -GTCAA.CTCGGTACTC-G-AATAAG...AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAAGCAGG----.--C-------A..T----------------..--T---..........................................  842
MAC.US.x.251_BK28             -GTCAA.CTCGGTACTC-G-AATAAG...AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGAGA..AACCGAAGCAGG----.--C-------A..T----------------..--T---..........................................  830
MAC.US.x.EMBL_3               -GTCAA.CTCGGTACTC..GGTAAT-AG.AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGAAAC..CGAAGCAGG----.--C-------TGAT----------------..--T---..........................................  326
STM.US.x.STM                  -GTCAA.CTCGGTACTC...........AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGC..TACTGAGGCAGA----.--C-------A..T----------------..--T---..........................................  507
MNE.US.x.MNE027               -GTCAA.CTCGGTACTC..AATAAT-AG.AAGACCCTGGT.C---TAGGACCCTTTCTGCTTTGGGA..AACCGAAGCAGG----.--C-------A..T----------------..--T---..........................................  324
DEB.CM.99.CM40                .............................................AATTCGGGTCGAGCGGGCTAGGCTCAAG.CG-G-G.GTC-------..T----------------..--T---..........................................  263
DEB.CM.99.CM5                 .............................................GTCGGGAGCTCAAGCGGCTAGGCTCTGGGCAGG-G.AAC-------..T----------------..--T---..........................................  265
DEN.CD.x.CD1                  ..........................................AGGAGGTTCGGGTTGTGAGCCTTACGCAGGGTGA----.C-C--C----..----------------..--T---..........................................  248
LST.CD.88.447                 .........................................................................................................................................................    0
LST.CD.88.485                 .........................................................................................................................................................    0
LST.CD.88.524                 .........................................................................................................................................................    0
LST.KE.x.lho7                 -C-CTG.-T-C-G-GATCTCT-ACT---GAGA--GGAG-C.T---.......................GATCCGGG-C-.GGG--TC---..----------------..--T---..........................................  813
SUN.GA.98.L14                 T....A.AACCCTGTTC-...........AC-GCTC-T-.ACTA--GAGCAATTACTGGCCGCCTA..GTGATCCGGGGCTG-C.GGC.------..----------------..--A---..........................................  792
MND-1.GA.x.MNDGB1             ..................................AAACCTGT--TGTTCTCATTTAGAGAACAGAAGGACTTCTAGTTA-CCCTAGAAGCCTTTCAG.-------------.-..--T---..........................................  295
OLC.CI.97.97CI12              A...........................................................................--T.--C-A-----..----------------..--T---..........................................  584
WRC.CI.97.97CI14              -CT-GAG-T-G-TCTTCAG.AGGAGCTGAAGACCCC.........................................-A-----T----..----------------..------..........................................  822
WRC.CI.98.98CI04              -CT-GAGCT-GGACTTCAGAAGGAGCTGAAGACCCC.........................................-A-----T----..----------------..--AA--..........................................  825
WRC.GM.05.Pbt_05GM_X02        .........................................................................................................................................................    0
MND-2.CM.98.CM16              ........AG-AACTCT.GGTTA-T-G.AGA-CCCTCAGATC-GT-GCAGAGATTTGATATCTAGTGGC..AGTAGAGAA----.--C-C-----..----------------..--T---..........................................  471
MND-2.GA.x.M14                ........AG-AACTCT.GGTTA-T-G.AGA-CCCTCAGATC-GT-GCAGAGATTTGATATCTAGTGAG..AGTAGAGAA----.----C-----..----------------..--T---..........................................  393
MND-2.x.x.5440                ...........................................................................................................T---..........................................   22
DRL.x.x.FAO                   .....................................................................................----------------..--T---..........................................   22
MON.CM.99.L1                  ...-GCTT-A--GAGACTTG-GAGTCTG----CA-CTGTG.CG-ATCAGGAGCGGGCTTGAGCCGGGGAG..GGGACAGCAG-C-.G-C--TCTCCA..G----------------..--T---..........................................  243
MON.NG.x.NG1                  .........................................................................................................................................................    0
GSN.CM.99.CN166               ..................................................AGTCCG.GGTGGACACC-G-.---CA-C----..----------------..--T---..........................................  219
GSN.CM.99.CN71                ..................................................AGTCGC..GGTGGATACC-G-.---CA-C----..----------------.G-T---..........................................  227
TAL.CM.00.266                 -CGC-CTC-G...................................GTCGAGGGAACTTGGTGGGTGAAGA--TTTA-C-ACCAGC...----------------..--T---GAACACCCAGGTTCTCTGTCCTGAGGGCCAAGCGGCTTCGGGGACC  314
TAL.CM.01.8023                .........................................................................................................................................................    0
MUS-1.CM.01.1085              .........TTG---AC-AAGG--AGT-TGC-CCTATT.CAT-CACACTGAGAGCCTATCTCAGCGTGGTGTGAGTCTGG---.CAC--C----..----------------..--T---..........................................  229
MUS-1.CM.01.CM1239            ...............................GC-CCTGTT.TAT-CACACTGAGAGCCTACCCCGGCGTGGTGTGGTTCCAG---.CAC--C----..----------------..--T---..........................................  229
MUS-2.CM.01.CM1246            AGTAGAC---TG-CG-CACAG--CG-...................ACCTCGCGGG..................GGGA--T-.----CTC---..----------------.G-T---..........................................  248
MUS-2.CM.01.CM2500            CTAGAAGT--...................................ACC.....................GATCCAGGGCT-C-.-AC--CC---..----------------..--T---..........................................  228
SYK.KE.x.KE51                 .--G-AC--G--AC.......................................................................G-G-CTT---..----------------..--T---..........................................  265
SYK.KE.x.SYK173               TG---A.--G-G-GC-.........................................................................-CTT---..----------------..--T---..........................................  624
RCM.GA.x.GAB1                 .....................................................................................----------------..--T---..........................................   22
RCM.NG.x.NG411                .....................................................................................----------------..--T---..........................................   22
GRV.ET.x.GRI_677              CGAG-CTCTAG-GGT--ACCT-TCTTACTG.......................GGTT..CTCCTG..TACCCAGGTGGG-G--ACTC-.-----..----------------..--T---..........................................  710
SAB.SN.x.SAB1C                CT...................................................GGTTACTAAGGA..TCCCTGATAGAA-----.--CT-----..----------------..--T---..........................................  790
TAN.UG.x.TAN1                 .-GGGAACCCTC.GTT-CTGG................................GGTT...TCTTA..T..CCAGGCTGAG---.----C-----..----------------..--T---..........................................  728
VER.DE.x.AGM3                 A--GGG.TAA---.TTCCT-A-TG.............................GGTTCTCTCTCA...ACCCAGGCGAG-G--.A---C-----..----------------..--T---..........................................  241
VER.KE.x.9063                 AG-AGG.--A-C-GTTCCT-A-TG.............................GGTTCTCTCTCT..GACCCAGGCGGG-G--.A---C-----..----------------..--T---..........................................  746
VER.KE.x.AGM155               TGAA-G.--AG-AGTTCCT-A-TG.............................GGTTCTCTCTCA..AACCCAGGCGAG-G--.A---C----..----------------..--T---..........................................  745
VER.KE.x.TYO1_patent          TGAA-G.--A-T-.TTCCT-A-TG.............................GGTTCTCTCTCT..GACCCAGGCGAG-G--.A---C-----..----------------..--T---..........................................  748
COL.CM.x.CGU1                 ..............................-GCCCT....C....AGAGCGGGGCCTTTTTGCTG......TCCGGGAAGG---.GGC--T----..---------G------..--T---..........................................  286
```

```
H1B.FR.83.HXB2                ..........  657
H1A1.UG.85.U455               ..........  126
H1B.US.90.WEAU160             ..........  657
H1C.ET.86.ETH2220             ..........  36
H1D.CD.84.84ZR085             ..........  171
H1F1.BE.93.VI850              ..........  8
H1G.SE.93.SE6165              .....AAGCGAAAGTTAACAGG  57
H1H.CF.90.056                 ..........  4
H1J.SE.93.SE7887              ..........  0
H1K.CM.96.MP535               ..........  0
H101_AE.TH.90.CM240           .....AAGCGAAAGTTAATAGG  221
H102_AG.NG.x.IBNG             .....CGGTAATAGG  195
H103_AB.RU.97.KAL153_2        ..........  0
H104_cpx.CY.94.CY032          .....AAGTGAAAGTTAATAG  20
H1N.CM.95.YBF30               ..........  215
H1O.BE.87.ANT70               ..........  687
H1O.CM.91.MVP5180             ..........  659
H1P.FR.06.RBF168              ..........  114
CPZ.CM.01.SIVcpzCAM13         ..........  198
CPZ.CM.05.SIVcpzEK505         ..........  198
CPZ.CM.05.SIVcpzLB7           ..........  202
CPZ.CM.05.SIVcpzMB66          ..........  204
CPZ.CM.05.SIVcpzMT145         ..........  201
CPZ.CM.-.SIVcpzMB897          ..........  201
CPZ.CM.-.SIVcpzDP943          ..........  199
CPZ.CM.98.CAM3                ..........  26
CPZ.CM.98.CAM5                ..........  331
CPZ.GA.88.GAB1                ..........  676
CPZ.GA.88.GAB2                ..........  4
CPZ.US.85.CPZUS               ..........  680
CPZ.CD.90.ANT                 ..........  9
CPZ.TZ.00.TAN1                ..........  206
CPZ.TZ.2001.TAN2              ..........  205
CPZ.TZ.2002.TAN3_1            ..........  205
GOR.CM.2007.SIVgor2139_287    ..........  199
GOR.CM.2007.SIVgorCP2135con   ..........  198
GOR.CM.2004.SIVgorCP684con    ..........  118
H2A.DE.x.BEN                  .....AGAGGACTGAGAA.G..CCCTG.  899
H2A.GW.x.ALI                  .....AGAAGACTGAGAA.G..TCTAG.  894
H2A.SN.x.ST                   .....AGAAGACTGAGAA.G..CCTTG.  345
H2B.CI.x.EHO                  .....GGAAGAGTGAGAG.T..CTTCG.  895
H2B.GH.86.D205                ......GAAGAGTGAGAG.T..CTTCAA  896
H2G.CI.92.ABT96               .....GAAGGAGTGAGAGT...CCTTT.  267
H2U.FR.96.12034               .......AGAGTGAGAGT...CTTCTG  390
SMM.SL.92.SL92B               .....AGACG...................  326
SMM.US.x.H9                   .....AGGAGAGTGAGAG....CTCCT.  344
SMM.US.x.PGM53                .....AGGAGAGTGAGAG....CTCCT.  787
SMM.US.x.SIVsmH635F_L3        ......AGGAGGTGAGAG....CTCCT.  874

MAC.US.x.239                  .....AGGAGAGTGAGAGA...CTCCT.  862
MAC.US.x.251_1A11             .....AGGAGAGTGAGAGA...CTCCT.  861
MAC.US.x.251_BK28             .....AGGAGAGTGAGAGA...CTCCT.  849
MAC.US.x.EMBL_3               .....AGGAGAGTGAGAAGA.ACTCCT.  347
STM.US.x.STM                  .....AGGAGAGTGAGAGA...CTCCT.  526
MNE.US.x.MNE027               .....AGGAGAGTGAGAGA...CTCCT.  343
DEB.CM.99.CM40                .....GCTACCCAGAGCTCGGCCCGCAGGC  288
DEB.CM.99.CM5                 .....AGCATCCCAGTGTTTGAGCCCGAAGGC  292
DEN.CD.x.CD1                  .....AAGAGAGTCCTGAAACAGGGACTTGAGTACCCAGAAGGCTCAGGCCCGA  297
LST.CD.88.447                 ..........  0
LST.CD.88.485                 ..........  0
LST.CD.88.524                 ..........  0
LST.KE.x.lho7                 .....AGGAGAAGAGGCGCTGGCAAGAACCTCTGGAGGGAGGAAAAGCGCGCAG........AAGAGGCCAGGAAGCTAGGACAGA  886
SUN.GA.98.L14                 .....GGGAGAGAGAGTGACGAAGAGACGTGGAGTGAAACCTTCAGAAGAAGTGAGCAGCTGCTGAGAAGAAGGAAGGTTTCGACCTGGTGGAA  881
MND-1.GA.x.MNDGB1             ..........  295
OLC.CI.97.97CI12              ..........  584
WRC.CI.97.97CI14              .....GGGGCGTCTAGAGTGAAAGAAGACCGGCCGGCGA  856
WRC.CI.98.98CI04              .....GGGGCGACTAGAGTGAAAGAAGACCGGCCGGCGA  859
WRC.GM.05.Pbt_05GM_X02        ..........  0
MND-2.CM.98.CM16              .....CAAGGGTTCGAGTCATTCCAACCTGTAAGGGACGAAGGCGGCAGCCAGCCGGACCGACCCACCCGGCGAAGTGAGTTAACCAAG  555
MND-2.GA.x.M14                .....CAAGGGTTCGAGTCGCTCCGACCTGTAAGGGACAAGCGGCGGCAGCCGGTCGGACAGACCCACCCGACGAAGAGAGTTAATCAAGGAGCCCCGACG  489
MND-2.x.x.5440                .....CAAGGGTTCAGGTCGCACCCACCTGTAAGGGACGAGCGGCGGCAGC.AGTCGGACAGACCCACCCGGCG.AAAGAGTTAATTGAAGAGCCCCGACG  116
DRL.x.x.FAO                   .....AGACAGTCGACCCCACCTGTGAGGGACGAA.GGCGGCAGCCATCCGGACCGACCCACCCGGAAAAGAGAGTGC..TAAGAAGCCCCGACG  109
MON.CM.99.L1                  .....GTAACGACCCAGCGACT  260
MON.NG.x.NG1                  ..........  0
GSN.CM.99.CN166               .....AACAGCCCAGGACTTCAAGCCCGAAGGCGGGACCCAAGGATTGCCCCT  267
GSN.CM.99.CN71                .....AGCATCCCAGTGACTTCAAGCTCGCAGGCGGGACCCAAGGATTGCCCCT  276
TAL.CM.00.266                 GCGGGTTTTGGGTCCCAAAGAAGAGAGAGAAACTCCGAGGGGTCGTGAGAGGACGACAGTTCCGGAGTCGAGCCCCGGATCGGCGAAGACGACGCGCGTCGCAAAGTGAAAGTAAGGTAAGGAGCCGATCGGTCACCCCGGGCGAGTAGCCGCCGTGAGCGAGTGCGGTG  484
TAL.CM.01.8023                ..........  0
MUS-1.CM.01.1085              .....TTAAGCCCGCGAGGAGGTAGGCGGTGTGACGGAGAGAAGG...CCCC.  273
MUS-1.CM.01.CM1239            .....TTAGCCCGCGAGGGGGTAGGCGGTGTGACAGAGAGAAGG...CCCC.  272
MUS-2.CM.01.CM1246            .....CTACATATCAACCCGTTCGGAGGTTCGCGGCCGGCAGTCAACAGGCGGTGGAGCAGTCCCAGAAGGCTCCATCCCGAAGG..CGAGACGCCTGAG  341
MUS-2.CM.01.CM2500            .....CGAACGTCAGGTAGGATCGGTGAGCAGGCAGTAGAG.CCTCCCAGGGACTCTGCGCTGAAGGCGACAGGCGCCGAGAG  305
SYK.KE.x.KE51                 ..........  265
SYK.KE.x.SYK173               ..........  624
RCM.GA.x.GAB1                 .....GAGTGGCT  30
RCM.NG.x.NG411                .....GAGTGACT  30
GRV.ET.x.GRI_677              .....GAAGA  715
SAB.SN.x.SAB1C                .....CTAAAGACGTCCCGGCCGCAGAGTAGACCA  820
TAN.UG.x.TAN1                 ..........  728
VER.DE.x.AGM3                 .....CTTGAGTGA  250
VER.KE.x.9063                 .....GGTCCGAGGACTT..GAGTGA  765
VER.KE.x.AGM155               .....GAGTGA  751
VER.KE.x.TYO1_patent          .....GTGA..GAGTGT  758
COL.CM.x.CGU1                 .....AAGAGGTCCCTGAGGGTCGCGGAAGAAGAAGAGGAGACTGATCGCCTCCTCAGAAGAGAGGTAAGTCGAAAAAGGGGACTTACGGGTGGAGTCAAGGAGGACTCTACGGTCTGCGCAAAAGTGA  410
```

```
                                                                        ⌐Packaging loops begin
H1B.FR.83.HXB2            .......AAGCGAAAGGGAAA......CCAGA....GGAGC....TCTCTCGACGC..AGGACTCGGCTTGCTGAA...GCGCGCAC.GGCAAGAGGCGAGGGGCGGC....GACTGGTGAGTACGCC....AAAAATTTTG.ACTAGCGGAGGC..............  776
H1A1.UG.85.U455           .......---------TT........--------...-A--T...-----------.-----------------G...-T--A---.A-----------A.-----...----------------..TA---TT-----.-------------.......CTAG....  246
H1B.US.90.WEAU160         .......---------.........--------...-----...-----------.-----------------G...-------.------------------...----------------G...------------.-------------.......CTAG....  778
H1C.ET.86.ETH2220         .......---------T--G-.....--------...---A...-----------.------------------..-T--A-T-.------------A.-----...----------------..AATTTTTA----.-------------.......CTAG....  159
H1D.CD.84.84ZR085         .......---------TAG--.....--------...---A...-----------.------------------..A------T-.------------------A--...----------------TG..A---TT-----.-------------.......... 293
H1F1.BE.93.VI850          .......---------TAG--.....--------...-A--A...-----------G.G-----------------..GTGCAA---.------------A.-----...----------------GAATTTTTT----.-------------.......... 131
H1G.SE.93.SE6165          GACTCGA---------TT........--------...-A--T...-----------.-----------------G...-T--A---.A-----------A.-----...----------------...-----T-----.-------------.......... 179
H1H.CF.90.056             .......T--T-----TA---.....--------...-A--A...-----------.------------------...-T--A---.A-----------A.-----...-----------------....-TTTTG---TG-------------.......... 123
H1J.SE.93.SE7887          ...............................T...-----------.------------------...-T--A---.-------------------G...----------------.AAA--T-T-----.-------------.......CTAG....  102
H1K.CM.96.MP535           ...........................................................................................................................................................    0
H101_AE.TH.90.CM240       GACTCGA---------TT........--------...-A--T...-----------.-----------------G...-T--A---.A-----------A.-----...----------------...-------------.-------------.......CTAG....  344
H102_AG.NG.x.IBNG         GACTCGA---------TT........--------...-A--A...-----------A.-----------------G...-T------.A-----------A.-----...----------------...-------------.-------------.......CTAG....  319
H103_AB.RU.97.KAL153_2    ....................................................................................................................................-------............    8
H104_cpx.CY.94.CY032      GACTCGA---------TT........--------...-A--T...-----------.-----------------G...-T--A---.A-----------A.----G...----------------...C---TT-----.-------------.CT..........  143
H1N.CM.95.YBF30           .......--A------TAG--.......--G--...--CTGAA..------------.---------C-T--....-T--A---.A--G---------C-GC-GA.....AGT-----------...--.------A.--G----T--.CC..........  330
H1O.BE.87.ANT70           .......--T-----T---.......--------...---AGAAAAC---C.-----AAC--G--------AGC-G-...-T--A-C-.-CT---------A--AACTC.ACAGAG--------...---------...--G----T--.CCA..........  805
H1O.CM.91.MVP5180         ......A---T-----T-G--.....--------...---AGAAAAC---C.-----AAC--G--------AGC-G-...-T--A-CT.-CT---------A--AACTCACAAGAG--------...---------...--G----T--.CCA..........  780
H1P.FR.06.RBF168          .......-CTT-----T---.......--------...T-TCTGAA.TC---C.-----A.C--G-------CAGC-G-...-T--A-CT.-CTG---------A--A-CTC..AC-GC--------...-----------.TC--T--T--.CTG..........  232
CPZ.CM.01.SIVcpzCAM13     .......---------TA-G-......--G--...--CTGAA..----C------A.C----------T--.....-T--A---.AA--------------G...----------------...--.T------.-------T--.CTA..........  315
CPZ.CM.05.SIVcpzEK505     .......---T-----TA-C-......--G--...--CTGAA..--C-----G---.---------C-T--....-T--A---.A--G----------.----G....GAAA-GTGAGTACG-.....--.-----A.--G----T--..........  310
CPZ.CM.05.SIVcpzLB7       .......--A------TA--.........-T-G..A-C--T...-----------G.G----------------..-T--A---.-----------------A---G....-----------------...CT--C--------------.....TA..........  318
CPZ.CM.05.SIVcpzMB66      .......--A------TA-G-......--------...-A--T...-----------.----------------..-T--A---.-----------------A---G....ACT-GTGAGTACG-....CT--T--------CG-T..............  321
CPZ.CM.05.SIVcpzMT145     ..............---T----GTAAAA--G--....--CTGAA..-----------A.C----------------..-T--A-G-.A------------C-GC-AC.TCA-G------------...--.------.-------T--.CTA..........  322
CPZ.CM.-.SIVcpzMB897      .......--A------T-T--.......--G--...-A--TT...CTCT.-----------------A-----------..-T------CA-----G-----------C-GC.GACTGGT-A-TACG-CT.......--.T-----..............  315
CPZ.CM.-.SIVcpzDP943      .......--T-----T-G--.......--------...--CTGA...CTCT--CGACGCAC--------------.....-T--A---.A-T-----------C-GACTCCAGAGG--AGTGT-TTTT.........CT------.-C-----T--.CTG..........  321
CPZ.CM.98.CAM3            ......A-GTT-----T-G--.......--------...--CTGAC..---TC------ACG-------------.....-T--A---.A--G----------C-AACTC...C-GA-------G.........T-----.-C-----T--.CTG..........  143
CPZ.CM.98.CAM5            ..............---T----GAAGAA------...--CTGAC..---TC------A.C-------------A---..-T--A-G-.A-T-----------C-ACT-C....TGA--------TTT.......T-----.-C-----T--.CTG..........  449
CPZ.GA.88.GAB1            ..............---T----GTAA..--G--....--CTGAA..----C------A.C----------------..-T--A---.A------------C-GC-AC...........-----------...--.T------.-------T--.CTA..........  790
CPZ.GA.88.GAB2            .......---T-----CA---.......--G--GGAA--CTGAC..----C------G.C-------A-------T-..-T--A--ACA--G-----------C.G-TAC..........-------A.......TTTTC---------A-T--.CTA..........  125
CPZ.US.85.CPZUS           .......--A------TAC--.......--------...--CTGAA..A--TC------.G----------C----..-T--A---.A--G-----------C-AACTC...C-GA---------A.......--.-----.-C---A-T--.CTG..........  795
CPZ.CD.90.ANT             ..............--AG........-T-C-.....-A-A....C--G-.------C.----------TC-----C-.......GT.A---GA-------.....A.....AG---------.C-.TTCATTTTG-----.T-GCTA-T-A..GCA..........  116
CPZ.TZ.00.TAN1            ...............-AAGC-......GGGA-....C-C-G...C-C--G---------.-------------CA.......AAT-..A--------------C-GACTC...CGG-----------......A--TTT-G------..T---T--.GCA..........  313
CPZ.TZ.2001.TAN2          .......GGAGC-GG-......................ACGCGA...C-C--.-GACGCA.------------TGAC-.......AAT-.ACA.----------C-GACTCCGGAGG--AGT-C.............--.T------.T-GT---T--.GCA..........  312
CPZ.TZ.2002.TAN3_1        .......G-AGC-GG-......................ACGCGG...C-C--.-GACGCA.------------TGAC-.......AAT-.ACA.----------C-GACTCCGGAGG--AGT-C.............--.T------.T-GT---T--.GCA..........  312
GOR.CM.2007.SIVgor2139_287 .......---T----CCA-............G...TTCTGA...CTC---CGACGCAC--G--------CAGC-G-...-T--A-C-.-CTG---------A--AA-ACTCAC-GAC-GTGAGTACA...........--.-----.-C-----T--.CTG..........  314
GOR.CM.2007.SIVgorCP2135con .......---T----CCA-............G...TTCTGA...CTC---CGACGCAC--G--------CAGC-G-...-T--A-C-.ACTG---------A--AA-ACTCAC-GAC-GTGAGTACA...........--.-----.-C-----T--.CTG..........  313
GOR.CM.2004.SIVgorCP684con .......---T----CCA-............G...TTCTGA...CTC---CGACGCAC--G--------CAGC-G-...-T--A-C-.-CTG---------A--AA-ACTCAC-GAC-GTGAGTA............C-.-----.-C-----T--.CTG..........  233
H2A.DE.x.BEN              GAACTCG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-GAAA-GCGCGG-CCGAG-TA-C..A.AAGGC.G-CGTGT--AGC-GG....-G--AAAGAGG.C-TCCGGGTG-...AG-T-AGTA-........CTACACCAAAAAC 1047
H2A.GW.x.ALI              GAACACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-GAAA-GCGCAG-CCAAG-TA-C..A.AAGGCCG-CGTGT--AGC-GG....-G--.AAGAGG.C-TCCGGGTG-...AG-T-AGTA-........CTACACCAAAA.T 1041
H2A.SN.x.ST               GAACACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-GAAAAGCGCAG-CCGAG-TA-C..A.A-GGC.G-CGTGT--AGC-GG....-G--AAAGAGG.C-TCCGGGTG-...AG-T-AGT--........CTACACCAAATAC  493
H2B.CI.x.EHO              GGACACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-T-AACCA-GACG--GTGCTCCTGTAAAAACGCAG-CC..G-TA-C.....AGGC.--CGTGA--AGC-GG....-GGA-AAGAGG..-TCCGGG-GT-AAAG-T-AGTA-........TCCAAA....... 1034
H2B.GH.86.D205            GAGCACA..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-T-AACCA-GACG--GAGCTCCTGTAAAAGCGCAG-CC..G-TA-C.....AGGC.--CGTGA--AGC-GG....-GGA-AAGAGG..-TCCGGG-GT..AAG-T-AGTA-........TCCAAA....... 1033
H2G.CI.92.ABT96           GAACACG..--AG-GT-A-GGCAGTAAGGGCGGCAG-ATC-AACCA-GACG--GTGTTCCT-GAAA-GCGCGG-TC..G-TA-C.....AG-C.--CGTGA--AAC-GG....-G--.AAGAGG.C-TCCGGGTGCGC.AG-T-AGT--........ACCAAAAATTTGT  412
H2U.FR.96.12034           AAACTCG..--TG-GT-AGGGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCTGTAAA-GCGCAG-CT..G-TA-C.....AAGC.--CGTGA--AGC-GG....-G--..AGAGG.T-TCCGGTTGC...AG-T-AGT--........AACAAAAAATAGA  532
SMM.SL.92.SL92B           GACT.CG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-T-AACCG-GACG--G....TTTGA.....CC-G--CC..A-T-G-...T-AGC-G-AGT-CA--TAAGA...-GGCC--T-GGACAGTAAGGTCC-AAAG-T..GT-A.......ATATCCTACCTACG  468
SMM.US.x.H9               GAGTACG..--TG-GT-A-GGCAGCAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-TAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGC-GG....-G-C-GAGAGG.C-TCCGGTTGC...AG-T-AGT--........AACAAAAGAGTCA  488
SMM.US.x.PGM53            GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCT-TAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGC-GG....-G-C-GAGAGG.C-TCCGGTTGC...GG-T-AGT--........AACAAGAAGAG..  929
SMM.US.x.SIVsmH635F_L3    GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GAGCTCCT-GAAA-GCGCGG-CC..G-TA-C.....AGGC.G-CGTGA--AGCAGG...-G-C-GAGAGG.C-TCCGGTTGC...AG-T-AGT--........AACAAAAA..... 1014

MAC.US.x.239              GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-C-AACCA-GACG--GTGCTCCT-TAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG....-GA--AAGAGG.C-TCCGGTTGC...AG-T-AGT--........AACACAAAAAAGA 1006
MAC.US.x.251_1A11         GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-C-AACCA-GACG--GTGCTCCT-TAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG....-GA--AAGAGG.C-TCCGGTTGC...AG-T-AGT--........AACACAAAAAAGA 1005
MAC.US.x.251_BK28         GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-C-AACCA-GACG--GTGCTCCT-TAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG....-GA--A-GAGG.C-TCCGGTTGC...AG-T-AGT--........AACACAAAAAAGA  993
MAC.US.x.EMBL_3           GAGTACGG.C-TG-GT-A-GGCAGTAAGGGCGGCAG-A-C-AACCA-GACG--GTGCTCCT-TAAA-GCGCAG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG....-GA--A-GAGG.C-TCCGGTTGC...AG-T-AGT--........AACACAAAAAAGA  492
STM.US.x.STM              GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-CAAACCA-GACG--GTGCTCCTGTAAA-GCGCGG-C.........AAGC.G-CGTGA--AGC-GG....-GGC-AAGAG-..-TCCGGTTGC...AA-TGAGT--........AACAAGTAGCCAG  663
MNE.US.x.MNE027           GAGTACG..--TG-GT-A-GGCAGTAAGGGCGGCAG-A-C-AACCA-GGCG--GTGCTCCT-GAAA-GCGCGG-TC..G-TA-C.....AG-C.G-CGTGA--AGC-GG....-GA-.AAGAGG.C-TCCGGTTGC...AG-T-AGT--........AACACAAAAAAGA  486
DEB.CM.99.CM40            GGTGTGCGGAGTCTG-AA-CCCTGAGC-G-TCTCTTCCC--CGGA-.AGCT-CGCAGC-AAGG-AA--GGA-GCCGGAGAATACAGG.-C-GC--ATAAG--C-G-TGT....GTG.--CTC-G-AGGCAGAGT-CCAC-GTGGAGCA.......AGAGAGGCTCCTGG  443
DEB.CM.99.CM5             GGAGTGTGGAGTCTG-AA-CCCTGAGC-G-TCTCTTCCC-TCGAG-GAGCT-CGCAGC-TAGG-AA--GGA-GCCGGAGAATACAGG.-C-GC---TG-G-AC-G-TGT....GTG.ACCTC-G-TGGCAGAGTGTCAC-GTGGAGCA.......AGAGG..CTCCTGG  446
DEN.CD.x.CD1              AGGCGGTTG-AACGTACC-CCCTCTTA--GAGCAAG--C-AGGAAGACGCT-CGCAGCGCA-AGGTAAGGATCTCGCCG-A--A---.ACT-GTG-CA-TA-T-T--GG....CTG-AA--TAG-AAGGAGGGCT-TC-GGAA--A-A........GGGCTAGATAAGGG  454
LST.CD.88.447             ...........................................................................................................................................................    0
LST.CD.88.485             ...........................................................................................................................................................    0
LST.CD.88.524             ...........................................................................................................................................................    0
LST.KE.x.lho7             GAAAGTCTGCTCGGG-CAGG-AGAGAAGACCTGCTT-AGA-GAGGATAAGAAGGA-TTGACTGCTA-AAG--G-C-GAA-GTAAGGA.-AGT-AGAC-TCACC-G-AGA....GTG-CA-CAGG-CGGGGGC-GCCT-G-GAG-AGT-........TGTAATTCTGATCT 1043
SUN.GA.98.L14             GGCGCCGGGAA-TT-A--T--GTGAGGG--CTTACCA-GCACTGTCAGGGGACA-GCC--A-AGTCCTAG----GGGGGAG-TC-CG.CTAGGTTA-TACT-TCAA-TA....GGCA-ACTAGTA-TACGTGG--T-AAA-AAGAAGT........GTGTGTATAACTCT 1038
MND-1.GA.x.MNDGB1         .........................AG---GGCACTGA-ACTTGAGGCA--GCACTCC-C--G-AAGAAGCAGGTTGAA-GAG-G.T-G-CTG-T-TGAA-A--C-A....GGA-------CA-TGGGACTG-CT--ACAAGAATT........AGTTGTACCCTAGT  425
OLC.CI.97.97CI12          ..........................................................................................-AAA-G-TTCGAG.T-CCTGA--A-GAAGTTGGTC-T-AGG--AGTG-C........CGGGGAACACGAGG  646
WRC.CI.97.97CI14          GGGATTC-CT------AAGTCGTCC...AGA-.GGTAAGAGGGAGGACT--TGGT-CCT-TTGAGAC--CTAGA-GGAATTAGAGG-.TAGCCTT--AAC-TA-A---TGAGT-G-TCAT-AGGGA-TACTGGG-G-CAGGAG-AAG-ACC---GGGGAGGC....TATA 1017
WRC.CI.98.98CI04          GGGATTC-CT---G--A-GTCAGCCTC-AGA-AGGTAAG-AGGAGGGCT--TGGT-CCT-TTGAGAC--CTAGA-GGACTTAGAGG-.TAGCCTT-AGAC-CATA---TGAGT-G-TCAT-AGGA--TACTGGG-G-CAGGAG-.GTAACC---GGGGAAGC....TATA 1023
WRC.GM.05.Pbt_05GM_X02    ...........................................................................................................................................................    0
MND-2.CM.98.CM16          GAGCCCCG-CGCGC---ACTCAAGGTAAGC-GTGCACCGAACAGT-G-G-GTGT-AGTCCAGGGTCCGC-AGA-C-AGC-A-AC-G-.C-AGGC-AC-CCA-TAG-CAA....-AG-AAA--AGG-GAAGTA-GGCCGAGGCA-AGT-........AAAAGAGGTCCTTG  712
MND-2.GA.x.M14            CGCAGGACTC.A-GGTAAGCGGTGCAC-.......GTATTGTGGTAG-GAGTC-.AGG-TCCGCTT-AGCAGGCG-GATCGCTAAGG.CAACCCCA-TAG-T.AAAAGA....-AA-AG-G-A-GTAAGGCCG--GCAAAGTG-AAGT........AAAAGGGCCTTCCG  636
MND-2.x.x.5440            CGCAGGACTCAA-GGTAAGC-GTGCACTGTA....ATTGATTGAGAGAGAGTC-AGGG-TCCGCTT-AGCAGGCG-GATCGCGAAGG.CAACCCCA-TAG-CAAAA-AA....-AGAAG-G-A-GTAAGGCCG-CGCAAAGTG-AAGT........AAGGGCTTCCTCCG  269
DRL.x.x.FAO               CGAGACCCGC-TCGGTAAGT-AAAAGGAA--TAAAATC--TGCAC-G-GAG-CGAAGGGTCTAGT--GCC-TGTG-GCA---GCT-G.-CA---GCAACTCC-TA---A....-AAA-AA--A-AAAGGGAG--GT-AGGCCTGT-GC........TGATAAGTGAGGCA  266
MON.CM.99.L1              CGAGCCCGCAG-CG-ACAG-TCCGCGAGA--CCTTGCCGAGTGCGC-CGA-A-A--TGC-C-GCAA-AGGTAA-G-GAGATCG-GCG.-T-CCCTA-G-T--CAGACTA....G-GT-G--CGGGAA-CTATCC--GAGGGAAGG-TC........CAGAGGTGCTCCAA  417
MON.NG.x.NG1              ...........................................................................................................................................................    0
GSN.CM.99.CN166           GGCCGAGTG-TCCGG-CTCC-CGAAATTTGCTGCGCA-CC-CTTG...AAGAGAAGGT-A-GAGACCGGACGG-TCTA.-AAG-GTG.-CAG-CTTCTAG-CCAG-TG..GGTTG....A--A-G-GTACCC-C-C-G..A-GGGAGAAACT-TG..............  409
GSN.CM.99.CN71            GGCCGAGTG-TCC.G-CTCC-C..AACTTGCTGCGCA-CA-CTAG...AAGAGAAGGT-A-GAGACCGGACGG-TCTG.-AAG-GTG.-CAG-CTTC-AG-CC-GT-G..GTTGA....---A-G-GTACCC-C-C-GAGG-G-GA-A-T-T-A..............  416
TAL.CM.00.266             CTAGGAAG--T-GC--A--GGCCTAGCTT-ACCCAG--G-GA.GGA-GGA--GA-GACTT-CAGGCAGGG-TCTGC.AG-AAGTTCGT-CGT-CCTC-AGACAAGCA-C....CGGTAGA-AGG..AGACTGCTC-GGAGGAGCAC--........GGACCTGCGGGTGG  638
TAL.CM.01.8023            ...............................CAG--G-GAAGG--G--T--G-GACGC-CAGGCAGGG-TCTGCGGG-AAGT--G.C-AC-CTTC-AGACAAGCA-C....CGGTAGA-AGGT-A-AGCTC-G-GGAGGAGCTC--........GGACCTGCGGGTGG  124
MUS-1.CM.01.1085          ..AGCCG..AGTGCGT-CC--G..CCA--G--GAACAAGTGCTGCG-AGC-ATAAGGT---TGA-TTGGACGG-TCTGG-AAG-GTG.-CAG-CTTC-TCA-AT--AGCA...GGCATA---AGAT--TGCGGG-GGGGAGAA---C.....................  410
MUS-1.CM.01.CM1239        ..AGCCG..AGTGCGT-CC--G.ACCA--G--GGACAAGTGCTGCG-AGC-A-GAGGT---TGG-TTGGACGG-TCTGG-AAG-GTG.-CAG-CTTC-TCA-AT--AGC....-GGTAC---A-AC--TGCGTG-G-GGAGAA--CC.....................  410
MUS-2.CM.01.CM1246        GATCGCCT---CG-GT-..GCCCGGCA-GGAGAAAAAAGAGCTGCG-AGCAA-A-AGGTAAGG-GTCGGGCGA-CCTA.-A-G-GTG.-CAG-CCTCTAG-T--GC-GGT....-GAC-A-GAG.GT--CCGCG-GG-GGAGAA--CGCT....................  482
MUS-2.CM.01.CM2500        AACGCCCC---CG-GT-...-GCACCGA--CGAGGAAATCGCTGCAGC.GT--A-AGGTAAGG-GTCGGGCGA-TCTA.-A-G-GTG.-CAG-CCTCTAG-C--GC-GT....TCA-A--AG.G-A-CGCG-G-G-GGAGAA--.....................  438
SYK.KE.x.KE51             ..............CT--GCTAGGGA--C--GAGA----ACGCCAG--AAAGG-ACCGCAGAGGAA-G---GC-CGAA-AAGAG-A.-CGGT-TCCGCCCTATGCCGA....T-G-CGACCGGATTGAGTAGGTG-G--CTCCGGC-........TCAGACTGCACGCG  407
SYK.KE.x.SYK173           ...........CT-G-CC-GGGACCCGAG--GAGGT-AC--CAGTCAAAGG--GT-CGGATC-AA-AGGGATC-GCTGTC--TC-TA.---CGACC-GTGCT--T---T....CG-----TAC--A-GAGACG-GGTAAG--AT-C--........CGCCAGACTGCCCA  770
RCM.GA.x.GAB1             GAGAGACCTC---GGCTA-GGTTCGGCG-G-CAGGTCACCGCGGGAG-GGAAC-TTGACCAGG-AA-AGCTGCCTGGTG--TTCA-A.-TGCTAGA-AA--T-AGCTAG....CGAA-GA--C-G-G-AACCCGGTCCGG-ACTGGGC........CTCTAGGGAAGGAG  187
RCM.NG.x.NG411            GAGAGACCTC---CGC--GTCCGGACCAG-AGAGTGACT-GTAGAA-G-GGAC-CAGTCCAGG-AA-AGCAGCCTGGTG--T--T-A.AT-T--C-AGA-AA-TGA-GG....CTA-TA---C-A-G-GACCCGGTCAGG-ACTGGGC........CTCTAGGGAAGGGT  187
GRV.ET.x.GRI_677          GGCATCGGCA-CG-CC-CTG-GTTGCTGAGCGTCGGA---GGACGA---AG-TA-GGTGA--GC-TA-GA-T-TTTTGCTAC-TAGT.CAGCGAGAAA-GCTA-GCCGC....GACA-G-GCGCG-GTCCCATT-GTGGCAACCAACC........CAGTTGGACGAAGG  872
SAB.SN.x.SAB1C            GCCTGAGG-CTCCTTT-ACTGAGTAAGA-C-GACGGAAGAAAGGCGGA-CGTG-CGGAGT-CAGGTAAGA-GC-CCTGC-GA-C-GA.-A--CCCTCG-GCT-TGT-TG....TG-....GAG-A-AGAGGA----GAGAA-TCAG-C........AGGGCTGACAAGGC  973
TAN.UG.x.TAN1             .......GTCA-----ACTTCACTGAAGGG--GGAA--GCATCACCAGTCGAGTT-CCC----G---TAA-TC-GGACGCG-G-AGG.C-A-GAG-ATCC-CCAG-AAA....-AA-TACT-CCA-GGAGGC-GGGG-GG--C-GGTA........GATCTAGGTGGGCG  878
VER.DE.x.AGM3             AGGCACGT-CA-CTGA-A-G-CGTCGGA-GCGAAGGAAG--GCGGGG-GCGACGTGAC.CA-GAA--GC-TGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....G-CTAGTTAG-G-A-TAGG--GGGCCG-AGC-ATA........ACTACTCTGGGCAA  406
VER.KE.x.9063             AGGCACGT-CA-CTGA-A-G-CGTCGGA-GCGTAGGAACCGCGGGGTG-GACGT-ACC..A-GAA--GC-TGGTG-GTG-GCTT-T-.-AGTGCC-.AA--AA--TCGA....G-CTAGTTAG-G-A-TAGG--GGGCC.-AGC-ATA........ACTACTCTGGGCAA  918
VER.KE.x.AGM155           AGGCACGT-CA-CTGA-A-G-CGTCGGA-GCGAAGGAACCGCGGGGTG-GACGT-ACC..--GAA--GC-CGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....G-CTAGTTAG-G-A-TAGG--GGGCCG-AGC-GTA........ACTACTCTGGGCAA  906
VER.KE.x.TYO1_patent      AGGCACGT-CA-CTGA-A-GGCGTCGGA-GCGAAGGAAGCGCGGGGTG-GACG--ACC-A--AGGA-AC-TGGTG-GTA-GCTT-T-.-AGTGCC--GA-AAA--TCGA....G-CTAGTTAG-G-A-TAGG-G-GGCCG-AGC-GTA........ACTACTCTGGGCAA  915
COL.CM.x.CGU1             AAGTAAA-GTGCGCGCTCGTCGCGCAGT---GAAGT--CAGACGAGAGAG-ACA-AGAG---G-GCT-GCCAGAG-AGGAGAA-AG-.TCATTAG-AG-TCTTCA---A....GGGA--AGCGC-C-TGGGC-GGTCGAGG-GGGCCT........GCAAGTGGAGTAAG  567
```

```
                                                                      Packaging loops end   Gag and Gag-Pol start
H1B.FR.83.HXB2             .............................................TAGAAGGAGAGAG.....ATGGGTGCGAGAGCGTCAGTA...TTAAGCGGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGT  893
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__G__K__K__K__Y__K__L__K__H__I__V
H1A1.UG.85.U455            ..............................................-AGGA----A.......------------------..--------AA-A--------TC-----G---------------------------C---------G-C-G------T----  360
H1B.US.90.WEAU160          ..............................................-AGGA-----.......-----------------G---..-------------------A-G---------------------------T-----------C-------------  892
H1C.ET.86.ETH2220          ..............................................-AGGA-----.......------------------A--..-----A--C-A-A-------GCC------------AA--------------------C-C----TGC-G------CC----  273
H1D.CD.84.84ZR085          ..............................................-----------.......------------------..------------A----G---GCT----------------C-------------------G-----GC------------  410
H1F1.BE.93.VI850           ..............................................-----------.......------------------A--..------------A-------GA-------------A--------G---------------G-------A-G------C--A-  248
H1G.SE.93.SE6165           ..............................................------G-----.......------------------..----C-------A-------GCT--------------------G-------------G----TC--------A--------C----  296
H1H.CF.90.056              ..............................................-----------.......------------------..---------C---A-------GCT-----G--------C------------------------GGC---------C----  240
H1J.SE.93.SE7887           ..............................................-AGGA-----.......------------------A--..-----T------A-------GAT-----------------G-----------G------------GGA----G---C----  216
H1K.CM.96.MP535            ..............................................-----------.......------------------..------------A-------GC--------G--------C------------------------C-G-------C----  104
H101_AE.TH.90.CM240        ..............................................-AGGA-----.......------------------..-----T------A-------GC-------------------GC------------GA---------GGC-G------T----  458
H102_AG.NG.x.IBNG          ..............................................-AGGA-----.......------------------..-----T------A-------GC------G-------------------------------------G-C--------C----  433
H103_AB.RU.97.KAL153_2     ..............................................-----------.......------------------..------------A-------GC--------------------------------G---------G-A-------CC----  125
H104_cpx.CY.94.CY032       ................................A..........G-AGGA-----.......------------------..-----T------A-------GC------G-GG------------------------------------G-C-G------C----  259
H1N.CM.95.YBF30            ............................AGAAA.........GTAGGA-----G.....----------------------G..----CA-------A-------A-------TC----TAT--G--A-----------------C-G-A-G------T----  451
H1O.BE.87.ANT70            ........GACCTAGGGGAA.GGGCGAAGTCTC..........---GG-AG--AG.....------------T-----T--G..--G-CA--AA-CA----G---GC-------C-----A------A------ATCT--A--G-------G-C-------CT----  945
H1O.CM.91.MVP5180          ........GACCTAGGGGAA.GGGCGAAGTCCC..........---GG-AG--AG.....------------------T--G..--G-CA---A-TA----G---GC-------CG----A-----------ATCT--A--GGC-----GGC--------T----  920
H1P.FR.06.RBF168           ........ACCCTAGAGGAAGGGGCGAAGTCTC..........---GG-AG-.AG.....----------------------G..--G-CA-----CCG---G---GC-------G-----A--C-T-A------AA-T-------------TGC-------------  372
CPZ.CM.01.SIVcpzCAM13      ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GG-----AG.....----------------------G..----CA-----GA----G-----C--------G-----C-T--A--C------G------------TGA-G------T----  455
CPZ.CM.05.SIVcpzEK505      ...........CCAGGA.........................G-AGGA-G---G.....------------------..----CA--A---A--------A-----------TAT--G--A-----------------C-TGA-G------CT----  432
CPZ.CM.05.SIVcpzLB7        ........GG.CTAG...............................-AGGA-----.......-------------------A-C..----CA--C---A-G-----GAT---------G-----C-------C------GA---G-------TGA-G------T----  439
CPZ.CM.05.SIVcpzMB66       ........GG.CTAG...............................-AGGA-----G.....---------------------T..----CA--------A------CA-------G-----TATC-G------------------------TGA-G------T----  442
CPZ.CM.05.SIVcpzMT145      ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GG-----AG.....-------------------A--..----C-------AG-------GCT------------A--C----A------------------C-TGA-G------T----  462
CPZ.CM.-.SIVcpzMB897       ..............................A.AAGC.......-----------G.....------------------..----CG--A---A-G-----GAT-----------TATC-CC-----------------------TGA-G------T----  438
CPZ.CM.-.SIVcpzDP943       ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GG-A---G.....-------------------A--..----CG-------CG-------GCT-----------A--C-------------------------TGA-G------T----  461
CPZ.CM.98.CAM3             ........GACCTAGGGGAA.GGGCGAAGTCTC..........---GG-A---G.....------------------..----CA-------CG-------GCT-----------A--C-------------------------TGA-G------T----  283
CPZ.CM.98.CAM5             ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GG-A---G.....------------------..----CG--A---CG---G---GCT-----G-----A--C-----M-----A---------------TGA-G------T----  589
CPZ.GA.88.GAB1            ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GG-----AG.....-------------------T..C---CA-------A-----------C--------G-----C-T-----C------GA----G----TGA-G------T----  930
CPZ.GA.88.GAB2            ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GT------G.....-----------------TA--..----C---C--GCG------GCC-----G-----A--C-C--A---------G-----G----C-TGA-G-------------  265
CPZ.US.85.CPZUS           ........GACCTAGGGGAA.GGGCGAAGTCTC..........---GG-A---G.....-------------------T..C-G-CA-------CG-------GCT--------G-----C-T--A--T------------G-----TGA-G------T----  935
CPZ.CD.90.ANT             ........ATCCTAGGGGAA.GGTCGAAGTCTC..........-G-AC--GA-.....----A---G-G-------T--T..--G--G--A-AGA-GC-----AC-------GT--CA--C-TC----C--T--C------G----C-TGA---------C-G--  257
CPZ.TZ.00.TAN1             ........ACCCTAGAGGAA.GGGCGAAGTCTCTAGG......---AC---G--A.....----------------------G..--G--G--A-ATA-GC-T--AC-------TCC--AA--C-T-AAT-CA-A--C-G-------TT-A---------C----  456
CPZ.TZ.2001.TAN2           ........ACCCTAGAGGAA.GGGCGAAGTCTC..........---GTAACAG-G-AA.....-------------------G..--G--A---ATA-GC-T--AC-------TCT--AA--C-T-AA--CA-A--C-G--G-------TT-A---------C----  455
CPZ.TZ.2002.TAN3_1         ........ACCCTAGAGGAA.GGGCGAAGTCTC..........---GTAACAG-G-AA..-------------------G..--G-AA--A-ATA-GC-T--AC-------TCT--AA--C-T-AAT-CA-A--C-G--G-------TT-A---------C----  455
GOR.CM.2007.SIVgor2139_287 ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GG-AG--AG.....-------------------G..--G-CA------CG---G---GC--------CG----A--C-T-A------AA-T--------G-----TGC--------G----  454
GOR.CM.2007.SIVgorCP2135con ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GG-AG--AG.....-------------------..--G-CA------CG---G---GCG------CG-------C-T-A------AA-T--------G-----TGC--------G----  453
GOR.CM.2004.SIVgorCP684con ........GACCTAGAGGAA.GGGCGAAGTCTC..........---GG-AG--AG.....-------------------G..--G-CA------CG---G---GC--------CG----A--C-T-A------AA-T--------G-----TGC--------G----  373
H2A.DE.x.BEN               TGTAGCC.AGAAAAGGCTTGTTATCCTACCTTTAGACA..GGTAGA---TT-T-G---.....-----C------AAC--C--C..--G--A---AA-A--GC---CGA--TA----G---A-----C----C--C-------------G--C-GG----------T--  1206
H2A.GW.x.ALI               TGTAGCC..GAAAGGGCTTGTTATCCTACCTTTAGACA..GGTAGA---TT-T-G---.....-----C------AAC--C--C..--G--A---AA-A--GC---CGA--TA----G---A-----C----C--C-------------------C-GC----------T--  1199
H2A.SN.x.ST                AGTAGCC..AGAAGGGCTTGTTATCCTACCTTTAGACG..GGTAGA---TT-T-G---.....-----C------AAC--C--C..--G--A---AA-A--GC---CGA--TA-----G---A-----C----C--C-----------------GGC----------T--  651
H2B.CI.x.EHO               ...........TTTGCAATTCGTCTACTCTAGGACAGG.GTACGG-AT-A-T-G---.....-----C-------GC--C--C..--GTCA---AAGA-GAC----GA--TA-----GG--A-----C----C--C----------G---C-TGC-------------  1183
H2B.GH.86.D205             ...........TTTTGC.ATTGTCTACTCTAAAGAGGGAGTAGGGCAT-A-T-G---.....-----C-------GC--C--C..C-GTCA---AAGA--AC----GA--TA------G--A-----C----C--C-------------G----TGC--------G----  1182
H2G.CI.92.ABT96            AGCCAAAAA...GGGCTGTGAAT.CCTACCTTGAATA.GTCAGGTAGA-GT-T-G---.....-----C------AGC--C--C..C-GTCA---AA-A--GCT----GA--TA------G--A-----C----C--C----------------C-TG----------G-T--  569
H2U.FR.96.12034            ACCAAA......GGGTTGTC.TGCCTCTGAGAAAGG.AGAGGTAG--T-TT-T-G---.....-----C------AAC--.-C..--GTCA---AAGA--G-----GA--TA------G--A-----C----CAAC------------C-TG----------G-T--  685
SMM.SL.92.SL92B            GGATAGGTAGAAGTAAAGCGGGAG..........................-----.....-----C-------GC--C--C..--GTCA---AA-A--GC----GA--TA------G--A-----C----C----G-------G--C-TGC-T--G-----TA-  596
SMM.US.x.H9                TA..........GGACTGAGTTCCCTACTTTT.GAGA.AAAGAGTAG--GA-T-G---.....-------T----AAC--C--C..--GTCA---AAGA--GC----GA--TA--------A-----C----C--C------------G--C-TGC-G--G---------  638
SMM.US.x.PGM53             .......TCATAGGACTGAGTTCCCTACTTTTGAGG..AAAGAGTAG--GA-T-G---.....-----C------AAC--C--C..--GTCA---AAGA--GC----GA--TA--------A-----C----CAAC----------G--C-TG--G--G-----G----  1082
SMM.US.x.SIVsmH635F_L3     .......GTCATAGGACTGAGTTCCCTACTTTTGAGGAAAGAGTAG--GA-T-G---.....-----C------AAC--C--C..--GTCA---AAGA--AC----GA--TA--------A-----C----CAAC----------G-----TG--G--G-----G----  1168
                                                                                        Gag and Gag-Pol start
MAC.US.x.239               AAT.........AGCTGTC.TTTTATCCAGGAAGG.GGTAATAAG-T-GA-T-G---.....-----C-T----AAC--C--C..--GTCA--AAGA--GC----GA--TA--------A--C--C-A--CAAC----------G--C-TG--G--G--G----  1156
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__G__K__K__K__Y__M__L__K__H__V__V
MAC.US.x.251_1A11          AAT.........AGCTGTC.TTTTATCCAGGAAGG.GATAATAAG-T-GA-T-G---.....-----C------AAC--C--C..--GTCA---AAGA--GC----GA--TA--------A--C--C-A--CAAC----------G--C-TG--G--G--G----  1155
MAC.US.x.251_BK28          AAT.........AGCTGTCTTGTTATCCAGGAAGG.GATAATAAG-T-GA-T-G---.....-----C------AAC--C--C..--GTCA---AAGA--GC----GA--TA--------A--C--C-A--C-C----------G--C-TG--G--G--G----  1144
MAC.US.x.EMBL_3            AA..........TAGCTGTCTTGTTACCAGGAAGGG..ATAATAAG-T-GA-T-G---.....-----C------AACG-C--C..--GTCA---AAGA--GC----GA--TA--------A--C--C-A--C-C----------G--C-TG--G--G--G----  642
STM.US.x.STM               A...........GGGCTGATTGTTCTACTCTT.GAGG.GAAGAGG-------T-G---.....-----C------AGC--C--C..--GTCA---AA-A--GC----GA--TA------G--A-----C----C--C----------G--C-TG--G--------  812
MNE.US.x.MNE027            AAT.........AGCTGTC.TTTTATCCAGGAAGG.GATAATAAG-T-GA-T-G---.....-----C------AAC--C--C..--GTCA---AAGA--GC----GA--TA--------A--C--C-A--C-C--G-------G--C-TG--G--G--G----  636
DEB.CM.99.CM40             TCCACTAGGGACGTTTGTGCAAGTAGGCACGTCAGGGTAAG....................-----G---G-GCG---G-C-CCA----C-----CG-CT-----CA---TT-------G-AA-A--G--A---AAA--C--A--G---TGA----G-CC-C--  591
DEB.CM.99.CM5              TCCACTAGGGACGTTTGTGCAAGTAGGCACGTCAGGGTAAG....................-----G---G-GCGA--G-CGCCA----CG----C--CT-----CA---TT--G----AA-A--G--A---AAA--C---------G---TTGA----G-CC-C--  594
DEN.CD.x.CD1               GTTGAAAAAAATCTAGTCGAGAGG.........................................-----GAGCG-GAGC--G--TCCCA-G-CA--T-CG-CCC-G---A---TT------G-AA-A-----A------A-T-----G--G--C-TGA-T--G--CC-GA-  585
LST.CD.88.447              .............................................................-----AT--G-GAAC--T--C..C----TA--CA-ATTGAG-GAGA--TCTGC-GT--CA--C----A--T-ACA-T--A--G-CT---C-GAAG-G-----T----  104
LST.CD.88.485              .............................................................-----AT--G-GAAC--T--C..C----TA--CA-ATTGAG-GAGA--TCTGC-GT--CA--C----A--T--CA-T--A--G-CG---C-GAA--G------T--  104
LST.CD.88.524              .............................................................-----AT--G-GAAC--C--C..C-C--TA--CA-ATTGAGA-AGAT-TTTGT--TG-CA--C----A--T---A-T--A--G-CT---C-GAA--G---G---A  104
LST.KE.x.lho7              AGTGTAACTGGGACACTAGTGGTAGCTTCACT.................................-----AT--G-TAAC--T--C..C-C--TA--CA-ATTGA-A-AGAT-TTTGT-GTG-CA--C----A--T---A-T--A--G-CT---C--AA--G---G---A  1179
SUN.GA.98.L14              GTGTCTCGGAAGAGTGTCTCAAAC.........................................-----A---G--AAC-CAC-..G-GGATA---A--TTG-GAGGA-C-TT----G-G-AGCTC-G-AA--T--A--G--A--G-CC--CC-GGC--G----G----  1166
MND-1.GA.x.MNDGB1          GTAAGGGGCAGCATAGTCAGAGCA........................................------AATG-GAAC--T-CC...-GTTA---ACT--T--G---AA--TT--G------AA-A----A-AG---T--T--A---TGT----G---G-----CC-CTG  553
OLC.CI.97.97CI12           TTTAGGTCCTACAGGATAGGATCGAGCGCTCCATAGGGTTGGAGGCTCGG-TGA----G...--------G--C--A-AGTAAG...A--GA-CCT-AGTT---GA-GAAG-TA--G-GGG-CA-A--GC-TGA-AC---G-----GTGT-----GGAGGGCG---G--AA  809
WRC.CI.97.97CI14           GAAGTCCCAGGCATTTGACTAGCAAAG......................................------AT-AG-GAATAGCA-CTCGC--GCACCCCATAT--GC--GG-TCTCC---GGG-G-AT---C-CAA--ACTC---------G---C-GAAG--G--C-----  1151
WRC.CI.98.98CI04           GAAGACCCAGGCATCTGAATAGCAGAG......................................------T-AG-GAATAGT--CTCGC--GCACCACATATT-GC--GG-TCTCC---GGG-G-AT--GC-TAA---CTC------G------C-GAA---G---G--A-  1157
WRC.GM.05.Pbt_05GM_X02     .............................................................------T-AG-GAATAGCA-CTCAC--GCACCA-A-AT--GT--GA-TCTCC---GGG-G--C--GC-AAA--ACTC--------C--------GAG-------C----  107
MND-2.CM.98.CM16           TGAAAGAGGACAGGGGGCAAGAAA.........................................-----C------C-------GT...C-T--G--A-A-A----G--CGAGCT------G---A-----C----CTCC-----A--G-----CC-GC-G------C---A-  840
MND-2.GA.x.M14             AAGGGAAGGGCAGGGGGCAAGAAA.........................................-----C------C-------GT...C-T--G--A-AGA----G--GA-CT------C---A-----C----CTCC----------G------CC-G---------------A-  764
MND-2.x.x.5440             AAGGGAGAAGCAGGGGGCAAGAAA.........................................-----C------C-------G-...C-T--G--A-A-A----G--GA-CT------G---A-----C----CTCC----------G------CC-G-----------G-TA-  397
DRL.x.x.FAO                AAAGAAGCCTAGG.................................-GTAAGA-A--.....-----A------C-------G-...C-T--G--A-A-A-GC-T---GCT-TT------GG--A-----C----CTCC-----A--G---------GC-C--G---C-T--  394
MON.CM.99.L1               GAGGGAGCATAAAATAAGGGCACG.........................................-----G-----GCAC---G-CGATGC-T--T--TACTA-------CAAG-AT--G--GG-G---C-T--A--GC-C--T-----G--G--CCTGA-T--G----------  548
MON.NG.x.NG1               .................................................................................................................................................................  0
GSN.CM.99.CN166            .............AAAGAGGAGTAACAGACTCCCACC......AGAG--TA-G-CAGA.....-----A-----GCAC---G-CGATGC-G-CA--TACTA-G--G---A-G-AT--G---G-G--TC-C--A--A-----------G--C-TGA-C--G--CT----  555
GSN.CM.99.CN71             .............AAAAGAAGTAACAGACTCCCACC......AGAG--TA-G-CAGA.....-----A-----GCAC---G-CGATGC-G-CG--TACTA-G--G---A-G-AT--G---G-GA-TC-C--A--GA-A-----------G--C-CGCGA-C--G--CT----  561
TAL.CM.00.266              TGTTAAGTGCACACACGGGGCAAG.........................................-----G--AC-GCAC--G-CGATGC-C-C---TA-G-CC------AA--AT--G---G-GA-AC-G-------GAAA--G--A--G--G-----TGG-G--G-----T--  769
TAL.CM.01.8023             TGCTAAGTGCGCACACGAGGCAAG.........................................-----G--AC-GCAC--G-CGATGC-C-C---A-G-CC------AAG-AT--G---G-GA-AC-------GAAA------A--G--G-----G-GG-G-G------G-G--  255
MUS-1.CM.01.1085           ...............GGAGAGGAAGAGACTCCCACA......GAGGAC---GCCAGA.....-----A-----GCAC--G-CGATGC-G-CA--TAACA-G--G---A-G-AT--G--GG-GA--C-C--A---AAA---------G-----CCTGA-C--G--C--TA-  553
MUS-1.CM.01.CM1239         ..............TGAGAGGAAGGAGACTCCCACA......GAGGAT---GCCA-A.....-----A-----GCAC--G-CGATGC-G-CA--TACCA-G--G---A-G-AT--G--GG-GA--C-C--A---AAA---------G-GG--CCT-A-C--G--C------  555
MUS-2.CM.01.CM1246         ..............AGTAGAGAAAGCGGACTCCCACC......AGAGA----G-C---.....-----G-----GCAC--G-CGATGC-G-AA--TACTA-G--GA--AA---AT--G--GG-CA-A---------TAAA----------G-----CCT-A-C--G--------  624
MUS-2.CM.01.CM2500         ..............GCGAGAGGAAGCGGACTCCCACC......AG-G--TA-G-C---.....-----G-----GCAC--G-CGATGC-G-CA--TAC-A-G--G---CAA--AT--G--GG-C---T--G--A---AAA---------G-GG--CCTGA-C--G--C------  583
SYK.KE.x.KE51              GCGGTCGGGGATAAATAGGGGGTC.........................................-----A--AG-G-GA--G-CGATCC---CA--TA--A-GC-T---AA--AT--G---------C-TC----GAA---G------GCGC---CCT-G-TCG-----T----  538
SYK.KE.x.SYK173            CAGCAGTCGGGGAAATAGGGGACC.........................................-----A--AGCG-G---G-C-ATA----CA--TA----GC-----CA---AT--G---------C-CC-A--GAAA--G-----GCGC---CCT-G-TCG-----T----  901
RCM.GA.x.GAB1              CGAAGTCCTAGAGAAGGGAGGAAA.........................................------------------C---T-G...--GTCA---AAGA-GC-----CGC--------TC-G--A-----C----C--C--G--A--G-----C-TGC-G--G---T-G--  315
RCM.NG.x.NG411             CGAAGTCC...........................T-GA--AG-GAGGAAA.....-----------G------A-T...--GTCA---AAGA-G------GCT------C-GG--A-----C----C--C----------------G-----TG----G--G--C-----  315
GRV.ET.x.GRI_677           GTTGGTAGGGGACGGGTCGGAGCA.........................................-----C-G-G-TCAC----C-..C-GTCA---A--AGCC-C--CACG-TC--G--G---A--C--C-T--GAAC--G--A--G--G--CC--A-T------T--A-  1000
SAB.SN.x.SAB1C             TGGCTTAGCAGTGGCAGAAGGCTAGGTCGA...............GCGG-A-AG--AA......-----------TAAC-----C..-----T--TA-GA--C----GC--TC----TC-G-CA-----C----CAAC-----------G--G------G-----G----C----  1120
TAN.UG.x.TAN1              ACAGATCAGCA.CCTGGGTGAGCA.........................................-----A---G--CAC--G-C-..C-GTCA---A-GA-T--G--CAC--TT--G-----A--C---C-C--GAAC-----------G---C---------CT--A-  1005
VER.DE.x.AGM3              GTAGGGCAGGCGGACGGTACGCA.........................................-----G---GCTA-C----C-..C---ATA---A--C-------CAA--TT--GC-T--A--AC-TC-C--GACC----------G--CC--A-T------T--A-  534
VER.KE.x.9063              GTAGGGCAGGC.GGACGGTACGCA.........................................-----G---GCTA-C----C-..C---AA---A--C-------C-A--TC--G-----A--A--C-C--GAAC-----------C--A-T---G---T--A-  1045
VER.KE.x.AGM155            GTAGGGCAGGCGGACGGTACGTA.........................................-----G---GCTA-C----C-..C-G-ATA---A--C-------GA--TT--GC-T--A--AC-TC-C--GAAC----------G--C--A-T------T--A-  1034
VER.KE.x.TYO1_patent       GTAGGGCAGGC.GGTGGTACGCA.........................................-----G---GCTA-C----C-..C---ATA---A--C-------C-A--TT--G-----A--AC-TC-C--GAAC----------G--CC--A-T------T--A-  1042
COL.CM.x.CGU1              CTGCGGCGCTCCAGGACCGGTAGAGGGAGGTCATTA..............................-----CAACGAGCA-GG-C-G...---G-GAA-AA-AC------AGACCTAC-G---G-C---C-G-A-AA----AAG--AGG-TGT-----GA--------G-TAG  707
```

```
H1B.FR.83.HXB2              ATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGA 1063
Gag                         _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__L__Q__T__G__S__E__E__L__R__S__L__Y__N__T__V__A__T__L__Y__C__V__H__Q__R__
H1A1.UG.85.U455             -----------------G---AA---A--C---C--------T--------G------A---CAG-------------AT-------G-T--C--------A------------------------------------------GT-----------A---------  530
H1B.US.90.WEAU160           ---------------------------------C--------------------------------------------AT----------------------------------------------------------------TGT------------------A-- 1062
H1C.ET.86.ETH2220           C-------A---------G---AA---T---C----C----A---T-----T-------C--------A-------A-TAA------------G-T------------A----G-------A-------T---------G-----T----------A-------A--  443
H1D.CD.84.84ZR085           ------------A-----G--------T---C-----------T--------------------------A-------A---A------------G-T---A--------------------------A--------------------------A---A------  580
H1F1.BE.93.VI850            ---------------------------T---C--G----------TC----------------------CA-A-----A-AA------------------------------------G----AG-------T-------------GT--------A---A---------G  418
H1G.SE.93.SE6165            -----------------G--GA----T---C----C----A---T--------G-------T---CA--------A--A-----T-G--------T--C---------A----G---A---A----------G-------------------------A---------  466
H1H.CF.90.056               -----------------G---A----T---C----C--C-----T----------C----------CT---G---A-A-A----A----G--G-TA--A--------A----------------A-------T----CT-------GT---------C--A----G--AA-  410
H1J.SE.93.SE7887            -----------------G--CA----T---C----C---------TC----GT--G--A----------CA----------A-T------C-----G-T--C----------A--C-----A---A-------G--------------------------C--A------G----  386
H1K.CM.96.MP535             ---------------------------T---C-C--C--------T-----G---A-----------C-G------A-AAC---AA-----G-------A-----A---------------GA---A----C-----------A-----GT---------T---A-------A--  274
H101_AE.TH.90.CM240         ------------A---T-----A--------C----C---A--T-T-----------G--------A---CA-------A-A-A----T----GT--A-T--CA-----------T--------------A---------T-------------------GG--C--A--C-------  628
H102_AG.NG.x.IBNG           -----------------G---A---------C----C--------T-----------G----G---A---CA----C--A---A----T------T--G-T--CAG-----------------------A---------T----------A---------T-GG--C-----------  603
H103_AB.RU.97.KAL153_2      -----------------G---A-------GC----C---A----T------------------A---CA---------------A----T----------A-T--CA------------------------A----------------------------------------A-------C---  295
H104_cpx.CY.94.CY032        ------------A---T-G---A--------C----C--------T-----------G---------A---CA----T---A---A----T------T--A-T--CA-A-----------------------------------------------TA--A----------GG--C--A---------A-  429
H1N.CM.95.YBF30             -----------------G---A-------TTG---C--A--T--CA-G--C---G-G--C------GCCA-GT----AAAT--AT--G-----G-T--CA-------G----------GC-C--T-------CG-TC-----GTT--T---------C---AGT----  621
H1O.BE.87.ANT70             ------------A-----G---A--------TG--------AG--AC--------TG----G--TAA-GAGA-GC-G--ACAG---T--G-G---G-T--CA-------G-----CAGC--GCAG---C-C-GG--CG--A---T-GTG----GG-----T--CA-C--AT 1115
H1O.CM.91.MVP5180           -----------------G---A---A---TG----------T--A---------TG-------TAC-GAG---C-G--ACAG---T--G-G---G-T--CA-------G-----G--C--G-A---TC-C-GG--CG--A-----GTA----GG--C--T--CA-C--AT 1090
H1P.FR.06.RBF168            ------------A-----G---A----T---TG--------AG--TA-G------G-----------GA-A--T--T-A-AG---T--G-----G-T--CAGA-----T--T--T-GC--GCAG--TC-T-GG--C--CC---T-GTTT-A-GG-----T--CAG---AG  542
CPZ.CM.01.SIVcpzCAM13       ------------A-----G---A--------TG---C--C--G--AA-G----GTG--------A---CTTA--T-G--ACAG---T--G-G---G-T--GA-A--C--C--G----GG--GC----T----T---C--TC------GTGT-A-GG-----T--CAGT-AC-  625
CPZ.CM.05.SIVcpzEK505       -----------------G---A-------GTG---C--A--T--CA-G--C--G--------T---GCC---T-----CA----AT--G-----G-T--CA-A-----G-------G---GC-C--TC---T---C--CC-G---GTT--T--C-------T---A-T-AC-  602
CPZ.CM.05.SIVcpzLB7         -----------------G--GA---T--TTG---C--C--G--TA-G-------A-----------GAG----T-GA-TAC----T--G-----G-T---A-------G-------A-C-----A----T----T---C--CA----TGGT-----GG--C-----------G--A-  609
CPZ.CM.05.SIVcpzMB66        -----------------G---A-------TTG---C--CTC---CA-G-------A-----------AG--GC-TT--CAG------G-----G-T--CA-A-----T---------G-----------T--G-T---C--CA---T-GT---T-GG--C--------------G--  612
CPZ.CM.05.SIVcpzMT145       ------------A-----G---A-------TTG---C-----T--AA-G----ATG--------T---GAGA-GT--T--AA-------G-----G-T---A-------G-------G---GC-C--T----T---C--TT-G---GT---T-GG------A---A-G--A-  632
CPZ.CM.-.SIVcpzMB897        -----------------G---A-------TTG---C--C------AA-G--C---A----------A---GA-----T--A-TAAC---T-G--G---G-T--CACA----------------G---A------T----T---C--CC------GT------GG--C--A-----------A-  608
CPZ.CM.-.SIVcpzDP943        C----------------G---A-------TTG---C-----T--AA-G------GG---G--A---GA---GC--T-AAA----T--G-----G-T--CA-A----------------GC--GC----GC-T-T----C--TC----GGTA--T-GG-----T---A----A-  631
CPZ.CM.98.CAM3              ------------A-----G--CA-------TTG---C-----T--AA-G-----GG-------A---GAG--GT--T-AA-G--AT--G-----G-T--CA---------------GC--GC----GC-T-T------TC-----GTG--T-GG-----T---A-G--A-  453
CPZ.CM.98.CAM5              -----------------G---A----T--TTG---C-----T--AA-G--G---G-T-----A---GA-A-TC----AAA----T--G-----G-T--CAG-------------G-GCT-GC----GC----T---C--TC----GGTA--T-GG-----T---A-G--A-  759
CPZ.GA.88.GAB1              -----------------G---A-------TG-G-C--C--G--AA-G----GTAAG-----A----CTA--T-GT-ACA----AT--G-G---G-T--CA-A-----C---------G---GC-G--C--G-T---C--TC-G---GTA--G-GG--CA-A---AGTGAC- 1100
CPZ.GA.88.GAB2              -----------------G---A-------TTG-G-C-----T--AA-G--------AA----------TAT--GT--T--AA---------G-----G-T---A-A-------------G-T-AC-G-C---T-------TT-G--TGTGT-A-GG-----T---GCT--A-  435
CPZ.US.85.CPZUS             T----------------G---A----T--TTG---C--A--T--AA-G------G----C------CTC----T--T-AAA----T--G-----G-T--AA-A-------G-----GC--GC----GC-C-T----C--CT-----GGT---T-GG-----C--CAGT---G 1105
CPZ.CD.90.ANT               T--------ATC------GC-G--T--T--GC-C-GCT-CTC---TC-----------------T---GA-A-GGCTA-CCAT--AT-GAGC--T---A-AG-A-T-A----CCCT---A-A-T---T-G-T---C--CA-TTGTGTT--G-GG--C--A---A--G--G  427
CPZ.TZ.00.TAN1              ------CG-A--C--A---C-G--T-----GA-G-----C--T--CA-G--G-ACGT---------CT-GA----CA-CCTC-----G--G--T--GG-AG-C-TT--T--TCC-----A-C-TT--T-G-T------CA-CTGTGTA-----C--C--A--CGC-G-AG  626
CPZ.TZ.2001.TAN2            ------CG-A--C-----GC----T-----GA-G-----C--T--CA-G--G-ACGT------T---T-GA----CA-CCTC---T------T--AG-AG-C-T---T--TCC-----A---TT--TC-G-T---C--CA-CTGTGTA---T--C--C--A--CGC-G-AG  625
CPZ.TZ.2002.TAN3_1          ------CG-A--C-----GC-G--C-----GA-G-----C--T--CA-G--G-ACGT------T---T-GA----CA-CCTC---T------T--AG-AG-C-TC--T--TCC-----A---TT--T-G-T---C--CA-CTGTGTA---T--C--C--A--CGC-G-AG  625
GOR.CM.2007.SIVgor2139_287  -----------------G---A----T---TG--------AG--TA-G-------G-C--G-----GA-A--T--T-A--G--A---G--T-G-T--CA-A-----G-----T-GCT-GC----TC-G-GG--C--TT-G-T-GTG--G-GG-----C--C-GG--A-  624
GOR.CM.2007.SIVgorCP2135con ---T-------------G---A-G--T---TG--------AG--TA-G-------G-Y--G-----GA-A--T--T-A--G--A---G-----G-T--CA-A-----G-----T-GCT-AC----TC-G-GG--C--TT-G-T-GTGT-G-GG-----C--C-GG--A-  623
GOR.CM.2004.SIVgorCP684con  W----------------G--GA----T---TG--------AG--TA-G-------G-C--G-----GA-A--T--T-A--G--A---G-----G-T--CA-A-----G-----T-GCT-GC----TC-G-GG--C--TT-G-T-GTG--G-GG-----C--C-GG--A-  543
H2A.DE.x.BEN                G------GCG-AT--AT-G--TAA-----G-T-GGCAGAGA-------G--GT--AA-----T--CCA-A-G--T--CA--GTTT--G-T----TAG-A-CA-----G------A-TT-A-A-AGCC-T-T------C--CTGCGT-A-T-GG--CT----CGC-GAAG 1376
H2A.GW.x.ALI                G------GCG-AT--AT-G--CA------G-T-GGCAGAAA-------G--GT--AA-----T--CCA-AG----T--TAA-GTTT--G-----TAG-G-CA------------A-TT-A-A-AGCC-T-T------T--CTGCGTAG-T-GG--C-----CGC-GAAG 1369
H2A.SN.x.ST                 G------GCG-AT--AT-G--CA------G-T-GGCAGAGA-------G--GT--AA--G--T--CCA-A----T--TAC-GTTT--G-T----TAG-A-C------G------A-TT-A-A-AGCC-T-T------T--CTGCGT-A-T-GG--A-A--CGC-GAAG  821
H2B.CI.x.EHO                -------GTG-AT--AT----GA----T-G-T-GGCAGAGA---G---G-G-T--AA-----A--CC-GA----T-G-AA-GTCT--GG-----TAG-A-CA-----T------A-TT-A-A-AGTC-T-------C--CTGCGT-A---T---CC----CGC-GAAG 1353
H2B.GH.86.D205              -------GTG-AC--AT----CA----T-G-T-AGCAGAGA-------G---T--AA-----A--CA-A----T--TAA-GTCT--GCT----TAG-A-CA-----G------A-TT-A-A-AGCC-T-T-----TC--CTGCGT-A-T-T---CC----CGC-GAAG 1352
H2G.CI.92.ABT96             G------GCG-AC--AT-G--CA------G-T-AGCAGAGA-------G---T--AA-----T--CCA-A----T--TTC-GTAT--GCT----TAG-A-CA-----T------A-TT-A-AGAGCC-T-------T--CTGCGT-A-T-GG---T----CGCCGAAC  739
H2U.FR.96.12034             G------GCG-AC--AT-G--CA----T-G-T-AGCAGAAA-------G--GG-CAA-----A--CA-A----TT-ATCGGT-T--GCT----TAG---CA-----T------A-TT-A-AGAGCC-T-------T--CTGCGT-A-T-GG---T----CGC-GAAG  855
SMM.SL.92.SL92B             -------GC---A--AT----CA----T-G-T-TCGGCAGAAA-------G----GCAA-----A--CA-AG----T-A-C-GTA--GCT----TAA-G-CA-----C--G---A-TT-A-A-AGC-G-T--G---T--CTGCGTAG-T-GG--C-----CGC-GAA-  766
SMM.US.x.H9                 -------GCA-AT--AT-G--CA----T-G-T-AGCAGAAA-------G--G-ACAA-----T--CA-A----T--TTC-GTTT--GCT----TAG---C------T------A-TT-A-AGAGCC-T-------C--T--CCGCGT---T-GG--CA-T--CGC-GAAG  808
SMM.US.x.PGM53              -------GCA-AC--AT-G--CA----T-G-T-AGCAGAAA-------G--G--CAA-----T--CA-A----TA-TTC-GTTT--GCT----TAG---C-G----T------A-TT-A-AGAGCC-T-------T--CTGCGT-A-T-GG--CA-T--CGC-GAAG 1252
SMM.US.x.SIVsmH635F_L3      -------GCA-AT--AT-G--TA----T-G-T-AGCAGAAA-------G--T-ACAA-----T--CA-A----T--TTCGGTTT--GCT----TAG---C------T------A-TT-A-AGAGCC-T--------TA-CTGCGT-A-T-GG--CA-T--CGC-GAAG 1338

MAC.US.x.239                -------GCA-AT--AT----TA----T-G-T-AGCAGAAA-------G--G-ACAA-----A--CA-A-------TTCGGTCT--GCT----TAG-G-CA-----C------A-TT-A-A-AGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAG 1326
Gag                         __W__A__A__N__E__L__D__R__F__G__L__A__E__S__L__L__E__N__K__E__G__C__Q__K__I__L__S__V__L__A__P__L__V__P__T__G__S__E__N__L__K__S__L__Y__N__T__V__C__V__I__W__C__I__H__A__E__
MAC.US.x.251_1A11           -------GCA-AT--AT----TA----T-G-T-AGCAGAAA-------G--G-ACAA-----A--CA-A-------TTCGGTCT--GCT----TAG-G-CA-----C------A-TT-A-A-AGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAG 1325
MAC.US.x.251_BK28           -------GCA-AT--AT----TA----T-G-T-AGCAGAAA-------G--G-ACAA-----A--CA-A-------TTCGGTCT--GCT----TAG-G-CA-----C------A-TT-A-A-AGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAG 1314
MAC.US.x.EMBL_3             -------GCA-AT--AT----TA----T-G-T-AGCAGAAA---C--G--G-ACAA-----A--CA-A-------TTCGGTCT--GCT----TAG-G-CA-----C------A-TT-A-A-AGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAG  812
STM.US.x.STM                G------GCA-AC--AT-G--CA----T-G-T-GGCAGAAA-------G--GT--AA-----G--CA-A----TA-AACGGTCT--G------TAG-G-CA-----T------A-TT-A-AGAGTC-T-T------T--CTGCGT-A-T-GG---A-C--CGC-GAAG  982
MNE.US.x.MNE027             -------GCA-AT--AT----TA----T-G-T-AGCAGACA-------G--G-ACAA-----A--CA-A-------TTCGGCCT--GCT----TAG-G-CA-----C------A-TT-A-AGAGCC-T--------T--CTGCGT-A---GG--CA-T--CGC-GAAG  806
DEB.CM.99.CM40              C------G-T-A---A-----CA----T-GCC--GCAG-A-CAT-A-------ATAAG---------CAGG----TT--ACTAG-T----G--T-TAGAAGCA-----CAGT--GA-TT-G-A---TC-----GGC-T---CTGTGT-A----GGGCATGC---GCC-ACC  761
DEB.CM.99.CM5               C--------T-----A-----CA----T-GCC--GCAGGA-CAT-A--G-----CAA-------A---CA-G-G--CT-AACTAGAT-------T-TAGAGGCAT-----CAGT----GC----AG--CC----CGGC-T----TGTGT-G-T-GGGCATGC--CGCC-ACC  764
DEN.CD.x.CD1                T---T---AG-A----A-G--G-----T-GCT--AGGAGAACAG---C--------C-AT----A--CCA-ACT--CA-A-AGACCT--GCT----TAGAA-CACA---CAGT--GAGT--G-A-AGCC---TC-GA---TGCTGCGTTG---GG---TGC--CAGGG-AT  755
LST.CD.88.447               G-------CA--A--AT----TA-G--T-G-T-AGGAG-ACA-T--C----G---G----T-----C-AGA----T--T---GTCTGTTGG---CTATACTCC-----C--GA-GA-TT-A-A-G--C-TGTAGGA--T---TGTGTG--G-G---CTGC---TT-G---  274
LST.CD.88.485               G-------CA-AA--AT----TA-G--T-GTT--AGGAG-ACA-T--C-G--G---G----T-----C-AGAG---T--T---GTCTGTTGG---CTATACTCC-----C---A-GA-TT-A-A-G--C-TGTAGGA--T---TGTGTG--G-G---CTGC---TT-G---  274
LST.CD.88.524               C-------CG-AA--A-----TA-G--T-GGC--AGGAG-AAA---TC--------G----G-------AGA-G--TT-----GTC-GTTGG--T-TATA--CT-----C--TA--A-T--G-A-G-TC--GTAGGG----GTGTGT-A---G---CTGT--C-T-G-A-  274
LST.KE.x.lho7               --------CG-AA--AT----CA----T-G-T-GGGAT-ACAA--C--G------G-------G----AGA----C--ATCTGTCTGCTGG--T-TATA-GCA--T--T--TA-GA-C--A-AGG-C----GTAGGG-----TGTGT-A---GC--CTGT--C-T-G-AG 1349
SUN.GA.98.L14               -------G-T-AA--AT----TA-G--T-G-C----CAAG-AA------AG----GTG--------CAGA-G--T---TC-GTCTGTTGG--T-TGTA-GCAT-T--T--TA-GA-C--G-AGG-C----GTAGG------TGTGTG--TGCC---TGT------G-C- 1336
MND-1.GA.x.MNDGB1           T---TGT-AAG-T--AT----TA----T-GCT--ATCGGA-AAA--CC-T------CAGC----A---GA-A----C--CTC-GTATGTTGG---TATA-G-CCA---------TA-T--A-A-G-T--GGTAGGG-----CTGTGTTG-AGCC---CA-A--CGC-G-T-  723
OLC.CI.97.97CI12            -------GTACAA---T-GA---ATC-ATGCA-AG----A-ATGATC---G---TAAG-----------A-AG---TG-A-A-A-AT---TG--T-TGGAAG--GG----A-TC-GA-GT-G-A-AGT--GCTG--C-TCTGTAT-GTTA-T-TG---T----AG--G-CC  979
WRC.CI.97.97CI14            ----TT-GTG-A---AT----GAAG--GTGCT-AG--T-A-CAT-AA-CAGT---G-------GGTG-AGA-GC-TA--TC---T---GCT--T-TA--G-CTGAT-CC--TA-G---G-A-AGAGCA-T-GGGGA-TGT--TGTGTGT-AGCA---C-C---G-G--A- 1321
WRC.CI.98.98CI04            ----CT-GTG-A---------GAA---GTGTT-AG-CT--ACAT-AA---C-----GA-AC---GAC--A-A---C-GA---A----TT--GC------TAG---CAGAT------A-----G-G-A-AGCC-T-GGGGA-TGA-GTGTGTG--TGCA---T-----GC-C-T- 1327
WRC.GM.05.Pbt_05GM_X02      T---CT-GTA-AA---T----GAAG--ATGTC-AG--GGGAAAT--A--TCT-GCCA------AGTG-AGA--C-CTGT-AG--TA--GCT--T-TGG-A-CAGA---G---A--AC-A-A---AGTA-T-GGGGA-TGA--AGTGTG--AGCA---CT-AT--A-T-AC-  277
MND-2.CM.98.CM16            -----T-----A---A-----CA----T-GCC--C--GAAAAG----------GTAAG------A--CGAGA----T--TA-CGTA--CTTT---CTAG---CT-----G------A-TT-A-TC--GC-G--C--C--CTGCTGCTG-G-A-GG--C--A---GCG-AAG 1010
MND-2.GA.x.M14              -----T-----A---A-----T-----T-GCC--C--GAAAAA--A---------TCAG------A--CGAGA----T--TA-CGTA--CTTT---CTAG---CT-----G------GA-T---CTTG-GC-G--C--C--CTGTTGTTG-A---GG--C--A---GCG-AAG  934
MND-2.x.x.5440              -----T-----A---A-----CA----T-GCC-AC--GAGAAGT---------GTCAG------A--CGA-A----T--TA-CGTA--CTTT---CTAG---CT-----G------A-TT-A-TT--GC-G--C--C--CTGTTGTTG-A-T-GG--C--A---GCG-AAG  567
DRL.x.x.FAO                 C----T-GCG-A---A-----T-------GCC-GC--GAGA-AT-A--------GTCAG------G--CCA-A----TT-A---GTA--CTTT---TCTAG---CT-----G------GA-TT-A-TT--TC-C-T------CTGCTGCGTAG-T-GG--C--T--CTCG-AAG  564
MON.CM.99.L1                ------TGC--A------T--TA------GTC-CTCGGACTCG--CC-G--G---CAG--C-------AGA-G--TT---A-GTTA-T-TT--TCTA-AG-C-----------C--GTCTA-C-A---TC-C-T-GG--TC-CTT-GGTAT-G-----CA-A---GCCG-T-  718
MON.NG.x.NG1                ..................................................................................................................................................................................... 0
GSN.CM.99.CN166             ------CTC--AA-----T--GA---AT--GC-GTCCGAC-CT--TC-T----ATCA---G--G---C---GG---T-A-A-ATTG-CTTC--CCTAGAA-CA-----G--------CT--C-A---TC-C-T-GGA-TC--GT--GTG---------A-C---GC-GAC-  725
GSN.CM.99.CN71              ------CTC--A------T--GA---A---GC-ATCAGGCA-T--TC-T---T-TAA------G---CA---GG--T---A-GTCACCTCC----TAGAA-CA-AT--G--C---TCT--C-A-AGCC-C---GGA-TC--GT--GTG---------A-C---GC-GAC-  731
TAL.CM.00.266               ------TTC--AA---A-G--GA-G--T-GCC-G-G-GAC-C---------G--CAA------G--CGTTA-G--CA-A-A-GCCA-TT-CATG-TAGAA-C--AT--C--------GGA-A-AG--TC-C-T-GGA-TC--GTGTGTT--A--C--CA-T---GCTG-T-  939
TAL.CM.01.8023              ------TTC--AA---A-G---A-G--T-GCC-G-G-GAC-C---T--G--G--CAA------G--CGTTA-G--CA-A-A-GCCA-TTCT--G-TAGAA-CA-AT--C---------G-A---AG--TC-C-T-GGA-TC--GTGTGTT--A--C--CA-T---GC-G-T-  425
MUS-1.CM.01.1085            ------TTC---------T--TA----T-G-C-CTC-GA-TC---T--G---T--AGG--T--G---CA-A-G--CA-A-AGGTCA-C-TC--C-TAGAA-CAG----G--C--GTCC--C-A----C-C-T-GGA-TC--GT--GTG----GG---A-C---CGC-GA-G  723
MUS-1.CM.01.CM1239          ------TTC--A------T---A-G--T-G-C-CTCGGACTC---T--G--G---AA---G--A---CA-A-G--CT-A-A-GTT--CAT-----TAGAA-CA------G--C---TCT--C-A----C-C-T-GGA-TCACGT--GTG----GG---A-C---GC-GA-G  725
MUS-2.CM.01.CM1246          ------TTC---A-----T--G--G----GTC-CTCAGA-ACG--CC----G-GCAAG--C--A--CCA--GC--C--A-A-GTAA-CTGC--C-TGGAAG-A--T--GAGC--G-CG--C-A----C-C-T-GGA-TC--GT--GTA----GG---A-C---GC-GACG  794
MUS-2.CM.01.CM2500          ------TTC--AA--A--T--CA------G-C-CTCAGACTC---T---------TAAG-----G--CCAG-GC--C--A-A-GT-A---T------TGGAG-C---C--G--C--GTCT--C-A----C-C-T-GGA-TC--GT--GTA-----------A-C---CGC-GA-G  753
SYK.KE.x.KE51               T------G-A-A------G--CA------GCC-CTCAGACCAG--TC-T----GTAA------T---GA------CA-AAA-ACCA-CTT-----TAGAAA-ACAT--C-------A-TT-A-A--------T-GGC-T-AC----GTAG---GGGCC--C--CGC-GAA-  708
SYK.KE.x.SYK173             T-------AG-A----T-G--CA------GCC--TCGGACCAA--CA-G--GT-TAA-----------GA-A----T--CAC-GTC----TT-----TAGAAGCT-AT--G------A-TT-A-A-AGCC-T-T-GGC-T-A-TT-GGT-G-A-GGGCC--T---GCC-AA- 1071
RCM.GA.x.GAB1               -------T---AAA-A---A-TAA---T-GCT-G-G-GA-CATT----------G--AC----A---GA-A-----T-A----GTC--G-TG--TCTAG---C------G-----G-GG--A-A-AGCC-C-T----TTGTGCTGCGTA----GG--C--A---CA-GGAAG  485
RCM.NG.x.NG411              -------T-------A-----GA----T-GCT-G-G-GA-ACTT---------GTG--------AATCCAGA----TA-A--GGTTA-T-TG--TCTAG---C------G--T--G-GCT-A-A---CC-C-T----TTGTGCTGCGTA--T-GG--C-----CA-GGAAG  485
GRV.ET.x.GRI_677            -------G-A-AA--AA-G--------T-GGT-AC--GAGAAA--T-----------AA-----------CA-A----CA-A-A-GTTT--ACC--G-TGGAA-C-------C--C--G-GG--A-A-G-TC-G-T----TTGTGCTGCGT-A-T-GG--CA-T--CGC-GAAC 1170
SAB.SN.x.SAB1C              -----------A---A-----T-----TAGCC--TCAG-AAAT--C-----G---AAG-----GGTCGTCA----T--TA-TGTA--C-TT--CCTAG---CT-----G------GA-TT-A-TTG-TC-C-T-----TT-TGCTGCGTGT-AGC-----A-T--CGC-GAA- 1290
TAN.UG.x.TAN1               -------G---AAA-AA-G--G--C--T-GGC-CC-CGAGAAA--C----------AA---G------CA-A-G--CA-A-A-GTAT-GAGC--GCTAGAA-CA--------------G---A-A-AGTC-G-T----TTGTGCTGCGT-A-T-GG---A-T--CGC-GAAC 1175
VER.DE.x.AGM3               -------G---A---AA-G--G--C----GCC-CC--GAGA-AT-AC-----T--GA-------A----AGA-G--CA-A-A-GTA--CT-C--GCTAGAA-CA-----G--G--G-GCT-A-A-AGTC-G-T---CCTT--GTGCGTAT-G-T---C--A---CA--GAT-  704
VER.KE.x.9063               -------G-T-AA--A-G---G--C----GCC-CC--GA-A-GT------------GA---T--A----A-A-G--CA-A-A-GTC--TTTT--T-TAGAA-CA-----C--G--G-GCT-A-A-AGTC-G-TC---CTT--GTGCGTG--A-----CA-C--C-GCGAAC 1215
VER.KE.x.AGM155             -------G---A-A--A-G--C--C----GCC-CC--GAGAAGT-A--G--G---GAG-----T----A-A-G--CA-A-A-GTT--CTCT--CCTAGAA-CA-----G--G----G-A-G-A-AGTC-G------CTG--GTGCGTAT-GCT---C--C--C---GAA- 1204
VER.KE.x.TYO1_patent        -------G---A----A-G--G--C----GCC-CC--GAGA-GT----G--G---GAG--G--G----A-AG---CA-A-A-GTC--CT-C--CCTAGAA-CA--------G--G-GCT-A-A-AGTC-G-TC---CTT--GTGCGTA--A-----CT----CA-GGAAC 1212
COL.CM.x.CGU1               ----ATGT---CA---G--AGC--C-GT-TTT-A-TATT--AG------A-GT--G--AC---AGTGGCT------T---A-A--G-GACG----TAG-GG-C-------------TGT-G--G--CC-----GGA...TGTT-TGTGTGT--C--CC----CAG--AAT  874
```

```
H1B.FR.83.HXB2              TAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAA...........................................................................AACAAAAGTAAGAAAAAAGCACAGCAAGCAGCAGCTGACACAGGA 1158
Gag                         I  E  I  K  D  T  K  E  A  L  D  K  I  E  E  E  Q  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  N  K  S  K  K  K  A  Q  Q  A  A  A  D  T  G
H1A1.UG.85.U455             ----TG----------------------A-T--A---------AT---.............................................................................--T--G-AC---C---GGA-------G---------A--------  625
H1B.US.90.WEAU160           -----G-------T----------------T-----------------.............................................................................--------C--------G-----------------------A- 1157
H1C.ET.86.ETH2220           -------------------------C-------------------A---.............................................................................---G-----C--C-----A-------G----G---A-CTGACA--  538
H1D.CD.84.84ZR085           ---------------------------------------------A---.............................................................................-----C----------GG-------------------G-------  675
H1F1.BE.93.VI850            C--G-G---C--------------------C----------A---.............................................................................---------C--C---GA---------G----------A---G  513
H1G.SE.93.SE6165            -----G-------------A------C----AG-AG-G--AA-GATA---.............................................................................--G------C--G-----AT------G------ATG--T-A----  561
H1H.CF.90.056               ----TG-G-------------G--------T-----------ATA---.............................................................................---------C--C-----A------------------T-AG-A-  505
H1J.SE.93.SE7887            ----A-------------T-----------G-----------ATT---.............................................................................-----G-AC--AC-GC-G------A------A-A------A-AA-  481
H1K.CM.96.MP535             -----G-------------------C--------AC----------A---.............................................................................---------C--CG---GA----A---...-A---A-CT......  360
H101_AE.TH.90.CM240         -----G------------------------T--A-----------TA---.............................................................................--T--G--CC--CG---GA-------G----------G-------  723
H102_AG.NG.x.IBNG           ----C-------------------C-----T--A--G-----ATA---.............................................................................--T--G--C---C----G-T-----GA---------C-------  698
H103_AB.RU.97.KAL153_2      ------------------------------T--A---------ATA---.............................................................................--T--G--C---C----GA-C--A--G------A---G-------  390
H104_cpx.CY.94.CY032        ----TG--C---------------------T--A---------ATA---.............................................................................-GT--G--C---C----GA-------G----------C-G-----  524
H1N.CM.95.YBF30             --C-----C-CA----AC----------G---------A-A--GA-A--GGAA.........................................................................C-GC-C-AGCCCG-GCC-AA-A-C-C--A-----GG-CAG-G-C-  719
H1O.BE.87.ANT70             ATA-A--TGG---T--GC--C-G--AA--C-A---T--A------TAATGGGG.......................................................................................-GC-GGAAGTCTGCG-AC--C---A-GGA--AC 1201
H1O.CM.91.MVP5180           -T--C--CCG---T--AC--C-G--AA--C-A---T--A------TAATGGCA.......................................................................................-GC-GGAAGTCTGC--AG--C---A-GGA--A- 1176
H1P.FR.06.RBF168            -G--T----GT-----AC--C-G--CA-TACA--ATGGA----G--AATGCAG.......................................................................................-G--G-AAGA-AGC-C--...............  613
CPZ.CM.01.SIVcpzCAM13       -C---G--G-------AC--A----A-----AC--C-TA--A--C-C--T......................................................GGAGAACAAGTA-GA----A-G----C-C--A-A-CA-T-A-AGCAGAC-GGGC-CT  732
CPZ.CM.05.SIVcpzEK505       --A--G--C-GA-T--AC-------------A--AT--AGA--GA-AATG...........................................................AAAGCAG-AC---AAG-ACCTG--C--G-A----------GG-CAG-G-C-  703
CPZ.CM.05.SIVcpzLB7         -C-C-----------------G--------T-CAC--ACA-C--TTAG-.....................CAGAAACAGCAAAAGGAAAAGGAA.........................................-----T--A-C-G-G-C-  701
CPZ.CM.05.SIVcpzMB66        -CCCAG-------T--A--A----------TC--C--C-A.........................GAAGCGCAGCAGAAAGGAAAACAGGAAGTGGCTGCAGCTACCGCCGACGGT.......................................  704
CPZ.CM.05.SIVcpzMT145       ----AG-GG-------AC-AC-G------C-G--AC-GA-A-CT-CCGTGAACCCA..........................................................................................GA-CAG  694
CPZ.CM.-.SIVcpzMB897        -TAGAG----------A--A--G----------AC----A-----T---GCAA.......................................................................-AA-GC---GTGAA-GTAA--CAG-A---C-CGG-G-C-  700
CPZ.CM.-.SIVcpzDP943        -TCCAG-GG-------AC--C-G------ACT--AC-GA-A---ACAGCGGCC...............................................................................CGR-A----G---TG--CCAACCACAGCCG  714
CPZ.CM.98.CAM3              -TTCAG--G-G-----AC--C-G------ACA--AT-GA-A----CAGTT.........................................................................-CGGCCTCCCG-G--CAGG-AGTG--CCA-C-GC-GCAGCAG  545
CPZ.CM.98.CAM5              -CC-AG--G-G-----AC--C-----C--TCA--AC-GA-A---ACAGTT..........................................................................GCCGCCC--G--CAGG-AGTG--CCA-C-GC-GCA-CAG  848
CPZ.GA.88.GAB1              -CACTG--G-------AC--A-----C----AC--C--A--CGGC-T--T.....................GGAGAACAACAGAGCAAAACTGAA.................-GT-ACTCAGG--GCCGTG-AGGG-G---CAG-C-AGGC-CT 1216
CPZ.GA.88.GAB2              ----AG--G----T--AC--C----C---C-A---T-GA-A---TTTGTGCAAAAG......................................................................G---CAGCA-C-GC-CAGG-AGC-------AAGC-A-GT-AG  533
CPZ.US.85.CPZUS             -GACAG-GG-G--T--GC--C-G--C----TG--AC-GA-A--G-TA..................GTGCAGTTGCAGAAACAACAGGAAGAAAAAGAA.............................---CAGCA-CAG--GG--A-T 1206
CPZ.CD.90.ANT               A-A-------------AG-AC----CG-TA-A-CAG-GA-AATGA-AGT-..............ATGCAGACACAAGCAGAAACAGGAAGTAGCCAAACCGCAAGCAGAGG--TGCT-CT-CGGCTGCTC-T-TT-AACAA-CAGTGGTGTCAG  567
CPZ.TZ.00.TAN1              A-AGAG-CC----T--GG-A-----AG-CA-A-TTG-GA-AATGA-A-T-......ACTGTACAGAAAAAT...AACTCCACAGCGACATCTAGTGGACAAAGACAGAATGCAGGTGAAAAAGAGG-A-C-GTGCCACCT-GT-GCA-TAC--G-AACA-A-GG-G--C-  787
CPZ.TZ.2001.TAN2            A-A-AG--C-T-----GG-A-----AG-TA-A-TTG-GA-A-TGA-A-T-AGC...........ATAAACAAAAACAACTCCGCAGCGGCATCTAGCGGACAAGGTCAGAATGCAGGAGGAAGGG-A-C-GTGCCATCT-GT-ACA-TACT-A-AATA-A-GG-G-AC-  783
CPZ.TZ.2002.TAN3_1          A-A-AG--C-------GG-A-----AG-TA-A-TTG-GA-A-TGA-A-T-AGG...........ATGCAGACAGACAACTCCGCAGCGACATCTAGCGGACAAGATCAGAATGCAGGAGGAAGGG-A-C-GTGCCATCT-GT-GCA-TACG-G-AAT--A-GG-G--C-  783
GOR.CM.2007.SIVgor2139_287  -T--TG--TC------TC--C----C-----A--AT-TA------TCATGTCA...........................................................................---AG-A--TC--ACAGCCGCCCT......  701
GOR.CM.2007.SIVgorCP2135con -T--TG--TC---T--TC--C----C-----A--AT-TA------CCATGTCA...........................................................................---AG-A--TC--ACAGCCGCCCT......  700
GOR.CM.2004.SIVgorCP684con  ----TG-CTC---T--TC--C----C-----A--AT-TA------CCATGTCA.....................................................................................AAAAGA-AGTC-  608
H2A.DE.x.BEN                AGA-AG-G-----T--TG-------AAAG...-AC---CAC-GAGA--TCTAGTG...................GCAGAAACTGGAACTGCAGAGAAA.............................AT--C-AATA--AG-AGAC--AC- 1477
H2A.GW.x.ALI                AGA-AG-G-----T--TG-A-G---AAA-...C-A---ATAC-GAGA--TCTAGCG...................GCAGAAATAGAAACAGCAGAGAAA.............................AT--C-AGCA--AG-AGAC--AC- 1470
H2A.SN.x.ST                 AGA-AGCG-----T--TG-A-----AAA-...C-A-AG-TAC-GAGA--TCTAGTG...................GCAGAAACAAAAACTACAGAAAAA.............................AT--C-AGTA--AG-AGAC--AC-  922
H2B.CI.x.EHO                AGA-AG-G-----T--TG-------AAA-...-------CAC-GCGA--TCTAGCA...................GCGGAC........ACAGAAAAA.............................AT--C---TATGAG-A-AC--A-T 1445
H2B.GH.86.D205              AGA-AG-G-----T--AG-------AAA-...-------CAC-GAGA--TCTAGCG...................GCGGAC........ACAGAAAAA.............................AT--C---TA--AA-A-AC--AC- 1444
H2G.CI.92.ABT96             AGA-AG-----C-T--TG-A-----AAA-...C-AG---T-C-GAGA--TCTAGTG...................GTAGAAACTGGAACTACAGAAAAA.............................GT--CT--CA--AGCAGAC--AT-  840
H2U.FR.96.12034             AGA-AG-G---C-T--TG-------AAAG...C------TAC-GAGA--TCTAGCG...................GTGGAAACTGAAACTGCAGAAAAA.............................AT--C---CA--AG-AGAC--AC-  956
SMM.SL.92.SL92B             -GA-AG-G---------AG--------AAG...--A-C--TAC-GAGC--TCTAGTG...................GTAGAAAGTGGAACTGCAGAAAAA.............................TT--C---TCA-AGCAGAC--AC-  867
SMM.US.x.H9                 AGA-AG-G---C-T--TG-A-----AAA-...C-A----T-C-GAGA--TCTAGTG...................GTGGAAACAGGAACAGCAGACAAA.............................AT--C----A--AG-AGAC--AC-  909
SMM.US.x.PGM53              AGA-AG-G---C-T--TG-------AAA-...C-A----T-C-GAGA--TCTAGTG...................GTGGAAACAGGAACAGCAGACAAA.............................AT--C-AT-A--AGCAGAC--AC- 1353
SMM.US.x.SIVsmH635F_L3      AGA-AG-G---C-T--TG-------AAA-...C-A----T-C-GAGA--TCTAGTG...................GTGGAAACAGGAACAGCAGACAAA.............................AT--C----A--AGCAGAC--AC- 1439

MAC.US.x.239                AGA-AG-G---C----TG-------AAA-...C------T-C-GAGA--CCTAGTG...................GTGGAAACAGGAACAACAGAAACT.............................AT--C-AA-A--AG-AGAC--AC- 1427
Gag                         E  K  V  K  H  T  E  E  A  K  .  Q  I  V  Q  R  H  L  V  .  .  .  .  .  .  .  V  E  T  G  T  T  E  T  .  .  .  .  .  .  .  .  .  .  .  .  M  P  K  T  S  R  P  T
MAC.US.x.251_1A11           AGA-AG-G---C----TG-------AAA-...C------TAC-GAGA--CCTAGTG...................GTGGAAACAGGAACAGCAGAAATT.............................AT--C-AA-A--AGCAGAC--ACT 1426
MAC.US.x.251_BK28           AGA-AG-G---C----TG-------AAA-...C------T-C-GAGA--CCTAGTG...................GTGGAAACAGGAACAGCAGAAACT.............................AT--C-AA-A--AG-AGAC--AC- 1415
MAC.US.x.EMBL_3             AGA-AG-G---C----TG-------AAA-...C------T-C-GAGA--CCTAGTG...................GTGGAAACAGGAACAGCAGAAACT.............................AT--C-AA-A--AG-AGAC--AC-  913
STM.US.x.STM                AGA-AG-G---C-T--TG-------AAAG...C-AG-G-TAA-GAGA--TCTTGTG...................GTAGAGACTGGAACTGCAAACAAA.............................AT--CT--CA--AG-AGAC--AC- 1083
MNE.US.x.MNE027             AGA-AG-G---C----TG-------AAA-...C------T-C-GAGA--CCTAGTG...................GTGGAAACAGGAACAGCAGAAACT.............................AT--C-AA-A--AG-AGAC--AC-  907
DEB.CM.99.CM40              -CTC-G-GG-T------G----G--CAA-A-GGCAG-GT-CAGGAG-AT-AAGCAGGTGGAACAGGAGGAAGAATTAGAAATGGCCCTAATCCAAAGACAGAAAGAGAAGGAAGTAGAACAGAAA...C--CAA---C--C-G-A---AAG-CAGCAG-AGCCTCAG-C-  928
DEB.CM.99.CM5               -CTC-G--G----T---G----G--AAA-A-GGCAG-G-CTAG-AG-AT-AAGCAGGCAGAGCAGGAGGAAGAGTTAGAGATGGCATTAATTCAGAAGAAAGCAGAAAAGGAGTTAGAGAAAAAG.....CAG--A-CTC-G........--------A-CAGACAC-  919
DEN.CD.x.CD1                -G--AG-GGC---T--AG----G--AC--AGAG-AG--AGAA-GAGAAT-GGTGCAAAAAAGAAGAACATAGTTATTGGCAATGCAGACCAAGGAGATGATCAGCCTCCACAGGGAGCCGCGGGA...GGTG-C-GCTCCGCG-A---AGG---CT-T--G-G-GG-A-T  922
LST.CD.88.447               --A-AG--GC-------C-A------A-GC-A------A-AAT---A-C-ACTCAGGAA.................................................................................................  333
LST.CD.88.485               --A-A---GC-A-----GCA------A-GC-A------A-AAT---A-C-ACTCAGGAA.................................................................................................  333
LST.CD.88.524               --AGAG-C--T--T--AC-------AA-TA-T--AG-GA-AATT-----GAACAAA....................................................................................................  330
LST.KE.x.lho7               --AGA---TC------TC-------CA-CAGT--AG--A-AAT---A-C-GCACCAGAAGCAGCAGGGAAG...AAA...................................................................................... 1423
SUN.GA.98.L14               -TCCAG----------AC----G---C--A-A---G-CAGACT-----C-GCAAAAAATGAAGAAGCACAA...GCT...................................................................................... 1410
MND-1.GA.x.MNDGB1           ----A--T--GAG---AC-A--T------A-A--AT--A-A-TCATAAC-AGAAAGGAAGAAAAGCAG.........GAGGATGAA.......................................................................  800
OLC.CI.97.97CI12            -TCCAG---G-A------TA--G--AAA-A-G-G-G-GA-A-TGA-AG--.......................................................................................................... 1029
WRC.CI.97.97CI14            -TCCATGT-G-A-T---C-A-----AAA---A---G-CA-T-T---A--GAAAAAACCA................................................................................................ 1380
WRC.CI.98.98CI04            -TCCCTGC-CCA-T---C-------AAA---A---G----A-TGA--G-TACAAAACAGAAAGAAAAGCCT.................................................................................... 1398
WRC.GM.05.Pbt_05GM_X02      -TC-AG---GT--T--TC-A-C---AA-TA-A--------A-TGAG--C-AGAGAAGAAAAGGCA..........................................................................................  342
MND-2.CM.98.CM16            AGA-AG---C---T--AG-------GAA---G--AG--A-AC--A---TCCATCTAGTGGCCGAAAAG.........GAAAATGCAGCATCAGAA...AAAGAACAAAGA...........................................G--AT- 1114
MND-2.GA.x.M14              -GA-AG---CT--T--AG-A--G--AAA-A-G--AG----AC--TGCTTCCATCTAGCGGCCAAAGGA.........GAAAGTGCAGCCTCAGAA...AAAGAAGAGAAA...........................................G--AC- 1038
MND-2.x.x.5440              -GA-AG---C---T--TG-------AAA---G--AG--A-AC--CG-T-CCACCTAGTGGTTGAAAGA.........GAGAATGCAGCCTCAGAA...GAAGAAAAAGGAGCA...........................................AC-  671
DRL.x.x.FAO                 -GA-AG---C------AG--------G-TC-A--AG--A-AC--C-CTGCCATCTA...................GTGGAT......AAA...AATGAGAATGCAGCT............................TCAA--AAT-A-AA--GGGA-AC-  668
MON.CM.99.L1                ----A--TG-G-----AG-AC-G---AAGC-GC-AG--A-AATCAG-TGTCACTTAGCAGGGGAACAGGGA...GAACAGAAA.........................................................GCT-----C--G-C-G-TCC-  819
MON.NG.x.NG1                ..............................................................................................................................................................    0
GSN.CM.99.CN166             --C--G-GG-G--T--AG--C-G---AAGC-GC-AG-CAG-AT-AGATGTCATCTA...................GAGGAGACCAAGGGT...AAAGAAAAGCAG.............................A-TA---ACC-CC-G-GGG-C-C-  832
GSN.CM.99.CN71              --TG-G--G----T--AG--C-G---AAGA-GC-AG-TC--AT-AG-TGTCATCTA...................GTAGGAACAAAAGAG...AAAGAAACACAG.............................A-CA---ACC-CC-A-GGG-C-C-  838
TAL.CM.00.266               ----C--TG-G--T--AG--C-G--CAAG-CAG-AG--AG-A--CGCTGCCACCTAGGGGAGAAAAAGGAA...ACAACAGAGGAGGAAAAA..................................................-T-AGAA-GG--AC- 1043
TAL.CM.01.8023              ----A---G-G--T--AG-AC-G--CAAG-CAG-AG--A--A--CGCTGCCACCCAGTGGAGAAAAAGGAG...GAGAGCAAGGGA.............................-C--C--AG---GG-----TCAGAAGTG--AT-  541
MUS-1.CM.01.1085            -GC-TG-GG-------AG----G---AAGA-GC-A--CAGA-C-AG-TGCCATCTA...................GAGGAGAAAAAGGGAGAAAAAGAA.................................................TC-  806
MUS-1.CM.01.CM1239          -G---G-GG-------AG--C-G---AGGA-GC-A--CAG-AT-AG-TGTCATCTG...................GCGGACAAGCAAGGGGAA...GAA.................................................AAG  805
MUS-2.CM.01.CM1246          --ACAG-G--------AG----G---AA-A-GC-A--TAG-AT-AG-TGCCACCTA...................GCGGAGAAAAAAGAAGACTCAAAAGGA.............................................---T-----  886
MUS-2.CM.01.CM2500          -GAGCG-GG-------AG----G---AAGA-GC-A--CAG-ATGAG-TGTCATCTA...................GGAGAGAAAAAAGAACAACAAGAACAG.......................................AGA--TCAGCC-  848
SYK.KE.x.KE51               A---CG-GG----T--GG-AC----AAA-C-G--AG-GA-A--G-CTTGCAATTGGGAAGATGAGGAAACA...GTGACATCTAGTGGCCAAAAGAAAATAGCAACGACACAGCG............-C-TC--GTGGC-G--A-GGAA--ATGCAGCTGCCAG---C-  863
SYK.KE.x.SYK173             AG--AG--G-------TG--C----AAA-C-G--AG--A-A---CATGTAATTGGAAAGATGACCCACCA...GCGACATCTGGTGGACAAAGTGAAAATAGCAGTCAAAACATG...GCTAGTG-G-C-TC--GTGGCC--AAGGTAGT-CAGCAG-AAA-ACA-AAG 1235
RCM.GA.x.GAB1               -GA-AG-G--------AG--------G---CA--AG-GA-A---TGCTGCCATCTAGTGGAAAAAGCA.........GAAAATACAACAGAA......AAA...GAAAAGGGAGCAACA..............................................  586
RCM.NG.x.NG411              -GA-AG-G-----T--AG-------GG-G-CGC-TG--A-AA-GC--TGCCATCTAGTGGACAAAGAT.........GAGAATGCAGGAGAA......CAA...GAAAAAGGAGCCACA..............................................  586
GRV.ET.x.GRI_677            AGA-AG-G--------AG--------G---TA-CAG-TA--C--C-CT-CCATCTA...................GTGGAC.....AAAAATGAGAAAGCAGCT...............................A--AAGAA-AA---G---AC- 1271
SAB.SN.x.SABIC              --A-AG-G-----T--AG-------AAA--CG--AG--A-A--G--AGTGCCAGCAGAAATGACAGAA.........AGTGCCACAGCGACATCTAGTGGCCAAACTAAG...................G--CTGCAGGCAA--AAGAA-AA---GC--AC- 1424
TAN.UG.x.TAN1               AGA-AG-G-----T--AG-------AG-G-TATTAG--A-AC--CGCTGCCATCTA...................GTGGAA......AAA...GAAAAAACTGCAGCA.........GCGCCATCTGGTGGCCAG------AATTACAACACAG-T-CG 1297
VER.DE.x.AGM3               AG--AG-G--------AG-A-----AG---CA-TAG--AGAC--TGCTGCCATCTA...................GTGGAG......AAA...GAAAGAAATGCAGAAAGAAAT-CA-C-GAG-CATCT-GT-G---AA-GAA-AAT-ACA-GGG--T-  835
VER.KE.x.9063               AGA-AG-G--------AG-------GG---CA-CAG--AGAC--TGCTGCCATCTA...................GTGGAA......AAG...GAAAAAAGAGCAACA.........G-GCCATCT-GTGG--A-A--A--AGTAACAGG--G---AC- 1337
VER.KE.x.AGM155             AGA-AG-G--------AG------------CA-TAG--AGAC--TGCTGCCACCTA...................GTGGAC......AAA...GAAAAAACTGCAGTT.........-CGCCACCTGGTGG-CAG---A--AATAACA-A-GAGG-AC- 1326
VER.KE.x.TYO1_patent        AGA-AG-G--------AG-------AG---CA-CAG--AGAC--C-CTGCCATCTA...................GTGGAA......AAA...GAAAAAAGTGCAACA.........G-G-CATCT-GTGG-CA-A--A--AAT-ACAAG-GA-T--C- 1334
COL.CM.x.CGU1               GGA-T--TG-G-----TC-A--G--AGA-A-G--AG----A----CTT-TAAAAAGCAGGCCATGATAGAA...ATGGCCAGCAAGGAGGAGGAGAAAGCAAAA..................................................  975
```

```
                                                  Gag p17 Matrix end_Gag p24 Capsid start
H1B.FR.83.HXB2            ...CACAGCAATCAGGTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTCAGCATTATCAGAAGGAGC 1325
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__E__K__A__F__S__P__E__V__I__P__M__F__S__A__L__S__E__G__A
H1A1.UG.85.U455           ..............AG---T-----------C--------A--TGCA--A-----CCA-----C------T----------G--C--G-----------G--------------C-----------------------A---------------------------G-----  780
H1B.US.90.WEAU160         ...A----A--C-------------------------------C-A-----A-------------------C--------------------------------------A---------A-------------------------------------------- 1324
H1C.ET.86.ETH2220         .........GGAA---A---T---------T-------------T--G---------------------C-------G----------G------------------G--A------------G--A-------------A----T-------------  699
H1D.CD.84.84ZR085         ...A------GC------------------T---------A------C-A--A--------------C------C-----------------G--C-------------A-------A----------A-----------A------------------------------  842
H1F1.BE.93.VI850          ................-----T-------------------A----TC-T-----A-----------C---T-TC------------------------G--GA-----------T----------A---------------------------------G--  668
H1G.SE.93.SE6165          ...A--------C--A--------------T---------------TGCA--A--------------C---------A---------------G--------------------A-----C----T------A---------------------------------  728
H1H.CF.90.056             ...A-AGA---CA-------T---------T---------A-----TGCT--A-----G---------C-----------------G--C---------------------------A-------T--------A---------------------------------  672
H1J.SE.93.SE7887          ...G---A--G---------T---------T---------------TC-G--A------CC------C------C------------------------------GA--------A--A-----------------------------------------------------  648
H1K.CM.96.MP535           ......GA---AGG------T-------------------A-----TC-G-----------------C------C----------------G--G---A----G----------------------A------------A--------------------------  524
H101_AE.TH.90.CM240       ...AG-----GCA-A-------------------------A--TGCA--A---------C-------C-TT----------------G-----------G--------------A----G---T-A--------A---------------C-------------G-----  890
H102_AG.NG.x.IBNG         ...........AGCAG------------------------A--TGC-A-A---------AC-------T----G-------------------------G-------A-------A----G---------------A----------------------------G-----  856
H103_AB.RU.97.KAL153_2    ...AG-----GCA-------T------------C------A--TGCA--A---------AC---C---T----G-------G----G---------G-------A-------A---------------------A----------------------G-----G-----  557
H104_cpx.CY.94.CY032      ...GGT-----GCA-T------------------------A--TGCA--A-----------------AG---T---------------G----------------A-------A---------------------A-----------C-----------------G-----  691
H1N.CM.95.YBF30           GCAACTGAT-GCA-TA----TAGG-----T---C----C----CTGCT--A--A---------------C-GC-GA----C-----C---------T------G------GA----G---------C--T--T--------A------A-----ATG--C--G--------G--  889
H1O.BE.87.ANT70           ACAAG-GCA-GG----CGG-T-------------------ATCA--TGCG-----A---------------------C--C--C--G-------------------------G-C-----------A--A--C--T-A---T---A-C--T--T------CATG-----G-----G----- 1371
H1O.CM.91.MVP5180         ACAAG-CCT-GG---ACA--T-------------------AACA--TGCA-----A--------------A------C--C--C--G-------------------------G-C-----------C--T-A---T---A-T--T--T------ATG---------------G-- 1346
H1P.FR.06.RBF168          GAGGG-----C-GGAACA--T-----C--T-----T--------TGC---------------ACC-----T-C-GC-G--T--C--G--GC-G-----C------G-G-C-----------------C----A---T---A-T--C--T------ATG--C--G---------GT-  783
CPZ.CM.01.SIVcpzCAM13     AGTGCTTC-TC-GGCACT--TGGG----------------A-----TGCT--A-----G----------ACTC----C---C--------C--------------------------G------------T-AT---T--G-A-----A-------------TC-G-----------  902
CPZ.CM.05.SIVcpzEK505     GCA-CTGAA-GCAGTA----TAGG--C--T---C-G--A--A--TGCT--A------------------C-GC-G-----C-----C---------T------------G----------A------C--T-AT--------A-----A-----ATG--C--G---------G--  873
CPZ.CM.05.SIVcpzLB7       AGGG---ATG-CAGT-GT---AG---C-------------ACA--TGCA--A------GCC-------C------C-C----------------------C------C---G---------------A------A--T------------A-----------CG----C---G-----------  871
CPZ.CM.05.SIVcpzMB66      ......-CT-GCACT-----TAGG--C-TT----------AGCA--TGCT---------G-------------AC------C--G---------------------------------------A------A--T------------A--T-----CATG--TC-G-----------  868
CPZ.CM.05.SIVcpzMT145     GGTG-A-CACCATCA-CT---GCG--C-T----G-----A-GA---GCT-----A--GT--ACG---A-TC-A---------C-----CC-------C-----------C---------------A--A-------AT-----G--C------A------CATG----------------G--  864
CPZ.CM.-.SIVcpzMB897      GAC-C----GGCGGC-AG---AG---C-T---G-------AACA--TGCT------A------C-----T-GC-C-----C---------------------------------------------A-------A--T----------------C---------ATG--TC-----------  870
CPZ.CM.-.SIVcpzDP943      CAG--AGA--GCA-T-----TAG---C-T----AG-----A-----TGCA--A------T--A-G------C-----G-----------------------T------T--G------G-------A---AA---T-AT----------------A------ATG--CC-G------G-----  884
CPZ.CM.98.CAM3            CAG--AGA--GCGCT-----TAG---C------AG-----A-----TGCA--A--A--GC-A--G-------C-----G-----C--G--C---------T-----T-------A--------A---AA---T-A-------G---------A-----CATG--CC-G-----------  715
CPZ.CM.98.CAM5            CAG--GGAA-GCAGCA-A--TAG---C------A------A-----TGCA--A--A--------A-------------G--------G--CC--------T-----C--G---A--------A---AAC--T-A-----------A------A-----ATG--CC-G--G--G----- 1018
CPZ.GA.88.GAB1            AGTGC-TCTGC-GGCA-T--TGG-----------CC-C--A--A--TGCT--A-----G------G-----------C------C------C--G--------------G------G--A-------A---A--T------T--G-A-----A-------------TC----------G----- 1386
CPZ.GA.88.GAB2            TCAA---C--C-G--A-A---AG-----------AG-G--------TGCA------A--G---A-----------------G------C-----------G--C------T---C----------A-A--C--T--------------C--T--A------CATG--T--------------G--  703
CPZ.US.85.CPZUS           GGC...--T--CATT-GT---AGC---------AG--A-A-----GCT------A--G--------------A--G--T--C-------------C-------------CG----------A------C--T-A---G----C--C--G------CATG--C--------------G-- 1373
CPZ.CD.90.ANT             CGA--TCTT-G-GGC-AAG-GAG---------C--CA-AGT-G-TGCAGGA--AAT-GCAAGG-------C-AC-GA----A------C---------C--------GTGT------------AAA------AT--------C--C--T--------T--T---------------G--  737
CPZ.TZ.00.TAN1            ACAG-G-CACC-AGT-GG--TAG-CTA-----AG-GA-AACTG-TGCA-----AGTTGCAAGG-------C-T---T-----------C--------C--------GG---A-------A--A-GG--T-AT-----------A------A--------C----------G---T--G-----  957
CPZ.TZ.2001.TAN2          ACAGCG-CACC-AGC-GA--TAG-TTA-----A---A-AACTG-TGCG-----AGTCGCAAGG-------C-TC-T------------C-C-----C---------AC---G--G--C--A-GG---AA---------A------A--------------C-----T-------  953
CPZ.TZ.2002.TAN3_1        ACAACG-CACC-GGT-GA--TAG-GTA-----A---TA-AACTG-TGCG-----AATCGCAAGG-------C-TC-T------------C-C-----C---------GAC---G--G--A--A-GG--T-AA----------A------A--------C-----GC-----T-------  953
GOR.CM.2007.SIVgor2139_287 ...........G-AACA--T--G--C--T--AG-G--AGCA--TGCT--A-----G-----C------C-GC-G--T--C--G--C-------C-----G-----CG-----------A------C----AT------A-C--C--T-----CATG-----------------GT-  859
GOR.CM.2007.SIVgorCP2135con ...........G--ACA--T--G--C--T--AG-G--AGCA--TGCT--A-----G-----C------C-GC-G--T--A--G--C-------C-----G-----C------------A------C----AT------A-C--C--T-----CATG-----------------GT-  858
GOR.CM.2004.SIVgorCP684con GACAG-C--CC-G-AACA--T--G--C--T--AG-G--AACA--TGCG--A-----G-----C------C-GC-G--T--A--G--C-------C-----------CG-----------A------C-----A------A-A--T--T-----CATG-----------------GT-  778
H2A.DE.x.BEN              GCA-CACCT-G-GG-AAA--AGG---C-----CG-GCAA--AGCGGGT...--CA-CTAT--C---GT-C-AC-GAGC--CC-----C--------------T----G--G--A---AAG---G-GG-------AG-G--AGGA---CAG---C-C--------CTG 1644
H2A.GW.x.ALI              GCA-CACCT-G-G-ACAGG-GGG---C-T---CG--CAA--AGTAGC-...--CA-CTACACC---GT-C-GC-GAGC--CC-----C---G----T---------------A---AAG---G-GG-------AG-G--AGGA---CAG---C-C--------CTG 1637
H2A.SN.x.ST               GCA-CACCT-GCGG-AA-G-AGG---C-T---CG--CAA--AGTGGC-...--CA-CTATACC---GT-C-AC--AGT--CC-----CC------T----------C--------G--A---AAG---G-GG-------AG-G--AGGA---CAG---C-C--------CTG 1089
H2B.CI.x.EHO              AAA-CA-CT-GC-GCC-A......GC---T--AG-GCA---A-TAGCT...--CA-TTATTCC--C-T-C-GC--AGC--CC-----C--------C--------T------------------AAG---G-GG-------AG-----AGGA---CAG---C-C--------GTG 1606
H2B.GH.86.D205            GCA-CACCT-GCGGC-GA.......-----T--AG-GCA---ACTGGCT...--CA-CTAC--C--C-T-C-GC--AGC--CC-----C--------T--------GT-------------A---AAG---G-GG-------AG-----AGGA---CAG---C-------------TG 1605
H2G.CI.92.ABT96           GCA-CACCT-G-GGCAGAG-AGG---C-----AG-GCA---AGTAGG-...--CA-CTAT--T----T-C-AT--AGT---C--------------------T-G---------A---AAG---G-GG---------G-G--AGGA---CAG--GC-G--------CTG 1007
H2U.FR.96.12034           GYT-CACCT-G-GGC-G---AGG---------GG-GCA-----TAGGT...--CA-TTAT-CC--C-TAC-TC--AGT--------AC-C------C--------C-G------G----AAAA--TG-AG-------AG-G--AGGA---CAG---C-C-----G--CTG 1123
SMM.SL.92.SL92B           GCT-CACCT-G-GGA-GA.......--C--T--AG-GCA---AGTAGGA...AATA-TTAT--G--CACTC-AC-G--C--AC-G--AC-C-----G-----C---C-G---------A---AAA--TG-AG-------AG-G--AGGA---CAG---C-C--------CTG 1028
SMM.US.x.H9               GCA-CACCT-G-GGCAGAG-AGG---------AG-GCA---AGTAGGT...--CA-TTATACC--C-TAC-TT--AGT--A-----A--------T---------T-G-----------AAAA--TG-GG---G--AG-G--AGGA---CA---GC---------CTG 1076
SMM.US.x.PGM53            GCA-CACCT-G-GGCAGAG-AGG---------AG-GCA---A-TAGGT...--CA-TTACACC--C-TAC-TT--AGT--A-----A--------T---------T-G-----------AAAA--TG-GG---G--AG-G--AGGA---CA---GC---------CTG 1520
SMM.US.x.SIVsmH635F_L3    GCA-CACCT-G-GGCAGAG-AGG---------AG-GCA---AGTAGGT...--CA-TTAT--C--C-TAC--T--AGT--A-----A--------T---------T-G-----------AAAA--TG-GG---G--AG-G--AGGA---CA---GC---------CTG 1606
                                                  Gag p17 Matrix end_Gag p24 Capsid start

MAC.US.x.239              GCA-CATCT-GCGGCAGAG-AGG---------AG--CAA--A-TAGGT...--TA-CTAT--C--C-T-C-AT--AGC--G-----A--------C---------T-GA---G--A---AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------TTG 1594
Gag                       _A__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__E__K__K__F__G__A__E__V__V__P__G__F__Q__A__L__S__E__G__C
MAC.US.x.251_1A11         GCA-CATCT-GCGGCAGAG-AGG---------AG--CAA--A-TAGGT...--TA-CTAT--C--C-T-C-AT--AGC--G-----A--------C---------T-GA---G--A---AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------TTG 1593
MAC.US.x.251_BK28         GCA-CATCT-GCGGCAGAG-AGG---------AG--CAA--A-TAGGT...--TA-CTAT--C--C-T-C-AT--AGC--G-----A--------C---------T-GA---G--A---AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------CTG 1582
MAC.US.x.EMBL_3           GCA-CATCT-GCGGCAGAG-AGG---------AG--CAA--A-TAGGT...--TA-CTAT--C--C-T-C-AT--AGC--G-----A--------C---------T-GA---G--A---AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------CTG 1080
STM.US.x.STM              GCA-CACCT-G-GGCAGAG-AGG---------AG--CA---AGTAGG-...--CA-TTAT--C--C-TAC-AT--AGC--A-----C--G-----------GT----G--G--A---AAA--TG-TG--------G-G--AGGA---CAG---C-----------CTG 1250
MNE.US.x.MNE027           GCA-CATCT-G-GGCAAA--AGG---------AG--CAA--A-TAGGT...--TA-CTATACC--C-TAC-AT--AGC--G-----A--------C---------T-GA---G-----AAA--TG-AG-------AG-G--AGGA---CAG---C-G--------CTG 1074
DEB.CM.99.CM40            GCA-C-CAG-CAACACCATCAGTT--C---------CT-AAGAC-GGGG...CAA---GT-T--G---AT-C-AT-AGC--C--G-TAG----AA-C---A----T-CT-----G--A---AAA--TTCT-----GA-AG-T--TC-T---CA--TC--GG----G--CTG 1095
DEB.CM.99.CM5             CAG--A-AA-CAA-CCAACCAGCT----TT--A--CC--AGAC-AGGA...CA----TAT------AT-C-TT--AGT-----G-T-G-G--AA-C---A----T-C---G--G--A---AAA--TTC------GA-AG-G---C-A---CA--T---GG-------GTG 1086
DEN.CD.x.CD1              TCCTTT--AG---CAAATTCT--------T--A--T--CAG--CAGCT-----A---TATCAG------A---T--AGTA-C--GGTG-----AA-----A-T----AG------G--C-GAAAA--TTCG-----GA-AG-G--AC-A-----G-TCC-CG-C-----GCT 1092
LST.CD.88.447             ...G-A--GC-G-CTAAA--TGGC-----T--AC-C--CAGAG-A-AT...CA-AG-TG------CACAC-AT-G------------AA--C-AA-T----------A----------T-GA-GA-GG-AG------ACAG-GG-A---------T--CC-CA-----AA---  497
LST.CD.88.485             ...G----GC-G---AAA--TGGC-----T--AC-CA-CAGAG-A-AT...CA-AG-TG------CAC-C-AT-G------------AA--C-AA-T----------A----------T-GA-GA-GG-AA------ACAG-GG-A---------T--CC-CA-----AA---  497
LST.CD.88.524             ......CCTGTG-CA-AA--TA------TT---T-GA-TAGAG-A-AT...CAAAG-TG---C---ACCC-TC-------C------AA-TC-AA------G----A----------T-G--GA-GG-AG------ACAG-GG-T---------------TC-GA-----AA---  491
LST.KE.x.lho7             .........C-A---ACAG-GGG------T--CC-GA-AAGAG-A-AT...CA-AG-TG------ACAC-TT----G--C------AA--C-AA------G----A----G------C-GA-G--GG-AA------ACAG--G-A---------------C-----A-----AA--- 1581
SUN.GA.98.L14             .................TCT-------TT--AG-GCA-AGAG-GGGA...CAAA-TTACA-T------C-TC-C--G--------AG-GC-AA-------G--GA----G--------------GA-GG-AA------ACAG-GG-A---------G----T---ATGAC------ 1559
MND-1.GA.x.MNDGB1         ..................-TA-G----T----G--CAAAG-G-TGCAGCA--A--GTATCAGT---ACTC-A---AGT------G-T-A--C-GA--------------AC----G-------A---AAG-GG-AA--G--G--C--C---TC-A--C------------GA----------  952
OLC.CI.97.97CI12          .......AG--AG-A-AGCAGTT-GCCATG--A---A-CAGA--TGCT-----AATCT-T-AG---GT-C-GC-GG-TA-------CC-GC-AAGT-----------C-T-G-----C-G---G---T-AAT-----A-AG-G--TC-A------T--G--TG-T--G---TT 1193
WRC.CI.97.97CI14          ..................TCT--G---ATG--AG-GC--AGA-CAGCA...CAAGGGCCTCA-TGGG-CC--T-G-------------AA--C-AA-------G---CATA-T--G--T-G--GG--TG-AG-----AGTG-T--AT-C---G-----TC-TA----CA-G-- 1529
WRC.CI.98.98CI04          ..................TCT-----CATG--AG--T--AG--CAGCT...CAAGG-CCTCA-TGGG-TC-TC-G--T--A-----CA--C-GA-T------G---CACA-----G--T-GA-GC--TG-AG-----TGTG-T--AT-C------T--T--GA----TA---- 1547
WRC.GM.05.Pbt_05GM_X02    ..................TCAGTG---ATG--AG-CC--TAGAGGACCA...CAAGG-CCACA-TGGG--C-TC----C--G--G--CA-TC-AA-----------CATA---------T-G--GA--TG-AG-----TGTG-T--AT-C------T--T--GA----TA----  491
MND-2.CM.98.CM16          GTGACACCT-G-GGCCG-TCAA-----------A---CA-ATA-TA-AT...CA-ACCCCA--C--------GA--T-----GC-C--GC-G-----G-----------TGCA-------------AAG--T-----------A-AG-G---------CAT---C--G--T------GTG 1281
MND-2.GA.x.M14            GCGACACCT-G-GGCCG-TCAA-------T-----TCA-ATA-TA-AT...CAAAC-CCA--G-----A-GA--T------A--------------C-----C---TGTA-T--G--A---AAA-----T------A-AG---T------CAT---CC-------------TG 1205
MND-2.x.x.5440            GCGACACCTGC-GTTCG-TCAA----------C---TCA-GTA-TA-AT...CAAAC-CCA-------------GA--T-----A---------C--G-------------TGTA----G-----AAAG--------------A-AG-G---T------ATT---C--G-----------TG  838
DRL.x.x.FAO               GCGACATCT-GCGGCAGGTCAAG---C-----C--TCA-GTTGTT-AT...CAA--GGCA-----C------T--T-----A-----C----------C------G--G---A-T--G--A---AAG-----TG-----------G-G--T------C-TG---CC-G----------CTG  835
MON.CM.99.L1              CCAACAG--GGCGTACC---TGG---C------G-C---AGA-CACAGGG---AGGCT-CCA----------AG-GGAG---C-GCTCC-C--GA---------C-G--CA-T--G------AAA--TGC------------G-GG--C-C--CCAG--TC-----------CAT  989
MON.NG.x.NG1              ..................................................................................................................................................................    0
GSN.CM.99.CN166           GGGGGACAGGCGGTA--T--T-----C------G-GA-TAG---TGC---A--A--GTATCA-----------C-CAATAG-----T-C-T--GA-T--------GCA----G--G-----AAAG--TGCT-----GACAG-TG--T-------CAG---T-------------CTG 1002
GSN.CM.99.CN71            GGGGGACAGGCGGTA--T--T--G--C------C-----TAGA--TGCT-----A--GT-TCA---C-------C-CAACAGC---GTCC-T--GA-T--------GCA----G------A---AAG--TGCT-----GACAG-CG-TC------CAG--TC--A--------CTG 1008
TAL.CM.00.266             GTGACATCT-G-GGCCAA-CAGG---------C---A-AAGA--TGCA---CA---GTATCA----------AC-G--------GCTGC----GA------A-TGCCACTA----G-----AAAA--TGC---T---ACAG-GG-GC-C--C--G---C--G-G------GTG 1213
TAL.CM.01.8023            GTGACATCT-G-GGCCAT--TGGG--------C---TA-AAGA--TGCA---ACAA--GTATCAG----------T-G------C----CTGC-G--AA------A--GCCACTA--------A---AAA--TGC--------GACAG-GG-GC-T--C--T----C---G-----------TG  711
MUS-1.CM.01.1085          CAAGCGCAGGCAGCT...--TGCC--C------AG-CA-AAG----GCA--A------GT-CCA-----------C-CAATG-------TA------AA-------GTC-A----G--G--A---AAA--TGCTG--------TG-GG-A----------CAG-----GG----G--G--  973
MUS-1.CM.01.CM1239        ACAGCACAG-----AACT--TGCT--------C---CA-AAG-----CCA------A---T-TCA---C----T--C-TAATG------TC------GA-------GTC-A----G------A---AAA--TGCAG-------AG-GG-A----------CA----C---G-----------A-  975
MUS-2.CM.01.CM1246        GCAA-AGC--CCACCAGTG----G--------AG-CA-AAGA-CAGCA------A--GTATCAG------T-TC-TAGC------GCTA------AA-C-------TC-AC------G------AAAA--TGC---------AG-GG--C-----CCAG--CC-------------TG 1056
MUS-2.CM.01.CM2500        TCAGGAGCAGGGGCAAG----------C------AG-GA-TAG--CAGCT--A---ACCTATCA--------T--C-CAGC--C-----T-------GA-C------GTCCAC-A-T----------AAA--TGC---------AG-GG-TT-----CCA----C-------------TG 1018
SYK.KE.x.KE51             ATG-CATCT-G-GGC-G-TCAGG-----------C--A-TAGA---CCA--AAAT---TG-A-T---GTA-GGG-CAACA-C--G---------AA-C-------G-G-C---GA-CAGC--AAAA--TGATG-CTC-C--G-C--AC-A---CCA-ATT---A---------TTT 1033
SYK.KE.x.SYK173           GCAGCA-CACCG-CACCT--AGG---C--T--AC-GC-CAGA---TCCT--AAAT---GTG-A-T---ACA-GAG-CC--GT----G-----C-----AA-------GG-G-CT--GA-CTCC--AAAA--TGATG--TCCA-AG-G--AT-A---CAG--TC--A---------TTT 1405
RCM.GA.x.GAB1             GCG-CACCT-G-GGACAA--AGG-------T------A-TACT-TA-AT...CA---GCCT-AG---A-TC-T----------A--G-----C--------C-----C--G--G------G-----AAAA---TCAG--------AGCG---------C--G---C-----------CTG  753
RCM.NG.x.NG411            GTGACATCT-G-GGACAA--AGG-----------A-----AACA-TT-AT...CA---GCCT-A--------C-A-----T---------A--------T-----------------G--A---AAA--TG-AG--------AG-G--T-----------G----C-G---------CTG  753
GRV.ET.x.GRI_677          GCG-CACCTGG-GGC-AATCAAG---------AG----AA-TC-G-AT...AATGCCTG------C---C-TT-G--T--GC-C--G----------G------C----TGC--G--G--A--AAGG-GGG-AG--------AG-C----------CCA----C-C------G----TG 1438
SAB.SN.x.SAB1C            GTGACACCT-G-GGT-G-TCG-G---------C--T---CAGTGTT-AT...AAT---TG---C--C---C-AT-G--C--A--G--AC-T-----T------------GA-T--G--A---AAG-----TG----G--AG------------AGC--CC--G----------- 1591
TAN.UG.x.TAN1             ACA-CT--TGGCAGA...CATGGT----------G----A--AC-G-AT...AAT---TG--------ACTC-TT-G-----AC-T--CC--------------T---AC---T--G--A--AAGG---G-GG-----A-AG----------CCA---CC-C----------GTG 1461
VER.DE.x.AGM3             ACAGTGCCACC-GGT-G---T-------T---AGC-CAA---AC-GGGA...AATGC-TG-A------GT-C--T-G------AC-C--C--------G------------C-------G------AAAA---G-AG------A-AG-G-----------CCAG--T-----------GTG 1002
VER.KE.x.9063             GCG-CACCTGG-GGC...--T-------T---AGCGCAA---AC-AGGA...AATGCCTG--------GTTC-AC-----C---C-C--A----------G----------C--------G--A---AAA--TG-AG----GA-AG-G--A-----CCA----C-C----------GTG 1501
VER.KE.x.AGM155           GCGACACCTGG-GGC...----------TT--CGC-CAA---C-AGGG...AATGC-TG---G---GTAC-AC-T------C-C--CC------------------C----------------AAAA--TG-GG-----A-AG-----------CCA---CC-C----------CTG 1490
VER.KE.x.TYO1_patent      GCG-CACCTGG-GGC...--T--G----TT--AGCGCAA---AC-AGGA...AATGCCTG--------GTAC--T-G-----GC-C--C---------G-----------C-------G-----AAAA--TG-AG-----A-AG-----------CA---CC-----------CTG 1498
COL.CM.x.CGU1             .......-AGG-AGCA-AG-AATTGG-CATG--A------AACAGGGCCT...CAAGGGCCA--G------AC--C-GAGC--C--G--G---GGA--C-----T--GTGT------G...GGA-GAA-TGCT--TTCCT--GCT---------C-T---T-AT--TACG--G-- 1133
```

```
H1B.FR.83.HXB2            CACCCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAATGGGATAGAGTGCATCCAGTGCAT...GCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTG 1492
Gag                       __T__P__Q__D__L__N__T__M__L__N__T__V__G__G__H__Q__A__A__M__Q__M__L__K__E__T__I__N__E__E__A__A__E__W__D__R__V__H__P__V__H__.__A__G__P__I__A__P__G__Q__M__R__E__P__R__G__S__
H1A1.UG.85.U455           ------------------T-TG-----G--TGT-----------C--G-----T-----------------T-----------------------G-----C--GT-A-----------...-----------C------------------------------------  947
H1B.US.90.WEAU160         ---------------------------------------------------------------------------------------------------T-A-----------...-------------------------------------------------- 1491
H1C.ET.86.ETH2220         ---------------------------T-------------------------------------------C-------------G------------C--GT-A-----------...--------G-------------A--------C------------  866
H1D.CD.84.84ZR085         ----------------------------------------------------------------------------------------------C--GC-A-----------G...--------G-------------------------------------- 1009
H1F1.BE.93.VI850          ---T--CAC------------------T-------------------------------------------C-----------------------C---T-A-----------...--------GCCC---------------G-----T-----G----  835
H1G.SE.93.SE6165          ------------------TCT----T---------------G----------T---------C----G--T--T-------------------G-----C--GA--------CAA--G...---------T--C-------------A-----------------  895
H1H.CF.90.056             -----------C-----TG-T--------T--------------------------G-----------T--A-----------------------C--G--A-----------...-----------C----------A---------------------C-  839
H1J.SE.93.SE7887          ------G-----------T---------------A-A--------C--------T-------------T--T-----------------------C--G--A--------A----...----------------------G------------------------  815
H1K.CM.96.MP535           ---T---------C---------------------------------------------------T--T--------T-----------------C--GT-A--C--------A---...-----------CC----------A---------G-----G----  691
H101_AE.TH.90.CM240       -----------------T-TG--------T-T-----------C--G-----A-----------------A-----------C-----------------G--A--C-----A---...-----------C---------------G-----------------  1057
H102_AG.NG.x.IBNG         ------G----------T-TG----------T-----------C--G-----A-----G-----------T-------------------------G--A-----------A---...-----------C---------------G-----------------  1023
H103_AB.RU.97.KAL153_2    -------------------TG-----G---T------------C--G-----T----------G-----T------------------T------GT-A--------CA--G...---------T--C------G-----------------------------  724
H104_cpx.CY.94.CY032      --------------G----TG-------T-T------------C--G-----A-----------------T-----------------------C-----C--GACA--------A---...-----------C-----------------------------------  858
H1N.CM.95.YBF30           A--G--CTC-----C----T--T----T-----T-----A----------G-----A-----G---C-G--G--AGT---------------A-----C---------GACA--------C-C-...-TG--A--AC-AC-C-----G--AC-------C--T--A--------  1056
H1O.BE.87.ANT70           A-TTT-CT-T---A-T--T--T--------TG-CA-A--A---------G----TT-A---G--C---G--G--AGTA----------------A-T---G----------ACT--C---CCA-CG...-T------GT-GC-------G-----A--G--------CA------- 1538
H1O.CM.91.MVP5180         TGT---CT-T---A-C--T---------G--TG-CA-A---------C----GG--TT-A---G----G--G--AGTA----------------A------------------ACT------CCAGCA...ATG------GT-AC-------G-----A--G--------CA------- 1513
H1P.FR.06.RBF168          T-TA--CG-T--CA-T--T--T-----C--TG---A--C---------G-G----AT----GG---G-----GTG--T------A--G------------------ACA-----T-CT-C-...-T---A--AT-AC--------G---C--T----------CA---G-G-  950
CPZ.CM.01.SIVcpzCAM13     TCTA--TTTG---G-T-----A-----G--TG-T--A-------------G-G-------------G-TC---G----GT---------A------------------C-GA-------T--T----...--C--C--C--A--T------A----C-A----------C----T---- 1069
CPZ.CM.05.SIVcpzEK505     A--G--C-----C----------T-----T-----A-----------G-----A-----G---C-G--G--AGTT--------------A------G------C--G-GA--------A---...ATG------A--AC-------A--AG-A--G--G------A-------- 1040
CPZ.CM.05.SIVcpzLB7       ------T------------T----------T-----A-----------C-G---------------G--C-----------GT---T-------------ACA-------------GC------TA-C---...ATG------A--A-------------A----------------------- 1038
CPZ.CM.05.SIVcpzMB66      ------T------C---------------------------------C-G-----A-----------C-------AGTA---------------------G-----C--GC-C---------T---...------A---G--C-----A--A------------------------- 1035
CPZ.CM.05.SIVcpzMT145     A--G--T-------CC-C--T--T---------T---A-A--A---------G-----------------TC-------AGTA-------T--G------T--G-----C--GACA--C--TACC---...-----------A--T-----A--A------G-----------A--C---- 1031
CPZ.CM.-.SIVcpzMB897      ------T--G---------------------T---A-A-------------G-----A--------------C-------AGTG---------------------G-----C--GACA--C--TA-A---...--------A---C-T------A--A---------G---------A------ 1037
CPZ.CM.-.SIVcpzDP943      G--T--C--G----G-C--T--T-----T--TG-TA-A--A--G--C--G-G---T-----GG------G--AGT----------------ACTG---------C--GACC------T-------...-----T--A--A--------A--AC------GG--C-A------- 1051
CPZ.CM.98.CAM3            A--T---------G-T--T--T-----T--TG-TA-A--A--G--C--G-G---T-----GG-C--G--G--AGT-----------------A----------------ACA--C--T------...-----A-AG-A--C-----A----C-A------GG--C-A-------  882
CPZ.CM.98.CAM5            T--T--C--G----G-C--T--T-----T--TG-C--A--A--G--C--G-G---T-----GG-T-----G--AGT-----------A------A-----G--------GC-A--C--T------...-----A--A--C-----A--AC-------GG-----A--G------ 1185
CPZ.GA.88.GAB1            -TTA--T--G---G-T---------T-----TG-----A-------------G----------G---C-----G--AGT---T-----------A-----C--C--G------C-CT-A---CACT---...-T--T--G--A--C-----A---C-A-G--G------T---- 1553
CPZ.GA.88.GAB2            A--A--C------G-G-----A-----G--TG-TA-A--A-----C---G---T----GG-AC-------GTA---------A-----A--G--G---------T-A--C--T-A---...-----A-AG-G--------T--A-------G------A--C----  870
CPZ.US.85.CPZUS           G--T--C------G-C-----T-----C--TG-TA-A------------G-G---A------G-T-----G--AG---------------------A-----T--A------ACA-----TC------...------C----A--------G-------------C--------- 1540
CPZ.CD.90.ANT             A--T--T--T-----------------T--TG----T----AC-------G-------GG-C-------AGTA------------------A--T--G--------GT-A--C--CACT---...-----A--AG-ACAGG----A--AT-A--G------CA------  904
CPZ.TZ.00.TAN1            A-----TT-T---C-----T-GT-----C--TG-T--T----A-----------A-----------G--G--AGT----------------A-----G-----C-----CA-----C-CT---...-----A--CCAGCA-G-----GAT-C-A-----G--C-CA--GGCA- 1124
CPZ.TZ.2001.TAN2          A--T--TT-T-----------T-GT-----C--TG-TA-T----A------G-----A-----G--A--G--G--AGT----------A-----A-----G----------CA-----Y-CT---...-----A--G--GCA-G-G--GAT-C-------G--TTCA--GGCA- 1120
CPZ.TZ.2002.TAN3_1        A--T--TT-T-----------T-GT-----C--TG-TA-T----AC------G-----A-----G-----G--G--AGT---------C--G--A-----G-----------CAA-----C-CT---...-----A------GC--G-----GAT-C-------G--T-CA--GGCA- 1120
GOR.CM.2007.SIVgor2139_287 T-T---TT-T---A-T--T--T-----G--TG-CA-T--A---------G-GG--AT-A--GG-A-----G--AGT----------------A-------------C--GA----C--G-CAGTG...-TG--A--C-GC--GTG--A--AC----G--G--GTCA--GG-A- 1026
GOR.CM.2007.SIVgorCP2135con T-T---TT-T---A-T--T--T-----G--TG-CA-T--T---------G-G---AT-A--GG-A-----G--AGT----------------A-------------C--GC----C--G-CAGTA...-T---A--C-GC--GT---A--A-----G--G--GTCA--GG-A- 1025
GOR.CM.2004.SIVgorCP684con T-T---TT-T---A-T--T--T-----G--TG-CA-T--A---------G-GG--AT-A--GG-C-----G--AGT----------------A-------------C--GCA---C-----CAGTG...-T---A--C-GC--GTG--A--AC----G-------TCA--GG-A-  945
H2A.DE.x.BEN              ---G--CT-T---A-T--TCAA-----T--TTGT-----C-AT--C---------T---------AA-C-G-----TT--T------A-----A-----C-------TCGCA---C---A-A-CA...---C--CT-AC---G----A---C-C------C------A--GTC-- 1808
H2A.GW.x.ALI              ---G--CT-T---A-T--TCAA-----T--TTGT-----C-AC--------------------AA-C-GG-----TT---------A-----A-----C-------GTT-CA-----CA-A-CA...---C--CT-AC---G----G---C-T------------A--GTC-- 1801
H2A.SN.x.ST               ---G--CT-T---A-T--TCAA-----T--TTGT-----C-AC-------------T---------AA-C-GG--A-TT--T------A-----A-----T-------CGC-CAA--C---A-A-CA...---C--CT-AC--G-G--G----C-C--G--G---------TC-- 1253
H2B.CI.x.EHO              ------TT-T---A-T--TCAG---T---TTGT--A--A-A-------G-------------TA-T-GG---TA--T------------A-----C-------CCA-CA--------TC--CA...---C--A--GC--G-----G--AC-C--G------A--GTCA- 1770
H2B.GH.86.D205            ------TT-T---A----TCAG-------TTGT--A--A-A-------G-------------TA-T-G---A-TA--------------A-----C-------CCAGCAA--C--GTCA-CA...---C--A--GC-GG----A--AC-T--G--C------A--GTCA- 1769
H2G.CI.92.ABT96           -CTA---T-T---A-C--TCAA-------TTGT--A--A-A-------------------------TA-T-G---A-TA--T--------G--A-----K-------CTTCAA------CAA--A...-----GCCGC--G-----A--AM-C--G------G-CA---TCA- 1171
H2U.FR.96.12034           ---G--CT-T---A-T--TCAA---T---TTGC-----A-A---------T--T--------TA-C-GG--A-TA--T------T------A-----T-------C-C-CA---C---AAC--A...---C--ACCAC--G-----G--AC-T-----G------A--TTC-- 1287
SMM.SL.92.SL92B           T--A--CT-T---CA-T--TCAG-----T--TTGT--A--T-A------G-------------TA-C-G---A-TA--T--------------A-----C-------CCT-CAA--C---AGAGGACAACAGCC-G-GCAAC-GG-T--AGG-C-A--G------TCA---TCA- 1198
SMM.US.x.H9               ---T--CT-Y---A-C--TCAA-------TTGT--A--A-A------G-------------TA-T-G---A-TT--A------A--G----T--T-------TT-CAA-C--RCA--CA......-T--A--AC--G-----A--AC-T------G------R---TCC- 1240
SMM.US.x.PGM53            ---T--CT-T---CA-C--TCAA-------TTGT--A--A-A------G-------------TA-T-G---A-TT--A------A---------T--T-------TTGCAA-C--GCA--CA......-T--A--AC--G-----A--AC-C------T------A---TCA- 1684
SMM.US.x.SIVsmH635F_L3    ---T--CT-T---CA-C--TCAA-------TTGT--A--A-A------G-------------TA-T-G------TT--A------A----------C--T-------TT-CAA-C--GCA--CA......-T--AC-AC--G-----A--AC-T------G------A---TCA- 1770

MAC.US.x.239              ------CT-T--CA-T--TCAG---T---TTGT-----A-AC---------G--T-----G--TA-C-G---T-TT--A--C-----G--------T-----CTTGCA---C---CAA-CA......-CT--A...CA--A---A--AC-T--G--G--GTCA--TCA- 1755
Gag                       __T__P__Y__D__I__N__Q__M__L__N__C__V__G__D__H__Q__A__A__M__Q__I__I__R__D__I__I__N__E__E__A__A__D__W__D__L__Q__H__P__Q__P__.__.__.__A__P__.__.__Q__Q__G__Q__L__R__E__P__S__G__S__
MAC.US.x.251_1A11         ------CT-T--CA-T--TCAG---T---TTGT-----A-AC---------G--T-----G--TA-C-G---C-TT--A--C-----G--------T-----CTTGCA---C---CAA-CA......-CT--A...CA--A---A--AC-T--G--G--GTCA--TCA- 1754
MAC.US.x.251_BK28         ------CT-T--CA-T--TCAG---T---TTGT-----A-AC---------G--T-----G--TA-C-G---T-TT--A---------G--------T-----CTTGCA---C---CAA-CA......-CT--A...CA--A---A--AC-T--G--G--GTCA--TCA- 1743
MAC.US.x.EMBL_3           ------CT-T--CA-T--TCAG---T---TTGT--A--A-AC---------G--T-----G--TA-C-G---T-TT--A---------G--------T-----CTTGCA---C---CAA-CA......-CT--A...CA--A---A--AC-T--G--G--GTCA--TCA- 1241
STM.US.x.STM              T-----CT-T--CA-C--TCAAC--T---TTGT--------AG---------------------TA-------------T---A----------------T--T-----C-TGCA------CAA-CA...C-G--T--GC-GC---G----G--AC-T------------C--GTCA- 1417
MNE.US.x.MNE027           --T---CT-T--CA-T--TCAG---T---TTGT-----A-A-----------T-----G--TA-C-G---A-TT--A--C-----G-------T-----CCTGCA---C---CAA--A......-CT--A...CA--A---A--AC-T--G--G--GTCA--TCA- 1235
DEB.CM.99.CM40            T-----CT-T--CA-T--TGGTC-A--T--TG--A-T--A-AT-T---G-GG--T-----G--CA-T--------TGTA--T------------A-----G-----TT-CAA--C---CAA--A...CAGCC-.....C----AG--G----C-T--G-------TCA---GCA- 1256
DEB.CM.99.CM5             T--T--CT-T--A--T--TGGGC------TG-CA-A--A-AT-TA----GG--A--------TA-------TGTG--T------------A-----G-----CTT-CAA--C---CA---A...CAGCCA......C---A---T---C-T--G-----TCA---GC--- 1247
DEN.CD.x.CD1              G-TA--TT-T---A----TCAAC-C-----TG--A-T--A-A----C----G-----------------G------------TGT---T-----A-------CAT--T-------TTG-CT-----G-CTA--TTACA-CA---GC--C---------CAGCACAA--G--C-----A--TTCA- 1262
LST.CD.88.447             TTTG---G-T---C-----TGT-----T---TG-CA-A--A-AC-------GG--A-----G--AA-T-----TCAT--TGTG--A----GG------------C--GAAA-----GCAA--A...CA-CCTG-ACAAC--GG---AGG-C-T---ACC--T-AT-C--CA-  664
LST.CD.88.485             TTTG---G-T---C-----TGT-----T---TG-CA-A--A-AC-------TG--A-----G--AA-T-----TCAT--TGTG--A----GG------------C--GACA-----GCAA--A...CA-CCTG-ACAAC--GG---AGG-C-T---ACC--T-AT-C--CA-  664
LST.CD.88.524             -TTG---G----C-----TGT-----T---TG-TA-A--A-AC-------GG--A-----G--AA-------TCAT--AGCG--A----GA--T-------------CA-----CCA---G...CA-CCAG-ACAAC--GG---AGG-T-A---ACT----AT-C--CA-  658
LST.KE.x.lho7             TTTG--TG-T-------TGTT---T---TG--A-T--A-AT-----G-GC--------------AA-------TCA---AGTG------G------G---------CA-----C--GCAA--A...CAGCCAG--CAAC--GG---GGGTT-A---ACT--T-AT-C--CA- 1748
SUN.GA.98.L14             A-TT---G----CA-C--TGTG-----G--TG-CA-A--A-----AG-------G---T-----G--A-------TCAT--AGTG------AG----------------C---A-----C--TCA---G...CAGCCTG-ACAGC--GG---AGGCC-A--GACT--T-AT--T-CA- 1726
MND-1.GA.x.MNDGB1         A-T-AGT--T-----G--T-T-------G--TG---A--A-AT-----G-G---A------G-C-----------TGTA--T-----C-----A----------CT-ACA-----TCAA--A...CA-CCAG-ACAAC--GG---AGGAT-A--GACC--TTCA--CTC-- 1119
OLC.CI.97.97CI12          GCTG---G----GA-C--T--C--T-G--TGT---T----ACT-------GG--T--------G---G-G--AG-------------------A-A---G-AT--C--GCA-------CA--CA...AT------AT-GC-GGTG--A--A--TCC-CCG-------CT---- 1360
WRC.CI.97.97CI14          ATTAT-T--T---A----T---C--T-------ACA-A--A-AT--------T--A----------AA--------C-TAG-A------A---ATACA---G----------ACA-------CA-GCA...-----------A---------A---A--------C-G------CA---TCA- 1696
WRC.CI.98.98CI04          TTTGG-T--T--CA-T-------C--T------T-ATA-A--T-AT------G--T-------------A------A-TAG-A----------AT-CA--C---------ACT--C---CAA-CA...----------A---A---------A--A------GC-G-----CA--GTCA- 1714
WRC.GM.05.Pbt_05GM_X02    ATTGA----T--CA-C--T--TC-AT-G--T-ATA-A--A-AT-----------------G--AG-T------A-TAG-G-------A---ATAA----------C---ACA--------CCA-CA...-T---A---C-GC-C-----GGCT------A-T-----CA---TCA-  658
MND-2.CM.98.CM16          -CT---CT-T---C-C--TGG-------C--TG-TA-T--A-AT-----G-G---GC-T-------AG-G--G--TGT----------------------C---A-CTTGAGA--C---CAAGTG......-T-----GC-T-A---TGTCT------A--C-----C-----C----- 1445
MND-2.GA.x.M14            -CT----T-T---CC-C--TGGT---T---TG-TA-T----A-------G-G---TC-C------AG-G--G--TGT----------------------------CTTAGA--C--CCAAGTA......-A----T-GC-T-A---GGT-C-----A--C-----CA------- 1369
MND-2.x.x.5440            --T---GT-T---CC-C--TGG------C--TG-CA-T--A-AT-----G-G---TC-C------AG-G--G--TGT----------------------------CTTAGA--C--CCCTGTG......-A-----GC-C-A---AGT-C-A---A--C----CA-----A-- 1002
DRL.x.x.FAO               --T-AGCT-T--CA-C---CAG---T---TG-TA-T----AT--C----GT---C-C------AG-------C-T---------------A--T--T---------GAGA-----TCAAGTA......-A----C-CC---A---GGT-C-T---A-C--CTCA--GTCA-  999
MON.CM.99.L1              G-TT--CT-T--CA-C--TCAGC-A------G-CA-A--T-AT-----G-GG--AC-C-----TA-T--G--TGT-----------------G--CA-T-----CTT------C---CA--CG...C--CAA--GCAAC-CAAT-CAGG-C--G------C----GGCC- 1156
MON.NG.x.NG1              -------------------------------------------------------------------------------------------------------------------------...-------------------------------------------------    0
GSN.CM.99.CN166           T-TA--CT-T---A-G---CAA-----C--TG-TA-A--A-ATT------GG--AG-A--G--TA-T-------TGT----------AC-G--A-----G-----CCTTT-A-----G--A-CA......CA---CAAC-GGTG-CAGG-C-T------C--T--C---GCA- 1169
GSN.CM.99.CN71            T-TA--TT-T---A-G---CAA-----C--G-T--T--A-ACT------GG--AG-A--G--TA-T-------TGT----------AC-G--A-----G-----CTCT-A-----GCA--CC...--G-CA--CCAAC--GT--CAGG-C-T------C--T--C---GCA- 1175
TAL.CM.00.266             --T---CT-T--CA-C--TCAA-----T--TG---A--A-AG------G-GG--A-----G--CA----G--C-T---T------C-G--AT-----G-----CAGA-------CCA-......AT---T--AC-AC--G-----A----T-A--G------GCA--G---- 1377
TAL.CM.01.8023            --TT--TT-T---A-C--TCAG-----T--TG---A--A-A------G-GG------G--CA-T-----T-T---A-----C-G--C-----C-------CAGA---C---CCA-......AT---T--AT-AC--G-----A---C-----G---G----GCA--C---- 875
MUS-1.CM.01.1085          A-TA---T-T--CA-C--TCAAC----G--TG-CA-A--A-A------G-G---T--A------TA------TGTA--T------AC----A-------------CCTGT--A-C---CA--CC...CA-CA---GCAGC-CAAT-CAGGCC--C--GA-C-----CA---TCA- 1140
MUS-1.CM.01.CM1239        A-TA--CT-T--CA-G--TCAGT-------TG-CA-A---AT------G-GG--T--A--G--CA---------TGT---A------C----A-------------CTGC-A--C---CA--C-...CAGCA---GCAGC-CAAT-CAGGCC--C-GAGC---TCA--TCA- 1142
MUS-2.CM.01.CM1246        --T---CT-T---G----TCAG-----C---G-CA-A--A-AC--C--G-GG--A--A---C-AA----------TGTGG--------C-G--A-----CA-A--CGTTT-A-----CCAA-CA...A--CA---ACAAC--AAT-CAGG-C-A--GT-C---TCA---GGC-- 1223
MUS-2.CM.01.CM2500        T-TT--TT-T---G-G--TCAA-----T--TG-------A-AT--C----GT--AG----GC-AA-T--G--TGTGG--------C-G-----T--G-------CTTT----C--CCAA-CA...---CA----CCAAC--AAT-CAGG-C-TC--T-T---CTCA---GCA- 1185
SYK.KE.x.KE51             T-TA--TT-T---CC-C--TGA------T--TG-CA-T--A-AT------G---A-------G--A-C------TGTG--T------------GGT-T---------CCT-CAA--C---CAA-C-...CA-CA---AGCAGC-TGT--CAGG-C-A------T--T--T-CCTCA- 1200
SYK.KE.x.SYK173           T-TT--TT-T---C-C--TGG------T--TG-T--A--A-AT------G-G-----------GG-AA-T------C-T--------------GG-G-C---------CTT-AA-----CCAA-CG...CA-CAA---GC-C---A--CAGGAT----G--T--CTCC-CCTCA- 1572
RCM.GA.x.GAB1             --TA--CT-T---A----TCAA--------TG-CA-A---A---C--G-GT--GC----G--AG---G--AGTG----------------A------C---------GCTAGA---------A-CA......--C--G--AC--G-----G--AC-T------------CA-------  917
RCM.NG.x.NG411            --TA--CT-T---G----TCAG---T---TG-CA-A----A----------GT---T--A-----AG----G--AGTA--------T--G--A-----T-------CTTGAGA-----CAAT-CA......-A--AC-AC--G-----T--AT-A--G-------CA-----C-  917
GRV.ET.x.GRI_677          TCT-T-CT-T---G----TCAG-----C--TGT-A-A--A-AC------G-GG--AT-A------TC-T--G--AGT---T------A-----A-----G-----C--GACA--CAG-CCA-CA...-T--C--GT-AC--G-----G--AC-C-----C--G-CA--GTCA- 1605
SAB.SN.x.SAB1C            T-T---TT-T---A-C--TCAA-----C--TG---A--A-A--A--C-G-GG--AC-A-----AG-G-----TGTT---------A-----G-----T-----CCTTAGA--C--GCC--CG...CA-CA---CCCA--T-AG--AGT-C-A--G--C----CAA--GTCA- 1758
TAN.UG.x.TAN1             -TTAAGTT-T--CA-T--TCAG---T---TGTCA-A--A-AC------GG--A-----G--TA-C-G--AGTA---------T-----A--CC-G---------T-ACA--C--CCCA-CA...-----A--GC-AC--G-----G--C--A--G--C-----A--TTCA- 1628
VER.DE.x.AGM3             ---A--CT-T--CA-C--TCAA-----T--TGTCC----A-AC------GG--GC-A-----AG---------A-T-------------A-CC-G-------T--CT--C---CCA-CA...-----A--AT-AC--G-----A--AC-C------C--T--A--CTC-- 1169
VER.KE.x.9063             T-----CT-T--CA-C--TCAA-----T--TGTGT--A--A-AC------GG--AC-A-----AG---------A-T-------------A-CC-G---------GT-ACA--C--CCA-CA...-----A---T-AC-------A---C-T------C--T-----TTCA- 1668
VER.KE.x.AGM155           --------T-T--CA-C--TCAG-----T--TGTCT-A--A-AT-----G-GG---T------AG-G-----A-TA--T------------A-CC-G---------GT-ACC--C--CC--CG...-----C--CT-GC--G-G--A----C-C--G--T--GG-----TCA- 1657
VER.KE.x.TYO1_patent      ---A--CT-T--CA-T--TCAG-----T--TGTGC-A--A-AT------GG--AT-A-----AG-G-----T---T-----A-----A-CC-G---------GT-ACA--C---C-A-CC...-----A--CC-AC--G-----A--C-C--G--G--TC-C--CTCA- 1665
COL.CM.x.CGU1             A-TAG-TT-T---A-G---TTG-----C--T-TTT---ATACT-------G-TT-CCT--GG-C--G--G--TGAG--T---A-AA-G--A-A---G-AT--CCT-C-A-----G--C......CAGCA---ACAACAG-A---AGCAT-A--GC-G--T-CA-C-TCA- 1297
```

```
H1B.FR.83.HXB2              ACATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAAT..............AATCCACCTATCCCAGTAGGAGAAATTTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGA 1647
Gag                         D__I__A__G__T__T__S__T__L__Q__E__Q__I__G__W__M__T__N__.__.__.__.__.__N__P__P__I__P__V__G__E__I__Y__K__R__W__I__I__L__G__L__N__K__I__V__R__M__Y__S__P__T__S__I__L__D__I__R_
H1A1.UG.85.U455             -----------------C---G----A---------------------GGC..............------------------G-----C--C---G-------------------------------------------GTT------T---------- 1102
H1B.US.90.WEAU160           ---------------------------C------------------GC..............--------C---------------C--------------------A------------------------GT------------------------ 1646
H1C.ET.86.ETH2220           ------------A--------------C------------------GGG..............--C------G-T-----G-----C--C-------------------------GC--------------GT-------T---------A- 1021
H1D.CD.84.84ZR085           -T--T-----------------------------------------GC..............-----------------------C------------------------------------------------GT-------T-------C--- 1164
H1F1.BE.93.VI850            -T-----T---------------------CA---------GGGC..............--C------G-------G-----C--C--------------C-----A---------------------T---GT-------T---------A- 990
H1G.SE.93.SE6165            -T-----------------T----G-------------ACT---------CGGC..............--C-------------------G---------------------G----------------------GT-------T-A--------- 1050
H1H.CF.90.056               -T----------------------G--------------C---------GGC..............------G-----------G-----C--C----------------------G-----------T---GT--------------------A- 994
H1J.SE.93.SE7887            -T----------------------C-------------------------GGC..............---------------------G--------------T------C---------------------------GT---T---T----T------ 970
H1K.CM.96.MP535             ---------------C--C------------------C-----------GC..............--C------G-------G--G----C--------------T------C-----------------------GT------T----------- 846
H101_AE.TH.90.CM240         -------------------------A--------------------------C..............---------C---------G-----C--C--------G--------------------------------GTT------T----------- 1212
H102_AG.NG.x.IBNG           -------------------------A-------------------------GC..............------------------G-------C-------------------G----------------G--------GT-------T----------- 1178
H103_AB.RU.97.KAL153_2      ------------G--G--------------A-------------------GC..............--C---------------G-----C--C------------------A-----------G---------------GTT------T-A--------- 879
H104_cpx.CY.94.CY032        -T------------------------A-----------------------GC..............--C------G-------G--------C-----------------T----G---------------C--------C-TT-----CT----------- 1013
H1N.CM.95.YBF30             -T-----------A-----C-----GGCA-----GG-G-CT--------TGC-..............-----T---G-T-----------T-------G------------G--------G-----C-G---T--G---------T---GT----------A--G--C-A- 1211
H1O.BE.87.ANT70             ----T--T--G--A-----C---AG--A--G-----TCAC----CT--C-GGCCC.............--C-A---------------------C--C----G--A-------G-GT-A---C-----C-----G----A-------C-----AGTG-----CT-A--T--T-AG 1696
H1O.CM.91.MVP5180           ----T--T-----A-----C--A-AG--A--G-----TAT-----CT--T-GAGGG.............GC-AACT--------------------C--C----G--A-------G-G--A---C-----C-----G----A-------C--T--AGTG-----CT-A--T--T--G 1671
H1P.FR.06.RBF168            -T--T--------C--Y--Y---AAA---------G---ACC-----A----GG..............---CAAT---G-A---------------T--C----G--AG------G-GT----RC-C------G-G----A-------C-----C-TY-----C-------------AG 1105
CPZ.CM.01.SIVcpzCAM13       -T--------G-----C--C--AT-G-----G-----T------GCA---GCA..............--C-----C------------------C--C----G-------G---T--A--GC----C-G-G------A-------T-T--AGTA-----C--T--T----AG 1224
CPZ.CM.05.SIVcpzEK505       -T-----------A----------GGC---G---G---C-----------TGCC..............------T---G-T--------------T----------G-----------G--T-------------------G--T---------G---------------A-----C-A- 1195
CPZ.CM.05.SIVcpzLB7         -------------C--------TG-A--A------------------T-GC..............--C--T-----A-------------------C-------------TA----GC-G-----GG----C-------C-------GTT-----C-------------AG 1193
CPZ.CM.05.SIVcpzMB66        -------------C--C--C---A-C--A--G--GG----G---------T-GC..............--C--T-----T----------T--C------------------TT----GC-------G-------A-------T-----GT------A--A-------A- 1190
CPZ.CM.05.SIVcpzMT145       -----------C--------CT----AG----G---G----G---------GCA..............--C--T-----T-----------------T------G---------------G------------------T-A------C--T---GTT-------T-------T-A- 1186
CPZ.CM.-.SIVcpzMB897        -----------G--A-----C---A--------------G-G---------C-G-..............--C--T----C---G--------C------C--G----------G-TT-A--GC-T----------------------------------GT---T------A--T----AG 1192
CPZ.CM.-.SIVcpzDP943        -------T-----C-----C--T--AGCA-----G--T-CT----------GCA..............--------CG-T-------------------C-----G-------G-GG-T--A--GC-T--C---G-G------------------AGT---T--C--A-----T-AG 1206
CPZ.CM.98.CAM3              -------T--G--------C--T--AGTA-----G--T-CT----------GCA..............--------CG-T--------------------A--C-GG-------G-GG-T--A--GC-T--C---G-------G-------T----AGTT--T--CT-A-----C-AG 1037
CPZ.CM.98.CAM5              -------T-----------C--T--A--A------G-T-CT----------GCA..............--------C-------------G-----C-----G-------G-GG-T--A--GC-C--C---G----G-A------CT----CGT---T------A--T--C-A- 1340
CPZ.GA.88.GAB1              -T--------G--C--C--C--A--G------------T--G----CA---GCA..............-------T--C-----------G--TG------G-------G----TT-A--G-----C---G-G--T------------T----AGTA--T--C---T-----C--- 1708
CPZ.GA.88.GAB2              ----T-----T-------------T-G--A--G--G---------------TCC..............----------AG----T--------G-----C-G-------G-GG-AT-A--GC-T-----GG-G----------------T----GTT------CT----------A- 1025
CPZ.US.85.CPZUS             ----T--------C--C--C---T-G--A------G-G--C----------GCA..............---C-----C--T-----------C--A------G-------G-GG----A--GC-T---------G------A------CT-T--CGTTG-T--CT-A--T--C-AG 1695
CPZ.CD.90.ANT               -T--------G--A--A--C--AG-G-----G--G--GCA--------T---CACCT......CAACAG---GG-GGAG------------G--C--C------G----------C---A----------------GG-G--C--G-N------T--AGT-----------A--G----A- 1068
CPZ.TZ.00.TAN1              -T--T-----G--C---------G--A--A------G---CTG-----------CCCCACAGGCA......C-AGG-GGAG-G-----------C--C---------G----------TT-A-------------T----C-------C-------GTT------T----------A- 1288
CPZ.TZ.2001.TAN2            ----T-----G--C---------AG-G-----------G---ATG----------C-CC......CCCCAGGCAC-AGG-GGAG-G-----------C--C-------------------T------C-G------T----G--------C--T---GTT--T---T------------A- 1284
CPZ.TZ.2002.TAN3_1          -T--T-----G--C-------T-AG-G-------------CTG----------C-CC......CCCCAGGGAC-AGG-GGAG-G-----------C--C-------------------T------C-G------T----G--------C--T---GTT-----T------------A- 1284
GOR.CM.2007.SIVgor2139_287  -T--T-----G--A--A-----T-AA--A------G---CT------A----GG...........AAT--CAAC---G-G-----G--T--C--A-----GG-A-------G-G--A--G-----C---G------A------CT-T--AGT---T--C--A--T--T-AG 1184
GOR.CM.2007.SIVgorCP2135con -T--T-----G--A--A-----T-AA--A------G---CT------A----GG...........AAT---AAC---G-G-----G--T--C--A--C-GG-A-------G-G--A--G-----C---G------A------CT-T--AGT---T--C--A--T--T-AG 1183
GOR.CM.2004.SIVgorCP684con  -T--T--------A--A-----T-AA--A------G---CT-----------GG...........AAT---AAT---G-G-----G--T--C--A--C-GG-A-------G-G--A--G-----C---G------A------CT-T--AGT---T--C-----T--T-AG 1103
H2A.DE.x.BEN                -------------A--A--C--AG-AG-T-----G--CCAG------TAT-GGCCA............C-AAAT--CG-A--G-----GA-C--C--C-G----------CCAAA-A--GC-GC-A--GTGT--C----A---C-A---A--T-A---CT-A--------A- 1966
H2A.GW.x.ALI                -------------A--A--C--AG-AG-A-----G--CCAG------TTC-GGCCA............CGGAAT---G-G-------GA-C--C-----G----------CCAGA-A--GC--C-G--GTGT--C--G-----C-A---A-----A---C--A--------A- 1959
H2A.SN.x.ST                 ----------G--A--A--C--AG-AG-A--G--G--CCAG------TTT-GGCCA............C-AAAT---G-A----------AGC--C-----G----------CCAGA-A--GC--C-G--GTGC--C--G-----C-A---A-----A---C--A--------A- 1411
H2B.CI.x.EHO                -------G-----C--C--C--AG-AG-A-----G---CAG------TAC-GACCC............C-AAAT---G--------G---A-C--C-----G---G------TCAGT-A--GC--C-G---TGT--CC-G------C-AT-----T-AT--A-----------AG 1928
H2B.GH.86.D205              -T-----------C--C--C--AG-AG-------G---CAG------TAC-GGGCC............C-AAAT---G--------G---A-C-----------G---------TCAAT-A------GC-G---TGT--CC-------C-AT-----------A---AT-A--------AG 1927
H2G.CI.92.ABT96             -T--------G--C---------AA-AG-A----------TCAG----CACAC-GGCAG............C-AAAC-----A------------A-T--Y-----G----------CCART-A---C-TC-A---TGT--C--G------Y-AT--A--A-A---C--A--TG---A- 1329
H2U.FR.96.12034             -T--T--------A-----C--AG-GG-A--G--G---CAG------TAC-GGCCG............C-AAAT--A--A--G--G--TA-C--C---CGG-------CCAAT----CC--C-A---TGT--T---------C-AT--A--T-AT--CT-A--T-----AG 1445
SMM.SL.92.SL92B             ----------G--A-----C----CAAGT-----G----A--------TAT-GGGCA............C-AAAC--AG---------G-----C--C------G---G-------CCAG-------C-CC-A---TGC--C--G--------A---A--A-A---C------TG-T-A- 1356
SMM.US.x.H9                 ----T--------------C--AS-AG-C---------TCA--------TAY-GGCAG............C-AAAC-----A--------CA-C--C--C-GR--G------CCAAT-A--GC-GC-G---TGC-------------AT--A--A-A-----T-----TG-G-A- 1398
SMM.US.x.PGM53              -T--T-----------------AG-GG-A--G-----TCAG------TAT-GGCAG............C-GAAC--C--A------G---A-C------C-GG----------CCAA--T--G---C-A--GTGT--C----------C-AT--A--A-A------------G---A- 1842
SMM.US.x.SIVsmH635F_L3      ----T-----------------AG-AG-----------CCA--------TAC-GGCAA............C-AAAC--C--A---------CA-C-----------GG--G------CCAAT-A--GC-GC-G---TGT----------------A---A--A-A-----T-A--TG-G-A- 1928

MAC.US.x.239                -T--T--------A-----T-AG-AG-T--------CCAG------TAC-GACAA............C-GAAC--C--A---------CA-C------C-GG----------CCAA-----G--GC-A---TGT--C------------A---A--A-A-----A--TG---A- 1913
Gag                         D__I__A__G__T__T__S__S__V__D__E__Q__I__Q__W__M__Y__R__Q__.__.__.__.__.__Q__N__P__I__P__V__G__N__I__Y__R__R__W__I__Q__L__G__L__Q__K__C__V__R__M__Y__N__P__T__N__I__L__D__V__K_
MAC.US.x.251_1A11           -T--T--------A-----T-AG-AG-T--------CCAG------TAC-GACAA............C-GAAC--C--A---------CA-C------C-GG----------CCAA-----G--GC-A---TGT--C------------A---A--A-A-----A--TG---A- 1912
MAC.US.x.251_BK28           -T--T--------A-----T-AG-AG-T--------CCAG------TAC-GACAA............C-GAAC--C--A---------CA-C------C-GG----------CCAA-----G--GC-A---TGT--C------------A---A--A-A-----A--TG---A- 1901
MAC.US.x.EMBL_3             -T--T--------A-----T-AG-AG-T--------CCAG------TAC-GACAA............C-GAAC--C--A---------CA-C------C-GG----------CCAA---A-G--GC-A---TGT--C------------A---A--TA-A-----A--TG---A- 1399
STM.US.x.STM                -T--------G--------CT-A-CAG-----------TCA--------TACCGGCAG............C-AAAT-----A--G------A-C--C--C-GG----------TCAA--A--GC--C-A---TGT--C-----------C-AT---AGTT-A-----T-A---G---AG 1575
MNE.US.x.MNE027             ----T--------A--------AG-AG-T--------CCAG------TAC-GACAA............C-GAAC--C--A---------CA-C-----------GG----------CCAA-----G--GC-A---TGT--C-A----------A---A--A-A------A--T------ 1393
DEB.CM.99.CM40              -------------A-AC--C--AG-GG-A-----G-----C---------C-GACCAGCAGGTCAAGGAC-AGG---C--AGAT-------C-----A-----G-------G-C--A--A--TC-GC-A-G-TGT--G-A-------C-AT--A-----A---CT----TG-T-A- 1426
DEB.CM.99.CM5               -------------A-AC--C--AG-GG-A--G--G-----CT---------C-GGCCAGCAGGACAAGGAGCAGGC--A--AGAT------C-G--A--C-GG-------G-C------A--C---C-G-G-TGT----A-------C-AT--A--T-A-----T----TG-T-AG 1417
DEN.CD.x.CD1                -T--------G--C--CTCCT-TA-AG-A--T--G---AC----------T-GACAA............C-AGGC--A--AAAT--------T--A------G--G---G-T--TA----------C-A--GG-G--C-AG-----CCAA--AGT-TC-----T---------AG 1420
LST.CD.88.447               -TG-C------GTC--CTCA--AG-AG----G--GC---CT----CC--TGCAGAC..............A-C---G-AGAT------A-G----------GAG---G----T-AA-CTA-GG-A---G-----T-----CC--CAA--AGTT--TG-AA-A------C-AG 819
LST.CD.88.485               -TG-T------GTC--CTCA--AG-AG----G--GC---CC----CC--TGCAGAC..............A-C---AG-GGAT------A---------------GA----G----T-AA-CTA-GG-A---G-----T-----TC--CAA--AGTT--TG-AA---------C-A- 819
LST.CD.88.524               -TG-G-----G-TC--ATCA--AG-AG-A-----GCA---CC----CA---GCAGAC..............A-----C--AGAT------GA-----C---------GA----G------A--CTA-GG-G---G-G--------AC--CAG--AGT---TG-AA---------A- 813
LST.KE.x.lho7               -TG-G------GGTC--CTCA--AG-AG-A-------C---CC----CT---GCAGAC..............A-T---G-AGAT------A----------------GA----G--C--T-A--CTA-GG-A---G-----C-----CC--CAA--AGTA--TG-CA--------C-A- 1903
SUN.GA.98.L14               ----------GGTA--ATCA--AG-AG-A--G--G---CAG-----CTATGGCCTT......AATGGAGCAA--AGAG-TG------CA----C--C-----------G-C----G-C--AGA----CA-GG-A-G-G----T---C-CC--CAGA------TCAG-AG-A--A--T--- 1890
MND-1.GA.x.MNDGB1           -T-----------------TC---AG-GG-A------T-G-C---------AT-TGCAA............C-AAATG-A---AAT------AC---C-----G--T-----T--A------CA-G----G-T-G---A--GTC--T-T--A-TA--T--AACA--TG------ 1277
OLC.CI.97.97CI12            -T----T------C---------ATT-G---C-G---C-G-CT---G-A-ATGCAGGA.......................AGAT---------C-C-C------G-------G-T---A-A--G--GC-G--GT-G--G-AG-------TCG--AGTT-CAG-G--A------C--- 1509
WRC.CI.97.97CI14            -T---------GTA--AGCA-GTG-GG-----TG--C---CC----T--G-TC-AAAT...................AG-AGAT-----GTC---A--------------G-CC-TAGA---C--C-G---G----TCAG--C---GTT--AGTA---------------A- 1851
WRC.CI.98.98CI04            -T--------GGTA--AGC-T-AG-AG-----TG--T-G-CC----T--G-TC-AAAT..................----A-AGAT-------C-G-G--C-----G---G-GC-TAG---GC--C-A--GG-----CA---T---GTG--AGTT------T-A--T-----AG 1869
WRC.GM.05.Pbt_05GM_X02      -T---------GTA--ATC---TA-GG------G--T---C-----C-AG-TC-GCAA....................A--AGATA------C---A--------------------TAGA---C--C-G--GG----GCA---T---GTA--A--ATCT--CT-A--T--C-A- 813
MND-2.CM.98.CM16            -T--T--T-----A--A--CT--A-AG-A--G--G-----A-----CT-----GGGAG............C-AGATG-AG-AAAT--G-----G---A---------CA--------G-TT-A--GC-TC-G--GTGT-----C--------AT--AGTA-A---C----------AG 1603
MND-2.GA.x.M14              ----T--------A--A--C-G-A-AG-A-----G-----A-----CA--T-GACAG............C-GGA--AGG-AAAT-----G-C---C------GCAG------G----A--GC-GC-A---TGT--G--C------C-A---AGTG-A----------------A- 1527
MND-2.x.x.5440              ----T--------A--A--CT--A-AG-A--G--G-----A-----CA--C-GGCAG............C-AGAC-AGG-AAAC--------GT--A--C---CA-------G-A--A--GC-GC-A---TGT------C------------A---GTA-AT---T-A-----C-A- 1160
DRL.x.x.FAO                 ----T--------C--C--CT-TA-AG-A--G--------AG-----CC--C-GAGCC............C-GGATT-C---AAT--G-----C---------G-CA----G-GG--T-----G---C-A-G-TGT----C-------C-AT---AGT--A---CT-----TG-G-AG 1157
MON.CM.99.L1                -T---------GTCT-GTCC--A-CC--AC-G--G-----AG------A--C-GGGCT..............--CAAC--CG-T-AG--CTC---C--C--C-G--A----G-C---T-----CC-GC-GCG-TGC--C-A-------C-A----GT--A------T-----T--C-A- 1314
MON.NG.x.NG1                ...................................................................................................................................................................................... 0
GSN.CM.99.CN166             -----------GT----TCA--T-CCA-T--G--G-----AG-------C-----GGCAG..............--C-A---G--AAT--GTC---T--A-----G--AG-----C--TT-----CC-GC-GCG-TGT----A---------A----GT--AT---T--------C-A- 1327
GSN.CM.99.CN71              -----------GTA--ATCA--T-C-A-T--G--G-----AG-------T-----GGCAG..............--CAAT--AG-TAAT--G-C---T--A-----G--A------C---T-----C--GC-GCG-TGT----A---------A----GT--AT---T---G----C-A- 1333
TAL.CM.00.266               ----------G--A--A--C--T-CCGCT------GG-G-A------CC--C-GAGCA............C-GAAC--CG-AGAT--G-----C---A--C---------G-C---A-----C---C-GCGCTGT----A-------C-AT---AGTT-A----A------TG-G-A- 1535
TAL.CM.01.8023              -------------A--C--C----CAGCA-------G-G-AG-----CC--C-GGCAG............C-GAAT--AG-AGAT--G--GTCC--A--C--G--G-----G-C---A-A--C---C-GCGCTGT----A-------C-AT---AGTG-A---C------TG-G-A- 1033
MUS-1.CM.01.1085            -----------GTAT-ATC------CA--A------G-----A------C--T-GGGCA..............--CAAC------AAAT---C-------C--C--G-AG---G-G----A----GC-GC-G-G-TGT--C-A-------C-AT---AGT--A---C--A--T--C-AG 1298
MUS-1.CM.01.CM1239          ----------GGTA--ATCC----CCA-T-------------A-------A--T-GGCAA..............--CGAT------AAAT--G-C---C--C-------G-A-------------A-----G---C-G-GGTGT--C-A---------A----GTT---T--AT----T--C-AG 1300
MUS-2.CM.01.CM1246          -TG-G--G---GTGT-ATCA--G-CAGCA-------------A-------------T-GACAG............C-AGAC--AG-AAAT-----GA-T--A-----G--A-------C---T-----CC--C-G-G-TGT--C-A-------C-AT---AGTA-A---CT----T--T-A- 1381
MUS-2.CM.01.CM2500          -T---------GGTAT-CTCC--A-CAGCA-------------AG-------C-GGCAA............C-AGAC--AG-GAAT-----------C--A-----GG-A------T---------C--GC-A-G-TGT--C-A---------A----GTA-A---CT-A-----C-A- 1343
SYK.KE.x.KE51               -------------A--C------A-C--A-----------AG-----A--T-GACAG..............--CAAC--A--T-AG--T---C-G--C--C-GGCAG---------T--A--------C-A--GTGT--GCA-G-C--C-AT----GT--AT---T--------C--G 1358
SYK.KE.x.SYK173             -------------A--A--CT-AA--GCT--G-------A-------T--T-GGCAG..............--CAAC--AG-A-A---------------A-----G----------C--T--A--GC-GC-G--GTGT--GCAGG-C--C-A-----GTTTCT--CT-A--------A- 1730
RCM.GA.x.GAB1               ----------G--A-----CT-AA-AGCA-----G----CT-----CC--C-GAGCA..............--CAAC--C--TAAT--G--CA-TC-G--C-G--AT----------AG-A--G---C-A---TGG----A-------C-AT---AGTG-A---C--A--T----AG 1075
RCM.NG.x.NG411              -T--T--------A--A--C-ATA--GCT---------CT-----CC--T-GAGCC..............--CAAC--A--TG----G---A-C--------GG-AT-------TG-G--T--G---C-A---TGT-G-A---------A---AGTA-A---CT-A-----C-AG 1075
GRV.ET.x.GRI_677            -T-----------------C--CT-AA-----------------AG----CCTTC---GCC..............---------AGA--AGAC-----G-C-CAA--C-G--A----G-T--TT-----C---C-A--GG----GCAG------C-AT---CCAA-AGG-C--A------TC-- 1763
SAB.SN.x.SAB1C              -------G--C--------C--AA-A--A--G-------A------CC--C-GAGCA............C-GAATG--G--AAT-------A-T-----------G---------C----------TC-CC-G---TGT--C-A-------C-A---AGTA-A-----T-A------T-A- 1916
TAN.UG.x.TAN1               -------G-----A--C--CT-TG--GCT--G-------AG-----CTTTT---GCT..............------AGAG-AGAT-------CGT--C--C-G-G-----G-T---------C-CC-A---TGT----A-------C-AT-----T-TCAG-C---------C--- 1786
VER.DE.x.AGM3               -------------C--C--C--AG-G--A-----GC-G-A-------ATAC-CAGCC..............------CAGAG-AGAT--G--T-CC--C-----G---G------T-----------G--GC-A---TGT----A-------C-A---AGTGTCTG-CT-A--------- 1327
VER.KE.x.9063               -------------C---------AG-G--A-----T---A-------TTAC-CAGCC..............--C--CAGAG-AGAT--T--G-CT--C------G-------------T------A--GC-GC-G--GTGT----A-------C-A---AGT-TCAG-AT-A-----C--- 1826
VER.KE.x.AGM155             -T--------G--C---------AG-G--A--G--GC---AG-----CTAT-C-GCT..............--C--AGGG-AGAT------G-CC--C---CG----------C-----A--G---C-A---TGT----A-------C-AT---AGTGTCTG--T-A--T--C--- 1815
VER.KE.x.TYO1_patent        -T--------G--C--C--CT-AG-A--A-----GT---A-------CTAT-C-GCT..............--C--C-GGG-AGAT------T-CC--C--CCGG--------T--T--A---C-TC-A--GTGT--C-A-------C-A---AGTATCAG-C--A------T--G 1823
COL.CM.x.CGU1               -T--CA----C-AC--A--CT-AG-AGCA-------G-G-CT---GGAGAG.........................--C--A.......CCA-C--------GGG-C-------G-A-A-A-TC-GG-A--GG--A-CCAG--TGCA--A--AT--TCTG-A--A--T--T--- 1437
```

```
H1B.FR.83.HXB2              CAAGGACCAAAGGAACCCTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAACCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACACT 1817
Gag                         _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__D__C__K__T__I__L__K__A__L__G__P__A__A__T__L_
H1A1.UG.85.U455             -----G-----A--------C--G--T-------TA-A---T--------C----T---------A-------T--------C-----------------C---------------T----C-----T-C------G-----A--G---G-------T- 1272
H1B.US.90.WEAU160           -----------A-------------T------------------------------------------A------------------------------------T--G------------------------------------------GA-------- 1816
H1C.ET.86.ETH2220           ----------------T------------------T-----CT------T--A------A---A--T-----------------C--------------------------------------A--C------G-----A--G---G---T--T- 1191
H1D.CD.84.84ZR085           -----------------------------------C----------------------------A--A------------------------------A--------------------C--------------------G-------T- 1334
H1F1.BE.93.VI850            -----G-----A---------------------A-A---T---GTC-------T-----------------C----GGG-----------C--A------------------------------------C-------------------A--G-------- 1160
H1G.SE.93.SE6165            -----G-----A--------------------TA------T---TG-T-G----T-----------------------T-----GGC----------C---------------------------------C--C---G-----A---A--GA---T---- 1220
H1H.CF.90.056               -----G-----A---------------------A----T-T-------T-------T--------CA-------T--G--G-----------------------------------A--T-------C-------A---G-----A---A--G---T--A- 1164
H1J.SE.93.SE7887            -----------A-----T----------------A------T---G-------T---------A-------T------------------T-----C-------------A--T-------C-----C---------------A---T---GA-------- 1140
H1K.CM.96.MP535             -----G-----A--------------------TA------T-----C--------T--A----CA-------A-----G--------------C---C-----------C--A------------------C-----------G--------G---T--T- 1016
H101_AE.TH.90.CM240         -----G-----A--------C--------------TA---------------C----G--A------A----------------C----------------C-A--------C-----T-----C-----T-C---------------A--A--GA------T- 1382
H102_AG.NG.x.IBNG           -----G-----A--------------T--------TA------T-------T-G----T--A------A-G--------------C----------------C-----------------T-----C-----T-C-------G-----A--A---G-------T- 1348
H103_AB.RU.97.KAL153_2      -----G-----A--------C--------------TA------T---------T----T--------A-------T----G--C--------------C--C--------------A--T----------A--C-------G-----A---T---G-------T- 1049
H104_cpx.CY.94.CY032        -----------A--------C-----T--------TA------T---TG---C-----A--A------A-C---------G------------------C--C-----------------T-----C-----T-C--C-----------A---A---G-------T- 1183
H1N.CM.95.YBF30             -----------A--------C----------------A-------C---------------A----G--AA-------A-----G-----------------AC-C--A--A---------A---------------ACAGC-CC---------A--G---GA-----CT- 1381
H1O.BE.87.ANT70             --G-------------A------------------T-------C-----AT-------T--------A-T--A--A------------------TC--C-T--T--------C--T---------C--ACAG-----G---T----A--G---GA-----CT- 1866
H1O.CM.91.MVP5180           --G--------A-----A--C-----T---------T-----T--C-----AT-------T--------A-T--A--A-----G---------------------C-T--T--G---T-A----------------ACAA---C-G--------A-------AA-----TT- 1841
H1P.FR.06.RBF168            --G-----T--------A--------T--------TA-A-----C-----C--C-----A--A-----AGT-----A-----G-G--------------C--A--AC-AA-A----------T--T--------------CAG-----G-----TC-----------G---C--CT- 1275
CPZ.CM.01.SIVcpzCAM13       -----------A--G-------A---T----------A-----T---------CA-G-----T-----------T-----A--G----C--------------C---C-T-----G---------T------------C---CAA---C----G--CC----T--G-G---CT-C-- 1394
CPZ.CM.05.SIVcpzEK505       --G---------A--GA-T---------T--C-----TA----T--C-----------------A-----G--CA-T-----A-----G---------------G--AC-C--A--A----------A--T--T---------A-A---CC-G-G------A---------GA--CT-C-- 1365
CPZ.CM.05.SIVcpzLB7         -----G-----A-----A--C-----------G---A----T--C-----AT-------A--A-----CA--------A--------------------C------C-T--A--G-----T--------------C--A-G---CC-G-G----------T----GA--CT-CT- 1363
CPZ.CM.05.SIVcpzMB66        -----G-----A--GT-A--C-----T-----G---A-A--T----------T----G--A--A------A-----A--A--------------------C--T--A-----G--------T--T----------------T-C--A----G------A--G---GA-----CT- 1360
CPZ.CM.05.SIVcpzMT145       -----------A--G--T--C--G-------------A--------C--G--CT----G--T--A-----A---------T--T--------------------C---C------A-A--G-----T--------C-----A-AC--C----G------A--------GA--C--T-- 1356
CPZ.CM.-.SIVcpzMB897        -----------A-----A--------T-----G---A-A--T--C----------------T--A-----AA-C--A--T--------------------G------C-A--G---------C--------------------CAA--------G---TC----T----G---A--CT- 1362
CPZ.CM.-.SIVcpzDP943        --G-----T-----G--T--C--G-----------TA-A-----C-----C--G--G--A--A--G--AA------------------CC---------T--T---C-CY----A---------T---------C-----A-A---C--G-GG--C--A--------GA--C--CT- 1376
CPZ.CM.98.CAM3              --G-----T--A--G--T--C--G--T--------TA-A-----C-----C-----G--A--A-----AA----A--A-------------------T--C---C-C--A--A--G-----T--T-----C-----A-AC--C----G---CC----------GA-----CT- 1207
CPZ.CM.98.CAM5              --G-----T--A--G--A--C--G--T--------TA----T---------T-------A--A--G--AA--------A-------------------T--C---C-T--A--T----------T--T-----C--C--A-GC--C--G-G---C--A--------GA-----CT- 1510
CPZ.GA.88.GAB1              -----G--------G-----------T-----G---A-A-----C------T----G--T-------------------A--G---------------T--C--TC-C--A--G--G-----A--------------------CAA---C-------------------T-G---C--C-- 1878
CPZ.GA.88.GAB2              -----G--G-----G--T-----------------TA-A-----C---GT-T-------A--A--G--------A--T--G-----C--------------G---C-CC----A--G-----C--------------------CAA--AC-T-GG---C-A-------GA-----C-- 1195
CPZ.US.85.CPZUS             --G-----T--A--G----------------------A-------C-----AT-------A--A--------T-CA--A-------CC---------------G---C-T----A-----C--T--T----------------A-----------G---C--------T-GA--C--C-- 1865
CPZ.CD.90.ANT               -----------A--G-----C-----T-----G--TA-A-----------AA-T-----A--A--G----------CCT--G---GCC------------------------AA-----------C--T---------C--ACAC--CC-G--G--T------A---GA--CT-CT- 1238
CPZ.TZ.00.TAN1              --G--------A-----A--C-----T-----------A-A-----C-----AA-C-----A--A-----A--T--ACCA-------C---------------------T--AC----A--------A------------------CA---C--------C----G-A--GA--A---T- 1458
CPZ.TZ.2001.TAN2            -----G-----A---T-A--C-----T--------TA-A-----C--G--CA-C--G--A--A----------T--ACC------G------------------T--AC----G--------A------------------CA---C-----G--T-----GA-T-GA--AT-CT- 1454
CPZ.TZ.2002.TAN3_1          -----G-----A-----A--C-----T--------TA-A-----C-----AA-C--G--A--A----------T---CCA-----G-C------------------T--AC----A--------A--T----------------CA--------------T--A--GA-T-GA--AT-CT- 1454
GOR.CM.2007.SIVgor2139_287  --G-----T--------A--CC----------------A-A-----C-----C--C-----A--A--G--AAGT--A--T-------------------------T-----AA-T--G-----T--T--T----------CAA---C-------CC-A--C----G---A--C-- 1354
GOR.CM.2007.SIVgorCP2135con --G-----T--------A--CC----------------A-A-----C-----C--C-----A--A--G--AAGT--A--T--G-----------------------T---C-AA-T--G-----T--T--T----------CAA---C-------CC-A--G--G-G---A--C-- 1353
GOR.CM.2004.SIVgorCP684con  --G-----T--------A--C-----------------A-A-----C-----C--C-----G--A-----AAGT--A--T--------------------------T---C-AA-T--G-----C--T--T----------CAA---C-------CC-A--C----G---A--C-- 1273
H2A.DE.x.BEN                --G--------A-----G--CCA-AG----------A-------C---GCT-G--G--A--A---A-AGAC-CA-CA------------------CC----GC--C-AA-A--G-----C--------C--C---TTAG-AC------G-C----GATGAATC-C--C-- 2136
H2A.GW.x.ALI                --------------G--A--CCA-AG--------TA-A-----C---GCT----G--A--A---A-AGAT-CA-CA-----G-------------TC----AC--C----A--G-----C--------C--C--ATTAG-GC-G----G---A--GATGAATC----CT- 2129
H2A.SN.x.ST                 --G------------G--A--CCAGAGT--------TA-A-----C--G-GCT-G--G--A--A---A-AGAT-CA-CA------------------CC----AC--C-A--G--G-----C--------C------TTAG-AC------G-C-A--GAT-AATC----CT- 1581
H2B.CI.x.EHO                -----G-----A--------CCA-AG---------TA-A-----C---GCT--C-G--A------A-AGAC-CA-CA-------------------C----AC--C--A-T--G-----T-----------C--ATTAG-GC-T--G-GC------ATGAACC-C---T- 2098
H2B.GH.86.D205              --G-----------G-----CCA-AG---------TA-A-----C---GCT--C-G--A--A---A-AGAC-CA-CA--G----------------C----AC--C--A-T--G-----T-----------C---TTAG-GC-T--G-GC------ATGAATC-C--CT- 2097
H2G.CI.92.ABT96             -----G-----A----------CAGTCT--------TA-A-----C--G-G---R-----A--A---A--GAT-CA-CA-------------------C-G--AC-AC--A-T--------C----------------ATTAG-AC----G-GG------ATGAATC----TT- 1499
H2U.FR.96.12034             -----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--------C--C--ACTCG-G-------GT-----TATGAATC----CT- 1615
SMM.SL.92.SL92B             -----T--------G------CA-AGT--------TA--------C----G-T-G----A-----GA-AGAT-C--CA--G--G------------C----AC--C--A-T--G--C--C-----T-----C--ACTAG-A-------GG-----CATGAATC-C--C-- 1526
SMM.US.x.H9                 -----------A--GT-A---CA-AG---------TA-A-----C--G-G----------A--A---A-AGAC-C--CAS-G--------------TC----AC--C--A-T--------C-----G--C----ATTGG-GC-C--G-GTC----YAT-AATC-C--TT- 1568
SMM.US.x.PGM53              -----G--G--A--G--G--CCAGAG---C-----TA--------C----GCT-------A--A---A-AGAT-CA-CA--G--G------------TC----AC--C--A-T--------T---------------ATTGG-GC-C--G-GTC----CATGAATC-C--TT- 2012
SMM.US.x.SIVsmH635F_L3      -----------A--G--A---CA-AG---C-----TA-A-----C----G----------A-----A-AGAC-C--CA-----G-------------CC----AC--C--A-T--------T---------------ATTGG-GC-C--G-GTC----TATGAATC-C--TT- 2098

MAC.US.x.239                -----G-----A--G--A---CAGAG-----------A-------C----G-T-------A--A--GA-AGATGCA-CA-----G------------TC----AC--C--A-T--------T------------C---CTAG-GC-G--G-GGC----TGTGAATC-C--C-- 2083
Gag                         _Q__G__P__K__E__P__F__Q__S__Y__V__D__R__F__Y__K__S__L__R__A__E__Q__T__D__A__A__V__K__N__W__M__T__Q__T__L__L__I__Q__N__A__N__P__D__C__K__L__V__L__K__G__L__G__V__N__P__T__L_
MAC.US.x.251_1A11           -----G-----A--G--A---CAGAG-----------A-------C----G-T-------A--A--GA-AGATGCA-CA-----G------------TC----AC--C--A-T--------T------------C---CTAG-GC-G--G-GGC----TGTGAATC-C--C-- 2082
MAC.US.x.251_BK28           -----G-----A--G--A---CAGAG-----------A-------C----GCT-------A--A---A-AGATGCA-CA-----G------------TC----AC--C--A-T--------T------------C---CTAG-GC-G--G-GGC----TGTGAATC-C--C-- 2071
MAC.US.x.EMBL_3             ---A-G-----A--G--A---CAGAG-----------A-------C----GCT-------A--A---A-AGATGCA-CA-----G------------TC----AC--C--A-T--------T------------C---CTAG-GC-G--G-GGC----TGTGAATC-C--C-- 1569
STM.US.x.STM                -----G-----A---A-G--CCA-AGT---------A-------C--G-G-T-G-----A--A-----AGAT-CATCA--G------------TCGG--AC--C-AA-T--------T--T-----C--C--ATTGG-A--G----GTC----CATGAATC-C--C-- 1745
MNE.US.x.MNE027             -----G-----A--G--A---CAGAG-----------A-------C----GCT-------A--A---A-AGAT-CA-CA-----G------------TC----AC--C--A-T--------T------------C---CTGG-GC-G--G-GGC----TATGAATC-C--C-- 1563
DEB.CM.99.CM40              -----G--C--------A--C-A---T-----T---A-----------G--CT-------T--A-----AGAT----CA-----------------CACC---C-CA--A-T--------A-----T-----C-G--TC--CC-G----G---A--T-AGAATC-C--CT- 1596
DEB.CM.99.CM5               --G--T--T-----------C-A-------------TA-A-----C------CT-G--G--T--A-----AGAT----CA--C--------------CACC---C-AA--A-T--G-----T--------------GA-TA---C------GG--A----AGAACC-C--TT- 1587
DEN.CD.x.CD1                -----G--C--A---T------AG--T--------TA-A-----C---G-----------A-----G--AGAT--ATCA--G--GCT---------CACTT-AC-A-----A--------A--T--------C---CAG---C-CG--------A-----TAACC-C---T- 1590
LST.CD.88.447               --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------------G--TGG-GA-------C--A---GA--G------TT-  989
LST.CD.88.485               --G--G-----A-----A--C-A--A----C--TA------T----G-AT-------A--AGG--GC-----T----------G-A-------A---G-AGA--C----A--------A------------G--TGG-AA-------C--A---GA--G-----GTT-  989
LST.CD.88.524               --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--C--C--A-TGG-AA----G--C--A---GA--GT--C--TT-  983
LST.KE.x.lho7               --------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--C--C-G-TTAG--A---------C-A---GA--G---AT-T-- 2073
SUN.GA.98.L14               --G--------A--G--T--C---------AC--TA------TC--GG--T-------A--ATC--GC---G---A--G----G-G-------A---G-AAA--C-A--G--------A-------------ATTGG-AA----G---C-----GA-A-TC-AT--T- 2060
MND-1.GA.x.MNDGB1           --G------------G-T----A-------------TA-A-----C--TGTAA-G-----A--A----------GGA--A-----G-TG------CAGC-GCATC--C-TA-AG-------A--------A--C---CAG-----G-G-AGC--A--GAA--GA--A--TT- 1447
OLC.CI.97.97CI12            ------G-T--A--G--T--CTC-------------A------T---GG-CT-G--G--A-------GG--C--------------CA-----------G-AGC----AA-A--G-----T--T-----C--C--AGGAG-C--G-----G---...GT-AAAC--T-G-- 1676
WRC.CI.97.97CI14            -----G--T-----GGAT--C-A--C---------TA-A---------GTAT-------A--A----GAG-C-CA--A------C-A-----------C-AGC-C--AA-A--G-----T--T--G-----C--A--A--C-------G---A...-A-CAAC-A--CT- 2018
WRC.CI.98.98CI04            --G------------GAT----A-ACT-----G--TA-A--------GGTC--G-----A--A----GAG-C-CA--A--G---C-G-----------C-AA--A---A-A--------T---------------AG----C-----G-G---A...-A-CA-C-A---T- 2036
WRC.GM.05.Pbt_05GM_X02      -----G-----A---GAT----A-TCT--------TA-A--T-T---GGTC--T-----A--A----GAGTT-CA--A-----GC-A-----------C-AG--AC-AA-A--------A--T--------C--ACAG-----G----GCC-A...-A-AGTC-A---T-  980
MND-2.CM.98.CM16            --------------G--G--C-AG-------G---A----T--C---G----GC-G--G----GGA--GAT-CAC-A--C----CC---------C----T---C-CA----G-----T---------------AT-C--CC-T----GGC-A--CATGAATC-CT-TT- 1773
MND-2.GA.x.M14              -----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------------------GC--C--G----G------CATGAACC-CT-TT- 1697
MND-2.x.x.5440              --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---------------AG---CCC-C----GG--A--CATGAATC-C---T- 1330
DRL.x.x.FAO                 -----T--T--A-----A--C-AG-----------TA-A--T------G-C--C--G--A--A--CA-AGATGCA-CA------------------C----AC--C--A-A--------A--T-----C-----AGTG---C-C----GCC-A--GATGAACC-C-GT-- 1327
MON.CM.99.L1                --G-----T--A-----A--C-A---T----------A-A-----C---TGC--C-----A--A---A-AGAT----CA--C--------------GC-G--GC--C--------------C-----------C--ACTC--CC----G--CA--...---GT----GCT- 1481
MON.NG.x.NG1                ..........................................................................................................-----G--C--C---CTC--CC-C------A--...---GA--C-GCT-   47
GSN.CM.99.CN166             -----C--T--A--G-------AG--A----------A-A---T---GTGCT---G----A-----AGAT---CC--G---------------C--T-GC--C----T--------T--------A--C---CTG--C--G-----CA--...----GA--C-ACT- 1494
GSN.CM.99.CN71              --G--C--T--------A----AG--A----------A-A---T----TGCT---G--T--A--G--AGAT---CA-----G------------C--TTGC--C----T--------T--T-----A--C---CTG--C--G--G--CA--...----GA--C-GCT- 1500
TAL.CM.00.266               -----C--C-G------A--C-A--------------A-------C-G-TG---T--G----A---A-AGAT---CA-----------------C-G--GC---------G----T--------G--C---G-C--C--------CA----C----GA--C--CT- 1705
TAL.CM.01.8023              -----C--C-GA--G-------AG--G--C------A-------C-G-TGC--T-----A--A---A-AGAT---CA--C---------------C-G--GC----A--G--G----C--------A--C---G-C---C-G--G--CA----T--CCA---G--C-- 1203
MUS-1.CM.01.1085            --G--------A-----A---AG--T-----G---A-A---T----TG------G--A-----GT-AGAC--A-C---C---------------CTC----C----A--A--G-----T--------G-----ACTC---C-G--GAGCA--...----GA--C--T-- 1465
MUS-1.CM.01.CM1239          --G--------A--G--A---AG--T-----G---A-------C---TGCT---G--G------T-AGAT--A-CA--T--G------------CTCC--T-----A--G--G-----T--------G------CTC--CC-G---AGCA--...----GA--C--CT- 1467
MUS-2.CM.01.CM1246          -----C--T--A-----A--C-A--T-----G--T-------T----TGCT-------T--A--G---GAC--A-CA--C----------------C-GT--C--C-A--G--------C-----C--G--C--ACTCG-C--G-----CA--...----G---C---T- 1548
MUS-2.CM.01.CM2500          --G--C--C-GA-----G--C-A--T-----G---A-------T----TG---------T--A-----CGAC---CT--C----------------C--T-T--------T--G-----C--T--T--G--C--ACTCG--C-G-----CA--...----GA--A---T- 1510
SYK.KE.x.KE51               --------T--A--A-A---AG--T-----T--TA-A------C-TTG---T-----A--A-----GGAC-C--CT-----------C-C---C-GT-A--AC-AA-T--G-----A-----T--G--C-GA-GC---C-G-------A---TCAA-...C-A--CT- 1525
SYK.KE.x.SYK173             --G--C-----A---ATT--C-A--------------A-A-----CC-TTGCT-------A-----G--GGAC-CATCA--C---GGG---C-A--TC---AC-----A-T--------C--------A--C-GACAG--CC-C--G---A---TCAA-...C----CT- 1897
RCM.GA.x.GAB1               -----------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--------C-----A-TGG-AC-C--G-GTC-----ATGAACC--T-TT- 1245
RCM.NG.x.NG411              --G--------A--G--T----AG------------TA-------C--GG-CT-G--G--A--A--G--AGAC-CA-CA--G--G------------C--T-AC--C--A-T--------T--T-----C---------AG-AC-T--G-GTC-----ATGAACC----C-- 1245
GRV.ET.x.GRI_677            --G-----T--A--------CCAG-----------A-A---------G-C--G-----A--A----AC-------T--T----------------C----T---C-TA---G-----C--T--G--------ATTG---C-G----G--------ATGAATC-A--CT- 1933
SAB.SN.x.SABIC              -----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------------------AG-GC-C----GG--A--CATGAATC-A--C-- 2086
TAN.UG.x.TAN1               -----GG-C-----G--T---AG-----C------A-A--T--C--G-C--T-----A--A---A-AC-C--A--T-----------------------AC--C-CA---G-----C-----------C---TTGG----G----G-C-T---AT-CATC-A--T-- 1956
VER.DE.x.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-----C-----AGTC--CC----G-GCC-----ATGCATC-C--T-- 1497
VER.KE.x.9063               --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------------C--AGTC--CC-G----G---A--ATGCACC----CT- 1996
VER.KE.x.AGM155             -----G--C--A-----A--C-A--T----------A-A----------G-AA-------A--A----------GGA--A--C---C-A-----------T-T---C-CA-T--G-----C------------C--AGTA-----G--G-GCC-A--GATGCACC-C--T-- 1985
VER.KE.x.TYO1_patent        --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-------------GTC--CC-G--G-GCC-A---ATGCACC-C--C-- 1993
COL.CM.x.CGU1               --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--C--C--A-GA---C-G----G-C--CAGAAG...C-AT-TT- 1604
```

```
                                                                    Gag-Pol fusion TF protein start
                                  Gag p24 Capsid end Gag p2 start              Gag p2 end  Gag p7 nucleocapsid start
H1B.FR.83.HXB2          AGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTTTGGCTGAAGCAATGAGCCAAGTAACAAATTCAGCTACC...............ATAATGATGCAGAGAGGCAATTTTAGGAACCAAAGA...................... 1950
Gag                     _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__N__S__A__T__.__.__.__.__.__I__M__M__Q__R__G__N__F__R__N__Q__R__.__.__.__.__.__.__.__
H1A1.UG.85.U455         --------------------C-------G-----------------A-----G----------G--------T------CA-C-GA--AGC..................----------------------------GG--CG....................... 1399
H1B.US.90.WEAU160       ------C---------------------G-----G-----------A-------------------------------------------A--------.........-G------------------------G--C----....................... 1949
H1C.ET.86.ETH2220       ----------------------------G--------T-C---C--A--------G--------G---------------AC---A--A-C..................-------------A-A---------A-GG--CT-A-....................... 1321
H1D.CD.84.84ZR085       ----------------------------G--G--G---A-------A-----------------G-----------C-----G-G-CT--G-TGCC.............-------------A-A-A--------A-GG-AC----....................... 1470
H1F1.BE.93.VI850        ----------------------------G--------TA--------------------C--G--------T----C--ATTCAG-C.......................-------------A-A-T-------A-GG-------....................... 1287
H1G.SE.93.SE6165        ----------------------------G---------A-------A---------A-----G----------G-CCT--GG-G----AG-AGCC.............-------------A---------A-GG--C----....................... 1356
H1H.CF.90.056           ----------------------------G--------TA-T-----A-----------------G--------------------A--AA---AGCC.............-------------A------C----A-GG-------....................... 1300
H1J.SE.93.SE7887        ----------------------------G--------T--T--------G----------------------------G--C---A-CAAC.................------------A-----T--C------G----T-A-....................... 1270
H1K.CM.96.MP535         ------G---------------------G-----G--TA-------A------A----------G---------G---------C---T-..................G-----------A------C----A-GG---T---....................... 1146
H101_AE.TH.90.CM240     --------------------C-------G--------TA-------A-----G-----------------------C-C-CA-C--G--A-...................----------------------C-A-GG---G....................... 1509
H102_AG.NG.x.IBNG       ----------------------------G--------T--------A-----G-----------G--------T------CA-C-GG-CAAC.................G------------------------GG---G....................... 1475
H103_AB.RU.97.KAL153_2  ----------------T---C-------G-----------------A-----G--------C--G--------T------CA----G--AAC..................-------------A-A-T---------GG--C-....................... 1176
H104_cpx.CY.94.CY032    ----------------------------G---------A-------A---------------A-----G-C-T-----G----AG-AGCC.............-------------A-A----A----A-GG-------....................... 1319
H1N.CM.95.YBF30         ------G-------G--C--C-------G--G-----A-CA--------------GC-A--A--G--T---TCA--G--GCAGC-GC--A-A--TAGT.............G-CT-TGCA--A--G--A--C----AAGG-AT---G....................... 1517
H1O.BE.87.ANT70         ------G------GT---C------------------AACT----------C-----GC-A--A---------GCTAC--CCCAGC-AGATTTG-AAGGAGGATACACAGCAG--T-C-----A-----GC-AAA-CCA-TTAGG-A-....................... 2014
H1O.CM.91.MVP5180       -------------GT---C-----A-----------G--AACT--C-------A-A-AC-A--A---------GCTTCT-CCCAGC-AGATTTA-AAGGAGGATACACAGCAG--T-C-----A-----GC-GAA-CCA--TAG--A-....................... 1989
H1P.FR.06.RBF168        -------------AT--C-----A---------G-----AACA-------C--G--C--A--A-----T---GCAGC--CT-ATC-AG-TAACCAAGAT......................TTA-A---AGGG-A--CA-CTGTTTTCATGCAGAGTGGACAG....... 1417
CPZ.CM.01.SIVcpzCAM13   --------------------C-----A--G------------A-C----------------C--A---------T---CAATG--CCA--G-CA---A-GAACA.........GAC--TT-CT-T----A---GCCAGGGGCA-C--C----....................... 1533
CPZ.CM.05.SIVcpzEK505   G--G----------------C--A--G--G-------A-CT-------------AC-A--A-----C----CA--G-C-CA--CAG--A---GT.............G-TT-T-----A------A-------AAGG-AT----....................... 1498
CPZ.CM.05.SIVcpzLB7     ---G----------------C-------C--G--T--A-C------A--C--G--C--A--A-----T-----T----C--ATGCAG--.............ACAGTC--T--------A--G--A-------AAGGA-C--A-....................... 1496
CPZ.CM.05.SIVcpzMB66    -----GGGATGA-C----C-TC--------G---------A-------A--T-----CC----A-----------T----C-CA-C-CAGTAA-GATGCA.............-AG-G-CAATTT-A----CCAAAA....................... 1484
CPZ.CM.05.SIVcpzMT145   --------------------T-----A--G-----------TCA---------T-----GC----A------C--------G-----AC--AA--...............GAA--CT-T------A-----GGGACAA-AT-GA-CTCCTAAT....................... 1489
CPZ.CM.-.SIVcpzMB897    --------------------C--C----------G--C---TCT-----A--C-----A-------A--------T--G-C--AC-CCA-T-T--TG..............-----------A------T---G---A--TT--G-A-....................... 1489
CPZ.CM.-.SIVcpzDP943    ------G-------------C--A-----G-------A-C---------T--G---------------T-----------CA----C--CA-A-GT...............G-GT-TT--------G------AA-G---GA-C--CT....................... 1509
CPZ.CM.98.CAM3          --------------------C----------G-----A-C---------T-----AC----------T---------C--CA---CC-TA-A...............GGTG-GT-CC----A--G--A---AA-G----A-C--CT....................... 1340
CPZ.CM.98.CAM5          ------------------C--C------G--------A-C---------T-----C-----------C------------CA------CT-A...............AATG-GT-T------------A---AA-G---GAGT--AT....................... 1643
CPZ.GA.88.GAB1          ---G----------------T-----A----------ATCT----------G---C-A--------C---TCAATG--TCAG---CA--GG-GAGCA.........GATG-TT-CT-C----AG--AC-AGG-GC-GG--C--A-....................... 2017
CPZ.GA.88.GAB2          ---G----------------C--C--A-----------C--G-CA-------T------C-A--A-----T------------A--CA-C-GA--AGC..............TCA---T-C------A----CAGGGGGA--A-CT-CTCCT....................... 1328
CPZ.US.85.CPZUS         ----------G-----------------A----G-----------CT-----------G--C--A---------C---T-----A-G-A---CC-CT-A..............AGCG--T-TC-A----A---A--CGC-G----G-C-G-G....................... 1998
CPZ.CD.90.ANT           ----------T-A-----T-----A-----------C--A-C---------------G------A-----T---GCTTCT-CT-AT---G--CAGGGAACC..........GCAG-CT-TC--------------GGA-AT-GAGG-G--AAA....................... 1380
CPZ.TZ.00.TAN1          ----------C-C-----C-----A-----G----------TCT-----------AGA--C--------------GC-TC--C----GCAGGG-GA..............GTA-AT---C------G----A--AAGACCACC-TT--A-....................... 1591
CPZ.TZ.2001.TAN2        ---G------C-C-----C-----------G----------TTCT----------AGA--C---------------GCAGC--C--AC-CAGG-......GGGGGAATAAAT..........----T-CA---GGGAAGA--AC-A-CTGT-AAAAAG....................... 1593
CPZ.TZ.2002.TAN3_1      ---G------C-C-----T-----------G-------------CTT---------A-A-CC--------------GCAGC--CG-AC-CCA--......GGGGGAATAAAT..........----T-CA---GGGAAGA--ACC--CCG-GAAAAAG....................... 1593
GOR.CM.2007.SIVgor2139_287  ----------------AT--T-----------G--G-----TACT--C--A---------C-----A--G-----GCAGC--CCCACTCAGATTTA-AAGGAGGATATACAGCAG-CT-T-----A-AG--GGGAAAC-ATC-AGGG-TTAAGAAG................ 1508
GOR.CM.2007.SIVgorCP2135con ---------------AT--T-----------G--------TACT--C--A---------C--R--A--G-----GCAGC--CCCACTCAGATYTA-AAGGGGGATATACAGCAG-CT-T-----A-AG--GGGAARC-ATC-AGGG-TTAAGAAG................ 1507
GOR.CM.2004.SIVgorCP684con  ----------------AT--C-----------G--------TACT--C-------------C-----A---------GCAGC--CCCACTCAGATCTA-AAGGGGGATATACAGCAG-CT-T-----A-A---GGGAAAC-ATC-AGGG-TTAAGAAG................ 1427
H2A.DE.x.BEN            ------G---C-A--C--C--C-----G-----C-----A-----G--A--C--GC-AA----------CC-A-AAG-G-CT-TGGGAC--AGCC-TATCCCA......TTTGC-GCAGCC--ACA-....................................... 2257
H2A.GW.x.ALI            ------G---C-A--C--C------------------G-----A-----G--A--C---T-AA----A-----CT-A-AGG-G-CC-TG-CAC-----C-TATCCCA......TTTGCGGCAGCC--ACA-....................................... 2250
H2A.SN.x.ST             ----------C-A--C--C-----------G-----T-----A-----G--A--C---T-AA----A-----CT-A-AGG-G-CC-TGGCAC----CC--ATCCCA......TTTGC-GCAGCC--ACAG........................................ 1702
H2B.CI.x.EHO            ----------C-A-----C--C-----GA-------C--A--G--G--------GC-AA--------G--TT-A-AAG---CCTTG-CAC-TT-C---AATCCG......TTTGCCGCCGCT--ACC-A-AGCAGGG................................. 2228
H2B.GH.86.D205          ---G------C-A--G--C--C--A--GA-------C--A--G--G--------GC-AA----C------CT-A-AAG-G-CCCT--CAC-T--AC--ATACCG......TTTGCTGCCG-T--ACA-AAAGCAGGG-A-............................ 2230
H2G.CI.92.ABT96         ---G------T-------C-----A--GR-----------A--G--A--A--C--GC--A-----------T-A-AAG---CCTT---CC-TA-AG--CTACCA......TTTGC-GCTGCT--ACA-AAA-CAGGAG-A............................. 1632
H2U.FR.96.12034         ------G---C-A-----C-----A---A----------A--A--A--A--T--GC-GA----A-----CT-A-AAG-G-CCTT-C-AC-TATGC-TATACCC......TTTGC-GCAG-C--ACAGA-AGGCGGCCCA............................. 1748
SMM.SL.92.SL92B         ------G---C-A-----C-----A--G---------C--A--A--G-----CC-CC-CA----A--G--C----AAG-C-CCCTC-CGGGGT--TTAGTG..............GCAGCA---TTTA-GGGAGCAGC---AGGCCA-............................. 1659
SMM.US.x.H9             ----------T-------C--------------R--G--A--A--A--AS-T---C-AA----A--G---T---AAG-T-C-CTG-CGCA--GGC-ACTCCCA......TTTGCRGCAG-C--ACAGAAAGGACAA.............................. 1698
SMM.US.x.PGM53          ----------C-------C----------------G--A--A--A--A--T--GT-AA----A--G---T---AAG-G-C-CTG-CGC---GGCAACTCCCA......TTTGC-GCAG-C--ACAGAAAGGACCA.............................. 2142
SMM.US.x.SIVsmH635F_L3  ----------C-------C---------A------G--A--A--A--A--T---T-AA----A------T---AAG-G-C-CTG-GAC---ACCAACTCCCA......TTTGC-GCAG-C--ACAGAAAGGACAA.............................. 2228
                                  Gag p24 Capsid end Gag p2 Spacer start          Gag p2 Spacer end Gag p8 Nucleocapsid start
MAC.US.x.239            ----------C----G--T-----A--------G--G--G--A--G-----T---T-AA----A-----CC---AAG-G-CCCTCGCAC---TGC-AATCCCT......TTTGC-GCAGCC--ACAGA-GGGACCA............................ 2213
Gag                     _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__A__P__V__P__I__P__.__.__.__F__A__A__A__Q__Q__R__G__P__.__.__.__.__.__.__.__.__.__.__
MAC.US.x.251_1A11       ----------C----G--T-----A--------G--G--A--A--G-----T---T-AA----A-----CC---AAG-G-CCCTCGCAC---TGC-AATCCCT......TTTGC-GCAGCC--ACAGA-GGGACCA............................ 2212
MAC.US.x.251_BK28       ----------C----G--T-----A--------G-----A--A--G-----T---T-AA----A-----CC---AAG-G-CCCTCGCAC---TGC-AATCCCT......TTTGC-GCAGCC----AGA-GGGACCA............................ 2201
MAC.US.x.EMBL_3         ----------C----G--T-----A--------G-----A--A--G-----T---T-AA----A-----CC---AAG-G-CCCTGCGAC---TGC-AACCCCT......TTTGC-GCAGCC--ACAGA-GGGACCA............................ 1699
STM.US.x.STM            ------G---T-A-----------A--G--T--G-----A--A--G--A--T---C-GA----A-----CT---AAG-G--CTTCC-AC---ACC-ACTGCCC......TTCGC-GCAGCC--ACAGCAGGGA............................... 1872
MNE.US.x.MNE027         ----------C----G--T--C--A-----------------A--A--A---------A-T-AA----A-----CC---AAG-G-CCCTTGCAC---GGC-ACTCCCT......TTTGC-GCAGCC--ACAGAAGGGACCA............................ 1693
DEB.CM.99.CM40          G-----G---C-ACAT-----C--A--G--G-----C--ACAG--A-----C---C---A----G--G--T---GC-TC---CCTT---AGAG---AGGGTCC..........CTAGGC---G-A--ACA-A-G-GAGGACCCCCTGGTTCT.............. 1735
DEB.CM.99.CM5           ----------T--CAT-----C-------------T---CAA--G--A--CC-TC-GA----G--G--T---GC-TCG-CCTT---AGA---AGG-TCC..........CTGGGG---G-----CA-A-AGGAGGACAAGGAGG-G-T.............. 1726
DEN.CD.x.CD1            G-----G---T--CA---C--C-------------G--A--A--A-----CC-TCCCAC---A--G---T-AGCTTC--C-TTTC-GAAGCAAGGGTTATGCCATGGTTCA-GG-AAGG-GCCCA-CC-CCAG-AGAA-GA......................... 1732
LST.CD.88.447           ------G-------GG--C-----A--------------A-C------A-G---GA-------A---------GCAAC--CC-TGC-AAGTCAA-TGAGA.........CAG-AT---G-A--AGTGACACCACCA-----TGC-CAG...................... 1128
LST.CD.88.485           ------C-------AG--C--------------------A-C------A-G--AGA-------A--G------GCAAC--CC-TGC-AAGTCAA-TGAGA.........CAG-AT---G-A--AGTGACACCACCA-----TG-GCAG...................... 1128
LST.CD.88.524           --------------A------------------------T-C------AAG--A-A-C--A--A--G------GCAAC--CC-TGC-GGGGCAG-TGAAA.........CATC-C---G-----CA-ACACCACC---A--TGC-CA-...................... 1122
LST.KE.x.lho7           ------G-------A---T--------------------T-CT-----A-G--A-A-A-----A---------GCTTC--CC-T-C-GCA-CAAGTAAGA.........CAA-AC---G-----GTCACTCCGC----A--TGCGCA-...................... 2212
SUN.GA.98.L14           ---G--G-------G---C--C--A--G--T--G--G--GA-------A-G--AGA-CC----A---------GC-TCT-CTTTT-GGCAGCAAGTAGGA......AGACAAGCT---G-----CA-AA-TTACCACCC-GAA-C--CCAG................ 2205
MND-1.GA.x.MNDGB1       ---G------T--GA------------------T--G--ACAA-----A--C---T-AA----A---ATG-----AAC---GGT-GGACA-.......TCA..........CAA-ATT-TG-----CAGA-AGGGCC-CAA-GA........................ 1574
OLC.CI.97.97CI12        G-----G---C-------C-----A--------G-----A---T-----A---------C--A--A-----GC----A-C---GCAG-TGGTGCAACAGAGACCAGCACCA............................................................ 1785
WRC.CI.97.97CI14        ------G---T-A-----------A-----------G--A--AT----AAGC-A-A-AA----AC----TT-ACAAG---CTTTG...CAT...T-ACAG...........GCC---G-A--ACAG-CAGGAGCA-AACCT................................ 2142
WRC.CI.98.98CI04        ----------T----T--------A-----------G--A--AT----AAGC-AGA-AA----AC----CT-ACAGG-G-CTTTG...AGC...T-TCAG...........GC----G----ACA-CAGCCAG-G-AACCT................................ 2160
WRC.GM.05.Pbt_05GM_X02  ----------T--------------------------T--A---T----AAGT-A-A-AA----AC----T--CAGG---CCCTG...--CCAAT-ACAG...........GC----G-A--ACAGCAGGCAG-C-A-................................ 1104
MND-2.CM.98.CM16        ------G---CC-CT---C--C--A--G--------C--TAAAT----A---CA-A-GA----A--G--T----AGG---CCCAGTCA........................GCTG------------ATTCGGGAGGGCCACCG-GGG-C...................... 1900
MND-2.GA.x.M14          ----------T-ACT---------A-----G--G-----AAAAT--------CA-A-GA----A--G--C---CAGA-GA-GCAGTCA........................GAAG-C----------ACTCGGGGGGACCACCG-GGG-T...................... 1824
MND-2.x.x.5440          G---------T--CTG--C--C--A--G-----------TAAAT-C------CA-A-GA----A--G--C---CAG----C-CAGGCAG-T...................G-----------ACTCGGGGGGACC-CCG-GGG-T...................... 1457
DRL.x.x.FAO             G---------C--TT---T--C-----T-----------A---------------G--CA----A-----T-----AG--CAGCAGGCAG-T..........................-----------CAGAATCCCCCC--AGGT-C-CCG...................... 1454
MON.CM.99.L1            G--G--C---C-ACAG--------------G-----T--AATG--C--AT-T--GA--C----A-----T---GCAGGG-C--T-GCCAAC............ATGCCCATG-AT---G-A---GCCA-AGGACCACCCC-A............................ 1611
MON.NG.x.NG1            ----------C-CCA---------A--G--------C----CA------AGC--GT-A--A--A--G--T---GC-AC--C--TT---AGCAAC-TGCCA...ATG.......-AC---G-A--AG-CA-AGGAGG-.................................... 174
GSN.CM.99.CN166         G---------C-T-----------C--A--G-----G-----AACT--C--AT-T--GC-C--A--A-----C---GCAAC--CCCTC--AGGCA---GTTCA.........TAC-AC---G------A----CCCCC-G-A-GG---.............................. 1630
GSN.CM.99.CN71          G---------C-T-----------C--A--G-----G-----AACT-----AT-T---C-C-----A--G--C---GCAAC--C-CTC--AGGC...-G-TCA.........TAC-AT---G----A------CCCCCAG-A-GG---.............................. 1633
TAL.CM.00.266           ----------C-CCAG--C-----A--G-------------CAG-------AGC--GC--A----A-----T---GCAG---CCTT---A...CAA-A-ACCCAACAAGTCATAGCC---G-A--ACAG--GC-AGGAG-CCCA............................. 1844
TAL.CM.01.8023          G---------T--CAG--------A--G--C--G-----TCAG------AGC---C-CA----A-----C---GCGG-G-CCTTG--A...CAG-A-ACCCAGCAAGTCATAGCC---G-A----...--AC-GGGAG-TCCG............................. 1339
MUS-1.CM.01.1085        G--G------T--CAG-----C----------G----------ACT-----AT-C--GC--A----A--G--TC--GC-TCG-CCTTC...AA-CAGG-AGGC...ACACTG...-AC---G----AG---C---GGGAGCC-GGGG---G......................... 1604
MUS-1.CM.01.CM1239      ----------T--CAG-----C--A--G-----------G--A-CA------T-C--GC-AA----A--G--T---GC-TCG-CCCTT...AG-CAA-A-AGT...CAGCTG...-AC---G----AG--.........GCC-GAGGG-A-......................... 1597
MUS-2.CM.01.CM1246      -A----G---T-ACAG--C--C-----GA-T-AG-------ATG--C---T-----GC-GA----A-----C----CTA-T-CTTTG...AGGCAA-A-ACCATTAACACCATC-AC---G-A--A---......CAG--TCC-AGG......................... 1684
MUS-2.CM.01.CM2500      G---------C-CCAG--T-----A-----T-----G--TATA--C--AT-C-A-T-AA----A-----C---GCATC--CTCT-...AG-CAG-A-ACCATTGAGACCATC-AC---G-T---G-G......CAAG-TCC-AG-......................... 1646
SYK.KE.x.KE51           G---------C-TCA---T--C--A--G-----------ACAAT-C--A----A-C-A--A--A---------GTAATGAC-CA-C-CAGCTTG.................GGC-----A--AG-GCCA-GACAAG-A-G-A-TCCC............................ 1655
SYK.KE.x.SYK173         G--G------C--CA---T-----A-----C-----------TTA--C--------A-C-CA----A--G--C---GTAATG-CCCAGC-G----TA...................-AT---G-A--AG-GCCATCAAAAG--.............................. 2018
RCM.GA.x.GAB1           ----------C-A-----C--------G-------------CAG-----A--TC-G--AC-A--A--G--C---CAAATGA-GCA--G-AAT.........................--C---GCT---CA-TCAGCAAAC.................................. 1360
RCM.NG.x.NG411          G--G--G---C-A--------C--A--G-C-----T--ACAG--C-----T--G--GC-A-------------CAGATG--GCA-.......................AGTAAT--C---GCC--ACAG--AGGACC---A--AGG-CCC.......................... 1372
GRV.ET.x.GRI_677        G--G------C-A-T---T--C-------------G--ACAA--------T-AGC-AA---TA---ATG-----TA-T-G-CAG............................-AT---G-C--AGTG--ACC-CAG-AA--GGGCCCC.......................... 2054
SAB.SN.x.SAB1C          ------G---C-T--T------C--A--CA-T-----G-ACAA----------C-TC-AA----G-----T----CAGCG-CCTTTC-GCA-CAA--AGTT......GGGAAC--CT-TG-----CA---GGCAAGACCC-GGGGTCCC......................... 2228
TAN.UG.x.TAN1           ------G---C-------T----------------------A--A-----A---AGC-CA---TA---------CAA---A-GCA-GGA.......................GTT-AC---G----AG---CACCGCGAG-AGGTAGGG-C......................... 2083
VER.DE.x.AGM3           ----------T-A--T--C-----A-----G---------AA-TT-C--A----A---GA----A---ATG---CAAA-TA-GCA--GCCAG.......................-AC---------ACAG---GG-CAG--AGGAAG-CC-......................... 1624
VER.KE.x.9063           ----------C----T--C-----A--------G--C--AA--T-C--A-----------A----A---ATG---CAAA-CA-GCAG-C-CA-......................-AC---G-----CA---GGCAGCAG-AGGAGT---G......................... 2123
VER.KE.x.AGM155         T---------C-------C-----A--G--G-----C--AA-TT-C--A--C-A---CA----A---ATG---CAGA-CC-GCAG-G-CAG.......................-AC---G-A---CAG--AGG-GGA---GGAAG-CC-......................... 2112
VER.KE.x.TYO1_patent    T---------T-A--G--T--------G--------C--AA--T-C--A----A---AA----A---ATG---CAGACCA-GCA----CA-........................-AC---G-----CAG--AGG-CCA-AA-GA---.......................... 2117
COL.CM.x.CGU1           G-----C---C-AG-------C--------G--------A-A---C-------AG--GC-A--A--G--T---CAG--GT-CCAGC-AGAGAGA...............ACC-AC-----TG-AGT-AAG-CAGCA-AATG-TTC-ATTGTCAGGGAATAGGGCACTTAG 1759
```


PLV Complete Genomes

```
H1B.FR.83.HXB2                 ......AAGATTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACACAGCCAGAAATTGCAGGGCCCCTAGGAAAAAG...... 2022
Gag                            ._.__.__.K  I  V  K  C  F  N  C  G  K  E  G  H  T  A  R  N  C  R  A  P  R  K  K__.__.__._
H1A1.UG.85.U455                ......-GA-GAA------------C-------------A--CT----A-----T------------------...... 1471
H1B.US.90.WEAU160              ......--A-C-A--------------------------A---T---------------------------...... 2021
H1C.ET.86.ETH2220              ......-GAGCAA---A-------C-------G--------CT--------------------------A...... 1393
H1D.CD.84.84ZR085              ......--A----------------C-----------------T--A-------T-----------------...... 1542
H1F1.BE.93.VI850               ......-GAG------A----T----------------A---T---------------------A------...... 1359
H1G.SE.93.SE6165               ......-GA-C-A-C--A----------------G----A--TCT--------------------------...... 1428
H1H.CF.90.056                  ......--AT------A--C----C----------G--A---T---------------------------...... 1372
H1J.SE.93.SE7887               ......-GA-----------------------C---A---T--A-A--C-------------A------...... 1342
H1K.CM.96.MP535                ......--A----------C----C-----------------T---------------------A------...... 1218
H101_AE.TH.90.CM240            ......--A-GAA-------C----C----T-G-----A--CT---------------------A---C--...... 1581
H102_AG.NG.x.IBNG              ......-GA-CAA-A----------C-------------A--CT-------------A-------------G-...... 1547
H103_AB.RU.97.KAL153_2         ......--A-GAA------------C-------------A--CT---------------------------...... 1248
H104_cpx.CY.94.CY032           ......-GA-C-A------------C-------G----A--TCT--------------------------...... 1391
H1N.CM.95.YBF30                ......--ACCCA---A-------------------G--C--TTTG--A-----C--T-A---------A-G-GGA...... 1589
H1O.BE.87.ANT70                ......GGA-C-A-A--A--------C----A----G--A--T-T---A--------TC-A--A-----------...... 2086
H1O.CM.91.MVP5180              ......GG-CCCA-A--A--C-----------A----G--A--T-T---A-A---C--TC-A--A-----A----G-...... 2061
H1P.FR.06.RBF168               .....AGA---CCA--------C--T--C-----A--AT---A---T---A-AG--C--T-A---A-------GG--A...... 1492
CPZ.CM.01.SIVcpzCAM13          ......-GA-AAA-A--A--C-AT--------A--G-----C--TCTG--------C--T-A---A--A----G----...... 1605
CPZ.CM.05.SIVcpzEK505          ......----CCA-C--A--C--T--------T-----G--C--TCT---A-----C--T-A---A-----A------...... 1570
CPZ.CM.05.SIVcpzLB7            ......-GAG--A---A-------------------------T---A--------T-AA-----C----G---A...... 1568
CPZ.CM.05.SIVcpzMB66           ......-GA-----C--A--C--T-----------G-----T----T---T--G-------A------C--A-G---A...... 1556
CPZ.CM.05.SIVcpzMT145          ......--A-AGA---A--C--T--------A--------A---CT---A--G--C---------A--CC-A-G----...... 1561
CPZ.CM.-.SIVcpzMB897           ......-GA-----G--A--C--T--C--------G-----T----T---A-----C--T--A---------A-G---A...... 1561
CPZ.CM.-.SIVcpzDP943           ......-G--AAA-C--A--C--T--------------------TTT---A-----C-----------A--A-GG---...... 1581
CPZ.CM.98.CAM3                 ......-G--AAA---A--C--T--C--------G-----A--TTT---------C-----A-----C--A-GG---...... 1412
CPZ.CM.98.CAM5                 ......----AAA---A--C-AT--------------------TGT---T-----C-----A-----A--A-G----...... 1715
CPZ.GA.88.GAB1                 ......-GA-AAA-A--A--C--T--------G--------T---CT---T---------T-AA--A--A--A-G---A...... 2089
CPZ.GA.88.GAB2                 ......-G--GA--C-----C------------A--G-----T--T-T---A--------T-AA---------A-G----...... 1400
CPZ.US.85.CPZUS                ......-GA-AAA-C--A-----T---------------G--A--TCT--------------------T--A--A-G---A...... 2070
CPZ.CD.90.ANT                  ......-GACC-C-C--A-----T--C--C--T-----G--C--T--T--A------------A------A--A-GG--A...... 1452
CPZ.TZ.00.TAN1                 .....AAGGGTCAGC-GC-------T--C-----G---T---C--T-----A--------T-----T--A--A--G--A...... 1666
CPZ.TZ.2001.TAN2               ......GGTCAGT-GC-------T--C-----A--GAT---C--T-----A----------------T--A--A--G--A...... 1665
CPZ.TZ.2002.TAN3_1             ......GGTCAGT-GC-------T--C-----A--GAT---C--T-----A----------------T--A--A--G--A...... 1665
GOR.CM.2007.SIVgor2139_287     ......GGACCA--A-----C-----------G---T---A--T-T---A-A------T-----A--C--A-G----...... 1580
GOR.CM.2007.SIVgorCP2135con    ......GGACCA--A--A--C-----------G---T---A--T-T---A-A------T-----A--C--A-G----...... 1579
GOR.CM.2004.SIVgorCP684con     ......GG-CCA--A--A--C-----------G--G-T---A--T-T---A-A------T-----A--C--A-G----...... 1499
H2A.DE.x.BEN                   .....AGA---GCAA---G--A--GG--C-----A--G--G--A---T-G--A---C-G---C-A--A-----A-G-C--...... 2332
H2A.GW.x.ALI                   .....AGA-G--CAA-------C-GG--------A--G---------T-G--A---C-A---C-A--A--C--A---C--...... 2325
H2A.SN.x.ST                    .....AGA-G--CAA-------C-GG-----C--A--G---------T-G--A---C-A---C-A--A-----A-G-C-A...... 1777
H2B.CI.x.EHO                   .....AAG-G--CA--G-CA--C-GG--C--C-----G-CG--A--T--------GC-G----A----------A-GGC--...... 2303
H2B.GH.86.D205                 .....AGAGG--CA--G-CA--C-GG--C---------C-G--A-----------GC-A---------------A-G-C--...... 2305
H2G.CI.92.ABT96                ....AAAAGG-GC-CAA-A--A----GG-----------G--G--A--T----TT---C-G-----A------C--A-G-C--...... 1710
H2U.FR.96.12034                .....AGG-G-CCA--C-GA----GG--C---------G-----A-----T--A--GC-G--T-AA--------A-G-C--...... 1823
SMM.SL.92.SL92B                ...GGGAATAAGCCT--CA-C-G------T--C-----A--GAC---A---T-T--G---C-G-----A--T--C--A-GG--...... 1740
SMM.US.x.H9                    .....RGA-----AA--W------RG--C---R-A--G-----A---T-T--A---C-A-T---A---------A-G-C-R...... 1773
SMM.US.x.PGM53                 .....AGA----CAA-------C-GG--------A--G--G--C---T-T--A-A-C-A-----A---------A-G-C--...... 2217
SMM.US.x.SIVsmH635F_L3         .....AGG-G--CAA-C-------GG--------A--G--G--A---T-T--A---C-A---------------A-G-C--...... 2303

MAC.US.x.239                   .....AGA---CCAA---------GG--------G-----G--A---T-T--A--GC-A-----A-----A--A-G-C--...... 2288
Gag                            .__.__.__.R  K  P  I  K  C  W  N  C  G  K  E  G  H  S  A  R  Q  C  R  A  P  R  R  Q  .__.__.__._
MAC.US.x.251_1A11              .....AGA---CCAA---------GG--------G-----G--A---T-T--A--GC-A-----A-----A--A-G-C--...... 2287
MAC.US.x.251_BK28              .....AGA---CCAA---------GG--------G--G--G--A---T-T--A--GC-A-----A-----A--A-G-C--...... 2276
MAC.US.x.EMBL_3                .....AGA---CCAA---------GG--------G--G--G--A---T-T--A--GC-A-----A-----A--A-G-C--...... 1774
STM.US.x.STM                   .....CGA-G--CA--A--A--C-GG--------G----C----------T--A-A-C-G--T-AA-G---G--A-G-C-A...... 1947
MNE.US.x.MNE027                .....AGA---CCAA---------GG--------G-----G--A---T-T--A--GC-A-----A-----A--A-G-C--...... 1768
DEB.CM.99.CM40                 .....AGA-GACAGA---GA--C-----------AC--ATT--A--TCTCCAA-AGG-C----AAAGA-----A--G-CC...... 1810
DEB.CM.99.CM5                  ....CCTAGA-GACAGC---GA--C--T--------AC-GATT-----TGTGCAA--GG----T-AAAAG-----A---GTT...... 1804
DEN.CD.x.CD1                   ......-G-GGAC---GA-----------------C--ATT--A---CT-----AGG----T--AAAG--C-A--G-GT-ACT...... 1807
LST.CD.88.447                  ...GGAAGGTTTGTAAGAACAGGAGGAGGAGGCCCTAGA---CC-T-A-CT--C--T--------A--GCC---T--T-----A----TG--T--ACAG-----AC-GG--...... 1236
LST.CD.88.485                  ...GGAAGGTTTGTAAGAACAGGAGGAGGAGGCCCTAGA---CC-T-A-CT--C--T--------A--GCC---T--T-----A----TG--T--ACAG-----AC-GG--...... 1236
LST.CD.88.524                  ...GGGAGATTTGTAAGAACGGGGGGAGGTGGCCCAAGA-GACC-C-C-CT--------------G--GCC---A--TT----A-A--TG------CAG-----AC-GG-A...... 1230
LST.KE.x.lho7                  ...GGCAGGTTTGTGAGAACAGGAGGAGGAGGTCCCAGG--ACCCC-C-CT--C-----------A--GCC---T-------A----TG-----ACAG--C--AC-GG-A...... 2320
SUN.GA.98.L14                  ...GGGAGGTTTGTAAGAATAGGAGGAGGAGGACCTAGG---CCCA-G-CT-----T--------T---CC---A--TCT---A-ATC-G-----A-AG--A-A----GGA...... 2313
MND-1.GA.x.MNDGB1              ......GGACCAGTTAGACAACCTACTGGA-G--AACC--TC--C-----C---AAT-----------TGT---A--GTTC-T--A---------A-GG--A...... 1670
OLC.CI.97.97CI12               ......GGAGGAGGGCCTCAGCAAGG--GGC-G-G-----AT--------T-G-TTT-----T--TAAA-AT-----TTCACAAAAGG-TCC-...... 1872
WRC.CI.97.97CI14               ......AAAGG-CCAA---GA-----T---------AC--ATT-----T-TG--A-A-G-A--TCCCAAG--A-A----CCAGTCCCTATGAGAAGA 2232
WRC.CI.98.98CI04               ......AAAGGACCAA---GA---------------AC-GATT-----T-TG--A-A-G-C--TCCCAAG--A-A----CCAGTCCCCTTGAGAAGA 2250
WRC.GM.05.Pbt_05GM_X02         ......AGAGGACC-C-A-GA-----T---------AC--ATT-----T-TG--A-A-G----T-A-AAA--A-AA-GGTCTCCTCAGCAAAGG... 1191
MND-2.CM.98.CM16               ......CCCCCGAGACAACCTCCT-GA-ACA-C-GA--CCCT--C-----A---TTT--T--TGGGCTG--GG-C--T-TAAG----C-A-------...... 1990
MND-2.GA.x.M14                 ......CCTCCGAGACAACCCCCT-GA-ACCCC-GA---CC---C-----A-GGT-T--A--TGTTCTG---G-C--T--ACT---C--A--G---...... 1914
MND-2.x.x.5440                 ......CCCCCGAGACAACCCCCT-GA-ACCC--GA--CCC------C--A--GTTT--A--TGT-CTG--GG--------A-----A--A--G-GA...... 1547
DRL.x.x.FAO                    ......CGGGGACCCAGGGGACCCCCTCCA--A-ACCC--GA---CCA--C--C--GC-GTTT--A--T--CCTG---C-A--T-CAAAG-----A-------...... 1550
MON.CM.99.L1                   ......AGGAAGGGACAACCC--A--C-----------A---TTT---------TG--G-A--------AA-----CCAA-GG-G-...... 1689
MON.NG.x.NG1                   ......RGTCAGCCACGCCGACAAGGCRW-CAAA-C-G-----AT--------G---TTT--A--TGTG--A-AG--C--T-CT--T--C--A--G-CA......  267
GSN.CM.99.CN166                ......GGA--A-CCCCA--A----AT--------AC--TTC--A---CT---A---G----TCCTAAA--C-AAG---GA...... 1705
GSN.CM.99.CN71                 ......GGA--A-CCCCA--A----A---------A--GTTT--A--T-TC--A---G-C--TCCCAAA--C-A-G---GA...... 1708
TAL.CM.00.266                  ......AGGGGGAGAGGGGGACCTAGAAGGAACCCCCCAGG-CAGA---G-----AT--C-----A---TTT--C--T-T---A-AG--C---CC------GCC--G-C--AAGGTG...... 1955
TAL.CM.01.8023                 ......AAAGGGAGAGGGGGACCAAGAAGAACTCCCCCAGGACAAA-C-GA----AT-----------GTTT-----T-TT--A-A---C--TCC---T---CA---GCCACGGGGG...... 1450
MUS-1.CM.01.1085               ......GGCAACGGCAGCCAGGGAAACCGAGGA-A-CCC-GA--C-A---C-----GC--T-T--C---GT------GG-----CCCAAA--C-----C--A...... 1700
MUS-1.CM.01.CM1239             ......GGCAGCCAGGGAGGTCCCAGGGGA-ACCC--G-----A---C-----GC-GTTT--A----TG--A----GC---CCCAAG--A-AA-CT-G-...... 1690
MUS-2.CM.01.CM1246             ......GGAGTAATGGGAAAAAAGAGGGAA--C-GCACACGT--C-A---C-----AC--TTT--A--TTTG--A---G-----CCCAAG----A-TCT-CC...... 1780
MUS-2.CM.01.CM2500             ......GGAGGGGCCGGAGGAGGGCCTAGA-G--CCCCACG---C-----C-----GC--TTC--A--T-T-----A-G-----CCCAAA--A--AGTT-G-...... 1742
SYK.KE.x.KE51                  ......AGGAGAGG-CCAACA-GA--C--T--------GC-GTT---A--TCTTCAA-A-G-C---CCTAGG--A-AA--GCTT...... 1733
SYK.KE.x.SYK173                ......AGATCA--GA----A--C-A---------GC--ATT-------TGCAA-A-G-C----AAAAG--ACTT---GCT...... 2093
RCM.GA.x.GAB1                  ......AGGGGGCCTCCAAGAAGATCAGGAGGAAATCCA-A-T-A-GA----A---------T--GCC---A----TTT-T---T----T-AA---------A-GG---...... 1462
RCM.NG.x.NG411                 ......CCTAAACTAGGGGGAGGACCT-G-T-CC-C--A--C-AT------------AC---T--T--T---T------AAA-AAG---A--G--A...... 1465
GRV.ET.x.GRI_677               ......CGAGG-CCGC-A--A--C--T-------------TTT--A--T-TGCAA--GG-A----A---A--A--AC-G-TC...... 2129
SAB.SN.x.SAB1C                 ......TTAGGGGGAAGAGGTAGACCTCTGAACCCA-A-A----A----A---------A--GCCT--T---CTG--T---TT---T-A-----------G-C-A...... 2327
TAN.UG.x.TAN1                  ......AGAGGCAGAGGCCCACCCCGC--C-----A-----AC--ATT-----T-T-CAA-A-G----TCCAAGAG-AG-ACC---C...... 2164
VER.DE.x.AGM3                  ......AGACCCCCA--A-------A---------A---TTT--C--T-TGCAA---C-A---CCT-AA--A--A--G-T-...... 1699
VER.KE.x.9063                  ......AGACAGAGACCCCCAC-A--A----A------C--A---TTC--A----TGCAA--GC-A----CA-AA--A--A--G-TA...... 2204
VER.KE.x.AGM155                ......AGACCCCCGCCA-------A---C-----A---TTT--C----TGCAG--GC-G--TCCT-AG--A--A----TA...... 2187
VER.KE.x.TYO1_patent           ......AGACCCCCAC-A-GA----AT--------A---TTT--C--T-TGCAA---C-A--TCC--AA--A-------CA...... 2192
COL.CM.x.CGU1                  CCAGAATGTGTCCCAAAAGACCAATAGGAGGAGCCGGTCGAGGAAGAGGCCGAGGACGGGGAGGCTTTAGAGGAGCTCCCAGA-G-CCA--G-G------T-CC--CAATC-G-----A----TGCAA---G----TCCC...... 1899
```

```
                                                                                                          Gag-Pol -1 ribosomal slip site
                                                                                                   Gag p7 nucleocapsid end Gag p1 start                                        Gag p1 end Gag p6 start
H1B.FR.83.HXB2              ...........GGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATTGTACT.........................GAGAGACAGGCTAATTT.TTTAGGGAAGATCTGGCCTTC.....CTACAAGGGAAGGCC........AGGGAATTTTCTTC 2137
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__.__.__Y__K__G__R__P__.__.__.__G__N__F__L_
Pol                                                                                                                     _F_#_F__R__E__D__L__A__F__.__.__L__Q__G__K__A__.__.__.__R__E__F__S__S
H1A1.UG.85.U455             ...........--------------C--G--A-------------------C--C---..........................-----------------.---------A--T-------.....-A-------G--------...........----------C-- 1586
H1B.US.90.WEAU160           ...........-------------------C---------------------------..........................-----------------.--------------------T-----.....-C-A-------------...........----------C-- 2136
H1C.ET.86.ETH2220           ...........-----------------------------------------C----C..........................-----------------.--------GAC-T-------.....-A------------------...........--------C---- 1508
H1D.CD.84.84ZR085           ...........------------------------------------------C---..........................--A--------------.---------A--T-------.....-C------------------.......G-----C---- 1657
H1F1.BE.93.VI850            ...........------------------GA---------------------C--C---..........................--A--------------.---------A--T-------.....-A-------G---------.......C--A-----C---- 1474
H1G.SE.93.SE6165            ...........------------------------T-----------C--C--G.........................-----------------.---------A--T-------.....-A-------G---------...........---------- 1543
H1H.CF.90.056               ...........------------------GA--------T--G--------C--C--A.........................-----------------.---------A--T-------.....-AG---A--G---------...........---A------C- 1487
H1J.SE.93.SE7887            ...........-----------------------------------------C--C---.........................-----------------.------------T-------.....-AG---A--G---------...........------C-----C- 1457
H1K.CM.96.MP535             ...........---------------G-----------T--G---------C--C---.........................-----------------.---------A----------.....-C-------G---------...........---------- 1333
H101_AE.TH.90.CM240         ...........--T----------C--G-----------T-----------C--C---.........................-----------------.---------A--T-------.....-A------------------.......G-----------C-- 1696
H102_AG.NG.x.IBNG           ...........---------------G--A--------------------C--------.........................-----------------.---------A--T-------.....-AG------G---------...........---A-------C- 1662
H103_AB.RU.97.KAL153_2      ...........------------------------C--T-----------C--C---.........................-----------------.---------GA--T-------.....-AG---A--G---------...........---A-------C- 1363
H104_cpx.CY.94.CY032        ...........-----C-----G------------G-----T-----------C--C---.........................-----------------.---------GA--G-------.....-AG---A--G---------...........---A-------- 1506
H1N.CM.95.YBF30             ...........---------------G-----GC-A--------T-----------------AAAATGAA.............GA--------------------.--------G----T--C-......--T---A--G--A--..........---A--C--C-CC- 1710
H1O.BE.87.ANT70             ...........--T--C-------------C-------T------------------C-GAAAT.............GA-A-------A-------.-------C--ATA-------C-.....GGGGGGCACG-----------.........---C-----A-G-G- 2204
H1O.CM.91.MVP5180           ...........--T--C-------------C-------T------------------C-AAAAT.............GA---------A-------.-----------TA-------C-.....GGGGGGCACG-----------.........---C-----A-G-G- 2179
H1P.FR.06.RBF168            ...........--A--C-----G-------C-------T--T---------------C--C-AATCA.............GA---------A-------.----A-------------C-.....TGGGGGCAAG-----------.........---C--C-A-G-A- 1610
CPZ.CM.01.SIVcpzCAM13       ...........-----C-------------GC----G-------------------C---T-A...............-G------A-TA-----.-------------GG----------.....-CG--GC--G-----------.........---C--C----G-A- 1720
CPZ.CM.05.SIVcpzEK505       ...........-----C-------------GC-A--G-----T--------G--C--C-GAAGT..........GGA--A--G-----G------.----------G---------CT.......-C---A--G-----------.........---A-----C-CC- 1691
CPZ.CM.05.SIVcpzLB7         ...........-------------------GC----------T-GG---C----------ACATG.............-G--AG...-T-------.-------------AC------------.....--C-C------------------.........---A--C-----C- 1683
CPZ.CM.05.SIVcpzMB66        ...........---------------G------GC-A-----G--T--G----G-A-----C--AAAT..............-A--------------------.-------------AC------------.....-AG----------------------.........------------C---- 1674
CPZ.CM.05.SIVcpzMT145       ...........-----C-----G-------C-----G-----T--G-----------------A-CAAAATAAT.........-GA--G...AT---------.-----A--A-AT------C-......--GGGG---G------------.........------C-A-G-G- 1682
CPZ.CM.-.SIVcpzMB897        ...........-------------------GC-A-------------CA-----A-C--CC-CGAT...............AT-CA--G---G--------.-------------A-------------C-......-CG---A--G-------------.........---------C--C- 1679
CPZ.CM.-.SIVcpzDP943        ...........--G--C-------------C-A--G--G--T--G--A----------------AGCT...........GGA--C-------------------.-----------AGGT---T--C-......G-GG-GC--GG-AT-.........-AA--GACCAGGGA 1702
CPZ.CM.98.CAM3              ...........--A--C-----G-----GC-A--G--G--T---------A-C--CC-AGCT...........GGA--A-------------------.-----------AGGT---T--C-GTG.....G-GC--GG-AT-AAAGAGACC------C--C--G- 1542
CPZ.CM.98.CAM5              ...........--A--C-----------C----C--------G--TA-C--------A-----AACTGGA..............GC--------------------.-----------GAT---T--C-GTG.....G-GC--G--AT-AAAAAGACC------C------CG- 1845
CPZ.GA.88.GAB1              ...........-----------CGG------GC-A--G-------------------C--C--A.................GA-------TG-------.-----------AGGT--------CCG......--GC--G-------------.........---C--C---G-G- 2204
CPZ.GA.88.GAB2              ...........--A-----------G--------A--------T--G-----G--C--CCAAAAT...........GAA-GA-------TG-------.-------GG-GG---------CCG......G-GC--G-------------.........---C------CC- 1521
CPZ.US.85.CPZUS             ...........--G--C-----G-----GC------------T-------------C--C--AGCAGGA..............A-C--------------------.-----------ACAT---T-CC-ATCGTG..G-GC--GG-AT-AAAGAGACC------C--C--GG 2203
CPZ.CD.90.ANT               ...........-----C-----G-------C-------------GC-T---A-C---C-AGCAACA..............AATACAG--A-A-TA-------.-----------ACCGACC--CA-.............-T-GT--GGGTGCAGACC------C---G-G- 1579
CPZ.TZ.00.TAN1              ...........--T--C----GG-------C-A--G-------T-------G--C--C--CACCAGAAAC...........AACAGC-CTGG--TA-------.-----------ACG-ACC--C-T..............GT--GGGTGCAGACC------C---G-G- 1793
CPZ.TZ.2001.TAN2            ...........--T--C----GG-------CC-A--------T--G-----------C-----CACA.........AAAAAATAACAATAGCCC-GGA-TA-------.-----------CG-ACC--C-T.....G-GGGG-T-C--A--...........G-----C---G-A- 1795
CPZ.TZ.2002.TAN3_1          ...........--T--C----GG-------CC-A--G-----T--G-----------C-----CACA.........AAAAACAACAACAGCCC-GGA-TA-------.-----------ACG-ACC--C-T.....G-GGGG-T-C--A--...........G-----C---G-G- 1795
GOR.CM.2007.SIVgor2139_287  ...........--A--C-------------GC-A--T--G-----------------C--C-GATCA..............GC--------G-------.-----------ATA-------C-.....AGGGGGC-AG-------------.........---C--C-A-G-G- 1698
GOR.CM.2007.SIVgorCP2135con ...........--A--C-------------GC-A--G--G-----------------C--C-GATCA..............GC--------G-------.-----------ATA-------C-.....AGGGGGC-AG-------------.........---C--C-A-G-G- 1697
GOR.CM.2004.SIVgorCP684con  ...........--A--C-------------GC-A--T--G-----------------C--C-GATCA..............GC--------G-------.-----------ATA-------C-.....AGGGGGC-AG-------------.........---C--C-A-G-G- 1617
H2A.DE.x.BEN                ...........-----C-----G-----C---CC-------ATC---GC-A-C--CC-G..................--A--------AGG---.-------TT-GG-CCA-GGGG...........AAAG-A---.........TC-C--C--C-CCG 2441
H2A.GW.x.ALI                ...........-----C-------------C---CC-------T-TT---GC-A-C--CC-A..................--A---------GG---.-------CTTGG-CCATGGGG...........AAAG-A---.........TC-C--C--C-CCG 2434
H2A.SN.x.ST                 ...........-----C-------------C---C--------ATC---GC-A-A--CC-A..................--A--------GGG---.-------TT-GG-CCATGGGG...........AAAG-A---.........CC-C-----C-C-G 1886
H2B.CI.x.EHO                ...........--A--C-------------C--C---------ATC---TC-A-A--CC-A..................--A--------GGG---.-------TTCGG-CCCTGGGG...........AAAG-A---.........TC-C-----C-CCG 2412
H2B.GH.86.D205              ...........--A--C-------------AAC----------ATC---TC-A-A--CC-A..................--A--------GGG---.-------TTAGGACCCTGGGG...........AAAG-A---.........TC-C--C--C-CCA 2414
H2G.CI.92.ABT96             ...........--G--C-----G-----G---CC---C---ATT---GCCA-A---C-A..................-------------AGG---.-------TTTGGACCTTGGGG...........AAAG-A---.........TC-C--C--C-CCA 1819
H2U.FR.96.12034             ...........--A--C-----------A----CC-------A-C---GCTAGC--CC-A..................--A-CCA---TGGG---.-------CTTGG-CCTTGGGA...........AAA--A---.........TC-C--C--C-CCA 1932
SMM.SL.92.SL92B             ...........--A--C-------------GG-A-------GTATTCA-GCCA-C--CC-A...............A-CCAGA----GGG---.-------CCTTGG-CCATGGGG...........AAAG-A---.........TC-C--C--C-CCA 1849
SMM.US.x.H9                 ...........-----C-----G---------A-C---C--TGTT---GCTA-A--CC-C..................--A--------GGG---.-------CTTGG-CCATGGGG...........AAAG-A---.........CC-C-----C-CCA 1882
SMM.US.x.PGM53              ...........-----C-------------AAC---C--TGTC---GCCA-A--CC-C..................--A--------GGG---.-------CTTGG-CCATGGGG...........AAAG-A---.........CC-C-----C-CCA 2326
SMM.US.x.SIVsmH635F_L3      ...........-----C---GGG--C-----AACG--T--TGTT---GCCA-A--CC--..................--A--------GGG---.-------TTTGG-CCATGGGG...........AAAG-A---.........CC-C-----C-CCA 2412
                                                                                                          Gag-Pol -1 ribosomal slip site
                                                                                                                      Pol start
                                                                                                   Gag p8 Nucleocapsid end Gag p1 Spacer start                          Gag p1 spacer end Gag p6 start
MAC.US.x.239                ...........--A--C-------------AATG-AC--TGTT---GCCA-A--CC-A.......................--C--------GGG---.------CCTTGGTCCATGGGG...........AAAG-A---.........CC-C-----C-CCA 2397
Pol                                                                                                                 F__F__R__P__W__S__M__.__.__.__.__G__K__E__.__.__.__.__A__P__Q__F__P
Gag                         _.__.__.__._G__C__W__K__C__G__K__M__D__H__V__M__A__K__C__P__.__.__.__.__.__.__.__.__.__.__D__R__Q__A__G__F#__L__G__L__G__P__W__G__.__.__.__.__K__K__P__.__.__.__R__N__F__P_
MAC.US.x.251_1A11           ...........--A--C-------------AATG-AC--TGTT---GCCA-A--CC-A.......................--C--------GGG---.------CCTTGGTCCATGGGG...........AAAG-A---.........CC-C-----C-CCA 2396
MAC.US.x.251_BK28           ...........--A--C-------------AATG-AC--TGTT---GCCA-A--CC-A.......................--C--------GGG---.------CCTTGGTCCATGGGG...........AAAG-A---.........CC-C-----C-CCA 2385
MAC.US.x.EMBL_3             ...........A-A--C-------------AATG-AC--TGTT---GCCA-A--CC-A.......................--C--------AGG---.------CCTTGGTCCATGGGG...........AAAG-A---.........CC-C-----C-CCA 1883
STM.US.x.STM                ...........--T--------------C-AC-G--C--T--G---GCCA-A--CC-A.......................--A-------TGGG---.------CTTTGG-CCATGGGG...........AAAG-A---.........CCAC--C--C-CCA 2056
MNE.US.x.MNE027             ...........-----C-------------C-AATG--C--TGTT---GCCA-A--CC-A.......................--C--------AGG---.------CTTTGG-CCATGGGG...........AAAG-A---.........CC-C-----C-CCA 1877
DEB.CM.99.CM40              ...........AAA----TT--------GC-------C---ATTGCA---A-----GGA.....................CAGACTCCGAG---A------.----------T-CTCCCTGGG-.............CTCC---AGACC......CA----C-----AG 1925
DEB.CM.99.CM5               ...........AAA--C-TC--------T-----G--C---AT-GCA--GA-C---GGA.....................CAA-TACC-AG---A------.----------TGCTCCCTGGG-.............CTCCC--AGACC......CA-------T-GG 1919
DEN.CD.x.CD1                ...CAACCAGGA-CG--C-TT------------CT---T--T---GCA--GA------GA...........AGTAACACTAATGATC--G------TA-----.----------TGGTCTTGGGAG...........-AAG-AT--TCCCAGGAACTTTGTACAGTACT 1949
LST.CD.88.447               ...........--A--C-----C-----T-GCA-G-A---TAGGT-TGCTC-A---C-AAAGCCT.....................A--G--A---TA-----.-------CT-TGGACCCTGGAG...........-A-TG-T--TCC.........---A--C----C-T 1354
LST.CD.88.485               ...........--A--C-----C-----T-GCA---A---TAGGT-TGCTC-A---C-AAAGCCT.....................A--G--A---TC-----.-------CT-TGGACCCTGGAG...........-AATG-T--TCC.........---A--C----C-T 1354
LST.CD.88.524               ...........-------------C-----TGCTA---AC---AG----GCCC-G--CC-AAAACCA.....................A--G-GA---TG-----.-------T-TGGTCCCTGGAG...........-AATG-A--TTC.........---C--C----CAA 1348
LST.KE.x.lho7               ...........--T---------C-----C-GTA---AG--TAG-T-TGC-C-G--CC-A.................AAACCAA-AG--A---TG-----.-------AT-TGGTCCCTGGAG...........--GCAA-CCT--AGG.....CAATT--CC---A- 2441
SUN.GA.98.L14               ...CCACCCGGCTC---C-----------------ATG-----TA--CAAGCTC-A--CC-A.................CAA-GA-A----TG-----.-------T-TGG-CCCTGGAA...........CA-GG----GCC...........------C----CAG 2437
MND-1.GA.x.MNDGB1           ...........--G--C-----T------GCAATG-AT--T--G-AAGCTC-A--CC--...............AAGCCAGCTCAGC--CAGAG--T------.-------T-TGG-CCTTGGGG...........TCCCTCCAAACC.......G-------A--C.. 1792
OLC.CI.97.97CI12            ...........ACA--C-AC--G-----G--ACC---C---AT-GCA---A----C-GG.........GTGGGAGGACAACAACAAC----CA----TG-----.-------ATGGA--GGAGG.........G--A-C-GA---TCCCCTCAACAT-C-ACAGAGGA 2011
WRC.CI.97.97CI14            GGACAGGGACCCACA--C---GG-------G-A-T-----TGT-CA--G-A-C---C-A.................AAAATTAGA-------TA-----.-------TGG-A-AGGGCA...........-ATGCCCT-GAACAATTATCCAGT-CAGAAG- 2368
WRC.CI.98.98CI04            GGACAGGGACCCACT--C---GG-------G-AATT-----TGTGCAA-G-A-C--CC-A.................AAAAGTAGA--G---AT-----.-------TGGGA-AGGGCA...........-ATGCCCT-AAACAATTACCCAGT-CAAAAGG 2386
WRC.GM.05.Pbt_05GM_X02      ...GGACCTCCCAC---C---GG------------ACC--C---ATCCA---A-C--CC-A.................GAAACAAGA--G---TA-----.-------TGG-C-AGGGCA...........-A-GCCCT-CAACAATTATCCAGT........ 1316
MND-2.CM.98.CM16            ...........--A--C-TC--G------G-TCT-------AT-----G-A----CC-A.............................-AGAT--TG-----.-------T-C-CCCTGGGG...........CA-TG-CAAACC......CA----C--C-C.. 2097
MND-2.GA.x.M14              ...........--A--C-TT--G-----GG-CAC------TATG----G-A-C--CC-A.............................-A-AT--TG-----.-------TGCTCCCTGGGG...........CA-TG-CAAGCC......CA----C-----C.. 2021
MND-2.x.x.5440              ...........-------TT--G------G-TCC------T-T-----GGA-C--C.........................CCA-AGAT--TG-----.-------T-C-CCCTGGGG...........CA-TG-CAAGCC......CA----C--C-C.. 1654
DRL.x.x.FAO                 ...........--A--C-TT-GG------GCTTT--AT--TATGT-A-G-A---C.........................CCA-A-----TG-----.-------T-CTCCCTGGGG...........CA-CG--AAACC...CAGGAATTT-CC--AGG 1662
MON.CM.99.L1                ...........AAA----AC--T-----CC-ACCG-------TCGCC------C--CC-GCAGCCCCCC............AAACAAA-C-A-GG--TG-----.-------TCCT-TCGG-C-...........CAAG-A-GGGCC.......-A----C--C-C-- 1816
MON.NG.x.NG1                ...........--T--C-TCCGC--C--C--------G--TRT-TCC---A-C--CC-A.....AACGGAGGACAAAACCAACCCAGA-ACAG--TG-----.-------TCCT-TTGG.............CCCCG-CAAGAG......G---CCAAGGAACT  397
GSN.CM.99.CN166             ...........AAG--C-TT--------C-GA-C---G---TTT-GC---CCA----GA...............ACAGGGACACCAATC-AGG--CT-T-----.-----ACCTTGG-C--GGGCG...........CC-CC-AG-CCC......CA----C-----CAG 1829
GSN.CM.99.CN71              ...........AAG---TT---------C----C-------TT-GCT-GGC-A--C-AAACA...............GGAACACCA-CC-AGG--CT-------.-----ACCTTGG-C--GGGCG.................-C-AG-CCC......CA----C-----CAG 1829
TAL.CM.00.266               ...CCGCCAGGCACT--C-AT--G-----C--ACC---C---AT-GCA---C-A--CT-C.......................TCAGC------C------.-----------TCT-TTTTAGG...........-AA--TC--ACC......G-TC-GAGGA-CGA 2076
TAL.CM.01.8023              ...CCTCCGGGGA------TC--G--C--GC--ATG-------GTGCCGC-C-G--C-GA.......................C--CC------A------.-----------GCT-TTTTAGG...........-AAG-TC--CCC......G-TC--GGGA-CGA 1571
MUS-1.CM.01.1085            ...........AAG----TC--------G-GA---------TTGGCA-G-C-A----GG...............TCTGATAATACC-A-GC--TG-----.-------G-TCC--TCGGCCT.......-C-G--AA-G------------......GA----C--C-CC- 1824
MUS-1.CM.01.CM1239          ...........AAA--C-TC-----------G---G-----TTGGCA---C-G--C-GG...............TCTGAAGGT-CC-A-TC-----------.-------G-TCCTCTTGGCCT.......-C-G-GAA-G------------......-A----C--C-CC- 1814
MUS-2.CM.01.CM1246          ...........A-A----TT--G-----T------------TTGGCC-G-C-G--CCGA...............ACCGACACAGGCA-A--CGCT-TG-----.-----------GATCTCGGC-T.......-CGGTCTC-G-A-----------......GA----C--C-CAG 1907
MUS-2.CM.01.CM2500          ...........AAA--C-TT--G-----C-G-------C---T-GCA---C-G--C-GA...............ACCAACTCAGACA-AG-CGCT-T-------.-----------AGGT-TCGG...........--T-C-CA-G-A----------......GA----C--C-C-G 1866
SYK.KE.x.KE51               ...........AAG--C-TT--C-------GGAACG------AT-GCA-GGC-G----GACAACCAAGAAAAGGACAAGGCAACCCTCCTCC-------C-----.-------AGGG---GGGAA.................CAGGAAACCCCC--CA-------C-G 1872
SYK.KE.x.SYK173             ...........AAA--C-TT--T---------AAC---------T-GCA-GG-CC----GGCAACCCAAGAGAAATCAAGGCCCTCCA-TTGC------------.-------AGGA---GGGGT...........-TCCA-G-----CCC........--CC--C--C-CAG 2235
RCM.GA.x.GAB1               ...........--A--C------------TCCCC--AC--T-TCT--------C--C--A.................................-AG--AATA------.-------GAC--CCCTGGGG...........TCAGG--AAGCC...GAGGAACTT-CC-T-GA 1574
RCM.NG.x.NG411              ...........--G--C---G--------G--------T--T-T------G-----CC-A.................................-A-G---T-------.-------TGCTCCCTGGAA...........-CAGG--AAACC...GAGGAACCT-CCCT-GA 1577
GRV.ET.x.GRI_677            ...........AAA--C-TT--G--C--C--AATT--C--TATGGCA------C--C-AG.......................A-TG--------A------.-------T-TGG-CATTGGGG...........A---GC-AA...ACCAAG-AATTT-G-G-AAT 2244
SAB.SN.x.SABIC              ...........-----C------------C-GCCC--AC--T--G-------------CAA.................................-A------T-------.-------TTTGG-CCCTGGGG...........CA--G--AAACC...CAGGAATTTCCC---GA 2439
TAN.UG.x.TAN1               ...........AAA--CCTA--------G---CC-------TT-GCA--------CCGC.......................AGTG--------C-------.-------GA--GCCAA-C-G.................CAAAGCC...GCGCAATTT---AGA.. 2277
VER.DE.x.AGM3               ...........A-A--C-T---------G--ACC---G--TTT-GCA--------C-GA.................................G------TA-------.-------T-TGG-C--TGGAT.................GC-AAACC...CAG-AATTT-CCCGCCG 1811
VER.KE.x.9063               ...........AAA--CCTT--------G---CT--------TGGCA------C--C-GG.................................G------TG-------.-------T-TGGAC--TGGA-.................G-CAAAACC...GAG-AACTT-C-CGC-G 2316
VER.KE.x.AGM155             ...........AAA---CTT-----------CC---G--TT-GCA------C--C-GG.................................G------TG-------.-------T-TGG-C--TGGAT.................G-CAAAACC...AAG-AATTT-CCCGCAG 2299
VER.KE.x.TYO1_patent        ...........AAA---CTA--G--------ATTG-------T-GCA------C--C-GG.................................G------TG-------.-------T-TGGAC--TGGAT.................GGCAAAACC...GAG-AATTT-CCCGCCG 2304
COL.CM.x.CGU1               .......................................................................................A-C-AG-----A-------.-------GCA-C-A--GGGGTGGA.................--TACAGACAGCCATCTTCCCC-C.. 1964
```

```
H1B.FR.83.HXB2              AGAGCAGACC.........AGAGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCTGGGGTAGAGACAAC........................................................................................... 2198
Gag                         Q  S  R  P  .  .  .  E  P  T  A  P  P  E  E  S  F  R  S  G  V  E  T  T  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
Pol                          E  Q  T  .  .  .  R  A  N  S  P  T  R  R  E  L  Q  V  W  G  R  D  N  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
H1A1.UG.85.U455             ----------.........-------------------C---A-T---TG--ATG----A-A---T........................................................................................... 1644
H1B.US.90.WEAU160           ---------T.........---A-----------------------------TCA---A-------------.................................................................................... 2197
H1C.ET.86.ETH2220           ----------.........--------------------GAGTCT-AGACCAGAGCCAAC--CCC--C-ACCAGA..................................................................................... 1575
H1D.CD.84.84ZR085           ----------.........-------------------C---AGCT-CGG-T-TG---AGGAGATA-C-........................................................................................ 1718
H1F1.BE.93.VI850            ----------.........-------------------C----------G---TCA-A-AG----T---........................................................................................ 1535
H1G.SE.93.SE6165            ---A---GA-.........---A-------------T-C---A---C--G---TC--A-AG----T-G-........................................................................................ 1604
H1H.CF.90.056               -------G--.........---A---------------C----------G---TC--A-AG----TG--........................................................................................ 1548
H1J.SE.93.SE7887            ----------.........-------------------C-------C--G--CTC--A-AG----T........................................................................................... 1515
H1K.CM.96.MP535             ----------.........-------------------C----------G---TC----AG----T---........................................................................................ 1394
H101_AE.TH.90.CM240         ----------.........-------------------C---A-A--GGG--ATG----A-----T---........................................................................................ 1757
H102_AG.NG.x.IBNG           ----------.........G--A--------T------C---------TG--ATG----A-----T........................................................................................... 1720
H103_AB.RU.97.KAL153_2      ----------.........-------T-----------C---A-A---TG--ATG----A-----T........................................................................................... 1421
H104_cpx.CY.94.CY032        ---A---G--.........---------------G--C-CG--AT----AGA-AGGAAA-AG---------..................................................................................... 1567
H1N.CM.95.YBF30             ---CA-C-A-AAGGAA...------C--------G---CT------T-ATG---T-CA--AG--............................................................................................. 1771
H1O.BE.87.ANT70             ----ACC-G-.........-C-C---T---------................................................................................................................. 2232
H1O.CM.91.MVP5180           ---AACA-GT.........GTCC---T---------................................................................................................................. 2207
H1P.FR.06.RBF168            ---AACA-GT.........-C-A-------------TAT.............................................................................................................. 1641
CPZ.CM.01.SIVcpzCAM13       ---A---GA-.........-------------------C------T-ATG---ACCAA-AGA-.............................................................................................. 1775
CPZ.CM.05.SIVcpzEK505       ---CA-C-A-AAGGAA...---------------T---AT------T-ATG---A-CA................................................................................................... 1746
CPZ.CM.05.SIVcpzLB7         ----------.........------C--------G---AT---------TG---T-CA-CC-A-............................................................................................. 1738
CPZ.CM.05.SIVcpzMB66        ---AG-----.........---A--------T------AT----------G---TCA---AG--............................................................................................. 1729
CPZ.CM.05.SIVcpzMT145       -AG----GAA.........G--A--------T-----TATG--ATC----G-A-T-CA................................................................................................... 1731
CPZ.CM.-.SIVcpzMB897        ---AG-----AGAACCACC---A----GT---------AT------T--TG---T-CA--A................................................................................................ 1740
CPZ.CM.-.SIVcpzDP943        -CTT-CTG-AGGGCAGGAAG--A--C-----T-----CAT------T--TG---A-CA................................................................................................... 1760
CPZ.CM.98.CAM3              --G----GAA.........G--A--C-----T--G--CCT------T-ATG---A-CA--A--CA---GG....................................................................................... 1603
CPZ.CM.98.CAM5              --G--G-GAA.........G--A--C-----T-----CATG-----T-ATG---T-CA--A-ACA----T....................................................................................... 1906
CPZ.GA.88.GAB1              ---A----A-.........---A-----G--------GAT------T-ATG---ACCA--AG--AGA.......................................................................................... 2262
CPZ.GA.88.GAB2              -A--G-AGGT.........G--A--------T-----TCC-----C----G---A--CA-AGC-----CA....................................................................................... 1582
CPZ.US.85.CPZUS             -A-A---GAA.........G-----C-----T------AT----GAT---G---A-CAA-A-------GT....................................................................................... 2264
CPZ.CD.90.ANT               ---AGGAGGAAGTAGT...G-----------T-----CAT-----T--A......-CA--AG---CA.......................................................................................... 1637
CPZ.TZ.00.TAN1              ---A--CC--.............AGAG-A--GGAA.................................................................................................................. 1815
CPZ.TZ.2001.TAN2            ---A--CC--.........----AA-GGGAAGAAT-AG--GC--GAACAG-CACAGAC-CC--CAGT-C-...ATCGGCCCCACCGATGACGGA.............................................................. 1877
CPZ.TZ.2002.TAN3_1          ---A--CC--.........----AA-GGGAAGAAT-AG---C--GAACAA-CACAGAC-CC--CAGT-GTGCCATCGGCTCCACCGAT...GGA.............................................................. 1877
GOR.CM.2007.SIVgor2139_287  ---AGCA-GT.........GC-A---------------AT............................................................................................................. 1729
GOR.CM.2007.SIVgorCP2135con ---AGCA-GT.........GC-A---------------AT............................................................................................................. 1728
GOR.CM.2004.SIVgorCP684con  ---AGCA-GT........................................................................................................................................... 1627
H2A.DE.x.BEN                T--C-CA.....................------T-AG-GGCT--TACCA-CAG-ACCTCC.......................GGCAGATCCAGCAGCGGAACTGTTGGAGAGATATATGCAGCAAGGGAGAAAGCAGAGGGAGCAGAGG 2548
H2A.GW.x.ALI                TA-C-C-.....................--TT--G-AG-G-TTA-CACCA-CAG-ACCCCC.......................AGCAGAGCCAGCAGCGGACCTGCTGGAGCAGTATATGCAGCAGGGGAGAAAGCAAAGAGAGCAGAGA 2541
H2A.SN.x.ST                 T-GC-CA.....................-AT---G-AG-GGCT--CACCA-CAG-ACCCCC.......................GATAGACCCAGTAGAGGACCTACTAGAGAAGTACATGCAGCAAGGGAAAAGGCAGAGAGAGCAGAGA 1993
H2B.CI.x.EHO                TCCAGGC.....................-C---A...G-GGATAGTGCCATCTG-GCCCCC...................GATGAACCCAGCATTCGGCATGACACCTCAGGGTGCGATTCCATCTGCACCCCCTGCAGATCCAGCAGAG 2519
H2B.GH.86.D205              T--C-CA.....................--TG--T-AG-G--T--CACCATCTG-ACCCCC...................GATGAACCCAGCAGAGGGCATGACACCTCGGGGGGCGACACCATCTGCGCCCCCTGCAGATCCAGCAGTG 2524
H2G.CI.92.ABT96             T--C-CA.....................--T---T-AG-GTCT--CTCCATCAG-ACCACC.......................................................................GATGGACCCAGCAGTA 1875
H2U.FR.96.12034             T-GC-CA.....................--T---G-AG-G-CT--C-CCA-CAG-CCC.....................................TCCAGCAGAACCAGCAGTGGATCTGCTGACCCCGACAGCCCCTCCAGCGGACCCAGCAGTG 2030
SMM.SL.92.SL92B             T-CAG-C.........................-A--TC-TT--CACCGTCAG-ACCCCC.......................................................................AGATCCAGCAGCCCGG 1899
SMM.US.x.H9                 T-GC-CA....................GATG--T-AG-GGCT--CACCA-CTG--CCYCC.......................................................................AGAGGATCCAGCTGTG 1938
SMM.US.x.PGM53              T-GC-CA....................GATG--T-AG-GGCT--CGCCA-CTG--CCCCC.......................................................................AGAGGATCCAGCTGTG 2382
SMM.US.x.SIVsmH635F_L3      T-GC-CA....................GATG--T-AG-GGCT--CACCA-CTG--CCCCC.......................................................................AGAGGATCCAGCTGTG 2468

MAC.US.x.239                T-GCTCA.....................--TG-AT-AG-GGCT--TGCCA-CTG--CCCCC.......................................................................AGAGGACCCAGCTGTG 2453
Pol                          H  G  .  .  .  .  .  .  .  S  S  A  S  G  A  D  A  N  C  S  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  P  R  G  P  S
Gag                         M  A  Q  .  .  .  .  .  .  .  V  H  Q  G  L  M  P  T  A  P  P  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  E  D  P  A  V
MAC.US.x.251_1A11           T-GCTCA.....................--TG-AT-AG-GGCT--CGCCA-CTG--CCCCC.......................................................................AGAGGACCCAGCTGTG 2452
MAC.US.x.251_BK28           T-GCTCA.....................--TG-AT-AG-GGCT--CGCCA-CTG--CCCCC.......................................................................AGAGGACCCAGCTGTG 2441
MAC.US.x.EMBL_3             T-GCTCA.....................--TG-AT-AG-GGCT--CGCCA-CTG--CCCCC.......................................................................AGAGGACCCAGCTGTG 1939
STM.US.x.STM                T-GC-CA.....................-ATA--T-AG-GGCT--CGCCA-CTG-CCCTCC.......................................................................AGAAGACCCAGCTGCG 2112
MNE.US.x.MNE027             T-GC-CA.....................-ATG-AT-AG-GGCT--CACCA-CTG--CCCCC.......................................................................AGAGGACCCAGCTGTG 1933
DEB.CM.99.CM40              -AG-A-TTGGGAAGTA...CATC--GGAGAG-A--G----TC--GTA-A-TC-A-.............................................................................................. 1977
DEB.CM.99.CM5               -AG-G-TTGGGAAATA...CATC---GAGAG-A--G----TC--GTT-ATTCAA-.............................................................................................. 1971
DEN.CD.x.CD1                -CCCTCAGATGAATCA...G--TTTGC--TGGAGC----TGAG-GT................................................CCAGCAGCAACCACAGTACCCACAAAATTTTCAAGAGGGG 2032
LST.CD.88.447               T--TGG-GGGGAA......T-CAGG-GT...AGT----TC--CTCC-CC.....................................................................AATGGAGAGGAGCCCAACACAAGCAGAG 1422
LST.CD.88.485               T--TGG-GGGGAA......T-CAGG-GT...AGT----TC--CTCC-CC.....................................................................AATGGAAAGGAACTCAACAAAAGCAGAG 1422
LST.CD.88.524               TC-TGG--GGGGC......T-CAGG-GT...GATT--TTC--C-CCACC...................................................................CATGGAGAAAGTGCCTCCAACAAGAGCAGAG 1419
LST.KE.x.lho7               T-G-AG--G-.........T-CAGGG-G...AATT---TC--CTCC-CC.......................................................................AATGGAGAGTGCTCCAACAAAAGCA 2503
SUN.GA.98.L14               TC-TGCCTT-TGCTCC...GCC....................................................TCTAGAGGATCTGACATTGGGGAACAGAATGACACCACCCCAGTCAAAGGCA 2508
MND-1.GA.x.MNDGB1           ................................................................................................GGCACAAGAGGTGACTCCAACAGCTCCACCATTAGAGGAGAAACCTCTG 1841
OLC.CI.97.97CI12            CAGAG-TGGACATCTT...TCCCT-TG--C-TA-C-TG--............................................................................................................. 2048
WRC.CI.97.97CI14            T---G--CTTGAACAA...TTTTGGGC--A-AA--A--ATGCCTTCTGCTCCTC-AAT-AG----GATC-AAACAT.......................................TCAGGAGCTGGAAAAATATTTATTG 2466
WRC.CI.98.98CI04            T---G--CTTGAACAA...TTTTGGGC-G--...-A--AT-CC-TCTGCTCCTC-AAT-AGG---GATC-AAACAT.......................................ACAAGAATTAGAGAAATACCTCCTG 2481
WRC.GM.05.Pbt_05GM_X02      .................................T-A-A--AT--ATGCAG--GTGAATACCTTTGGGC-AGTGGT.......................GCCAACAGCACCAGTGGATCCTACCATGCAAGAACTA 1395
MND-2.CM.98.CM16            ...................G-C-ATGC--TTGA-C---AC...-GC-CC.............................................TCCAATGCCAGGGATGGAAGACCCAGCAGAGAGAATGCTGTTA 2167
MND-2.GA.x.M14              ...................T-CCATGC--CTGA-C---TC--CACC-CC..................................................GATGCCAGGGATGGAGGATCCAGCAGAGAAGATGTTATTA 2091
MND-2.x.x.5440              ...................T-CCATGC--TTGA-C---TC--CACC-CC..................................................GATGCCAGGGTTGGAAGACCCGGCAGAAAAGATGCTGTTA 1724
DRL.x.x.FAO                 CCTC.........................................................................GCTGACCCCATCAGCTCCTCCATTGCCAGGATACCTCCAGGAGGATCCGGCCGAGAGGATG 1727
MON.CM.99.L1                TC-C-TCTGTGCAGCC...CAGCG--C-GA-T-TG----C............................................................................................................. 1853
MON.NG.x.NG1                TTCCAGT-A-CTCCTT...GC-------------G-G---CCA-C..........................................................................CCARGGGGCYTACAAGGTG  459
GSN.CM.99.CN166             TA-CTTTC-TGCTGCC...-ACAG-CC-G--GGA...G--G--AGAACC-CA-CA.............................................................................................. 1878
GSN.CM.99.CN71              TA-CTTCC-TGCTGCC...-TCAG-CC-G--GGA...G--G-T-GAACC-CA-CA.............................................................................................. 1878
TAL.CM.00.266               G--ATTTT--AGTGGT...GC-A---T-G--A--G---AT.............................................................................AGAGCCTCTCACAGCA 2129
TAL.CM.01.8023              G--A-TTC--AGTGGT...GC-A--C--C--A--G--GAT.............................................................................AGAGCCTCTCACAGCA 1624
MUS-1.CM.01.1085            ---C-TCT-TCTTGCC...-ACTG-TC-GC-AGA--AG--G-G-GAA--GGTCAA-CTCTCTCCAG--GT............................................................................... 1891
MUS-1.CM.01.CM1239          ---C-TCG-TTTTACC...-AC-G-CC-GC-AGAG-AG--G-GAGA-C-GGTCAA-CTCTCCCC-C--GT............................................................................... 1881
MUS-2.CM.01.CM1246          TT-CATCG-TCATGCC...TAC-G--C-CC-AGATA---C----CCGG-G-AC-TCTCCCC...ACT-GTAGGGCCCAAGGA................................................................... 1983
MUS-2.CM.01.CM2500          TT-CATCG-TCATGCC...-TCAG-TC-CC-AGATA---C-----TAG---AC-TCTCCCC----GTGGGTCCACGGAAGGAGAA................................................................ 1948
SYK.KE.x.KE51               T--TG--GGAGACACC...-ACAG-CC--C-AG-GGAG--..........................TTGGCCATGGCAGCAACAGACATGGGGAAATTATGCTTCACCCCAACAGAAGCACATAGCAATGGCCACCCCGCTACAGCCACAG 1994
SYK.KE.x.SYK173             T---G--CGAGCCATC...T-C....................TCC-CCTCT-GA...................GGACATCGAAGATGGGCCATGGCTCACATGGTCAGCACAGATGAGCCAACAAGCACAAGCGAAGGCACAG 2336
RCM.GA.x.GAB1               CTTC-TTGA-TCCCTC...T-CT---GGGATGGAGAGCA-TT-CGA-CC.......................................................................TGCAGAAGAGATGCTAAAGAAT 1642
RCM.NG.x.NG411              CTTC-TTGAGACCTAC...................TG-TCC--G.................................GAGAGAGGAAGGATGGAATCCCCAGTACGATCCAGCAGAAGAGATGCTCAGGAAA 1657
GRV.ET.x.GRI_677            -C--AG--GACACAGT...T-GT-TGGA-C-AA--G-CCCCCC............................................................AATGGAAACAGCTTACGATCCAGCAAAGAAGCTC 2318
SAB.SN.x.SABIC              CCTC--TCAGACCAAC...--CT--CC--ATGGA--GG--TT-C---AG.........................GCCGGAGGAAAATTGGTATGCAGACAGACCTCCGACCAGAGGACCAGGGCCAGACGATCCAGCAACAGCCCTGTTAAAGCAG 2567
TAN.UG.x.TAN1               .-CAAG-CGGAGCCGT.........TC-GA-GG-T--CCCGAT-CCTGCGCACGGATTCCC.........................................................CACAGGTTCGCCAGCTGCG 2347
VER.DE.x.AGM3               CT-CTCTTGGGGTGGA...GCCAA-TG-GC----T---CCGAGTCC.........................................................ATACGACCCTGCAAAGAAGCTC 1876
VER.KE.x.9063               CC-CTCATGGAGTGGA...-CC-AGCG-GC----T---CCGACACC.........................................................ATACGACCCAGCAAAGAAGCTC 2381
VER.KE.x.AGM155             CC-CTCTTGGGGCGGA...-CCAAGTG-GC----T---CCGA-C-A...................................................CTCTACACCTTACGACCCAGCAAAGAAGCTC 2373
VER.KE.x.TYO1_patent        CT-CTCTTGGAGCGGA...-CC-AGTG-GC-T--T---CCGAGCG--AC--CCC-................................................ATACGACCCAGCAAAGAAGCTC 2378
COL.CM.x.CGU1               ..................................................................................................................................................... 1964
```

```
                                                                                            Gag end
                                                      Gag-Pol TF end Pol protease start   Gag p6 end
H1B.FR.83.HXB2              ....................................AACTCCCCCTCAGAAGCAGGA...GCCGATAGACAAGGA......ACTGTATCCTTTAACTTCCCTCAGGTCACTCTTTGGCAACGACCCCTCGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTCTA 2321
Gag                                                         T  P  P  Q  K  Q  E  .  P  I  D  K  E  .  .  L  Y  P  L  T  S  L  R  S  L  F  G  N  D  P  S  S  Q  *
Pol                                                         N  S  P  S  E  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
H1A1.UG.85.U455             ....................................G--CT-----GC---A-----..--T--A-----G--.....-A-AC-------GT--------AA-----------------------T------G----A-----A--A--G--G-TA--------- 1767
H1B.US.90.WEAU160           ....................................-------T--------------..---A-----------------G-----------------AA--------------------------------------A-----------G------------ 2320
H1C.ET.86.ETH2220           ................GAGCTTCAGGTTCGAGGAAGC---A--TT---C---------..--T--A-----G---.............AG-C------------AA------------------A--T--T--------A--------A--G--------G-----C 1713
H1D.CD.84.84ZR085           ...................................CCCT-C-AGAAA---G-A---A-...-GAC-A---------ACT..........-C---G---------AA------------------G-T---------------------A--G-----A--------- 1843
H1F1.BE.93.VI850            .....................................C---T---C-----------...--A--A----GG---ACTGTA...CCC---C---G---------AA------------------T-A-----------A--------A--GA-------G-------- 1658
H1G.SE.93.SE6165            .....................................C---T-C-C-----------...-AT--AG--A-------.........G-TA-ATC-C--------AA--------------G---------A------G----A--------A--G----TA-----C--- 1718
H1H.CF.90.056               .....................................C---T---C---------GCA--T--AG--------ACC.........---C---G---------A---------------G--------T--T---G----A----A---A--GT---G---------- 1668
H1J.SE.93.SE7887            .....................................C---T-C-C---A--------.......AG--------ACT..........-----C-----------AA--------------G--------T--------GA-----------G--G-G---------- 1635
H1K.CM.96.MP535             .....................................C---T---C--G------...-A-C-A---------ACAGAG...CCC----------------------AA-------------------A-T------------AG--------A--GT---GA--------- 1517
H101_AE.TH.90.CM240         ........................GGGGGAAGAGAT---CT---TTA-C----------...--A--A---------ACATCC...TCC-------GT--------AA-------------------------T------G----A-----A--A--G--G--A--------- 1895
H102_AG.NG.x.IBNG           ...................................-C-----T---CAC--------...A----GG-------GGCTATA...TCC-----------------AA-------------------T-A--T---G---GA----A---A--G----TA-----C--- 1846
H103_AB.RU.97.KAL153_2      ...................................---C---T-C-T---A------...A-A--AG----G---ACAGCA...TCC-----C--T--------AA-------------G-----------T------G---GA----A--A--G-----A--------- 1547
H104_cpx.CY.94.CY032        .....................................CT---T---T----------...A----GG---------ACT.....A-------------------AA--------------G-----------T------------AC--------A--GA---G---G-----T 1687
H1N.CM.95.YBF30             ......................GAAGAGCACACAGGGG-AGGAGATG---G--A-CC-GGA-AG--C---G--CTCTCTGTA...CCCA----------------------A-------------------G--A------G-----A------AAAG--G---GA--------T 1915
H1O.BE.87.ANT70             ..................GATGGAGGAGGAAGTGAAGGGA-AGGAGA-TC--G-AC-GAAAGG-GGCCCG--C--GCT......C-----G--TG-C-------AA--C---------G-CA----AA-A--T---GC--GAG-T-----C--C---TGT-----T-T-G 2376
H1O.CM.91.MVP5180           ..................AATGGAGGAGGCAGTGAAGGAA-AAGAGA-TC--AGTC-GAA-GG-GATC-GG-A--GCT.........--C--A--TG-C-------AA--C---------G-CA----AA-A-------GC-----G-T-----T--T---TGT--G---T-- 2351
H1P.FR.06.RBF168            ..................GGAGGAGGACGAGATGACCCGGGA-AAG-----G-G----AAA-GA-GAC--G------..........A--C---------C-----------A---------------G-CA----AG-AA-AC--G---GAG---------TA-TTGT--G---T-- 1782
CPZ.CM.01.SIVcpzCAM13       ..................................GGAGAAGA-CC-AG----ACA-GAAGAG-GG--CAGACTATA...CCC----C----------------AA-----------------G----------A--C--G---GA----AA-------T--TGT-----C--G 1901
CPZ.CM.05.SIVcpzEK505       ..................GGAGGAGAAGACCACACAGGGGA-AGAGAG-G-AG-AA-GGA-AA--C---G-GCTCTCTGTA...CCCA--C-------------AA-----------------G---------A--C-T---G----A-----A--AG--A-----A--------T 1893
CPZ.CM.05.SIVcpzLB7         ...........................AACCACCCAG-AA-AGGAGG-C-G-G-A--GGACAA--C---G--CA-GCTGTA...CCC---C------------AA--T-------------------------A-T-----------AG-C-----CT-CT-G--A-G---A--G 1876
CPZ.CM.05.SIVcpzMB66        ...........................GGAACCTGCCCAGAAGGAGG-AGT-A--AGCGA-GA--C---G--CA-GCTGTA...TCC--AGC-----------AA--T-----------------------G-T-------G---GA----A---A--GT----A-----A--G 1867
CPZ.CM.05.SIVcpzMT145       .................................GGAGGAGAAGAAAGTC---G----GGT-AG-GAG--G-GCA-GCCGCC....-C-----------------AA-------------------------------A--C--G---A--T-AA--A----T----TT-----CT-G 1863
CPZ.CM.-.SIVcpzMB897        ..................TCAGAAAATAGATTCTCTG-AA-AGRAGGTC-G-A-C--GGAR-A--C---G--CAGACTGTA...CCC------GGA--------AA--T---------------------G-T---G--GC--GA-----A--AGTTA----A--G-TA--- 1887
CPZ.CM.-.SIVcpzDP943        ..................GGAAACAACAGTGAACAA-CAGGGAGAAA----AG--ATGGA-GT-GAGA-AG--CCCCTGTA...CCC--AA-----C--------AA--------------G-----------C--GG-G----A----A-------G--G-TA---------G 1907
CPZ.CM.98.CAM3              ..................................AAAGAGA---G--G----GAG-GA--AG--A-----ACCCCTGTATCC---G-----C----------AA--------------G-----------C--GG-G----A----AA-------G----TA---------G 1732
CPZ.CM.98.CAM5              ..........................................-G--A-AAA-AA-GAG--G-----ACCCCTGTATCC--AA-----C----------AA--------------G-----------C-TGG-G----AG---AA-----GT---TA---------T 2023
CPZ.GA.88.GAB1              ..................................GAAGAGC---G--A-...GAAAGA-GG---G-GCAGTCT......A-----ACC---C----------AA--------------G-----------A--C--G----AG---AA---------TGT------T-G 2382
CPZ.GA.88.GAB2              ......................GACAGG--AGAGT-AGA--GG-G----GAGTA-A-C---A-----GAG......CCC--TC-ATC--------------AA--------------G-----------C--C--G---GAG---CA--A-------TGT------C--- 1714
CPZ.US.85.CPZUS             .........................................GA-A---G--A-AC-GGGAAA-GA-A-GG-...GCC......C-T--AGC-------------AA--------------G-----------C---G-------AG---CA-----GA--GTA---------T 2381
CPZ.CD.90.ANT               ..................................CAAGAGGACTC--A---GTCTCAA-GGG--GG----ACT......ACC---C-CGTA------G-AA--C------------A----AA-GA-GGA-G-TCTC---CA---A---AA-TGTC-G-------- 1760
CPZ.TZ.00.TAN1              ..................GGCTCAGGAGCAGGAGAC-G-A-AGA-A-CAGT-GT-CCAACTG-CCC-CCACT---GATGACGA--A-AGGCGGGTTC-------A----A----------G-----AA-GA-G---G----AG-CCA---A---G-CTGTC-------T 1965
CPZ.TZ.2001.TAN2            ............................................................................TCTGCCACT...-A-AGACGGGTTC-------AA---A----------G-----AA-GA-G---G-----G-GCA---A---G-CTGTC-G--C--T 1970
CPZ.TZ.2002.TAN3_1          ............................................................................CCTACCACT...-A-AGACGGGTTC-------AA---A----------G-----AA-GA-G---G----AG-GCA---A---G-CTGTC-G-----T 1970
GOR.CM.2007.SIVgor2139_287  ..............................GGAGGAGGGAGAAAAGG---T-AG-A-ACCAGAA-AG--GG-ACT.........T------C------------AA-------------G-CA----AG-AA-AC-TG-G-------------A--T---TGT--G---T-G 1858
GOR.CM.2007.SIVgorCP2135con ..............................GGAGGAGGGAGAGAAAG-A-T-AG-A-ACCAGA--AG--GG-ACT.........T------C------------AA-------------G-CA----AG-AA-AC-TG-G-------------A--T---TGT--G---T-G 1857
GOR.CM.2004.SIVgorCP684con  ...................GCAACCAACAGCCCCACC--TGGAGGAGG-AG--A----AAC-AG--A-CCAGGAA-GGAGGA---T------C------------AA-------------G-CA----AG-AA-AC-TG-G-G-----------A--T---TGT--G---T-G 1777
H2A.DE.x.BEN                GAGAGACCATACAAAGAGGTGACGGAGGACTTGCTGCACCTCGA............-C-GAGAGA--C-CCTC-CAGAGAGGAGACAG-GGAC--GCTGCA-----AT--T---------A--A----AG-A------GCGT-C--C-A--AT--G-CGGTA----TCT-- 2706
H2A.GW.x.ALI                GAGAGGCCATACAAAGAGGTGACGGAGGACTTACTGCACCTCGA............-C-GGGAGA--C-CCAC-CA-AGAGGTGACAG-GGAC--GCTGCA-----AT--T---------A--A----AG-A------GCCT-C--T-A---C--G-C-GT-----T-T-- 2699
H2A.SN.x.ST                 GAGAGGCCATACAAAGAAGTGACAGAGGACTTCCTGCAGCTCGA............-A-ACAAGA--C-CCATGCAGAGAGACGACAG-GGAC--GCTGCA-----AT--T---------A--A----AG-A------GC-C-TG-T-A---C--G-C-GTA----T-T-G 2151
H2B.CI.x.EHO                GAGATGCTAAAGAACTACATGCAACTAGGGAAGAAGCAG-AGGA............-A-CAGAGA--G-CC-T-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T---------AG-A----AG-A----A-GC--CT--T-A---T---TC-GTA----TAT-- 2677
H2B.GH.86.D205              GAGATGCTGAAAAGTTACATGCAGATGGGGAGACAACAG-GAGA............-AGCCGAGA--G-CC-T-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T---------AG-A----AG-A----A-GC-TGT--C-A---T--GTC-GTA----TAT-- 2682
H2G.CI.92.ABT96             GACCTACTGAAGAATTACATGCARCTGGGGAGGAAGCA--AGGA............-C-GAGAAAC-A-CC-T-CA-GGAGGTGACAG-RGAH--GCTGCA------C--T---------AG------AG-A-----TGC-T-C----AAA-T--G-C-GTA----TA--- 2033
H2U.FR.96.12034             GATCTGCTAAAGAGCTACATGCAACAGGGCAAGAAACA--AGGA............-A-CAGAGA--G-CCAT-CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T---------A-------AA-A-----TGC-T-C----A---------C-GTAA---TA--- 2188
SMM.SL.92.SL92B             ATAGTCAA...GGAGTATCTAGAGAAAGCACAGAGGGA--AGA-............AAGGAG-AGC-GGCCTT-CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T------C--AG-A----AG-A----G-GC-CTA----A---T--G-CCCTA--G-T---- 2054
SMM.US.x.H9                 GATCTACTGAAGAATTACATGAAGGTGGGCAGAAGGCAG-GGGA............-A-CAGAGA--G-CCTT-CA-GGAGGT-AC-G-GGA---GCTGCA-----AT--T---------AG-A----AG-A-----TGCCT-C--T-AA-AA--G-CC-TA----TAT-- 2096
SMM.US.x.PGM53              GATCTACTGAGGAATTACATGAAGATGGGCAGAAAGCAG-GGGA............-A-CAGAGA--G-CCTT-CA-GGAGGTGAC-G-GGA---GCTGCA-----AT--T---------AGGA----AG-A---G-TGCCT-C--T-AA-AA--G-CCGTA----TAT-- 2540
SMM.US.x.SIVsmH635F_L3      GATCTACTGAAGAATTACATGAAGATGGGCAGAAAGCAG-GGGA............-A-CAGAGA--G-CCTT-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T---------AG-A----AG-A-----TGCCT-C--T-AA-AA--G-CCGTA----TAT-- 2626

                                                                                      Gag end
                                                          Pol Protease start     Gag p6 end
MAC.US.x.239                GATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG-GAGAAAAG----GAG-AAGCAGAGA--AGCCTT-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T---------AGGA----AG-A-----TGCTC-T--T-AA--A--G-CTGTA----TAT-- 2623
Pol                         C  G  S  A  K  E  L  H  A  V  G  Q  A  A  E  R  K  A  E  R  K  Q  R  E  A  L  Q  G  G  D  R  G  F  A  A  P  Q  F  S  L  W  R  R  P  V  V  T  A  H  I  E  G  Q  P  V  E  V
Gag                          D  L  L  K  N  Y  M  Q  L  G  K  Q  Q  R  E  K  Q  R  E  S  R  E  K  P  Y  K  E  V  T  E  D  L  L  H  L  N  S  L  F  G  G  D  Q  *
MAC.US.x.251_1A11           GATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG-GAGAAAAA----GAG-AAGCAGAGA--AGCCTT-CA-GGAGATGACAG-GGA---GCTGCA-----AT--T---------AGGA----AG-A-----TGCTC-T--T-AA--A--G-CTGTA----TAT-- 2622
MAC.US.x.251_BK28           GATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG-GAGA............AAGCAGAGA--AGCCTT-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T---------AGGA----AG-A-----TGCTC-T--T-AA--A--G-CTGTA----TAT-- 2599
MAC.US.x.EMBL_3             GATCTGCTAAAGAACTACATGCAGTTGGGCAAGCAGCAG-GAGA............AAGCAGAGA--AGCCTT-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T---------AGGA----AG-A-----TGCTC-T--T-AA--A--G-CTGTA----TAT-- 2097
STM.US.x.STM                GATCTGCTGAGAAGTTACATGCAGCTGGGCAAGAAGCAG-GAGA............-AGCAG-AA--C-CC-T-CA-GGAGGTGACAG-GGA---GCTGCA-----AT--T---------AG-A----AG-A------GCCC-T--T-A---T--G-CTGTA----TA--- 2270
MNE.US.x.MNE027             GATCTGCTAAAAAACTACATGCAGTTGGGCAAACAGCAG-GAGA............AAGCAAAAG--AGCCTT-CA-GGAGGTGGCAG-GGA---GCTGCA-----AT--T---------AG-A----AG-A-----TGCTC-T--T-A---A--G-CTGCA----TA--- 2091
DEB.CM.99.CM40              ..............TACCAGGGTACCAGTACAGGGGCAGGAGAA-AAC-CCC--AG---GGAAA-A-CGAGG-GA-GGAAGTA...CCCGGA-----TGGGG--G-AA--T---------AG-A----AA-GAC----G--C-C----AA--AAG-AA-GTCC-G-TG--- 2130
DEB.CM.99.CM5               ..............AACCAGGGTACCTGTGCAAGGGGA--GGAATG---CCC--G----CACAA----CCAGG-GA-GGAAGTA...CCCAGA-----TGGGG--G-AA--T---------AG-A----AA-GAC-G--G--T-TG-G-AA--AA--AGGGTCC-G-TG--- 2124
DEN.CD.x.CD1                GGAAACAGACAATCTTGTCAGGAAGACAAGAGGCAGAAC-AG-AGT-.........CC-TCA-GA-GAC--GG--ACAACGAGGGA-AG--T--ATC----GT-G-A---T---------AG------AA-A-GGCAG--CTG-G-TCA---A---AGGGTA--G-TC--T 2193
LST.CD.88.447               AGAGCCTTAGAAACATACAGGAACCTCGGACAACAGCTG-GGAG..............GCAA-A-CAGAT-CCAC-GAAATGTG--G-CGAGCC-TGCCTGAGTTTC-TCT-----CCAG------AA-A-AA-AC--TCT-G-CAAT-----G-C-GT-TCT-----T 1577
LST.CD.88.485               AGAGCCTTAGAAACATACAGGAACCTCGGACAACAACTG-GGAA..............GCAA-A-CAG-TTCCAC-GAAATGTG--G-CGAGCC-TGCCTGAGTTTC-TCT-----CCAG------AA-A-AA-AC--TCT-G-CAATA----G-C-GT-TCT-----T 1577
LST.CD.88.524               AGGGCAATAGAGACATACAGGAACCTAGGACAACAACTG-GGAA..............GCAA-AACA--TTCCAC-GAAGTGTG--G-CGAGCCTTGCCTGAATTTC-TCT-----CCAG------AA-A-AG-A-G--CT-G-CAAT-------GA-GTAACT--C--- 1574
LST.KE.x.lho7               GAGAGAGCACTGGAGACTTACAGGACTTTAGGACAACAGCTGAAAAG.........-C-GCA--A-CA--T-CCAC-GAAATGTG--G-CGAGCCCTGCCTGAATTCACTCT-...CCAG------AA-A-AA-GT---TT-G-CAAT-------C-GT-TC---A--G 2661
SUN.GA.98.L14               GAGAGAGCTCTGGAGACCTACAGACTGTTAGGACAGGGCCTCAG............A-CACA--A--AGAGG-----AAGGGGGGAA-G--AGGAGC-A-G-T-G-ATATG------CCAG-A-----T-A-AG-AGG-CTTC--TAAT-------CTGT-ACT---T-- 2666
MND-1.GA.x.MNDGB1           CAGAAAACTCTGAGCACTTATCAGAAATTAGGGAGAGGGCT...............CAGGCA-AA---GA-GG-----GAAGAG-GA-G-GGA---TCA------G--TA-T------CAAG-A----AA-A-AAGAGG-CTCAG-G-AT--TGTCACT-TAAG-------- 1996
OLC.CI.97.97CI12            ..............ACAGATTTTGGAGAAGGGAACAAG..................---AAA-AGAGAGATG-----GGAGAAGAAACC--T--ATC-C-----AGCA--T---------AG-A----AA-GA-AGAGC-GT-T----AA--A--G-A-GTCACT--A--G 2186
WRC.CI.97.97CI14            AAGGGAGCACAATTGAGGGAACAACAGGAACACAAGAA..................-ACAGA-GGACAGA-G-G---AGTGGGGG--CCGGAA-GGC-C-------AA--T-----C-AACCA-----AA-GA-ACA-TGC-TA----AA--A--TAGTGTCTC---AT-- 2618
WRC.CI.98.98CI04            AAAGGAGCTCAATTGAAGGAGCAGGCACAGGCAAA.....................AA-GGAAGA--AGA-G-GA--AACGGGG---CCAGAG-GGC----------AA--T-----C-AA-CA-----AA-GA-ACA-TGT-TA----AA--AA-TAGTGTCTC---CT-G 2630
WRC.GM.05.Pbt_05GM_X02      GAGCAATATCTTCTGAGAGGAGCTCAGATGAAGAGGGA..................-A-AGA-AAAGA-ATGC--A-GGAGGGGAAAGTAGAA-ATC-C-------AA--T-----C-AACCA-----AA-GA-ACA-TGT-TC--T--TC-AG-GAGTGTCTCT--GT-- 1547
MND-2.CM.98.CM16            GATTACATGAAGAAAGGGCAACAGCAGAAAGCGGAAAGC-A...............AC-GGAAAA--A---G-G--GTCCATACGA-GCAG---AC-AC-------T--T---------A-CA----AACTACAGT-G--G------AA-----AA-GT---G--CT-- 2322
MND-2.GA.x.M14              GACTATATGAGAAAGGGGCAACAACAGAGGGCAGCAAAGGAGA-............CA-GCA-GA--AG-------GTCCCTACGA-GCAG---AC-AC-------T--T---------A-CA----AACTACAGT-G--G-A----A---A---AAGGTA-----C--- 2249
MND-2.x.x.5440              GACTATATGAAGAAAGGGCAACAACAGAGGGCAGCAGC-GGAG-............CC-GAAAGA-GA-A-G----GTCCCTACGA-GCAG---AC-AC-------T--T---------A-CA----AACTACAGT-G--G-A----A-------AAGGTA--G--A--T 1882
DRL.x.x.FAO                 CTCCAAAAGTACATGGAGCAGGGAGCTCAGCAGAGAG-CAA-AGAG.........AC-ACAA-A-CAGA-G-G--GCCCCTACGA-G-GG---AC-AC-------T--T-----C---TCA-----AACTACA-T-G--G--G---AA------AAGGTA--G---T-- 1888
MON.CM.99.L1                ..AGGGACTGCGGAGCCAGTCAATCTCAACCAAGAGATTGG-T----.........-A-GAC-GG--CCCCA-----GACGAGG-G-G---T--ATC-C---T-AGCA--T---------GG-A----AA-AAGAT-TG-GCT----------A---AAGGT-TCCATG--- 2012
MON.NG.x.NG1                AACTTCTCACCTCCAGTGGGACCACCAGAGACCCC.....................A-CGCC-A-AGCCCCG-----GACGAGGAA-G---T--ATC-C---T-AGCA--C---------GG------AG--AGA---G--G-A--C-CT-----GTCCGT-TCCATGT--  608
GSN.CM.99.CN166             ......................GCTCAATCTCTCGAC-C-AGTAGGGG-ACCC---AGCAA-GGACAGA-GG----GAGCAGGAACAC--TC-ATC-C-----GTCA--T---------GG-A----AA-GAAGGA-G-CTCC---------A---G-GGT-TCCATGT-- 2025
GSN.CM.99.CN71              ......................GCTCAACCTGTCTCCGC-GGTGGGGG--CCC---A-CAA-GA-CAGA-GGCA---AAAGAGGAACAC--T--ATC-C-----GTCA--T---------GG-A----AA-GAGAGA-G-CTCC--G------A---G-GGT-TCCATGT-- 2025
TAL.CM.00.266               TCATGGCAGGAGGGGACTACCACGTGGACCCCACCAGCC-AGAA..............GCCC-AAGA-ACA-CCACAACAAGGGA-AG--TC-ATC-C--TT-GGCCAGT---------AG------AA-A----A-G-GCT-----AA--A---GAGGT----G-TA--C 2284
TAL.CM.01.8023              TCATGGCAGGAGGGGACTACCACATGGACCCCTCCGACC-AGGA............T-CCGCAG-ACCCA-GC-AC-GACGAGGGA-AG--T--ATC-C--GT-AGCCAGT---------AG------AA-A--A-AGG-GCTC-----AA---AAGGGAGGT---G-TA--C 1782
MUS-1.CM.01.1085            ........................................................--GTCCCGG-GAG-GG--AC-GGACAGGAA-AG--TC-ATC--------GTCA--C-------CAGG-------AG--AGAC--G--G-------AA------GG-GGT-TC-ATG--G 2005
MUS-1.CM.01.CM1239          ........................................................--GTCCCAA-GAT-GAGC-AGCGACAGGAA-AG--T--ATC--------GTCA--T-------CAGG-------A-TAGAGA-G---T-----AA--A---G-GGT-C-CATG--G 1995
MUS-2.CM.01.CM1246          ........................................................A--GCCA--C-AG----G-A-AGACAG-AA-AC--TC-ATC--------GTCA--T---------GG-------AA--AAA---G--G-C--TAAT-------G-C-GT-TCCATG--C 2097
MUS-2.CM.01.CM2500          ........................................................CC-AACC--CCA--C-CCACCGGACAGGAA-A---T--ATC--------GTCA--C---------GG--------A--AGA---G---T-G-CCAT--T--G-CGGT-TCCATG--C 2062
SYK.KE.x.KE51               TCCTCTCCAAAGACAGTCAAGGACCTCTTGGTCCCTGGG-AA--TT-.........A--GAA-AAAGA---G-T-A-GGAGAAGGGACC--T--ATC-C-----TCAA--T---------AG------AA-GA-GGA------TG---AA---ACCAAGGTAAC-ATGT-G 2155
SYK.KE.x.SYK173             AACTCTCCCTCGAAGAAACCCCCCACCAACAGGGAAGTGCTCT-G--.........-A-GGA-AGC-GT-GG-----GGAGACCAA-AG--TC-ATC-C--T------C--T---------AG-A----AA-GA-AGA-G--G-TG---AA--TG-C--GGT-C--ATGT-- 2497
RCM.GA.x.GAB1               TATCTGAGGAGGGCAGGGGAACAAAAGAGACAACAGAGGCAGGA............A--GAGCAA--AGAGAG---GAGCATAT-A-G-AG-C-----C-------AT--G---------A-G-----AACTACA-T-GCTC-A----AA-------GAA-GT----G-TC--- 1800
RCM.NG.x.NG411              TACCTAGCATTGGGGAGACAGCACAAACAGGAGCAGAGG-AAGA............-A-CAA-GA--A--TGGG-AGAGCCTACGA-G--G-G--GT-C-------AT--T---------A-G-----TG-AACATGGG--G-A-----A-------AAGGT----G--CT-- 1815
GRV.ET.x.GRI_677            CTCCAGCAGTATGCAGAGAAGGGACAGCGCCTGAGAGAGGAGAG............A--ACA-A-A-GGA-AC--A-GGAGAA-GAAGTGGAGGATGT----T-G--C--C---------AGGA----AA-GAAACG-G-C-TC----AA--AACG-C-GT-C-----CT-G 2476
SAB.SN.x.SAB1C              TATGCTGTTCAGGGGAAACGGCAGAAACAGCAGTGGCA..................AA-CCACT--CC-C-AC--AGCCCCTACGA-G-GG---AC-GC-------------T---------AG-A----AA-GAAA--TG-GT-C----A---A--GAA-GTCACT------- 2719
TAN.UG.x.TAN1               GGAGCATACGACCCAGCCAGGAAACTGCTGGAGCAGTATG-CAAAAAGGG......A--TCAATT--G-A-GC--A-GGAGAAGGAACTGGAGGACTA----T-G--T--C---------AGGA----AA-GAGT-AG--CT-C---CA---C--G-C-GTAC---T---G 2511
VER.DE.x.AGM3               CTGCAGCAATATGCAGACAAGGGGAAGCAGTTGAGGGA-CAAAGGAA.........AA-ACCA--AGC--TG--TCCCGATTGGACAG-GGGA-ATT---T...G-AC--C---------AG-A----AA-AAAA---G-TT-C----AA---GTC-CC-TCAG---AT-- 2034
VER.KE.x.9063               CTGCAGCAGTATGCAGAGAAGGGAAAACAGTTGAGGGA-CAAGGGAA.........-AGAACA--ACC-ACA--TCCAGACTGGGC-G-AGGA-ATT---T...G-AC--C---------AG-A----AA-AAAA---G-C-TT--T-AA--AGTC-CTGTAA---TC--G 2539
VER.KE.x.AGM155             CTGCAGCAGTATGCAGAGAAAGGGAAACAAATGAGAAATCAGAA-AG.........AA-CCCC--AGCGA----TCCAGATTGGAACG-GGGA-ATT---T...G-AC--C---------AG-A----AA-AAGG--CTGT-TA-----A---AACTGCCGTTA-G--AT-- 2531
VER.KE.x.TYO1_patent        CTGCAGCAATATGCAGAGAAAGGGAAACAACTGAGGGAGCAAAAGAG.........-A-TCCA---GC-ATG--TCCGGATTGGACCG-GGGA-ATT---T...G-AC--C---------AG-A----AA-AAAG---G-GT-T----AA---GTC-CC-TTA-G--A--G 2536
COL.CM.x.CGU1               ................................GAAGATG-GCAAAGA.........CCTGCCA--C-GGAGGG--A-GGGGGAGAG...--TC-ATC-C---T-G-A---T-----C--AG------AA-A--T-T-G--G-AG---AA--A---AAGTGCC----G--G 2090
```


PLV Complete Genomes

```
H1B.FR.83.HXB2          TTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTG...........CCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATAAAGCTATA.......... 2468
Pol                      L  D  T  G  A  D  D  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  K  A  I  .  .  .  .
H1A1.UG.85.U455         -------------------------C--------C--A-A----...........------A--------------A---------------------------------------------------------T-------T-----A-A--GA-----.......... 1914
H1B.US.90.WEAU160       ---------------------------------C----A----...........-----G-A---------------------------------------------------G---C-------------------------------------.......... 2467
H1C.ET.86.ETH2220       -----C------------------------------A-A----...........------A--------------------A----------------T---------------A--A---------T-----A-A--G------.......... 1860
H1D.CD.84.84ZR085       ---------------------------------C--A-A----...........------AG-----------------------------------------------C---------------------------G------.......... 1990
H1F1.BE.93.VI850        ---------------------------------C--A-A----...........------A-------------------------------------------A---------A-C-------------T--------C--G------.......... 1805
H1G.SE.93.SE6165        -----C--------------C---------------A-A---A...........------A-------------------------------------------------A-------AG---CT--------TA-----A-A--G------.......... 1865
H1H.CF.90.056           ---------------------------------G--A-A----...........--G---A----------------------------------------------------G--AG--GC---------------A-A--G------.......... 1815
H1J.SE.93.SE7887        ---------------------------------C--AGAC---...........---C--A---------------------------------------------------A-CG--G---CG-------TGAG---A-A--G------.......... 1782
H1K.CM.96.MP535         ------------------------------------A-A----...........------A-------------------------------------------------A-------AG----T--------T---------A--G------.......... 1664
H101_AE.TH.90.CM240     ---------------------------------T--A-A----...........------A-------------------------------------G----AG--A--------------T----------------A-A--G------.......... 2042
H102_AG.NG.x.IBNG       -----C---------------------------C--A-A---A...........------A-------------------------------------------------A-------A-----T--------T-----A-A--G--C---.......... 1993
H103_AB.RU.97.KAL153_2  ---------------------------------C--A-A----...........------A------------------G------------------G----------------------------T--------T-----A-A--G------.......... 1694
H104_cpx.CY.94.CY032    ------------------------------------A-A----...........------A-----G---------------------C-----------------------A-------------CT--------T-----A-A--G--C---.......... 1834
H1N.CM.95.YBF30         -------------T-----------A------GC-ACAA--A...........GAG---A-----------------------A-----------A------------G-----A------A-T---ACAG----C--ACAG---AGA-----AG-T.......... 2062
H1O.BE.87.ANT70         C-G----------------C------C--A-CA-C--ACAA---...........GA-----A-------------------------T--A-----------A-------A-G-A------A-TG-GACAG-------AGAA---AGGG-G-TACAG.......... 2523
H1O.CM.91.MVP5180       C-G--------G-----------------A-TA-C--ACAA--A...........GA-----------C---------------------T--A-----C-----A-------A-G-----A-CA-TG-GACAG-----G-ACAA---A-GG---TACAG.......... 2498
H1P.FR.06.RBF168        -------------T-----C------G----TA--T-ACC---A...........GA-----------R--------------A-----A--G-------A-------A-G-A-T---AG-AG-TAAGG-------AGAA---AGAG---T-TAC.......... 1929
CPZ.CM.01.SIVcpzCAM13   -----C-----G--T------------A--C----C--TCAA--A...........-AG----TT------------------------------C-----T-----G-A---A-------A-TG--TCA-----G--AGAG--GA-A-----TGCA-.......... 2048
CPZ.CM.05.SIVcpzEK505   C------------T-------------A----------ACAG---...........GA-----A-------------------------------G--------------A-------A-TG-GA-A-----G--ACAA--GA-A-----AG-T.......... 2040
CPZ.CM.05.SIVcpzLB7     -----------G------------GC--------TT-AGAA--A...........GGG----G-----------G---------A--------------------------AG--A------A-T----CA------AG-A---A-A--G--AC--.......... 2023
CPZ.CM.05.SIVcpzMB66    -------------------C-----GC------------AGAA---...........GG---T-------------------------------------C--A-------------A-------A-TG-TACA-----C--A-------A-A-G---AG--.......... 2014
CPZ.CM.05.SIVcpzMT145   C----C-----G--T-----C------A-T-----CT-A-A---...........-------A-----GT-------------------------------T-------A---A-T---AAGAG--AA----------AGAA---A-A--GATA-CC.......... 2010
CPZ.CM.-.SIVcpzMB897    -----C-----G----------------C----G------AGAA--A...........GAG----A-----GG-----------------------------A-----------------A----------C----------T------A-A-G---A---.......... 2034
CPZ.CM.-.SIVcpzDP943    C----C-------T-----------GA----TA-TG-ACAA--A...........A----------------------------------------------------------A---A-------A-C---GCA-------TGAA---CAGA-----A-C-.......... 2054
CPZ.CM.98.CAM3          C------------T-------------A----TA-TG-ACAAC-A...........A------A-----G-----------------------------------T-A---A------CA-C---GCA-------TGAG---AGA---A-A-C-.......... 1879
CPZ.CM.98.CAM5          C-G--------G--T-----C------A----TA-TG-ACAA---...........A------AG----G------G----------------------T------T-A---A-------A-C---G-A--------AGAG---CAGA---A-A-C-.......... 2170
CPZ.GA.88.GAB1          C------------T-------------A----GAG---ACAA--A...........-A----CTT----------------------A--------------------C-A---A-T----A-TG---A------G--AGAA---GAGA-----TAG--.......... 2529
CPZ.GA.88.GAB2          C----------G--T--------------------TA-T--ACAAC-A...........GA---G-A-----G-------T-A-----A--C---------A-----T-------A---A-------A-TG--GAA--------AGAG---A-A-G--TAGA-.......... 1861
CPZ.US.85.CPZUS         -----C-------T-------------C----CA-C--ACAAA-T...........GA---G-C-----G-----------------G-----------------------A------A---A--------CG-CAAT--------AGAG---AGA-----ACAG.......... 2528
CPZ.CD.90.ANT           -------------T-----C------G----G-G---TCA----...........-A-----AC-----G-------CA--T--T--------G--------TTCC---CA--------A-CA-AG---C----C----TG-A-ACAGG-C--TAC--.......... 1907
CPZ.TZ.00.TAN1          --------T----------C-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---TC-G-.......... 2112
CPZ.TZ.2001.TAN2        --------T--G-------C-GT--G--TTGTA-C--T-AG--A...........AAG---CAG----C----------CA---------------A-----A---G-ATC---T--TG-A--CT--A-C--C-CAG--C----TG-CAATA-AG---T--G-.......... 2117
CPZ.TZ.2002.TAN3_1      --------T----------C-GT--G--TTGTA-C--C-AA---...........AAG---CAG----C----------CA---------------A-----A---G-ACC---T--TG-A---T-CA-T--T-CAG--C----TG-CAATA-AG---T--G-.......... 2117
GOR.CM.2007.SIVgor2139_287 -----C-----G--T-----------G-----C-CT-ATC---A...........GA---T--G----------G---------A--C--A-----A-----T-------A-G-A-----AG-TG--GAA---------AGAACATAGG----TAG--.......... 2005
GOR.CM.2007.SIVgorCP2135con -----C-----G--T-----------G-----C-CT-ATC---A...........GA---T--G----------G---------A--C--A-----A--C--T-------A-G-A-----AG-TG--GAA---------AGAACATAGA----TAG--.......... 2004
GOR.CM.2004.SIVgorCP684con -----C-----G--T-----------G-----C-CT-ACC---A...........GA---T--G----------G---------A--C--A-----A-----T-------AGG-A-----AG-TG--GAA---------AGAACATAGA----TAG--.......... 1924
H2A.DE.x.BEN            C----C-----G--T-----CT--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-G--TA-G-.......... 2853
H2A.GW.x.ALI            C----C-----G--T--C--CT--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-G--TA-A-.......... 2846
H2A.SN.x.ST             C----C-----G--T--C--CT--A--G---C--GCG-AGAG--A...........GGGA-C-AT-AT-GT-----G--AG-----------A--G--A--C--A--TACC-A-G-A---A-AA-TG--GAA---AG-G-A-TAAATA-A-G--TA-G-.......... 2298
H2B.CI.x.EHO            C----C--------T-----CT--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-G--TA-G-.......... 2824
H2B.GH.86.D205          C----C-------TT--C--CT--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-G--TA-GG.......... 2829
H2G.CI.92.ABT96         -----C-----G--T-----CT-TA--G---C--GG--AGAA--A...........GG-A-T-AT-AT-GT---------AG-----A-----A-----A-----T--TACC-A-G-A--CA-AA-TG--GAA--TA--G-GGTA---A-A-G-ATACAG.......... 2180
H2U.FR.96.12034         C-GA----C--G--T------T--A--G---C--GG--T-AA--A...........GA-CC-TAT-AT-CT-------CTG-------A---A-A---G--C--T--TAC--A-A-A-T-C-CA-TG--AA---A--G-A-TG--CA-A-C-AT--G-.......... 2335
SMM.SL.92.SL92B         C-G--C-----G--T-----C--CT-CA--G---C--GG--AGAG--A...........GGTCC-G-G-AC-C-------G---TG-------G--------A--C--A--TAC--A-G-A----AGAACAG---AAA-------G-T-TA----A-GGT-ATA-G-.......... 2201
SMM.US.x.H9             -----------G--T-------T--A-TG---C--GG--AGAA---...........GGTCC--AT-AT-CC--T------AG----A-----A-----C--C--T--CACC-A-G-A---A-AG-TG--AAA---A--G-A-TA---CA-AGT-AT--AG.......... 2243
SMM.US.x.PGM53          -----------G--T-------T--A-TG-G-C--GG--AGAAC-A...........GGTCC--AT-AT-CC--T------AG----A-----A-----C--C--T--CACC-A-G-----A-AG-TG--AAA---A--G----TA--CA-GGT-AT--AG.......... 2687
SMM.US.x.SIVsmH635F_L3  -----------G--T-------T--A-TG--AC--GG--AGAA---...........GGTCC--AT-AT-CC--T------AG----A-----A-----C--C--T--TACC-A-G-A---A-AG-TG--GAAA---A--G----TA--CA-GGT-AT--AG.......... 2773

MAC.US.x.239            C-G--------G--T-------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-GGAT--A-.......... 2770
Pol                      L  L  D  T  G  A  D  D  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  T  K  E  Y  K  N  V  E  I  E  V  L  G  K  R  I  .  .  .  .
MAC.US.x.251_1A11       C-G--------G--T-------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-G-AT--G-.......... 2769
MAC.US.x.251_BK28       --G--------G--T-------T-TA-TG--AC--G---AGAG--A...........GGTCC-CAT-AT-CC---------AG----A-----A------------T--TACT-A-G-A--CA-AA-TG--AAA-------G-T-TA---CA-A-GGAT--A-.......... 2746
MAC.US.x.EMBL_3         --G--------G--T-------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-GGAT--A-.......... 2244
STM.US.x.STM            --G--C-----G--T--C--CT--A--G-G-CG-GG--AGAGC-A...........GGGCT-CA--AC-CC-----GG-AG----T-----A-----A-----A--TACT-AGG-A-T-A-GA-TG--AA------GG-A-TA---CA-G---ATA-A-.......... 2417
MNE.US.x.MNE027         --G--------G--T-------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-GGAT--A-.......... 2238
DEB.CM.99.CM40          A----------G-----------CA-TA-TTC-----AAGA-A-A......GATTTAGAG-C-CCT----CC-------CAG--------T-A---------C--A--T--C-A-TGT--CCCAGG----GAA--T-C---GGC-AATA-GGT---AG--.......... 2283
DEB.CM.99.CM5           A----C-----T-------------A-TA-TTC---G-AAGA---A......GATTTAGAT-C-CCT----CC--G----CAG-G-------T-A--G--A---------T----A-TGC--CCCAGG----GAG--TA----GGCA-ACA-GGT----G-T.......... 2277
DEN.CD.x.CD1            A----------G---------C--CA-TA---C----CAAGA-A-A...AATTTGGGAA---AGG------TC-------CAG-------------A-----A-A---T--T--TCAGGCA--CCCAGG----GAA--TA-GC-TATG-ACA-AGT---GG--.......... 2349
LST.CD.88.447           C-T--C--T--G-------------TA-C--TTC-----ATTCAG-T......AGAATAGAG---CCT-AT-C---T-G-G-AG----------A--G--ACAA--T-G------A-G-A--CA-GG-TG-CT-T---------AGCA---A-A-C-A-AG--.......... 1730
LST.CD.88.485           C-T--C--TA-G-------------CA-C--TTC-----ATTCAG-T......AAAATAGAG---CCT-AT-C---T-G-G-AG----------A--G--ACAA--T-G------A-G-A--CA-GG-TG-CT-T---------AGCA---GA-G-C-A-AG--.......... 1730
LST.CD.88.524           --G--C--T------------------A-C--TTCT----ATTCAG-T......AGAATAGA----CCT-AT-CT----G-G-AG-------T--A--G--ACAA--T--------A-G-A-TCAGGG-TG-TT-T---C-----AGCA---GA-G-CCA-AG--.......... 1727
LST.KE.x.lho7           --G--C--T--G--T-----C---A-T--TTC-----ATTCAG-A......AGAATAGA----CCC-AT-CT----G-G-AG------C--A--G--GCAA--T-GG------A-G-A---AGAG-TG-CT-T---------AGCA----A-A-CTA-AG--.......... 2814
SUN.GA.98.L14           --G--C--G-----------C---A--G-GTCT---GATTC-G--......TCCATAGA---TCAC-----G--C-GGG-AG-G--T--T--A--G--ACAA--T-G------A-G-A----CAG-TG-CT-TG--------TA-G-ATA-A-C------C-.......... 2819
MND-1.GA.x.MNDGB1       C-------------T---------CA-C--TA-T----GA-A-A-A......AAATTAAA-----AT---C-G-----------T-----------A--T--AAACT-A-G-----A------------A-TG--TATG--------AA-A--GA-GGG-A-AT-T.......... 2149
OLC.CI.97.97CI12        -----------G---------C------A-TA-G-GG-A-GA-A-C......TACATGGA---T-ATG---TT--T-G---T------------A--G--AGC---T-GG----TG--A---AGATGT-GGA-A-G-ATTT-T...-AGA-AG---ATACA-.......... 2336
WRC.CI.97.97CI14        C--------------------------A----T-T--ATTTA...............AATAAG-AT---C-G--T-G---A--T--A--G--A-----CAC---G--T----A-TTAC--AGA--AG--AAA---AG---A-TA---A-A-C-ATAGA-.......... 2762
WRC.CI.98.98CI04        ---------------------------A----TATCCATTTAAAT...........AATAAG-AT---C----T-G---A--T--A-----A-----CAC---G--T--T-A-TTA---AGGA-A----AA---AG---T-TA---GA-A-C--TAGA-.......... 2777
WRC.GM.05.Pbt_05GM_X02  --------------------C--T---G----T-TTCAATTA...............GA-----AT---C------G----A-----A--T--A-----AAC---GC-------A--TA---A-AG-TG-CAGA--A----T-TA---A-GG---TA-G-.......... 1691
MND-2.CM.98.CM16        C-G--C--------T------------A-CA----TC-AGA---A...........AA---T-AT---------C-G--T--T--A----------A-CA-----T----A-----T-TTCA-TTGTAAAG--AC---AGCA--CA-A-CTA-ACAC.......... 2469
MND-2.GA.x.M14          C-G--C--T------------------A-CA----TG-AGA---A...........A------AT--------CC---T-----A-----------AGCA--T-GG--G-A-------TTCA-TTGTAAA---AC-G-AGCA---A-A-GTA--CAT.......... 2396
MND-2.x.x.5440          C--------T--------C--------A-TA----TT-AGAG--A...........A------AT------G---CG---T-----A-----------AGCA----G--------------TT-A-TTGTAAAG-------AGCA---A-A-TTA--CAT.......... 2029
DRL.x.x.FAO             C----C---------------------A-TC---G-T--GAA--A...........A------A------G---C-G--T-----A-----------AAC---T--T--G-----A----TT-A-TTGTGAA--TA-G--AGCA---GA-A-TTA--CAT.......... 2035
MON.CM.99.L1            --G--------T--C--C--C---A-TA--C----TCA-CACG-A......AATTTAGGG-AG-AT----C---T---G-AG----------A--G--AA-G--AG----T-A---A--CA-AGCA----AAG-GAT-T-TGAA-ACAGAG-G-T-TGG.......... 2165
MON.NG.x.NG1            --G--C--G--R--G--C-----CA-CG-CC----RCAAGACA-A......AATCTAGG--A--AT----C---C------G-------------A----C--A--AC-G--T-AG--A-TCA-AGG-G--ACGG--TCCT--CAG--TA-A-C--TG-G-.......... 761
GSN.CM.99.CN166         -----C-----G--G--C--C--CA-CG-CC----CTCAGCAA-T......GAATTAGATAC-CC-----C----------G-------------A--G-----A--AC--------GG--C--A-A--TG--GAAG--ATCT-TAA--AGA-A-G-ATC-A-.......... 2178
GSN.CM.99.CN71          -----C-----G--G--C-----CA-TG-CC----TCAGCAA-T......GAATTAGACAC-CC-----C--------C-G--------G--A--G-----A--AC------AGG-AC--AGA--TG--GAAG-GATCT-TAA--A-A-A-G-ATC-A-.......... 2178
TAL.CM.00.266           C----C----------------C---A-TA--C---T-GGAGAAA-A......GAATTGGA----GT-CCCC-C-A-----CAG---------T-G-C---A-AC--AC----GCA-AGT---AG-GCC---GAA---C-GTGGGAAAACA-A-G---A-GG.......... 2437
TAL.CM.01.8023          C----------G---------C---A-CA--C-----GG-GACA-A......GAATTAGA----GTGCCCC-C-A-----CAG---------T-G-C---A-AC---C-G---CA-AGC---AC-GGAG--GAA--C---TGGGAACACA-A-G-----GG.......... 1935
MUS-1.CM.01.1085        C-C--C--C-------------C--CA--A-CC-G--TGCCCA-A-A......CACCTAGA----CC------T--C---TCAG-----------A--G--AAAC--ATC---C-A-G-A---AGGGGAG---AGG-TTGTT--AA----A-A-CTAT--GG.......... 2158
MUS-1.CM.01.CM1239      C-C--C-----G---------C--CA--G-CC-G--TTCC-ACA-T......CAGATAGATAA-CC------T--C---T--G-G---------C--G--AAAC--ATC---C---G-A---AGAGG-G---AAG--TCTT--AA---AGA-A-C-AT--AG.......... 2148
MUS-2.CM.01.CM1246      C-G--C--C-----G------C--CA-CA-CC----CTCAGA-A-A......AAAATTAGT--GGA--T--CT--C---G-AG-G--A--TT-A--G--GC-G-----T----A---A---A-GGG-G-GA-AG--A--T--AA--AGA-A-G-ATAG--.......... 2250
MUS-2.CM.01.CM2500      C-G--C-----G---------C--CA-TA-TC-G--GTCTGAAA-A......GAAATAGATAA-GA-----CT------G--G---------T-A--G--A--A--T--T----AG--A---A-AGCAG--AAAG----GT-TAA--ACA-A-G-ATC---.......... 2215
SYK.KE.x.KE51           C-G--C-----G---------C--CA--A--AG----CAAGACA-C......CAGTTA-ATCA-CCT----CC---------AG---------T-A-----AAAC--A-GG---------A----A-C-----CAGAT-TC----AGAAA---CAG-T-GG-G-ATAAAGGAAGT 2319
SYK.KE.x.SYK173         G----------G-----------TA--A-TAG----GAAGA-A-T......CAATTA-ACCA-CC----TC---------AG---A---T-A-----AAAC--T-C---T--G--A---AGAA-T--TAGAT-CACTG-TGT-AA---CCTC--GA-A-AGAAAGCAAGT 2661
RCM.GA.x.GAB1           -----C----------------C------A-T----G---AGAA--A...........GG-AATGAT----CC---------A-----A-----A--G--A-A---T--T-----A---A----A-AA-CTGTGAA--T------AGC----A-A-GGA--CAT.......... 1947
RCM.NG.x.NG411          ------------------G------TA--C--A-TT-AGA---A...........GG--A--AT-----G--------T-----A-----A--G--A-A---A--T--T-A---A-T-A-C--TTGTAAA---ACT--AGCA--GA-G-CTA-ACAT.......... 1962
GRV.ET.x.GRI_677        --------------------C--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-TCATC-CC.......... 2635
SAB.SN.x.SABIC          -----C----------------CT------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-CTTG--A-.......... 2866
TAN.UG.x.TAN1           --------------T---------CA--A--A-G---GAAGA-A-C......CATTTAGA----C-C----CGG--T---G-AG----A--T--A-----GGGA---TC-----A---A-----CA-T--TT-AG-GA-T--AGA----A-AG--A---G-.......... 2664
VER.DE.x.AGM3           --------G--G-----------CA-TA--A-----GCAGA---A......CAATTAT-------C--------------A----------C--------GGGAC----T----A-G-----AG-G-T-GGGAAG--AG-T-GGAA-ACA-A-TTTTG-G-.......... 2187
VER.KE.x.9063           -----------G---------C--CA-CA--A------CAGACA-A......CAACTAA-----CCT----G-------G-A------A-----------AGGAC-T--T----A-G-----AG-A-TG--GAAG--A-----AGAA-ATA-A-TTTTG-G-.......... 2692
VER.KE.x.AGM155         -----------G---------C--TA--A--A-G--T-CAGA---A......CAATTAAGG---TC-----G----------AG----A---------G--AGGGT-A--C---A-G-A-------A-TG--GAAG--C--T-GGAA-ACA-G-T-TTA-G-.......... 2684
VER.KE.x.TYO1_patent    C----C-----G---------C--CA--A-TA------ATGA---A......CAATTAT----TCC------G---C-----T---------C--A-----AGGCC-T--T----A-G-A---A-CG-C-GGGAAG--A-----AGAA-ATA-A-TTTTG-G-.......... 2689
COL.CM.x.CGU1           -------------------------T--TA-----GAAGA-G-T......CAATTAA----CGCT----C-G--GT-CAA--TCA------A--G--AGCA----G-----A--------AGAG-T-A-A-TC--A-G--GG-A---A-AG--TAC-G-.......... 2243
```

```
                                                                                                       Pol Protease end Pol p66 and p51 RT start
H1B.FR.83.HXB2              ....GGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTCCCATTAGCCCTATTGAGACTGTA.........CCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATGGCCAT 2625
Pol                         __.__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__.__.__.__P__V__K__L__K__P__G__M__D__G__P__K__V__K__Q__W__P__
H1A1.UG.85.U455             ....------------G---------------------------------G--A------------------T-------------A-----T--------A-----.........--------------A---A-------------G--------------- 2071
H1B.US.90.WEAU160           ....------------------------------------------------------------------C-------------------------T--------A-----.........------------------------------------------------- 2624
H1C.ET.86.ETH2220           ....---------C-----------------------------C-----CA-----------C----AC----A-----C-----A-----T--C-----A-----.........-----------------------------------C------------- 2017
H1D.CD.84.84ZR085           ....--------------------C--------------------------T------C-------C----------C-----A-----T--------A-----.........---------------------------------------------G- 2147
H1F1.BE.93.VI850            ....-------G-------------G--------------------------A-------------------T--------------AG----T--------A-----.........-------------G-------------------G--------------- 1962
H1G.SE.93.SE6165            ....--G--GA------------------A-------------G------A-----------------------------------A-----T--------A-----.........-------------------------------GG--------------- 2022
H1H.CF.90.056               ....-----------------------------------T----------G--A-A--------A-----------C------------A-----T--------A-----.........--------------------------------G--------------- 1972
H1J.SE.93.SE7887            ....-------G---A---------------------------------G--CA-----------C-------T---------------A-----T--------A-----.........---------------------------------A-------------- 1939
H1K.CM.96.MP535             ....-------------------------------------------C----C-------------------T---------------A-----T--------A-----.........--------------------------------G--------------- 1821
H101_AE.TH.90.CM240         ....-------------------------------------------C-----A-------------------T-------------C--A-----T-------C-----.........-------C-------------------A----G--------G------- 2199
H102_AG.NG.x.IBNG           ....-------------------------------------------C-----A-------------------T---------------A-----T--------A-----G.........--------------------------------G--------------- 2150
H103_AB.RU.97.KAL153_2      ....-----G-------------C--------------------------A-----------C-------T---------------A--A--T--------A-----.........-------C------------------------G--------------- 1851
H104_cpx.CY.94.CY032        ....--C-----G--------------------------------C----CA-----------C-------T---------------A-----T--------A-----.........--------------------------------G--------------- 1991
H1N.CM.95.YBF30             ....--------------------A-------T--T--T--A-----------T--A--C--------C--T---------------A--A--T--------A-----.........--------------A---------------------G--A-----------T- 2219
H1O.BE.87.ANT70             ....--A--------G--G--------T-----T--T--T-------------A-A-----AGGAT-A-----T--AC-----C--C--T--A-----C--A-CCC-A--G.........---------C---A--------------A--------A-----------CC 2680
H1O.CM.91.MVP5180           ....--A--------G--G--------T-----T--T--TC----G-----CA-A-----AGGAT-A--A--T--AC--------C--T--A--T--C--A-CCC-A--G.........--------GC---A--------------A--------A-----------CC 2655
H1P.FR.06.RBF168            ....--A--T--------R--C--A--C--A--A--Y-----A---------A-T-----ATTA-----A--T-------G---C-------T-C--A--A-TA---.........-------------------G--------------G--T--G--A--G-------C- 2086
CPZ.CM.01.SIVcpzCAM13       ....--G-----G--G--G-----A-----------T--------G------A-T--A--A-----A--G-----AC-GGTA--C--T--C--TT-A-----A-----C.........-----C-----G--A-----------C--G----G-A-A--G--------T- 2205
CPZ.CM.05.SIVcpzEK505       ....--C-----C--------G--A-----------T--T--T-------------T-TC----A--A--------T---C-----------A--A--T-----------A--CA--.........-----G-----G--------T-------------G-G---A-----G-----T- 2197
CPZ.CM.05.SIVcpzLB7         ....--------T-----------C-----A--A--T--C--A-----G--CA-T--A-----A-----A--T---C-----------T----------------A-----.........------C-------------------------G--------A-----------C- 2180
CPZ.CM.05.SIVcpzMB66        ....--------T-----------C--------A--T--C--A-----G--A-T---------A-----A---------------C-----C--T-----------A-----.........-------GT-----------------------G----GG--C-----G-------- 2171
CPZ.CM.05.SIVcpzMT145       ....A----T--------------A--C---------T-----A---------G-CC-A--AA-AT-A--G--T---C--------C--A--A--T--A--A--C--A---.........AA---GTTTC----------G-----C--G--C--G--A-----G-----TC 2167
CPZ.CM.-.SIVcpzMB897        ....--------------------G--C--------G--T--T--A--G-------A-C--A-----A-----G--T---C--------------------A---A--A-----.........--------------------------G-----C--T----G----------G-----CC 2191
CPZ.CM.-.SIVcpzDP943        ....--C-----G--G--G-----A-----A--G--T--------C------A-A-----A--------C--T-----------C--T-----------C--T--A--T--C----GA-----G.........-------------------------------T-G---A-----------C- 2211
CPZ.CM.98.CAM3              ....--C--G-----G--G-----A-----A--G--T--------C------A-A-----A--A-----C--T----------------T--A--T--C-----A--A--G.........-------------------------------T-G---A-----------T- 2036
CPZ.CM.98.CAM5              ....--C-----------------A-----A--A--T--------C------A-A--A--A--A--------T---C----------T--A--T--CG-------A--G.........--------------------------------G---A-----------T- 2327
CPZ.GA.88.GAB1              ....--G-----------------A--------A--T--------------A-T-----A--AT-G-----T---C--GTG--C--A-----TT-A-----A-----C.........-----C-----G--A-------------T----G--A--G-------C- 2686
CPZ.GA.88.GAB2              ....A-C-----C----G-----A--T--A--T--T--T--A---------G----A--C--AT-A--A--T---C-----------A--A--T--A---ACTC----G.........--------------------------CA-----A---GG--A--G--------- 2018
CPZ.US.85.CPZUS             ....---T----T-----G-----A-----A--A--T--T-----C--G---A-T--A--A--A-----A--T----------------T--C---T-----A-A-------.........-------G-----A-------------------T-G---A--G--------- 2685
CPZ.CD.90.ANT               ....-CA------C-GT-G-----A-AC--A--A----T--A--T------G-T--ATG--TTT-A--A--T--------------A-----TAAAG---A--A--G.........--C--T-------AGA-------------A---G---A-----------GC 2064
CPZ.TZ.00.TAN1              ....-CC--T--CC----G---GAA--C--CA-T--T-----A--T------A-T--A-AG--AT-A--A--T--C----------T--C---------T--------A--A---GTA---.........AA---C-------AGA-------------G-------A--G--G-----CC 2269
CPZ.TZ.2001.TAN2            ....-CC--T--CC--------GAA--T--CA-T--T-----A--C------A-T--A-AG--AT-A--A--T--C----------T---C---------T-----T--AG----GTA---.........AA--GC-GC---AGA-------------A--------A-----------TC 2274
CPZ.TZ.2002.TAN3_1          ....-CC--T--TC--------GA---T---A-T--T-----A--C------G-T--A-GG--AT-A--A-----------------------------T--AG-A---GTA---.........AA---C-G----AGA-------------A--------A-----------TC 2274
GOR.CM.2007.SIVgor2139_287  ....--G-----TC----T-----A--C-----A--T--T--A----G--A-T--A--AA-A-----A--T--CC-------------------A---T-C--A--TGTA---.........--C---TCT-----------G--------A--C-----C--G--------T- 2162
GOR.CM.2007.SIVgorCP2135con ....--G-----T-----T-----A--C-----A--T--C--A----G--A-T--A--AA-A-----A--T--C---------------------A---T-C--A--TGTA---.........--C---TCT----A-----G--------A--C-----C--G--------T- 2161
GOR.CM.2004.SIVgorCP684con  ....--G-----C-----T-----A--C--G-CA--T--T--A---------G---A--AA-A-----G--T--CC-------------------A-----C--A---GTA---.........--C---TCT-----------G--------A--T--------G--------C- 2081
H2A.DE.x.BEN                ....-CC--CA--A-GAC----GA---C--AA-------CT----C------A-TC----AGCCT-A--CATGT-A--------A--AG--GC-AAG--A---C-AA--.........AA-----C---G--------G-AA-----A---GGC-G-----------CC 3010
H2A.GW.x.ALI                ....-CC--CA--A-GAC---TGAC--C--AA----T--TT----C-----CA-T-----AGCCT-A--CATGT-A-----CC-A--AG--GC-AAG--A---C-AA--.........GAG----G-----------------AA--C--G------T-A-G---------C- 3003
H2A.SN.x.ST                 ....-CC--CA--A-GAC---TGA---C--AA-------TT----C-----CA-TC----AGCCT-A--CATGT-A------C-A--AG-CGC-AAG--A--AC-AA--.........AA-A---TGC-G-----------A------A------C-G-G---------C- 2455
H2B.CI.x.EHO                ....-CA------A-GAC---GGAC--C--AA-A-----TT----C------A-T--A-A-AGCT-A--CATG---C--------C--AG-AGCAAGG--A--AC-A---.........AA---CC-G--------T-A--AA-----G-----A-C-G---------CC 2981
H2B.GH.86.D205              ....-CA--TA--A-GAC----GA---C--AA-A-----TT----C------A-T--A-A-ACCT-G--CATG-----------C--AG-GGCAAAGG-A--AC-A---.........AA---TG-G-----A--T----AA-----G-----GA-C-G---------TC 2986
H2G.CI.92.ABT96             ....-C---T--TA-GAC---GGA---C---A-A--T--CT--------G--CA-T--AGT-A--C-A--CATGT--C-----------A---GCAAAGG-A---C-A---.........AA-----GGC----A--T--C--------G------A-C-GG---------CC 2337
H2U.FR.96.12034             ....-----CA--A-GAC----GAC--C--AA----T--TT----T--G-----T-----AG-AC-G--CATGT--C-----C-----A--AG-AAAAGAA-CACTA--C.........ATT-GGGG-C----A-----C-A------A------T-A-G---G-----T- 2492
SMM.SL.92.SL92B             ....-AA---C--A-GAC----GA---C--GA-A--T--TT----C------A-TC-AG-AGCAT-A--GGT-T---------AC--AG-GGCAAAGG-A--ATA-AC-.........AA-------G--G---GAG------------G----G-T-A-----------C- 2358
SMM.US.x.H9                 ....--A---A-TA-GAC----GA---C--AA-T-----TT----C----------C-A--AGCT--G--CATGT--C-----C-C-----AGCTAAGG-G--RC--A--.........AA-----C-C----A--------AA-----A-------T-A-G---G-----GC 2400
SMM.US.x.PGM53              ....--A---A-TA-GAC----GA---C--AA-T-----TT----C-------------A--AGCT--G--CATGT-C------C-C---G-AGCTAAGG-G---C--A--.........AA-----C-C----A--------AA--C--A-------T-G-G---G-----GC 2844
SMM.US.x.SIVsmH635F_L3      ....--A---A-TA-GACG---GA---C--AA-T--T--TT----C------T--C-A--AGCT--G--CATGT-----------C-----AGCTAAGG-G---C--A--.........AA-----C-C----A-----G-AA--A--A---------G-G---G-----GC 2930
                                                                                                       Pol Protease end Pol p66 and p51 RT start
MAC.US.x.239                ....--G---A-CA-GAC---GGAC--C--GA-T-----TT----T------T--C-A--AGCTC-G--GATGT--C--------------AGCTAAAG-A---C------.........AA---CGCC---------------A------A------T-G--G--G------- 2927
Pol                         __.__K__G__T__I__M__T__G__D__T__P__I__N__I__F__G__R__N__L__L__T__A__L__G__M__S__L__N__F__P__I__A__K__V__E__P__.__.__.__V__K__V__A__L__K__P__G__K__D__G__P__K__L__K__Q__W__
MAC.US.x.251_1A11           ....--G---A-CA-GAC---GGAC--C--GA-T-----TT----T------T-AC-A--AGCTC-G--GATGT--C-----C--------AGCTAAAG-A---C------.........AA---CGCC---------------AA-T---A------T-G--G--G------- 2926
MAC.US.x.251_BK28           ....--G---A-CA-GAC---GGAC--T--GA-T-----TT----T--G---T--C-A--AGCTC-G--GATGT--C-----C--------AGCTAAGG-A---C------.........AA---C-CC---------------A--T---A------T-G--G--G------- 2903
MAC.US.x.EMBL_3             ....A-G---A-CA-GAC---GGAC--T--GA-T-----TT----T--G---T--C-A--AGCTC-G--GATGT--C-----C--------AGCTAAGG-A---C------.........AA---CGCC---------------A--T---A------T-G--G--G------- 2401
STM.US.x.STM                ....--A--TA-TA-GACT---GAC--C--CA-T-----TT----C--G--C--A-----AGCTT-G--GATGT-A-----------AG-AGCTAAAG-A--AC-A---.........AAG--G-C------A--------AA-----A------A-A--------------- 2574
MNE.US.x.MNE027             ....--G---A-CA-GAC---GGAC--C--GA-T-----TT----T------T--C-A--AGCTC-G--AATGT--C-------C-----AGCTAAGG-A---C------.........AA---C-CC---------------AA-----A------T-G-GG--G------- 2395
DEB.CM.99.CM40              ....TCAGATC-----------GA---C--AA-T--T--TC-A---------TAT--AG-AA-A--G--GGTAT--A----------AG-ATTAGA-TA-ACCC-A---.................---------AGAG--GGCA-----A-----GA--------------C 2434
DEB.CM.99.CM5               ....TCAGAT--G--------TGA---C--AA---------T-G-----G--TAT---G-AA---G--GGTAT-CA-----------AG-ATTGGACT--ACAC-A---.................--G--G---GAG---GCT----------GGA--------G-----CC 2428
DEN.CD.x.CD1                ....-AC-----------------------CA-A-----CT-A--------CTATC-T-G-A--T-A--AGTA--AC----CA-GG-AG-ACAG-AA--GCG-C--CA-...GTCACC----------------A------T-T--------T-G-A-A-----------T- 2512
LST.CD.88.447               ....--G---------T-G-----------A--AG----TG-A---------A-T--A---A-TG-G--AGCA-GA---GTAC-AT-ACAATTAAG-GACA-A-TACCC...ATCACAAAG--C-----G--------CTGT-----T--C-G---A-----G-----T- 1893
LST.CD.88.485               ....--G------C--T-------------A--AG----TG-A---------A-T--A---AGTG-G--AGCA-AA---GTAC-AG-ACAATTAAG-GACA-A-TACCT...ATCACAAAG--C-G---G--------CTGT-----T--C-G---A-----G-----T- 1893
LST.CD.88.524               ....--G-----T--GC-------A--C--A--AG-T---G-A--------CA-T--A-AC-CTC-A--GGCT-GA---GTG--AG-ACAATTAAG-GAGA-A-TACCT...ATAACAAAG--GC-T--G--------GTGT-----G---G---A-----G------C 1890
LST.KE.x.lho7               ....--G-----C---T-----------C--A--AG-T--TG-A--G--G---A-T--A---GCA--G--GGC--AA--G-T---AG-TCAGTTAAG-GA-A-A-T-CCT...ATTACAAAG---TCT-----A--T--GTGT-----A---G---A--G--G-----TC 2977
SUN.GA.98.L14               ....--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-...ATTACAAA---T--G-----AGA---GGCC-C---A--------G-----------CC 2982
MND-1.GA.x.MNDGB1           ....--G---------GA-----------T--AA-AG-T-----A--G-----CA-AA--GAAA-AT-A--AG-A-AA----TA--GG-ACAATTGT--GA-A-A-TACC-...ATAACAAA---G--------A------G-A-----A--C-G-A-A-----------T- 2312
OLC.CI.97.97CI12            ....-AGG----------G---GA------A--T--TG-GG-A-----G---G----AG-AGCAG-A--GATACG-C-GGTC-GCG-ACAGCTTT-AGA-A---TCCCT...ATTACTAAG------------AGAG---CACA-GCCT---C-GA----G--------T- 2499
WRC.CI.97.97CI14            ....-CAGACA-TA-GAC---GAAA------A-A--T--T--A-----G---G-A--A-A-A-A--G--AATGCA----GCC---G-ACAATTAT-AGACA---T-CCT...ATAACAAA---G-GC--G--A------TGCTCA--A--T-G---A-----G-----C- 2925
WRC.CI.98.98CI04            ....-CAGATA-TA-GAC----AAA--T--AA-T--T--T--A--T------G-A--AGAAA-A--G---ATGCA---GGCC---G-ACAATTAT-AGA-A---TACCC...ATAACAAA----GT--G---------TGCACA--A---G---G--G--G-----T- 2940
WRC.GM.05.Pbt_05GM_X02      ....-CA-ATT--A-GAC----AAA------A-A--T-----A---------G--C-AGAAA-AT-G--AATGCA-C--GCA---G-ACAATTAT-AGA-A-A-T-CC-...ATAACAAA---CTC------A--T---TGTC-A--A---G-A-A--G--------TC 1854
MND-2.CM.98.CM16            ....-C-T----CC----G--C--C--------A--T--TG-A--T------G--C---AGA-AT-A--A--T--AC-----------AG----TAAAG-A--A--A---.........AAG----C-C---A--------CT--C--A--T---A-A-------------- 2626
MND-2.GA.x.M14              ....-C-T----GC----G--C-----C-----A--T--T--A--T------G-TC-T-AGA--T-A--A--T---C-G--------T---C-G--------T-----TAAG--A-----A---.........AA----C-C---A----------------A---G-A-------G------C 2553
MND-2.x.x.5440              ....-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---.........AA----C-C---A-------CA-----A--G-G-A-A-----G------C 2186
DRL.x.x.FAO                 ....-C-T-G--G--G-----C-----C-----A--T--T--A--T------G-AC-CTG-A-AT-A--AGCT---C-T--------TG----T-AAG-A--AC-A---.........AA----C-C---A----------------A-----GA-A-----------CC 2192
MON.CM.99.L1                ....-CA-----------T-----A--T---A-A--T--TT-A--G------G-C--AG--A----G--GGTA--CC-C--CA-GG-TGC-G-GGGAGACCTCCAACCCCCATCTTTAGAG--C-CT-----AG-------AA--G--A-----GT-G-----G------- 2331
MON.NG.x.NG1                ....-CW-----T--GC-G-----R--C--AA-T--T--TT-G--G-----CG-CC-CT-AAG---G--GATA--CC-G--CA-GY--CAGG-AATCGCCC-AT-CCC-...TCCTTGA----G-CCC-T--A-------AA-----G------T-G-----------CC  924
GSN.CM.99.CN166             ....-C---T---C-G--T-----C-----A--A--T---T-A---------TGC--A-G-A--T-A--GGTA--------CA-GGTGCAGCAGAAG--A--AC-CA--.........GA---GC-CC-----G-------AA--G--A-----GA-C-----G-----CC 2335
GSN.CM.99.CN71              ....-CC--T---C-T-----G--C-----A--G--T---T-G---------TGT--A-GCA--T-A--AGTA--C-----CA-GGTGCAGCAGAAAC-A---C-C---.........GA---GC-CC---AT------AA--A--A-----GA-------------C- 2335
TAL.CM.00.266               ....-AC-----T--GA-T-----A--G--AA-T-----TC-A--G------T-T--AG-AA-AT--A--GTT--AC-T--CA-GGT--GCTCAGAC--ACC--TAACC.........AAG--T--------AGAG--T-A---ACC---C---A-C--G-------C- 2594
TAL.CM.01.8023              ....-AC----------A----G--------AA---------T-A---------T-CC-AG-CA-AT---GGTA--AC-T---A-GGT--GC-CAGAA--ACCAGTAA--.........AA-------G--G---GA-----A---ACC------GA-C--G-------C- 2092
MUS-1.CM.01.1085            ....-CA---------------G--A-----A-A--T-----G--G-----CTGT--A-GCA-AT---CATT---C-----CA-GATACAGG-AAAAT-A--AC-AA--.........AA---G-CCC----AGA----GCAA-A--G----TG--C--G--------T- 2315
MUS-1.CM.01.CM1239          ....-CA-----G--G--G-----------CA------------G--G------TGT--A-G-A-AT---AAT---C--G--CA-GATACAGGAAAAA--A--AC-CA--.........AA---GGCC------GA---GGCAA-A--A----TG--C------G-------C 2305
MUS-2.CM.01.CM1246          ....-CC-----C--G--GA-T--------CA----T--CT-A--G------TGC--AT-AA-AT-G-AAATA--AC----CA-GGTA--GGCAGGACAGC-ATTGAA-...CCCACTAA---T-GCC-T--AGA-A---AAA-A--G----TG--A------------C 2413
MUS-2.CM.01.CM2500          ....-C------TC----------------AA----T--TT-A--------CTGC--AT-AA-AT-G---ATT--C--G--CA-GGTAGGAG---AAC-G--AC--ACG.........AA---TGG-C-C--AG----G-AA--A--G----TG--A--------------- 2372
SYK.KE.x.KE51               AGAG----GTC-TC----------------A--A--T--TT-A--------CA-A----GCA-AT-G--AGCA-AA---GTAA-GATAGCACAAAAAT-A--AC-AA-G.........AA---CC--C-----------G-A---GCTA--T---T-A-----G-----TC 2480
SYK.KE.x.SYK173             AGAG--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.........AA-----GTC-G--A----AC-AA--ATTG--CCGGC-A--------------- 2822
RCM.GA.x.GAB1               ....-C-CAT--GC----G-----A-----A--G--T--T--A--G------G-T--A-AGA-AT-A--AGC---AC----------A---AC----------A--A----AA--A--A---A---.........AAG---G--------T-T---CAA-----A---CG---A--------------- 2104
RCM.NG.x.NG411              ....-C-CAT------G---------A-----A--A--T--------C--G---G------GCA-AT-A--AGC-T--C----------A---C-T--C----------A--A--TAAAGCA--A--AA--.........AAG---G-----------------CAA-----G--T-G---A-----------T- 2119
GRV.ET.x.GRI_677            ....--CT--A-TC--A-----AG------AA---T--T--A----------A-T--AG-----GCA--CATG-AA---GT-A-GGGAG--CTAT---G-C---T--AG...GAAACAAA----C--C----AGA---G-AA-----A--T---T-G-----------C- 2798
SAB.SN.x.SABIC              ....-C---CA-T--G------GAA-----A--T--T--T--A---------G-C--AG-A--AT-A--AGTT---C-------GA-ACAA--AGAG--A--AC--A---.........AA---CC-C-------------CAA-----G----GGA-A-GG--------T- 3023
TAN.UG.x.TAN1               ....----AT---C--A-T-----A-----AA-A--T--TG-A---------A-TC-AG-A-CTT-A--AGCA-GA----TAA-GGGACAGCTAAA-GAGA-A-TACCC...ATAACCAAG---C----G--A--T--G---TCA--A--G---T-A-----------CC 2827
VER.DE.x.AGM3               ....--G--CA----GA-----AGC--T--CA-A------------------A-A--AG-A-CAGCA--AGC--AA---GTAA-GGGTCAACTGT-AGAAC-A-T-CCC...ATTACC--T--G--------AGA---GGCTAGA--A--TTTCT-A-----------CC 2350
VER.KE.x.9063               ....--C-----TC-TT-----G-A--C---A----T---------------T-AC-AG-A-CAGCA--AGCT-GAC--GTGA-GGGACAATTGT-AGAAC-A-T-CCC...ATCACC-----C---C-G--AGAT---GCAAGG--A---C-TA-A-----------TC 2855
VER.KE.x.AGM155             ....--A-----CC-CA-----G-A--T--CA----T--C--A--------CT-T--AG-C---GCA--AGC--AA---GTGA-GGGGCAATTGT-GCAGACA-TACC-...ATCACC--G---CGC------GA---GGCCAGA--A---CG-T-G--G---------C 2847
VER.KE.x.TYO1_patent        ....--A---A----GT-----G-A--T--CA-T--T-----A--T------T--C--G-C-C-GCA---GC-CGG---GTAA-GGGACAATTAT-AGAAA-A-T-CCT...GTCACA--T--C-----G---GA---GGCTCGG--A--CTGT--A-G---------TC 2852
COL.CM.x.CGU1               ....--AGATA-C---A------AC--C--CA-A--T---G-A--T------A---A-AAG-AGGGAAA-TAGTAG--GC-CAATTATCAGATAGG--CCCAGTCACC.........AAG--G-CG-----AGAG--C--------G--------A--G--G-----C- 2400
```

```
H1B.FR.83.HXB2       TGACAGAAGAAAAAATAAAAGCATTAGTAGAAATTTGTACAGAGATGGAAAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTCAGAGAACTTAATAAG 2795
Pol                  L__T__E__E__K__I__K__A__L__V__E__I__C__T__E__M__E__K__E__G__K__I__S__K__I__G__P__E__N__P__Y__N__T__P__V__F__A__I__K__K__K__D__S__T__K__W__R__K__L__V__D__F__R__E__L__N__K__
H1A1.UG.85.U455      ---------G-------------AC---------AT---------------A--------------------------------------------T----------G----C----------G--------------------C----A 2241
H1B.US.90.WEAU160    ---------G----------------------------A------------A-----------------------------T-------------------------------------------------------------------- 2794
H1C.ET.86.ETH2220    ------------G----------AC--C------GA--A----GC------A---------G-------------C-----T--------------------A-G-G-----------G--------------------G----C----A 2187
H1D.CD.84.84ZR085    -----------------------AC-------------T------------A---------G---------------------------A----------------------------G------------------------------- 2317
H1F1.BE.93.VI850     ------------G----------AC-----A--CT---A-------A----A--------------------------------------------------------------G--------G--------------A----------A 2132
H1G.SE.93.SE6165     -----------------------AC---------A---A-----G------A-------------------------------C-----A-----------------G--------------------G------------G--C----A 2192
H1H.CF.90.056        -----------------------ACG--------------------A----A-----C---G---A-------G-------GC------A------------A-G-G-T---------------------G-------------C----A 2142
H1J.SE.93.SE7887     -----------------------AC-C-------G---A-----G------A---------G-G-----------------T-C----------------------------------------------G-----T-------C----A 2109
H1K.CM.96.MP535      -A---------------------AC-------------A-------A----A---------------------G-------T-C-------G--------------------------------------------T-------C----A 1991
H101_AE.TH.90.CM240  -----------------------AC---------A---------G------A-----C--------------------------------------T----------G----C--C-------G-----------------G--C----A 2369
H102_AG.NG.x.IBNG    -----------------------AC---C----------------------A----------------A--------------------------------------G--T---------------------------------C----- 2320
H103_AB.RU.97.KAL153_2 -A-------G-------------AC---C-----AG---------------A--------------------------------------------------------------------------------G-----------C----- 2021
H104_cpx.CY.94.CY032 --------------------C--AC---G--A------C------------C-G-----------------------------------A------T---------------C-------------------------------C----A 2161
H1N.CM.95.YBF30      -----AC--------G-G-----AG-------------A------------A-------T-G---A-------G------T--C-----A-T----T-----A--G---T--C--------------------------G---T-A---A 2389
H1O.BE.87.ANT70      -AT-TA----------G----C--GAC--C---A---CAG--A------C-A-----A--------G--A--A----------T--T-----A--TA-C-----T-----A--G-----TG------------------G--------T--G---T-A------ 2850
H1O.CM.91.MVP5180    -AT-TAG--G------G----C--ACT-C---A---CA--A-------C--------A-----C-----G--A--A----------T--T-----A--TA-T-----T-----A--G-----T--C-----G--------G-----C---------T-A----A 2825
H1P.FR.06.RBF168     --T--A-G---------G----T---AG--CT--Y---CA-----------C--------A-----AA----------------------T--C--C--CA-C-----A-----A--G-----T--C-GC--R-----G---------C----------A----- 2256
CPZ.CM.01.SIVcpzCAM13 -AT---C---------T-----C--GAC------A---CAG--A----------------G--A--T-----A--A-------C----------------TA-T------A--T--A--G-----T---T-A--------------------------T-A----A 2375
CPZ.CM.05.SIVcpzEK505 ---------------------G--C--AC------------------A-------A-----A-----A--------C-G---A-------------------T----------A-T-----T-----A--G-----T--C----------------------GT-A--C--A 2367
CPZ.CM.05.SIVcpzLB7  -A-----G---------T-----T---AC------C---------A-------C-----A----------G---A---------C-----T---------T--G-----T-----A--G-----T--C--A------------------T--------T-A----A 2350
CPZ.CM.05.SIVcpzMB66 -A----------------G--------AC------------------A--------------------G-G-A-----------------------C---A-T-----T-----A--G-----T--C--------------------------GT-A-----A 2341
CPZ.CM.05.SIVcpzMT145 --T--A-----------G-G------ACT------A----A----A-------------A------------A-----A--------T---------A--TA-T-----A--T-------G-----C--C--------------C-T--T---------------------- 2337
CPZ.CM.-.SIVcpzMB897 -A---------------------GAC---------------------------------A----------G---A----A--------T------A---A-----------T--T-------G--T--C--A------------------C----------T-A----A 2361
CPZ.CM.-.SIVcpzDP943 -A--------G-----C-----T---AC-----------CAG--A-------A-----A--G-----------A-----A--------T--T-----A---A-T-----------A--G-----TG--GAA-------------G--------T--------T-A----GA 2381
CPZ.CM.98.CAM3       -A----C----------T---------AC------------CAG--A-------A-----A--G-------G--A-----C-----------T-------A--A-T------T------A--G------T------A------------C-G-----------G---T-A----A 2206
CPZ.CM.98.CAM5       -A----C----------C--G--C---AC------------CCAG--A-----------G--A-----------G--A--C--A--G--C--T--T--C--A---A-T------T-----A--G----------A--------------C-G--------T------T-A--C--A 2497
CPZ.GA.88.GAB1       -AT---C----------T------C---AC------------CA---A------------------A--G-----A--A------------------TA-T------A--C--A--G-------------------------------C-----------A----- 2856
CPZ.GA.88.GAB2       --T--A-------G---G-G------AC------------GG--A---------A-----A-----------G--A--T---------------T----------CA-------A----G--G------------A--G---------------------T--------T-G----A 2188
CPZ.US.85.CPZUS      ------C-----------T--G------AC------------CCA---A---------A-----A--G---A--G-G-A----A--------T---------A--A-T-----T--C--A--G-----T--C----------------------C--T-------T-A----A 2855
CPZ.CD.90.ANT        -CT--A----G------G-G--CC--AA-----------GATA--T-A---GCA---AAT--G-----T-G---A-----A--T--C-----------A---A-T------A-----A--G-----------C-T-A----------GC-----T----------T-A----A 2234
CPZ.TZ.00.TAN1       -CT-CA-G--G-----TG-G------AC-------A----AGACAT----------------A--------TGC-G----A--A-----C-----T--C--A---A-T---------T---------G--T-CCT-------------------------------G----A 2439
CPZ.TZ.2001.TAN2     -AT-CA---G------TG-G---C--AC-------C----ATACAC-A---------------------GC---A-----A-----C------------A--TA-T------A--C--------------TC-T------------G-----------------GT-A------ 2444
CPZ.TZ.2002.TAN3_1   -AT-CAG---G-----TG-G------AC-------------ATACAT-A------A-------A--------TGC---C-----A-------------------A--CA-T------A--C-G-------G----CCT------------G----------C--T-----GT-G------ 2444
GOR.CM.2007.SIVgor2139_287 --T--A-------------G----C---AC--CT-------CAG--A-------C---------C--G--A--------------------------T------A--TA-T-----------A--G-------------------------G---------T--------T-A----A 2332
GOR.CM.2007.SIVgorCP2135con --T--A-------------G----C---AC--CT-------CAG--A-------C---------C--G--A--------------------------T------A--TA-T--C--------A--G------------K-------------G---------T--------T-A----A 2331
GOR.CM.2004.SIVgorCP684con --T--A-------------G----C---AC--CC-------CAG--A-------C---------C--G--A--------------------------T------A--TA-T-----------A--G--------------C-----------C-G-------T------GT-A----A 2251
H2A.DE.x.BEN         -A---A----G------G-----C--AA---G--C---GA-A-A----------A--G--CC-GC-AGA-G-GGCACCT--AACT-----T--T-----C--CAC--------A--T-----------G----AG-AC---------G-TGC-GA---------T----------A------ 3180
H2A.GW.x.ALI         -A---A----------G-G---C--AA------C---GA-A-A-C------GA------CC--T-AGAGG-GGCACCT--AACT--C--T--T-----C--CAC--------A----------G--G----AA-AC-----------TGC--A-----------T--------T-A--C--- 3173
H2A.SN.x.ST          -A---A----------G-G---C--AA---G--C---GAGA-A---------GA--G--CC-GC-AGAGG-GGCACCT--AACT-----T--T-----C--CAC--------A--C-----------G----AA-AC-----------TGC--A-----------T-----------A--C--- 2625
H2B.CI.x.EHO         -AT-CA----G------CT---CC-CAA------C---GA-A-A-----------A--G--AC-GT-AGA-G-GGCGCCT---ACT--------------T-G--CACC--C----------A--G-------AA-AC--------G-TGC--A---------------------A--C--- 3151
H2B.GH.86.D205       -AT-CAGG-----G---CT---CC-CAA------C---GA-A-A--------------G--AC--T-AGA-G--GCACCC---ACT-----------C--A--CACC----------------------G-----T-AG-AC-----------TGC--A-----C--T------GT-A--C--- 3156
H2G.CI.92.ABT96      -AT-TA-G----------C------C-CAGG----------GACA-A-------C-----------C--T-GGA-G--GCACCC--CACC-----G--T--C--A--CACC---A-T--C-------G--G----AA-AC-----------TG---A-----C-----------------C-G- 2507
H2U.FR.96.12034      -AT-TA-G------T--G---------AA---G--A---GAGA-A-----------G--AC--T-GGA-G--GC-CCC--CACT------------C-----CACC------------------G--G----AA-AC-----------TG---A-----C--T------GT-A--C-GA 2662
SMM.SL.92.SL92B      -AT-TA-G------G--C----------AC---G--C--CGAGA-A-----------G--AC--T-GGA-CG-GCCCCT---AC-------T--T-----A--GACC------T--T---------GA-T-AGGAC--G------TGC-CA---------------T-A--C--A 2528
SMM.US.x.H9          -AT-GA--------G----TT--------AG-------C---GA-A-A------G--A--T--CC-GT-AGAGGR-GC-CCT--AACC-------T--C--C--CAC-------T--------R-GR------G-A-A---------G-TGC--A-----------T--------T-A--C--- 2570
SMM.US.x.PGM53       -AT--A------------GTT--------AG-------C---GA-A-A----------A--T--CC-GT-AGAGG--GCCCCT--AACC-------T--C--C--CAC-------T-------A--G----------AA-A-----------G-TGC--A-----------T--------T-G-----G- 3014
SMM.US.x.SIVsmH635F_L3 -AT--A-----------G----TT--------AG-------C---GA-A-A----------A--T--CC-GT-AGAGG--GCCCCT--AACC-----G--T--C--C--CACT-------T-------------G------AA-A----------------TGC--A-----------T--------T-A------ 3100

MAC.US.x.239         -AT--A-----------G---GTT--------AG-------C---GA-A----------------T--TC-GT-GGAGG--GC-CCC--GACC-------------C--C--CAC-------T---------------------G--T-AG-AC--------------TGC-GA--------T--G------A----G- 3097
Pol                  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__A__P__P__T__N__P__Y__N__T__P__T__F__A__I__K__K__K__D__K__N__K__W__R__M__L__I__D__F__R__E__L__N__
MAC.US.x.251_1A11    -AT--A-----------G---GTT--------AG-------C---GA-A----------------T--TC-GT-GGAGG--GC-CCC--GACC-------------C--C--CAC-------T-----------------------T-AG-AC--------------TGC-GA--------T--G------A----G- 3096
MAC.US.x.251_BK28    -AT--A-----------G---GTT--------AG-------C---GA-A----------------T--TC-GT-GGAGG--GC-CCC--GACC-------------C--C--CAC-------------------------------T-AG-AC--------------TGC-GA--------T--G------A----G- 3073
MAC.US.x.EMBL_3      -AT--A-----------G---GTT--------AG-------C---GA-A----------------T--TC-GT-GGAGG--GC-CCC--GACC-------------C--C--CAC-------------------------------T-AG-AC--------------TGC-GA--C-----T--G------A----G- 2571
STM.US.x.STM         -AT--A-----------G---G-G----C--AA----G--C---GA-A-A------G-------T----C-GC-AGA-G--GC-CCT--AAC-------T-----------C--CACT------T-------------------G-----AA-AC--G-------G-TGC-GA-------------------A-----A 2744
MNE.US.x.MNE027      -AT--A-----------G---GTT--------AG-------C---GA-A----------------T--TC-GT-GGAGG--GC-CCC--GACC-------------C--C--CAC-------T---------------------G-----AG-AC--------------TGC-GA--------T--G------A------ 2565
DEB.CM.99.CM40       -TT--AGG---------G-G---C---AA---T--C-TA-AGCCC---AT-G---C---AC-G--AGTCCCGGCA-C---CACC--C-------T--CT-----------G------TA------------G--G-----AA-G---G---C---TGC--A------C--A------CCT-A--C--- 2604
DEB.CM.99.CM5        -TT--AG----------G-----C---AA---G--CATA-GGCCT-----T-G-A-C---AC-G--AGTCCCC-CA-A---CACT--C-------T-GCT------T--G------TT------------G--G-----AG-A-----------------TGC-TA------CC-G-------CC-------- 2598
DEN.CD.x.CD1         -A--TAC---G---T-------T---CA-C-G--AGTG---C---------------G--A-GGT-GGA---GGCGCCA--CACC--C--T-----------A--T--G-----TA--TCCA--G---------AA--AC---AT--G-TG----A-----------GA--T-A----G-- 2682
LST.CD.88.447        --T-TA-------G---G---GCC---CA--C--------GATAGAT-A------A---------------A-------GCA-ATG-A-GG--C-------T--C--A---A----------------------G--------AA-A-G-G------------------A-T-------T---C-----A---GCT 2063
LST.CD.88.485        --T--A-------G---G-G-GCC---CA--C--------GATAGAT-A------A---------------A-------GCA-ATG-A-G---C-------T--C--A---A-C--------------------G--G------AA-A-G-G-------------C-GA-T-------T---C-----A---GCT 2063
LST.CD.88.524        -AT-CA-----------G---G-----CA--C--------GATAGAT-A----------------------A--CCC-G---AT--A-G-------T------C--C--A--C------T------------------G-----AA-A-G----------------G--GA--------------A--T-A----G-- 2060
LST.KE.x.lho7        -AT--A----G------G---GCC---CA--CT------GATAG-T-A--------A-------A-------A--TCC-G-G-AT--A-GG--C-----------------------A---------T--C--------G-----AA-A-G------------G---A-T------------A-G--C---G-A 3147
SUN.GA.98.L14        --T--AGG--G------G---GCC---CAGA--------GATAGAT-----GCA------A--G---AGC-GGGCA-AC--A-GG--C--T--T--C--A---A-T------T-----------------G---------AA-A-G------------G---A-------T------T----A---GGT 3152
MND-1.GA.x.MNDGB1    -A-GTA----G------GTT-GTC-TCAGA-----A---GATAGAT-A--GG----------A------AGT-GGG-A-AT--A-G----AAT--------A--TA-C----------------------G--G--T-AA-A-G-------------------A-----C--T------T-A--C--- 2482
OLC.CI.97.97CI12     -A-GTA-G----G---G-----T-----AGA-----A---GACAGAT------G-A------T--G--ACAG---GCA-A-G--TC------A--GG------C---A------TG-------A--G--G-----AA-A-G-G-AT------G--GA--------T--G-CT--A---GCA 2669
WRC.CI.97.97CI14     --T--A--A---------G-G-GGC--CAGA--------GATAGAC-A---GCA------A-------AGC-G-GCA-AT--A-G----------T--------------TA---------T--T--A--G-------AA-A-G--------------------GA------------A--T-A---G-A 3095
WRC.CI.98.98CI04     -AT--A----G------G---GGC--CAGA-----C---GATAGAT-A---GCA------A-------AGT---GC-TCC--A-G----------T--------------TA-----------A-----------------G------AA-A-G-------G--GC-CA----------T---A--T-A---G-A 3110
WRC.GM.05.Pbt_05GM_X02 -CT--A----G------TG-G-GGC--CA-A-G--A---GATAGAT-A---GCA------A------AAGT-GGGCA-AAG-AA-T------T----------A---A-C------A----GA--G-------T-AA-A-G----------------------A----------T-----G---A---GCA 2024
MND-2.CM.98.CM16     --T-TA-G-----G--TTT---C----CA------A----GTC-A-------G------------CC-G--C---T--G--A--T--A---------T-------C--A--G--G---TG---C--A--G-------TG-A--C-GC---------------G-----------T----C-GT-A--C--A 2796
MND-2.GA.x.M14       --T-TA-------G--TTT---C---CAGAC---A----GAT--A-------------------C--------C--T-G---A--T--A---------------T-----A-------------TGT-----A--G-------TG-A--C-GC-----------------------------T--GC--T-A----GA 2723
MND-2.x.x.5440       --T-TA-------G--CCT---C---CA-------A--C-ATC-A-----------A------C--G--C---T-G-G-A--T--A-----------------C--C-----------CTGT-----A--G--------TG-A--C-GC-----------------------------T----C------G--C--A 2356
DRL.x.x.FAO          -CT-CA-G--G------TG-----C---GAA-------A----GAGC-A-------G-----------TC----AAGT-G---A--T--A--------------------C--A--G--G---TG---T------G------TG----C--------------------------------GC--T-A------ 2362
MON.CM.99.L1         -AT-CAG-------G---G-G--TC-TAC-C------AC-CA-------CTT--ATT---AC--T-GGA-CC--CC-A---CA-T--C------T--CT-A--CTGT-------------------------G------AG-GC-------------TGC-CA---------------------------C--- 2501
MON.NG.x.NG1         -CT-TA--------G---G-G--TC-CACCACC----RCCCGG-------AT-G-ATT---CC--T-GGA-G--GCA-A---AACC--------T----T----TTGC--C--G------G-------G--T-AA--C-----------G-TGC-CA-------------G----T-A--C--A 1094
GSN.CM.99.CN166      --T-TAGG-----G---G-G---C---ACCC----AACACA---------CTT-GATT---TC--T-GGA--------A--A--A-------------------CT-A-------C---------------A--G--G----AAT-AC----------TGC--A------------T----A--T-A---G-- 2505
GSN.CM.99.CN71       --T-TA-G-----G---G-G---C---ACCC----AACACA---A---CTC---TT----CC--T-GGA---G--A----A-----------------------CT-G-------C------------------A--G------AAT-AC----------TG---A------------------GA--T-G---G-- 2505
TAL.CM.00.266        -A---CAG------------G----GGC--CA-ATC--CATAGAC--C-----TG------C----AC-GT-AGAGG-GG----A--G-----------------T---T-----------G--C--T--C-GA--G--G----AA---C-------------TGC-CA-----C---C-C-CC--A---C-A 2764
TAL.CM.01.8023       ----CA-G------------G---GGC-GCAG-CC--CATAGAC--C-----TG------C----AC-GT-GGAGG-G----------G------C---------CT-----------G------T--C-GA--G--G----AG--AC------------TG--GA----------C---CC--G---C-- 2262
MUS-1.CM.01.1085     ----CC-G------G--C-GG--C--G-AG--C--AACCCAGC------CT----A-T----CC------AGAGG-----A--A--A--G------T--T---T------TGC------A--T-G--------G----AAT-----G---------TGC--A------------GC--T-A---G-- 2485
MUS-1.CM.01.CM1239   ----TC--------G--T--G--C----A--GC--AGT-CAGC-A---TT----ATT--ATC-G--AGA-G-G--A------A--C------T------CT------TGT------A--C-G----G-------T-AAT--------------TG---A---------T----C-G--A---G-- 2475
MUS-2.CM.01.CM1246   -CT-TA-G--G--G---G-G---C--AA----------AC-CAGA-A---AT----ATT----TC--C-GGAGG--GCA-------A-T--C-------T---T-A------------C---------GA--G----------AG--A------C----TG---A------------A-G----------G-- 2583
MUS-2.CM.01.CM2500   --T-CA-G------------G-----TC---A-------AC-CAGAGA---TT-----CTG--AC--T-AGA-G---CA--------C--------------------GC----C--------T----G-------G------AAT-C-------------TGC--A------C-------A-G--C---G-- 2542
SYK.KE.x.KE51        -----AG-------C-TG-G---C-GCG--C---AGT-T-----T---CTTG-AA-G---C--T-GGA---GGCCTCC---AC------T--T------C-------T-----TA---C-A--G--G----AATTA-----------G-TGCCTA-G------T--GA----C---G-A 2650
SYK.KE.x.SYK173      -AT-G-T---G---T--G-G--CC-CAAG-C---AGTAGA---C---CTC------C---AC--T-GGAG---GC-A-T---ACC-------T--T-----C--------------TA--T-G------G-----AA-AG-----------G-TGC-CA-T-------T----A--T-A---G-- 2992
RCM.GA.x.GAB1        --T--A-------------G----T---AC------------C-AT-CA------G---------A--------------------A-----A-----C--C-----C--A---A----CTGT--T--A--G-------TCC--A-----------------G-----------T----------A------ 2274
RCM.NG.x.NG411       ---GTA----G------TG----T---AC---G--------AT-CC------------G--A-----------T-G---A--A--A-------------T----------------A-----TGT--C--A--G-----TTCC--A---------------GC-------------------T-G----A 2289
GRV.ET.x.GRI_677     -AT--AG----------TG----T---AC--------A--C-A-C-A-------G----G--A---T-A--T-GG-A--AGGA----------T--T------A-----G-------------------------G-T-AA--AC-----------TGC-T---------------G-----A--C--A 2968
SAB.SN.x.SABIC       --T--A-------------G-G---C---AAG-CC--------GA---CT-A------AC----AC-TT-GGA--G---A------A--------------T--C--C------C-------A-----G---G------T-AA---C-G---------T-C-CA-G--C------GC-GT-A-----A 3193
TAN.UG.x.TAN1        -AT---------------------CC-GAC-------AAC-CAG-CTC-A-----------A----------T--GG----AGGA--G-----T------------C--G---TGT--CCG-------------AA--AC-G-----G-TG---A----------T----CT--C-----A 2997
VER.DE.x.AGM3        -CT-CA-----------------C----CAG------A---GACC-AT-A--G--A------A-------AGC---G--A--AGGA--G---G-----------C----------G---TG---------------------AGT-AC-----------------TG------------T--G------A--C--A 2520
VER.KE.x.9063        -AT--A-G--G---------G--C---AC---G--A---GA-C--T-A------A-------------CAGC--G--A--AGGA-------G-T--T-----C---A-C---TGT--C-GA--G--G----AA---C----------G-TG--------C-----------T-A-----A 3025
VER.KE.x.AGM155      -CT-TA--------------T---CC-GCA-----------C-A-ACAT-A--GG-A------A---T-AAGC-GGG-A---GGA--C--G------------A-----------CTGT------G-------------AAT-AC-G---------TGC-G--------------G-----C--C--A 3017
VER.KE.x.TYO1_patent -CT-TA----G--G--TG-----T---CAG------A---T-CC-AT-A--GC------A------CAGT-G-G-A--AGGA-------G---------------C---A------TG------------G--G----AAT-CC-G-------G-TGC--------C--T------GT-A------ 3022
COL.CM.x.CGU1        -A-GCA-----------G---GT----CA-A-G--A---GA-AGAT------GCA--G--A------AGA----GCA-AATTA-GG-------T-----A---A-----TG---T-G----------T-AA-A-G---------G-------A-----------G--G--A--C--- 2570
```

```
H1B.FR.83.HXB2                 AGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAAATCA..................GTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCATTTACCATACC 2947
Pol                            _R__T__Q__D__F__W__E__V__Q__L__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__T__I__P_
H1A1.UG.85.U455                --------G--------------------------C--G---A-A--G--TC-------------------.................-----------A--------G--C-----------------T---------AG---T--A-----------G--C-------- 2393
H1B.US.90.WEAU160              --------------------------------------------TT-----------------------..................--------------------------------------A--------------------------C----------------- 2946
H1C.ET.86.ETH2220              --------------T-----------------G--------C--A------------------------..................--G--------A--------G-----------C--------T---------GT------A--A-----------C-------- 2339
H1D.CD.84.84ZR085              --------------------------------G-----------T-----A-----G--------G---..................A-------------------C---------------A--------TG---------T--------C----------------- 2469
H1F1.BE.93.VI850               -----------T--------G-----------------------T--------------------G---..................--------------------G---------------------------A-G--T----A------C--------C-------- 2284
H1G.SE.93.SE6165               --------------------G--C--------------T-----T------------------------..................-----------A--------G-----------------------------------T--A--------------C--T----- 2344
H1H.CF.90.056                  --------------------------G--------------C--A------------------------..................---T----------------G-----------------C--T------A----A-----A--------------C-------- 2294
H1J.SE.93.SE7887               --------------------------G-----------------A------------------------..................--C-----------------G-----------------C--T---T-------T--------A-----------C--T----- 2261
H1K.CM.96.MP535                --------------------------------------------A-----A------------------..................--------------------G-----------------C--T------A-------------------------C--T----- 2143
H101_AE.TH.90.CM240            --------G-----T-----------------------G-----A-----T------------------..................-----------A--------A--------------------T---------AG---T--A--------------C-------- 2521
H102_AG.NG.x.IBNG              -----------T--------G--C--G-----------T--------G--A------------------..................--------G--A--------G---------------------------A-------T--A-----------------T----- 2472
H103_AB.RU.97.KAL153_2         -----------------------------------------C--T-----------------------T..................------------------------------------------------C----------A----------------------- 2173
H104_cpx.CY.94.CY032           --------G-----------------G-----------G--C--A-----------G------------..................---------T----------G---------------------------CC---G-----------C--------C-------- 2313
H1N.CM.95.YBF30                --G--C-----T--T--------G--GC-------T--------A-----A-----GC-----------..................--G-----TT--------A-A-----T---------TG------G--CA----T--T--A--------A--T----------- 2541
H1O.BE.87.ANT70                -----A-----G--------G--A--GC----T--C--------G-GG--T--G--GC-A--GC----T..................--T-----CT-A-----A--A-----T-----C---TG---------CCC---T--T--A--A--------T--C--T--T-- 3002
H1O.CM.91.MVP5180              -----A-----T--------G--G--------T--T--------A-GG--T-----GC-A-GGC----T..................--T-----CT-A-----A--A-----T-----C---TGC--T------CC------T--A--A--C-----C--C--T--T-- 2977
H1P.FR.06.RBF168               -----A-----T--T--------A--G--G-----C--T-----A-G---TC-T--G--A-GG------..................--G------T-A--------G-----T-----C--CTG---T------CC---T--T--A--A---R-A--T-----T----- 2408
CPZ.CM.01.SIVcpzCAM13          -----C-----T-----------A-----G--C-----C--C--A------------G----G------..................-----------A--------------------C--TTG---TC-----A-G-----------------A-----C--A--T-- 2527
CPZ.CM.05.SIVcpzEK505          -----------T--T--------A-----------C--------A------C----G------------..................--------G--A-----A--A-----T--C------TG-------------A-T--T--A--------A--T----------- 2519
CPZ.CM.05.SIVcpzLB7            --------------T-----------------------T--C--T--G--A-----G-------GG---..................--C------T-------A--A-----C--C------TGC---C--------A----T--A--A-----------C-------- 2502
CPZ.CM.05.SIVcpzMB66           --------------T--------C--GC----------C-----T------------C-----------..................--------TT-A-----A--G--C--T--C------TG----C-G------A----T--------C-----T----------- 2493
CPZ.CM.05.SIVcpzMT145          --G--------A--T--------A---C-G--C--------C--A--T--------G------------..................--------CT----------A-----C---------TG---TC-T--CCC------T--A--A-T---------C-----T-- 2489
CPZ.CM.-.SIVcpzMB897           -----C-----T--T--------C--GC-------------C--T--G--AC-G---------------..................--------G-----C--A--G-----T---------TGC---C----------G--T--A--A--C-----C--C-------- 2513
CPZ.CM.-.SIVcpzDP943           -----A--G--------------C--------C--------C--G-----A--------------G---..................-----------A--------A--C--T--------CTGC--AC--------AGT--T--A--A--C--A-------------- 2533
CPZ.CM.98.CAM3                 -----A--G-----T--------C--G-----C--------C--G-----T--------------G---..................--------G--A--------A-----T--------CTGC--A---------A-T--T-----A-----A--------T----- 2358
CPZ.CM.98.CAM5                 -----A--G--A--------G--C--------C--------C--A-----AC----G--A--G--G---..................-----------A--------A-----T-----C--CTGC--A---------A-T--T--A--A--C--A--T-----T----- 2649
CPZ.GA.88.GAB1                 -----A--------T--------G--G-----C-----T--C--A------------------------..................--G------T-A-----A--A-----C--C--C--TTG----C-G---A----T--------A--------------A--T-- 3008
CPZ.GA.88.GAB2                 -----A-----T--T--------A--G-----C--T-----C--T-----T-----G--A--G------..................--------G-C---------A-----T--------CTGC------------A----T--A--------A--C--C-----C-- 2340
CPZ.US.85.CPZUS                --G--A--------------------G-----C--------C--A-----A-----G-------G----..................--------C--A-----A--G-----C--------TTG---AC-G---A-G--A--T--A--A--C--A-----------C-- 3007
CPZ.CD.90.ANT                  -----A-----T--T-----GA-A-----------T--T-----A--C--A-----GC-A--G------..................--G------T----------A-----C-----C--CA-A---------CG------T--A--A-----A--T--C--A--T-- 2386
CPZ.TZ.00.TAN1                 -----------T--T-----GT-A--GC----------C-----G-----------G--A-G---TATG..................--G--------------A--G-----C--C-----CA-----C-G---CC---------A--------A--T----------- 2591
CPZ.TZ.2001.TAN2               --------G--T--T------C-A-----------C-----C--AT----------GG-A-GG--CATG..................--------G--------A--G-----C--------CA-----C-G---CC---------A--------A--T--C-------- 2596
CPZ.TZ.2002.TAN3_1             -----------T---------T-A--G--------C-----C--AT----------GG-A--G--TATG..................--G-----G--------A--G-----C-----C--CA-----C-G---CC---------A--A-----A--T--C-------- 2596
GOR.CM.2007.SIVgor2139_287     C----A--G--T--T-----------G-----G--T--T--C--A-G----C-T---------------..................-----------------T--G--C--T--------TTG---------CCC---T--T--A--A-----A--T--C--A--T-- 2484
GOR.CM.2007.SIVgorCP2135con    --G--A-----T--T-----------G-----G--T--T--C--A-G----C-T----------G----..................-----------------T--G--C--T--------TTG---------CCC---T--T--A--A-----A--T--C--A--T-- 2483
GOR.CM.2004.SIVgorCP684con     C----A-----T--T-----------G-----G--C--T--C--A-G----C-T---------------..................-----------------T--G--C--T--------TTG---------CCC---T--T--A--A-----A--T--C--A--T-- 2403
H2A.DE.x.BEN                   GTG--------T---ACA---A----GC-------T-----C--G-----AC-GCC---A--G---AGG..................A-CT-TA--T-A-----A--G-----C--------CA-A--AC-C--------T--T--C--------------------C-- 3332
H2A.GW.x.ALI                   GT---------T---ACA--GA----G-----G--T--------A-----A--GCC--------G-AG-..................A-C--T--G--------A--G-----T--C-----CA-A--AC-GC--GAG--T--AC-------------------TC---- 3325
H2A.SN.x.ST                    GT-------------ACA---A-C--G--------T-----C--A-----AC-GCC------G---CG-..................A-T--T--C--A-----A--G-----T--C-----CA-A--AC-C-----G--T--T--AC----------------TC---- 2777
H2B.CI.x.EHO                   GT---C-----A--TACA--G--C--GC-G--T--T--T--C--A-----AC-GGC-TCA--G---AG-..................A----------A-----A--A-----C--C--CAGT--C--AC-----CC------AC-A--------A--------TT-G-- 3303
H2B.GH.86.D205                 GT---C--G------ACA-----CA-C-.G--T--T--C--C--G-----TG..GC-G-A----GGAG-..................A--------A-A--------A-----C--C--CAGTA-C--AC-----CC-A------AC--------A-----C--T--G-- 3305
H2G.CI.92.ABT96                GT---C--G-----TACT-----C--G--G--C-----T--C--T-----T--GGC-G-A----GGAG-..................A--------------A-A--G-----C--C--CAGCA----TC------T---T-----C-----C--A--T--C--TT---- 2659
H2U.FR.96.12034                GT---------T--TA-T-----C--G-----T-----C--C--A-----AC-GCTG--------GAGG..................A-------T--A-----A--G-----C--C--CAGCA-C--GC-GTG------------C--A--------T-----TT---- 2814
SMM.SL.92.SL92B                -TG--A-----A---ACA-----A--GC-G--T--C--------G-----CC----GG-A-TGG--AGG..................A-------G--A--A--A--G-----C-----------A---C----CCCG--A--T--AC-A--C--A-----C--T----- 2680
SMM.US.x.H9                    GTC--------T--TACA-----A--G-----G--------C--TR----AC-GC------G--G-AGG..................A-C-----GT-------A--------------C--CA-A--T-----------A-----C--A--C-----C------T---- 2722
SMM.US.x.PGM53                 GTT--------T--TACA-----A--G-----G--------C--T-----AC-GC------GG-G-AG-..................A-C------T-------A--------------C--CA-A--TC----------A-----C--A--C-----T-----TT---- 3166
SMM.US.x.SIVsmH635F_L3         GTC-----------TACA-----A--G--------------C--T-----AC-GC------G--GGAGG..................A-C------T-------A--------------C--CA----TC----------A-----C-----C-----C-----TT---- 3252

MAC.US.x.239                   GTC-----G-----TAC------C-----------------C--T-----AC-GC----A-GG---AG-..................A-T------------A-A--------------C--CA-A--TC----------A--T--C-----C-----C-----TT---- 3249
Pol                            _R__V__T__Q__D__F__T__E__V__Q__L__G__I__P__H__P__A__G__L__A__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__T__L_
MAC.US.x.251_1A11              GTC-----G-----TACA-----C-----------------C--T-----AC-GC----A-GG---AGG..................A-T------------A-A-----C--------C--CA-A--TC----------A--T--C-----C-----C-----TT---- 3248
MAC.US.x.251_BK28              GTC-----G-----TACA-----C-----------------C--T-----AC-GC----A-GG---AGG..................A-T------------A-A--------------C--CA-A--TC----------A--T--C-----C-----C-----TT---- 3225
MAC.US.x.EMBL_3                GTC-----G--AC-T-ACAG-AG-C-A-TA-----------C--T-----AC-GC----A-GG---AGG..................A-T------------A-A--------------C--CA-A--TC----------A--T--C-----C-----C-----TT---- 2723
STM.US.x.STM                   GTG--A-----T--TACA---A----GC----T--------C--T-----AC-GC------GG-G-AG-..................A-C------T-------A--G--------------CA-A---C-------G-G---T--C-----------T--C---T---- 2896
MNE.US.x.MNE027                GTC-----G-----TACA-----C-----------------C--T-----AC-GC----A-GG--GAGG..................A-C--------------A-----C--------C--CA-A--TC----------A--T--C-----C-----T-----TT---- 2717
DEB.CM.99.CM40                 GCC--A--G--GG-G-----G--G--GAC---T--------C--AT-T-CAC-CC-C--A-TGG--CAG..................A-C--T--CT-A---C-T-CA-----C----A----A-AG-AC----CCC-ACA--TGCA--A--C--A-----C--AG---- 2756
DEB.CM.99.CM5                  GCT--C--G--TG----------G---AC---G--C--G--C---T-G----GCC-C--A-TGG--CAG..................A-------C--A--CT---CA-----C----A----A-A-----G---CC-C-G--TGCACGT-----A-----C--A----- 2750
DEN.CD.x.CD1                   GCC--C--G--------------A--------T-----------T-----A---C--C----GG--CAC..................A-C--T--GT-A--CT--AAA----------A-AGT--G--T--G---CC---T--TGCAC-A--C--A---------AG-G-- 2834
LST.CD.88.447                  TT---A-------TCAT----A-------C--T----C--T----AA---G--ATGC--GCAG..................--------GG-A---A-A--A-----C------AGCA-----C-G---CC-----A---AC----C--A-----C--AG---- 2215
LST.CD.88.485                  TT---A-----C-TCAT----A-------C--T-------T----AA---G--ATGC--GCA-..................A-------GG-A---A-A--A-----C------AGTA-----C-G---CC----AT--AC----C--A-----C--AG---- 2215
LST.CD.88.524                  TT---A-------TCAT--GC-A--G-------------C--A----AA---G--ATGC---CA-..................A-T------G-A---A-A--A-----C-----CAGCA-A--T------CCT--T-A---AC-A--------G--C--T----- 2212
LST.KE.x.lho7                  TT-----------TCAT--GC-A-------T--T----C--A----TA-C--G--ATGC---CAG..................A-T------G-A--CA-A--A-----C-----CAGTA-C--T------CC-A-T-A---A---------A-------T----- 3299
SUN.GA.98.L14                  -TG----------TCAT--GT-A--G----G--T--T----A----A----G---TGT---CAG..................A-T------G-A--------G---------C--C---A-A--A--G--CC---T-AT--A---------A-----C--A----- 3304
MND-1.GA.x.MNDGB1              TT---A--G--T--TCAT---T-A--G----T-----T--C--A----AA-------TGT---AG-..................A-------C--A---A-A--G-----C------AGTA-A--TC-G---CC---T-AT--ACCC---------C-----GG---- 2634
OLC.CI.97.97CI12               GCC--A----T-------G--A-------------T----A----------G-A-GGG--CAG..................A-T------T-A---A-A--A-----T-----C--TA-A--GC-----CC---A-AT--AG-A-----A--------A--T-- 2821
WRC.CI.97.97CI14               GTT--A--G----T-TT----A--G----------T--C--T-G--AC-G-G-----T--GAAC..................A-TT-----G-A---A-A--G-----------C--TA-A--A------CC---A-AT---C-A--------T-----AG---- 3247
WRC.CI.98.98CI04               GT---A--G-------TT----A--G--G--C--T--T--C--A-G--AC---G-----T---AAT..................A-TT-----G-T---A-A--G-----C--------CA-C--AC----CCC-----AT---C-------A--T-----TG---- 3262
WRC.GM.05.Pbt_05GM_X02         GT---A----------TT--G--A-------C--T--T----A-G------C-T------G-AAT..................A--------T-A---A-A--A-------------TTA-C--A------CC---A-AT--AC----C--A--------AG---- 2176
MND-2.CM.98.CM16               GT-----G-------TT--G--G--G--G----C-----C--T-G---TC-C---C-ATGTG-GCA-..................A-T-----------CA-A--G-----C------C---TG---T--G-----G-----T--A-----C---------C-----T-- 2948
MND-2.GA.x.M14                 GTG----G--T---TT--G--G--GC-------C-------G-G---C------C-ATGTG--CA-..................A-C---A--T-----A-A--A-----C----------TG---T--G-----GAG---T--A-----C----------------T-- 2875
MND-2.x.x.5440                 GTC----G--T---TT--G--G--G--------CA-------A-G---CC---GC-ATGTG-GCA-..................A-T------T----CA-A--A-----C----------TG------G----TG-----T--A-----------G--C-----T-- 2508
DRL.x.x.FAO                    GC---A--G-------TT--G--G--G--G--G-------C--G-G---C------GCATGCG--CAG..................A-C--T------------A---------C------C---TGC--TC-G---CCG-----T--A-----C---------C-----T-- 2514
MON.CM.99.L1                   GC-----G-------TT--G--G--G--G----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--CT-A----- 2653
MON.NG.x.NG1                   GC---A----T----TT--G--A-----------T-----C--G--T--AC-CC-G--C-TGGC-AAT..................A----TA--G----CA--AAA-----C-T--ACAGCA-C--TC-G------A-T--T-----A-------------T-A----- 1246
GSN.CM.99.CN166                GCT--G--------GGCT--G--C--G--G----T--G--C--A-----AA-TCCGC-CCCTGC-GG-CTAGAGCAAAAGGAGCAT-----TA--A-A---A--AAA-----T----A-AGTA-A--A--GT----G--T------------C--C--T---T-AG---- 2675
GSN.CM.99.CN71                 GCT--G--------GGCT--G--C--G--G--G--T--G--C--A-----AC--.................GAACAAAAAGAGCAC-----TA--A-A---A--AAA--C--T----ACAGT--A--A--GT-------A-----A-----C---------T-AG---- 2657
TAL.CM.00.266                  GC---C---------GC------A---CAG--G-----T--C---T----CC-GCCTC---TG--GCA-..................A----G--C--A--CA--AAG-----C----ACAGCA-C---------CCC-------TCAC---C--A--C--C-----C-- 2916
TAL.CM.01.8023                 GC------G-----TGCA--G--C---CAG--G--------C---T-T--AC--CC-C-------GCAG..................A-------CT-A---A--AAG-----C----A-AGCA-C--A------CC-------GCCCGT--C-----T--C-----C-- 2414
MUS-1.CM.01.1085               GCT--C-----AC--GCA---------C-------C--------A-----A---.................ATAGAGCATGAGCAC--T---A--G-----A-CAAG--C--C-T--ACAGT--C---C-----CC------T--------C--A--C---T-GC-G-- 2637
MUS-1.CM.01.CM1239             GCC--C-----G---ACA--G--A--GC---G--C--------A-----A---.................GCGGAACATGAGCAT--C---A--G-A---A-CAAG--C--C-T--ACAGC--C--A------CC--CA-----A--A--C--A--C--CT-GC-G-- 2627
MUS-2.CM.01.CM1246             GTG--A-----A--TGCA-----C--GAC---G--C---------T--------GCCC-A--G........GCA.........CAT--C--TA--G-A---A--CAG-----C-T--A-AGT--C--A-----CA----A---C-ACCC--C--G--C--CT-GG-C-- 2735
MUS-2.CM.01.CM2500             GCC--A-----------CA-----C--GAC---G--C--G--C--AT-----C-TTCT--A--G........GCT.........CAC--C--CA--G-A--CA--CAA-----C-T--ACAGTA-C--AC-CT----G--G--T---CCT--C--A--C--CT-AG-T-- 2694
SYK.KE.x.KE51                  GC---A--G-----T--------A--G-----------T-----A-G----C-C-GT-CA----G-..................A-G------T-A--CT--AAA--C--C----ACA----A--TC----------G-----AC--------A---------AG-G-- 2799
SYK.KE.x.SYK173                GCC--A-----T--T-TC--G--A-----G--------C--C--A-G---AC-----CG-C-----..................C-----A-TA-A---T-AAAA-----------AC-----C---C-G--CA----G--T---CCA--C--G---------AG-G-- 3141
RCM.GA.x.GAB1                  -----A--G-----T-TT--G--G--GC-G--------------A-G---A-----GC-ATGTG-GAG-..................A----G---T-A--CA-C--G-----------------TG--TTC-GT----GCCT--T-----A-----------------A----- 2426
RCM.NG.x.NG411                 -----A-----T----TT--G--A--GC-G-----C--G-----A-G---A------C-ATGC-GGCA-..................A-T-----CT-A---A-T--G-----G---------TGC---C--T----GCCA--T--A-----C-----T--C--G---T 2441
GRV.ET.x.GRI_677               GCT-----------T-TT--G-----GC-G----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-----A-------------C-- 3120
SAB.SN.x.SABIC                 --T-----------TCA------C--G-----G-----C--C--A-----C---C-GC-A-GGG-GCAG..................A-T-----GT----CA-A--A-----C------AGCTGC--T------CC------TCA--A-----A-----C-----C-- 3345
TAN.UG.x.TAN1                  GC---A--G-------TT--G--A-------T--T--C-----A-----A--G----GA--G---CAG..................A-C---------A---A-A--A-----C----A-AGTA-A--A--G---CC---A------CCC-----A--T---------C-- 3149
VER.DE.x.AGM3                  GC---A-----T--T-TC-----A--G-----C-----T-----AT-------CG-------TG-CGGA-..................A--------T-A--CA-A--G-----C----A----A-A--A----CCC---G--T--A--------C--T---------T-- 2672
VER.KE.x.9063                  GT---A-----------TT------------G--------T--C--G--------GC-G--A-TG---CAG..................A--------T-A---A-A--G------T----A----CA-A--A-----CC---G--T--A--A--C--A--------A--C-- 3177
VER.KE.x.AGM155                GCT--A-----------TT-----C--------T-----C-----A-----------G--A-TG--GCA-..................A----CA-TA-A---------G----------A-AGCA-A--AC-G---CCT--G--T--A--A--C--A--T--C-----C-- 3169
VER.KE.x.TYO1_patent           GC---C-----T----TT-----G--------G-----C--C--A-----A----G-----TG-G-CAG..................A-------TT-A-----A--A--C--C----A----CA-A--A--G---CC-A-T--T-----A--------T-----T--T-- 3174
COL.CM.x.CGU1                  -----A---------AT------C--------G--------C--A-G---TC-T-T-G-A--G-CCCAT..................A--------T-A---A-A--A-----C-----C---A-------G---CCG--T-ATCAA-----C--A--T-----A--C-- 2722
```

```
H1B.FR.83.HXB2              TAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAG 3117
Pol                          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  Q  N  P  D  I  V  I  Y  Q  Y  M  D  D  L  Y  V
H1A1.UG.85.U455             ------------------------AG-C--G--------------------G---------------------T----------G--------------------------C------TC----C---------------------------------C--------G- 2563
H1B.US.90.WEAU160           ---------------A--------------------------------------------------------------------------------------A------------------------------------------------------------------- 3116
H1C.ET.86.ETH2220           -----C---------A--------------------A--T-----C--C------------------------C----------G---------C-CC-------------C-----GGCC-CC--C-----A------------------T--------C----------- 2509
H1D.CD.84.84ZR085           ---------------------------------------------A-------------------------------------------------T---------------C------------------AG------------------------------------- 2639
H1F1.BE.93.VI850            ----G-C--------------T--A-----G--C--------------------A------------------------------T-------------------------C------TGA---C-----------C-----C------------------------- 2454
H1G.SE.93.SE6165            ---------T---------------G-------------------------------------------------------T--G-------------G------------C------GC-A-T---------A--G--G-----C-----------------A------- 2514
H1H.CF.90.056               ---------------------------------------T--------------------------------------------G----------------------C--C-------G---------T--A--G-----T--C-------------------------- 2464
H1J.SE.93.SE7887            ---------------------------C--------------C-----A--------------------C-----------T--GT--------------------A-A--C------G--AG--C-----A------------C--G----------C---------G- 2431
H1K.CM.96.MP535             ------T--------A---------G----------------------A-----A-----------------------------CA--------------------------C------T-A----------A--G-----A--C-----------------A------- 2313
H101_AE.TH.90.CM240         ------------------------A--C--------------------G-----------------------G-----------G--------------------------C------T-A----------A-G-----------------A------C--------- 2691
H102_AG.NG.x.IBNG           ----G----T-----------------------C--------------------------------G-----G--------T--GGCA-----------------------C------C-A----------G-----G-----C-----------------A------- 2642
H103_AB.RU.97.KAL153_2      -----C------------------------------------------------------------------------T------------------------------------------------------G----------------------------------- 2343
H104_cpx.CY.94.CY032        -----CC-----------------A--------------------------------C---------------------------T-------------------------C------TTCA-----C-----A-----C--A--C------T------------------- 2483
H1N.CM.95.YBF30             ---------------------T--T--------C-----T--------G-----A--C--------G-----------T--T--G----CA-----------TC----A--A--C--G-GA--C-------G--A-C--T--C--G------------CC-C-----G- 2711
H1O.BE.87.ANT70             ----G-G--T--------C----A--A------C-----------C-C--G--A------------T--------T------------TCA-----------TC----T--A--C--G-G-G-C--C----AT---AA--T-G---G-----------CC-A------- 3172
H1O.CM.91.MVP5180           ----G-G-----------C----AG-A------C-----------C-C--G--A--G---------T--------C-----T--G--TCA--------G--TC----T--A--------AGC--C-----AG--AA--T-----G------A------C--A------- 3147
H1P.FR.06.RBF168            ----G----T--------------A--------------T-----T--G--T-----------R-----R-----C--C--Y--GTA-TCA-----T--G-----------A--C----GG-GAGT---------AG--AA---------------------C-A------- 2578
CPZ.CM.01.SIVcpzCAM13       ----G----T-----A--T--------C-----------T-----TT-A-----A-----------G--T----T--C--T----------T-----------TC-G--A--C--C--GC-A----C--T--T--CAC-G----C-------------C--A------- 2697
CPZ.CM.05.SIVcpzEK505       ----G----T-----A--T-T--------C------------------G--A--C-----------------------C--T--G--C-CT-----------A-----A--A-----G--GA-C------AC----C--A-----G------------C-C------- 2689
CPZ.CM.05.SIVcpzLB7         ---------------A--------A-C---------A--------T-A--G------------------------T-----T----------------G--T------A--C----G---G--C-----T---A-T-------G----------------A------- 2672
CPZ.CM.05.SIVcpzMB66        C---G-------------------A--------C---T----------A-G--A--G---------G--------------T----------------G---T----------C-------G--------A--A---------------T----------A------- 2663
CPZ.CM.05.SIVcpzMT145       ---C-----T-----A--T-----------------A-T--------A-----T--A--------G-----T--TT-C--C--T--G-------------G---------A--C------C-A-GC-----A--A--T--C--G-------------C--------G- 2659
CPZ.CM.-.SIVcpzMB897        ----G----T--------------A--A--G--C-----------A--C--G-----------------------T--T-------C--T------T--G--TC-G--A-----A-------G--------A------C-----C--G----------C-A------- 2683
CPZ.CM.-.SIVcpzDP943        ----G-------------------A-----------A-T------A--------A-----------G--------T--T-------C--------------------A--A-----C----G------GT-GACA-T--C--G--------------C-T------- 2703
CPZ.CM.98.CAM3              C---G----T-----A--------A-----------A-T------T-A------A-----------G--------T-T-------C---------T-----TC--ACC--A--C--C-G--GC----T--A-A--T--C--G--T------------C-T------G- 2528
CPZ.CM.98.CAM5              C---G----T--------------AG-------------T-----T-A--R--A------------G-----G--C--T-------C---------------------A----T--A-C---GC--------TC--A-A------C------T-------CC-T------- 2819
CPZ.GA.88.GAB1              C--------T--------C-----AG----------A-T------TT-G-----A-----------G-T--T-T-C----------C-T------------TC----A-C-C---G--A-G---T----TAC---T--C--G------------CC-A------G- 3178
CPZ.GA.88.GAB2              ----G----T-----A-T------T--A--G--C-----T-----T-A------A-----------G-----------C-------C----------GG-T-----C----C---C-GA--C------AT-G--A--T--C--------------C-A------- 2510
CPZ.US.85.CPZUS             ----G-------------C-----A--------C-----T-----T-A------A-----------G--------T--------G--C--T-----------TC-----T-------------C------TG-TA-A------------T------C-C------- 3177
CPZ.CD.90.ANT               A--CG-G-----------C----A--A----TGT--T--------T-A------A-C----------------T--T--GCA-------C--G--------CA-A-----GG-TA-GT-----CAG---AA--T-----------------------C-C------- 2556
CPZ.TZ.00.TAN1              C---C-C--T---A-C----------AA-----T-----T-C---T-A--T--A--T------G-----T-----T--T--G--C--T-----------------C----T-----C-----G--C-C---TG-G-AC--T-C-------T-------C-T--CA--- 2761
CPZ.TZ.2001.TAN2            ----C-G--T--CA-C--------A-AA--G-T------T-----T-A--T--A--T------G-----C-----T--T-----------T-----------T--G--T-----C---------G--C----TG--AC-----C--G--T---------C-T--CA--- 2766
CPZ.TZ.2002.TAN3_1          ----C-G--T--CA-C--------A-AA--G-T------T-----T-A--T--A--G------------C-----T--T--G--C-----------------T--G--T-----C---------G--C----TG--AC----------G----------CC-T--CA--- 2766
GOR.CM.2007.SIVgor2139_287  ----G-G--T---------------G-A--G--------------A--A--C-----------------T-----T--------GCACTC----------------AC-T--T-----C-------GA----C---G---AA-------G------------C---------- 2654
GOR.CM.2007.SIVgorCP2135con ----G-G--T---------------G-A--G--------------A--A--C-----------------T-----T--------GCACTC----------------AC-T--T-----C-------GA----C---A---AA-------G------------C---------- 2653
GOR.CM.2004.SIVgorCP684con  ----G-G--T---------------G-A--G--------------A--A--C-----------------T-----T--------GCACTC----------------AC-T--T-----C-------GA----C---A---AA-----C--------------C---------- 2573
H2A.DE.x.BEN                AGCAG--------AT-GA------A-AA-----ATA--T--A--CT-G------A--------G-----------T-T---TAC-CA---GGC--G---------A-------C-------GC--C-----TG-CA--C--ATC--G-----------A-C-TAA--- 3502
H2A.GW.x.ALI                ATCAG---------CAGA------A-AA-----ATA--T--A--CT-A--G------------G-----------T-T---CAC-CA---G-C-G-------------A-C---------GGC--C-AG--TG-CA--C--AT--------------A-C-TAA--- 3495
H2A.SN.x.ST                 ATCA----------CTGA------A-AA----CATA--T--A--CTCA---------------G-----------T-T--GTAC-CA---GGC-GG----------A-A-C---------GC--C--G--T--CA--C--AT---G-----------A-C-TGA--- 2947
H2B.CI.x.EHO                AGCAG----T----CAGA------A-AG--------TT-------A--C--A-----------G-----C-----T----GTAC-C---G---GG-AC----------C-------C-------GCC-CAAT--TG-CAC---AATC--G-----------CA-TCTC--G- 3473
H2B.GH.86.D205              ATCAG---------CAGAG-----A-AG-----CATT--T--A--T--A-----A--G-----G--G--C-AGT---T-GT---TACTC-----G---GG-A---------T-------------GGCC-CAGC--TG-CA---AAT--G-----------CA-CCT-A--- 3475
H2G.CI.92.ABT96             ATCAG---------CAGA------A-AA-----CATC--G--C--A--G--A-----------G--C--------T--C-----CA--CA---GG--GG-AC------C--C--C-------GCC---GAG--TR--A---YATC-----T-------CA-TCT----G- 2829
H2U.FR.96.12034             ATCC--T--T----CAGAG-----A-AA-----CATC---G--T-G--T--A--C-----------------A---T--T--GTA--CA---GG--CC-AC-G--A---------------GGC---GA--TGCT--C--AAT--G--T-----------CA-TCT-A--- 2984
SMM.SL.92.SL92B             ATCAG-G-----CC-AGA------A-AG--G--ATT---G-C--G---------A--------G-----T--T---------GGCA-C-----GGC-GG-A------A--C-C---------GCT--C-----TG-GC--C--GTC--G--T------------ACTAA-T- 2850
SMM.US.x.H9                 ATCGG---------CAGA------A-AA-----CATC--T--G--AT-A--T--A--G-----------------T-T--T--RCA--CTR---G---TG-A-----A-------------GC-------TG-GACCC-GATC--------------CA-C-TAA--- 2892
SMM.US.x.PGM53              ATCAG---------CAGA------A-AA-----CATC--T--G--AT-A--T--A--G--------G--------T-T--T--GCA--CT---G---TG------GA------C-------GC-------TG-GACCC-GATC--------------CA-CCTAA--- 3336
SMM.US.x.SIVsmH635F_L3      ATCAG----T----CAGA------A-AA-----CATC--T--G--AT-A--T--A--G-----G--G--------T-T--T--GTA--CT---G---TG-A-----A-------C-------GC-------TG-GACCC-GATC--------------CA-C-TAA--- 3422

MAC.US.x.239                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--GGC-----------TG-GACCT-AGTC--G--T-------CA-C-TAA--- 3419
Pol                           P  S  V  N  N  A  E  P  G  K  R  Y  I  Y  K  V  L  P  Q  G  W  K  G  S  P  A  I  F  Q  Y  T  M  R  H  V  L  E  P  F  R  K  A  N  P  D  V  T  L  V  Q  Y  M  D  D  I  L
MAC.US.x.251_1A11           ATCAG----T----CAGAG-----A-AAC----CATT--T--G--T--G--T--A--------G--G-------C--C-----TAC-CT---G-C-TG-GC---A--C--C--G--GGC-----------TG-GACCT-AGTC--G--T-------CA-C-TAA--- 3418
MAC.US.x.251_BK28           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--GGC-----------TG-GACCT-AGTC--G--T-------CA-C-TAA--- 3395
MAC.US.x.EMBL_3             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--GGC-----------TG-GACCT-AGTC--G--T-------CA-C-TAA--- 2893
STM.US.x.STM                ATCAG---------CAGA------A-AA-----CATC-----G--T--G--------------G-----------T--T---TA--CA---G---T--A--------A--C--G-G-GCC--C-------TG--AC-C-GATC----------------A-C-TGA--- 3066
MNE.US.x.MNE027             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--GGC-----------TG-GACCC-AGTC--G--T-------CA-C-TA---- 2887
DEB.CM.99.CM40              ATC-G-C--T---ATACA------AGAA------G---T--GA--A--G--T---------GC-CT---------C--T--T---GCCTCAG-AGGG-G-CAG---C--ATC-----GG--A-G--------GC--A-A--AGTC--G--T-------------ACTA---- 2926
DEB.CM.99.CM5               -TCGG-G------ACACA------AGAG------G---T--GA--A--A--T--A--C----GT-C---------C---------GGCA-CAG--GGC-GGCAGC-GA-A-T-----AGG--A--T-C--------T--A----AGTC-----T-----------CC-A-TG---- 2920
DEN.CD.x.CD1                GTCC--T--------AGG------ACAA--G--CATT-T---A--CT-A-----A-----GT-------G--TA-C--T--T--GGCC-CAGCAGAT-----GC-TAGC-AA-----G--GGCCC----T--TG-GA--T-AATC--G--T-----------C-C-TG---- 3004
LST.CD.88.447               ATC---T--T--CC--GA---------GA------------T-----A--A--C--A--------G--GAGT---TGC---T--T---GC---CAG-AG---GG-C--C--TCAGAA----------TT----C--T------A-CT-G-------------------A-TCA--- 2385
LST.CD.88.485               ATC---T--T--CC--GA---------GA------------T-----A--A--C--A--------G--GAGT---TGC---T--T---GC---CAG-AG---GG-C--C--TCAGAA----------TT----C--T------A-CC-G-------------------ACTCA--- 2385
LST.CD.88.524               GTCAC---------C--GA------A-AA--------------T------T-G-----A----------G--GAGT--TTGT--T--T---GAA-CAG-T-GTGG-C--C-T-CAGAG--C--G---TT----------------C-CC----C--G------------------TCA--- 2382
LST.KE.x.lho7               -TC-C-C--T---C-AGAG------A-AA--------------------G------A--G------G--GAGC--TTGT------T--GG-A-CAG-AG---GG-C-TC-CTCAGAG-----G---TT--------------GA-C---T--C-----T-------------A-TCA--- 3469
SUN.GA.98.L14               GTCAG----T----C--G-------A-AA-------GT---T------C--A--------C---------AGT--TTGC--C--T--GG-A-CGG-AG-G-GCC---------AAA---------G----GT-------AG-TCAAT-G--C-----------------ACTCA--- 3474
MND-1.GA.x.MNDGB1           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-----------GA-CC------ACAG--CAGT-A-------------------T----- 2804
OLC.CI.97.97CI12            ATCAG----T------T-G-------A-GA--G---ATA--T------T-A--T--A--------G--G-------TGT--C------G----CAG--G-TC-TT-GA-G---------GG--G--GG-------T--G-----A-T-A--C--G--T---------C--------G- 2991
WRC.CI.97.97CI14            ATCAG----T------T-T-------AG-----GTA--T-------A--G--T--------------------T--TTGT--T--------TCA-CTG-AG--TCTC-A-----AATA-------GG-A-G-------A-T---CAAG-TCT-------T----------C-A---A-T- 3417
WRC.CI.98.98CI04            ATCAG----T------AT-T--T----AG--G---GTC--T------A-----T--A--G-----G--T--C--TTGC--T------GCA-CAG-AG---GTT-G------TGTA--C-AG-----G-----TACT--GCAG--TCTC--G------------------A---A--- 3432
WRC.GM.05.Pbt_05GM_X02      -TC-------T---------A------T-CC-----GTA--T-----A--C--T--A-----------------G--T--TTGC------T----GCA-CAG--G-TTCTT-GC-G----ATC---------GA--C-----TA-AG-GCAG--AGTC-----T----------C--------- 2346
MND-2.CM.98.CM16            ATCGG-G--T----C--GGC------A--C--G--C-----T------C--A-------------------G--C--C-------C--T-----GGCA-CAGCAGAT----------C-----C--C---G-GAG-C--------TG------G-------------T----------C-C-T---T- 3118
MND-2.GA.x.M14              ATCGG------T----C-AGGC-----A--C--G--C-----T------T-------------------------C--T--G---------T---GCA-CGGCAGAC-----TC--A-AA-A--C-A-G-GG--T-C-------AG--T-A--T-----------------------C---T---G- 3045
MND-2.x.x.5440              ATCGG-G-----CC--GGC-----A--C-----C-----T------C--A---------------------------C--T-----------T----GGCA-CAGCAGAC-----TC--A--G-------AGG--A-GC--------AG-CT-A----------G--T----------C---T---G- 2678
DRL.x.x.FAO                 ATCGG-G--T----C--GGC------A--C--G---------T------C--C---------------------------C------T------C------GC---CAGCTGAC--G--------C-----A--C----G-GA------------TG-GA----A-------G----------------C---T----- 2684
MON.CM.99.L1                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---A--T-C------GG--AC-T-AATA--G-------------CC-ACTGA--- 2823
MON.NG.x.NG1                -TCAG----T-----CAGA-----CCG-G-----------T--GA---T-G--C-----R-----------C--A------------T-GTACTCAGCAT--G-GC-GC-RA-AGRAA-C-G-G-A--T---------GG-CACCC--ATC--G--------------C-ACTCA--- 1416
GSN.CM.99.CN166             -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--GG-G-----TG-G--AT-AATA--G-------------------ACTGA--- 2845
GSN.CM.99.CN71              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--GG-G-----TG-G--A---AATA--G----------------C----TGA--- 2827
TAL.CM.00.266               GTC-C-G--------T-GG-----A-GA--------A-T---G-----A--C--A---------G--T--G--CA-C--C------CA--C-TC-GATC--T-A--GAG-AAA--C--GG-CA-GT-C--TC-GG---CAC---T--------------CC-A-TG---- 3086
TAL.CM.01.8023              GTC-C-G--T-----C-G--------GGC-G--C-----T---G--A--C--C--A--G---------------CA-C-----T---TAC-CTTCAGATC-GC-G---AG-AAG-----GG--A--T-C---C-GG---CGC---TC--G------------CC-ACTAA--- 2584
MUS-1.CM.01.1085            -CAGG-------CTT-GG-----CA-GA--------A-T---A-----A--------C---------G--T--TA-C--C-----GTTC-CAGCAG-C---C-------AGAGA---AG---GC---C-----TG-CAC-C--ATC--------------CC-CCTGA--- 2807
MUS-1.CM.01.CM1239          CCAGG----T--CC--GG-----CA-GA--G-------T---A-----G--------------------C-CA-C--C-------TAC-CGGCAG-----T-A---C--GAAA-C--GG-GC-----------T--CACCC--ATC--G------------C-C-TAA--- 2797
MUS-2.CM.01.CM1246          -GCAG-G-----CAT-GG-----CT-AA----TC--A-TT--G---T-G--------G--------G--------CA-C--T--T---GCC-CA-CAG-C--GC-TC------AAG----G--GG-C-----AG---GCCT-A--AGTC-------------CC--CT-A--- 2905
MUS-2.CM.01.CM2500          ----G----T--CCTAGGG--C-C--AG--G-----A-T---G--C--G--C--A--G-----G--G--C--CA----T--T---GCA-CA-CTG------T-A--------AAA-C--G---G------AG--TGC-T-A--AGTA-----T---------CC-ACTCA--- 2864
SYK.KE.x.KE51               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--A-G---T----T---A-A--AGTA--G-----------CC-T-TG---- 2969
SYK.KE.x.SYK173             A--C------T----CCT-C--G--AGAA---------A-TT-CA---T-A-----------------------T-CA-T--C---------C-CA--A-ATC-G--AC--C-----A---------GA-GT--T----TC-CAC-C-AATA--G----------C--A-TGA-T- 3311
RCM.GA.x.GAB1               AGCAG----T----C-AGG------AG-G--G-----A--T---------G--T-----------G------T--C--C--C--T--GGCATCAGCC-AT--G-------C----A--------GG--G-G---------TG-CA-C--T--C--G------------C-C-T----- 2596
RCM.NG.x.NG411              AGC---T--T---C-AGG---T--C-------------A--T---------T-G--C--------G------T------T------CA-CAGCT-AC------T-----A--A-----G--G-G-----------TG----C---------G--T-------------T----- 2611
GRV.ET.x.GRI_677            CTCAG----T---ACAGGG---------A--G-------T----TGT--G--T------------------T--TA----T------G-A--CGGCAG-----C--T------GAGA-C-A--GG--C-C---T-GGT---AA--TGTC-----------C--------GGT-G- 3290
SAB.SN.x.SABIC              ATCAG-C--T---AG-GAG-----C--A-----------T--G------C---------G----------C----CA----T--T--G-CA-CAGCC-AC-----TC-GC--GAA-----GC--A-G--C------TG----A---A--------------------A---TAA--- 3515
TAN.UG.x.TAN1               ACAAG----T----TAGGG--T--A-----G-----A-T----TGT--A--T--A------------------CA-C-----T--GCAC-CAGCACAGCG-G--C--A-AGAAA-AG-----A-GGT--A---TT---AA--AAT---G------------CA---GGA--- 3319
VER.DE.x.AGM3               ATCAG----T---C-AGGG-----T-C------------T----CTGT---------A---------G------C----A-T--T--T--G-AC-CAGCAG-TTCC--TC----AGAAA-A-A---GG-GTTAAA-CC-C--ACC--TGTG---------------CC-A-GG---- 2842
VER.KE.x.9063               -TCAG----T---C-AGG------TG----G-----A-T----TGC---------A--C--------G--T---A-T--T--T--G-A--CAGCAG-------T------AAAA-A-A-G-GG--TTAAA-CCAT-GACA--AGTG--G-----------------GG---- 3347
VER.KE.x.AGM155             --CGG-------------GG------C--A---------A-TT---TGC--A--G-----C------G------C--GA----T--------AC-CAGCAT--------TC----AGAAA-A-AG---G--TTAAA-C-GC-GACG--TGTC--G--------------CC-C-GG---- 3339
VER.KE.x.TYO1_patent        C-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-A-CTTG----CAC--ACC--TGTA--------------------A-GG---- 3344
COL.CM.x.CGU1               ATCAG-G------TT-G---------C-------GT------A-----A--G-----T---------G--T---A----C-------G-A-CAG-AG---G----C----A--G--C--G--GAG-GGA-A-TTG...CAA--AGC---G---------------A---C--- 2889
```

```
H1B.FR.83.HXB2            GATCTGACTTAGAAATAGGGCAGCATAGAACAAAAATAGAGGAGCTGAGACAACATCTGTTGAGGTGGGGACTTACCACACCAGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTG 3287
Pol                       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  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
H1A1.UG.85.U455           -------T--------------A------G-------------AT-A---GCT-----A-----C------T-C-TT--C-----------G------------------T----------G--------T--------C-----------T---------CA---- 2733
H1B.US.90.WEAU160         -----------------------------------------------------------------------T-----------------------------A-----C-----------------------------------------------------AA---- 3286
H1C.ET.86.ETH2220         ----------------------A------G-CCC------A---T-A---G-------A--A-A-------T---------------G----------G--------------------T----------G-------T--------C------------------CA---- 2679
H1D.CD.84.84ZR085         -------T----------A--------G-----------A-AT-A---G-----------------G-----------------------------------------T-----------------------------G-----------T-----ACA--- 2809
H1F1.BE.93.VI850          -G------------------------------------T-A---G-------AC----A-----T----T-----------------------------------------------------GC---------------------G---------CAAT-- 2624
H1G.SE.93.SE6165          ---------------------------G-------------T-A---G-------A----AA-----GT-------------T------------------------------------------A----G-----------C-------G----A------CA---- 2684
H1H.CF.90.056             ----------------------A------G--------------T-A---GCT---T------AA------T------------------------------------------C-----T-----------A-----------------C---------------A--G--AAA--- 2634
H1J.SE.93.SE7887          ----------G-------AA---------G-------A---------GG-------A----A------T------------T-----------------T-----------A----G--T--------C-----------C--A------CAA--- 2601
H1K.CM.96.MP535           -G--------------------A-C--------------------A--A---G-------A----AA------T-------------------G-------------C-----T-----------A--------------------C-----------G---------CAA--- 2483
H101_AE.TH.90.CM240       -------T------------------C----------------------A---GCT-----A-----C-----T----T-----------------G------G---------------------A--------------------C-G-----------C-----------AA--- 2861
H102_AG.NG.x.IBNG         -------T-------------------G-----------T-----GG------AC---AA-----T-------------------G----------------T-----------------------------------------------C------G---AA--A 2812
H103_AB.RU.97.KAL153_2    -------------G------------------G----------A-----G--------C-----------T-------------------------------------------------------------------------------T-----------------T-- 2513
H104_cpx.CY.94.CY032      -G--------------------A------G-----------A-----A---G-G-----A-----A-----T-C----------------------------------C-----T----------G------------------C-----------G---------CAA-C- 2653
H1N.CM.95.YBF30           --------------C---CA--A------GAGGC-G----A--C--T---G-T-----T----A------CT----G--C--T-------------------G-----C--G-----C--------A------------A--C----------C-----A---AA-T-A 2881
H1O.BE.87.ANT70           ----A--T---CCCC-GACAG-A-------A--GG--T--ATT---T---G----C--A-ATCA-------T-C--T--C--T-------G-----A--G--------C--T-A--------G----G-------A--C-----------------T-C--CCAAT-- 3342
H1O.CM.91.MVP5180         ----A--T---CC-T-G-CAG-A-------AG-GGG-C--ATT---T--GG-----T-A-ATCA-------T----T--C--T--T-----G--------G-------C--TT-A--------A----G-----C--A--C--G---------------C--CCAAT-- 3317
H1P.FR.06.RBF168          -G--A--T--GCCTT-GTCTG-A------CAGCGGG-----A-A--T--GG----C--C-ATGTA-----GT----A--T--T-------------A--R--G-----T-----C--------A----G-------A--C--G-----G--G--A--C--CAAAT-A 2748
CPZ.CM.01.SIVcpzCAM13     -G-----TC-T--C--TACC--A--C---CAG--GG----A---T-A--G-------T--C-TCAC-----GT----A--C-----T-----G-----A--------A--C---T-A--------G----G--T--C--A----------------G-----A---AA---A 2867
CPZ.CM.05.SIVcpzEK505     -C-----TC---------ACA--A------GAGGC-G-G--AAGA--T--GAGT-----T----CT------T----A--C--T-------G--------------A----TT-G--------A--------------A-----------------C---A-A---CAAT-A 2859
CPZ.CM.05.SIVcpzLB7       -G--A--T--GA-------A---A--C--G-AC--GG----A----------GCT--CT--C-A-A------GT-------C--T----------------------G----A----T--C--------A--------------A--C-------------------C---CA-T-A 2842
CPZ.CM.05.SIVcpzMB66      -C----------A-------A--TA------GA----G-G--A--AT----GGCT--C--A----AA-----TT-C-----T--T-------G-----A--G--G--A-----T--C--------A--------------C--A--C------------T----------CA-T-A 2833
CPZ.CM.05.SIVcpzMT145     -C--A--T---CCT----A-G-A-------AG-G-G----A-----C---GC----------A-AA-------T----A-----------G--G-----------------C-----T--C--------A-----GT-A-----A-----G--------C---AA---AAAT-- 2829
CPZ.CM.-.SIVcpzMB897      -------T---C-T----A---A-------GA---GG-G--A--AT-A---GCT-----A--A-A-------CT----T------T-------G-----A--------A------------------A-----G--------A--C-------------C-----C---CA---- 2853
CPZ.CM.-.SIVcpzDP943      -C-----YC-----------AG-A-------A---GG-----TTA--T------------AC-C-CT-----TT----T--T--T--T--G--G-----A-----G--A--T--TT-G--------C-----G--T-----A------------------G--A--A---CA-T-A 2873
CPZ.CM.98.CAM3            -------T--G---C--A-AG-A-------GA---GG-----CTA--C------G-----AC-T-CT-----GT----T--C--T--T-----G--------G--G--A--T--TT-G--------C-----G----------------G-----------G-----A---CA-T-A 2698
CPZ.CM.98.CAM5            -------C-G--G-----AA-A--------GA---GG-----ATTA--C-----------C-T-CT-----TT----T--C--T-------G-------A--G--------A--C--TT-G--------A-----G--------A--C-----------------G-----A---CAAT-A 2989
CPZ.GA.88.GAB1            -G-----TC-T-----T-AT--A-------A---GG-G--A--A--A----------T--C-T-AA------GT-C--A--C---------------------A--G--G--A--C--TT-A--------A-----G------C--A--C---------------C------TCAAT-A 3348
CPZ.GA.88.GAB2            ----A--T---TCTT-G-AAG-A--C--G-A--G-G----AC-A--C---G-----T-AC-C--A------TT-------C--T-----G--G--C--A--G--A------T--T--C--------A-----G--T-----A--C-----------------G-----A---CA---- 2680
CPZ.US.85.CPZUS           -G--A--TC--A-CT---AAA-------GGA---GG----ACT---C------T--T--C-T-CT-------T-C--T--C--T--T--G--G------G-G-----A-----------T-A--------C-----------------A------G---------G-----C---CA-T-A 3347
CPZ.CD.90.ANT             -------TA-G------TACTGCA------GA--TG-----AA----T---------T-ACA-GTC--------AGAG--T--T-------G-------------------------T-AA--------A-----GT-A-----A--C-------T------AAA---AA--A 2726
CPZ.TZ.00.TAN1            -T--A--TC-TA-TGA--A-G-A------G-A-CTG---A--A-----------G---------A-CA-------T-AGAG--C--T-------GT------G--A-----------A-G--------C-----G--A-----AA-----------------T--AAA---CACAT-A 2931
CPZ.TZ.2001.TAN2          -T--A---C-CAGTGA--AAG---------A-CT----A--A-----------G---------AGTA-------T-AGAG--C--T--T-----T-C----G--A-----------A-G--------------G--G-----AC-------------------G--AAAC--CACA--- 2936
CPZ.TZ.2002.TAN3_1        -T--A---C-CAGTGA--AAG-A------G-A-TTG---A--A-A---------G---T-----ATTA-------T-AGAG--C--T--T-----GT-C----G--A-----------A-G--------C-----G--A-----AA-----------------G--AAAC--CACA--- 2936
GOR.CM.2007.SIVgor2139_287 -G--A--T---CCT----ACAG-A------GCAG--GG----AA-A--C--GG------T-A-ATGCA------T-C-----T--T-------G-----A--G--------C---TT-G--------G-----G--G--C--------G-----------G---------AA-T-A 2824
GOR.CM.2007.SIVgorCP2135con -G--A--T---CCC----ACAG-A------GCAG---G----AA-A--C--GG------T-A-ATGCA------T-C-----T--C-------G-----A--G--------C---TT-G--------G-----G--G--C--------G-----------G--A------AA-T-A 2823
GOR.CM.2004.SIVgorCP684con -G--A--T---CCC----ACAG-A------GCAG---G----AA-A--C--GG------T-A-ATGCA------T-C-----T--G-------G--G--A--G--------C---TT-G--------G-----G--G--C--------------------G---------AA-T-A 2743
H2A.DE.x.BEN              CTAG----AGGAC-GGTTTAG-----GAC-A-GTGG-CCT-C----A-A-G---T---AAATG-CCTA--GT--T-T--T-----TG-G--GTTC--A-----C---------T-AA--------C-G------ATGG--AAC--------AGC-G---AAAC--CAA--- 3672
H2A.GW.x.ALI              CTAG----AGGAC-GATTTAG-A---GAC-AGGTGG-CCT-C--T-A-AGG---TC--AAATG-CCTA---T--T--------C-----TG-G--GTTC--A-----C-------A-AAA--------C----G---GTGG--AAC--------AGC-G--AAAA---CAAT-- 3665
H2A.SN.x.ST               CCAGC---AGGAC-GATTTAG-A---GAC-G-GTGG-TCT-C----A-AGG---T---AAATG-CCT----T--T--------C-----TG-G--GTTC--A-----C--------A--AA--------C--------GTGG--AAC--------AGC-G---AAGA---CAAT-- 3117
H2B.CI.x.EHO              C-AG----AGGAGCGATCT-G-----GAC-GGGT-G-GTCTC-A--A-A-G-G-TAT-AAAT-ACAT---T-CT-T--T-----AG----GTTC--A-----C-------------AAA--------G-----G---TGG--AA-G-------A-C-G---AAAA---CA---A 3643
H2B.GH.86.D205            C-AG----AG-AGTGATCT-G-----CGAC-GGGT-G-GTCCC-A--A-A-G-GTTAT-AAATGACAT---T-CT-T--C-----AG----GTTC--A-----C-----G---AAA--------------G---TGG--AA-A--G----A-C-G---AAAA---CAA--- 3645
H2G.CI.92.ABT96           CCAG----AG-ACTGATTTAG-A---GATCA-GT-G-TTT-C----A-AGG---TCT-AAAT-AAACR--GT-CT-T--C--T--TG-----GTTC--A-----C---------AT-AA--------A-----T-RTGG--AA-A--G----A-C----AAAG--TAAC--- 2999
H2U.FR.96.12034           CCAG----TAG-A-TGATTTAG-----GATCAGTT-G-TTCAC----C-AGG-G-TA---AAT-A-ACA----T---G---C----------G-----TT---------C--A---CA-AAA--------C--------ATGG---A-A--G----A-T-G---AAA--T-A---C 3154
SMM.SL.92.SL92B           -CAG-A--AG--GCC--ACAG-----GAC-A--TGG--ACAC--T-A---G-CATG--CAAC-ATCT---GT-C-GT--C-----AG-C--GTT---A--GA-T--A-----A-AG--------A---TTGT-AT-C--CA-G--G----A-T-----AAA----AA--C 3020
SMM.US.x.H9               CTAG---TAG-AC-GATTTAG-A---GAC-GGGT-G-TTTAC-AT-A-AGG---T----AAC--TATA---T--T--------C-----AG-G--GTTC--A-----T-C---------AG--------A-----T-GTGG--CACC--------A-C-G---AAAA---A-T-- 3062
SMM.US.x.PGM53            CTAG---TAG-AC-GATTTAG-----GAC-GGGT-G-TTTAC--T-A-AGG---T----AAC--CATA---T--T--------C-----AG-G--GTTC--A-----T-C---------AG--------A-----T-GTGG--AACC--------A-C-G---AAA----A-T-- 3506
SMM.US.x.SIVsmH635F_L3    CTAG---TAG-AC-GATTTAG-----GAC-GGGT-G-TTTAC--T-A-AGG---T----AACG-CATA---T-CT-------C-----AG-G--GTTC--------T-C---------AG--------A-----T-GTGG--AACC--------A-C-G---AAA----A-T-- 3592

MAC.US.x.239              CTAG----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC--TCA-AGG---TCT--AAT--CATA--GT--T-T--C-----AG-G---TTC--A-----T-C-------T-AA---------G--C---T-GTGG--AACA--------AGT-G--AAAG----A-T-- 3589
Pol                       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  P  E  E  K  F  Q  K  D  P  P  F  Q  W  M  G  Y  E  L  W  P  T  K  W  K  L  Q  K  I  E
MAC.US.x.251_1A11         CTAG----AGGAC--ACCT-G-A---GAC-GGGT-G-TTTAC----A-AGG---TCT-AAAT--CATA--GT--T-T--C-----AG-G---TTC--A-----T-C-------T-AA--------G--C---T-GTGG--AACA--------AGT-G--AAAG----A-T-- 3588
MAC.US.x.251_BK28         CTAG----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC----A-AGG---TCT-AAAT--CATA--GT-CT-T--C-----AG-G---TTC--A-----T-C-------T-AA--------G--C---T-GTGG--GACA--------AGT-G--AAAG----A-T-- 3565
MAC.US.x.EMBL_3           CTAG----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC----A-AGG---TCT-AAAT--CATA--GT-CT-T--C-----AG-G---TTC--A-----T-C-------T-AA--------G--C---T-GTGG--GACA--------AGT-G--AAAG----A-T-- 3063
STM.US.x.STM              CCAG----AG-AC-GATCTAG-----GAC-GGGT-G-TTTAC--T---AGG---TC--AAAT-AT-TA---T--T-----C--T-AG-G---TTC--------G--T----------T-AA--------A-----GT-ATGG--AACA--G----A-C-G---AAAA----AAT-A 3236
MNE.US.x.MNE027           CTAG----AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC--T-A-AGG---TCT-AAAT--CATA--GT--T-T--C-----AG-G--GTTC--A-----T-C-------T-AA--------A-----T-ATGG--AACA--------AGT-G--AAAG----A-T-- 3057
DEB.CM.99.CM40            -G--A--T---AGG-A--AT--A--CCT--AG--GG----CAC-AT-A-----GTTC--CC-AGA-A------T-G-GG------CCTG-G--GT------G----C--A--C--T-AC--------G-----GT-A--C----A-A------CGGT-GAT---AG-TACC--A 3096
DEB.CM.99.CM5             ----AA-T-----T-AG-TCTCA---TT-GA-C-GG--A-AC----T--GA--TT---CC-AGAAAA---GT-A-GG------CCCGG-G---T-C----G-G--T--C----A--A-----------G-----GT-A------AA-G------CACT-GTT---A-----A---A 3090
DEN.CD.x.CD1              C---A--TAG-TCCC---AA-----C--G-A--TGG-TA----A--A---G-CTTCT--GCAGTA------CT--GAG--C--T--TGCC---TT---AG--AC---A--CA-A-AA--------G-----T-A--C--CA-AC-C---CA-A------AAA----AC--- 3174
LST.CD.88.447             ----A--TAGG--C-AGAAAGG----GATCAGGC-G--A-A--AT-A-----G-T-TCT-AG---CA---AATT-A-AA---------G--G---TT---AGC---------C-AT-A-----------G-----T----T-----A----GG-----GA-A--G--AAA-CTAAA--- 2555
LST.CD.88.485             -G--A--TAGG--C-AGAAAGG----GATCAGGC-G--A-A--AT-A---G-G-TCT-AG---CA---AATT-AGAA--------A--G---TT---AGC---G------C-AT-A-----------G-----T----T-----A----GG-----GA-A--G--AAAG-TAAA--- 2555
LST.CD.88.524             -T--A---AGG--G-A--AAGGA------TCAGGC-G--A-A--AT-A---G-----T-T-AA-A-CT---AACT-AGAA--T-----G--G---TT---AG-CA-----A--C-A---A-----------G-----TC--G--C--A-----GG-----GA-A--G--AGAG-AAA--- 2552
LST.KE.x.lho7             ----A--TAG---G-G-AAAGGA---GATCAGGC-G--A-A--A--C---G---T---TA---CA---AACT-AGAG--------A--G--GTTC--AGC----G--A--C-AT-A-----------------T----G------------GG-----GA-A--G--AAAG--TAAAT-A 3639
SUN.GA.98.L14             -------T-AT--G-A-AA-A-----GA-GA--T-G-GA-AC-AT-A---G-G-TA--CA--GAA----AATT-AGAA--------G------T------GG---G--A--C-ATAAG--------G-CAT-T-A--C--A--C-G----GAGA--G--AAAG--AAAT-A 3644
MND-1.GA.x.MNDGB1         -G--A----ATAC-GC--AAG-----GAG-A-GCT----TA--AT-A--GGCTTTAT-AA---CA---AACT-AGAA------T--A--G---T--------------C--T-A--------------G-----GT-A--C--A-A------G-----AGA--G-AAAGG-TCAA--A 2974
OLC.CI.97.97CI12          ----A-----GCC--G--CAG-----TT-GAG---G--A-AC--------GG-G-TCT--AG---T-ATAAC--AGAA--T--TCCA-----GT-----G-G--T--A--T-ATGAG--------G---AC-T-AT-C--A-G-G----AGA--GCAAGA---AA-T-- 3161
WRC.CI.97.97CI14          ----A-----G--C-AGACAG-A---CT--AC--GG----A--AT-A---A-G-TA-------CA----AATT-GGAG--T-----G-----G-TA---TT----------C-GTAAC--------A---AT-T-G-----A-----G----CA-A-----GAG---AA-T-- 3587
WRC.CI.98.98CI04          --AG--------A--CCAAG------TT-GAG---G----A---T-A-------TG------CT---AATT-GGAA--T-----A--G--G---TTG--ATT---G--G--C-G-AAG--------------ATT--A-----------G----CA-A-----GAA---AAA--A 3602
WRC.GM.05.Pbt_05GM_X02    ----A--T-AT----AGAAAG-A---TTGCA--TG--TC-----T-----G-T-T---T--A-CT----AAC--GGAG--T-----A--G--G-TA--ATT-----------GTAAA--------G---ATCT-G-----A--------G-----GA-A-----GAG---AAAT-A 2516
MND-2.CM.98.CM16          -GAG---TAG--TTGCCCCAG-A--C--C-G--TG--TC-A---T-A---G-C--C--C---TTT-----G--CGAG--C--------G-GTT---A--G-----A--C--TGAG--------A-CAT---G--C---A-G------------G---AAAG--CA---T 3288
MND-2.GA.x.M14            --AG---TC-GAGTGCGTCAG-----C-AC-A--TGG--A-TAG-T----GG-----T-AAG-TTC-----G--CGAG--C----------GTTC--A----------T--TGAG--------G----TG--G--C--AC-G------------C---AAAG--CA---- 3215
MND-2.x.x.5440            -GAG---TC--AGTGCCTCTG-A--C--T-A--TG----TA--AT-A---GG-----T--CA-TTC-----G--CGAG--C----------GTTC--A--G-------C--TGAA--------A----TG--G--C--AA-G------------T---AAA---CA---A 2848
DRL.x.x.FAO               --AG---TAG-CC--A-CATTTA---GAGCAG-TG--TA-A--AT-A---AC----T-AAA-TTC-----G--CGAG--C--G--------TT----G-CA-T-----T-ATGAA--------G--CC-G--G----AA-G--G-------G---AGAA--TCGCT-- 2854
MON.CM.99.L1              ----AA-TCACTCTC-GCCAG-A--C--G-A--T-G---G--TCA-A----GCGACC--AC-A---AAA--GA--CAA-----CCCAG-G---TT---G-T--CTAC--GG-G-AA---T-A--G-----G--A--C--ATCCC-T-----A-A-TGCCAAA---A---A 2993
MON.NG.x.NG1              -T--A---CATTCCC--CCTG-A--C--G-AG-T-G--ACA-CCA-T---ACG-TR--A-CAGAAAA---GA--CAA-----ATAG-G---T-----G-CA-GTAC-----G-AG---T--G--------------C--CACCC-T-------C--AAAG-----AA--- 1586
GSN.CM.99.CN166           ----A--TAGGA--C-G--AG-A--C--GCA-GT-G-GC-CA-AA-A---A-TTTA--AACATCC-ATAACA-CCAA--C--C-GGCC------C-ACC---TTA---GG--AAG---C-C--G-------GT-G------A-GGG------GGC-C-----GG-G-AA--C 3015
GSN.CM.99.CN71            ----A--TAGGA--TC-G-CAG-------GCA-GT-G-GC-CA-AA-A---A-TTTGT--AC-TCC-ATAACA-CCAG--C--C-GGCC------C-AC----TTA---G--AGA---C-C--G-------GT-G--C---A-GGG------GGC-C--A--AG-G-A---C 2997
TAL.CM.00.266             --AGCA-TGAGAGCT-G-A-G-A--------GG-T-G---GA--A--A---GCT-TCT-AAAC-A-GTA--GA-CCAG--C-----GGCT----T-C--AG-CA----G--CA-TAAA---C-A--G--------T-GT----A-----G-----A------------T-AAT-- 3256
TAL.CM.01.8023            --AG---TGA-AGCT---AA----------G-G--T-G-G-GA----T---GCC-TC--TAAC-A-GTA--GA-CCAG------G--GGCC---TTC---G-CA-G----TA-TAAG---C-T--G--C--G--GT-C--A--G--------A------AAA-----A---A 2754
MUS-1.CM.01.1085          ----CA--AG----T--TCAAGC--C--G-A-GT-G---CAC--A-T---A-CATG--C--A-A-CTT--GA--CAA--C--TC-AG-----TT-----G---CTAC----G---AA---T-A--G-----G--A--C--CCGGGGC---A-A------CG----CAT-- 2977
MUS-1.CM.01.CM1239        ----CA--AG---G---TC-GGA-----G-G-GT-G---CAC--A-A---GA-CATG--C----AACT---GA-ACAA--C--G--AG-T--GT-C--------CTAC---G---AG---C-A--G-----GT-A--C--CAGAGGG---A-A------AG--ACA--C 2967
MUS-2.CM.01.CM1246        -G--A---AG----------AG-A-------GGCT-G-GA-AA--A-C---G-T-TC--C-CA--TAA---GA-ACAG--C-----AG-C-----C---CC---CTAC--GG-AGAA---C-A--G--C--GT-A-----GA-AGG-----G-A-CA-A--AG-G-A---- 3075
MUS-2.CM.01.CM2500        -G--A---AG-----------G-A------G-GGCTGG-CA-AAGAG-T---C-G-C-TC--AACA--CAAA--GA-ACAA--C-----AG---G--------C---CTAC---G-AGAG---C-A--G-----GT-A--C--A-GGGGT----G-A-C-G----AG---AA--A 3034
SYK.KE.x.KE51             -G---A--AG-TC-GA-AA-G-A---G-TCA--TTG--AGT--AA-TGT-A--AG---A--AGCAGTA----T-CT-T-TT---CCAG-----TGG---G-CA--TTC--C--G-AA---T-A--G--CACG--A--C--------G----GTC----AAAGG-TCAA--A 3139
SYK.KE.x.SYK173           -CA----TAG-TC-GAGAA-GCA---CAGGAG-T-G--C-AC-AA-AGTGAC-GCA--A--A-AAGT----T---AGGTC---A-AG------TGG---G-CC--TA----A-G-AG---T-A------CACT--T-----A--C--G----CAGT-G---AAA---AA--C 3481
RCM.GA.x.GAB1             -C--A--TAG-AC--AGTT-G-A---GAC-A--TG--CA-AC-AT-A---G-T------AC--TTC------CT--GAG--C-----------TT-----G-TA--------------ATT-G--------G-----G-G--C--AA-A-GT------------GAG--CAA-T-A 2766
RCM.NG.x.NG411            -G--A--TAGGAC-GC-CATG-A-----TCAG-TG--CA-AAG---C---G-G------C---CC------TT--GAG--T-----T--G---TT-----G-CA-G-----C--TGAA--------C-----TG--------A--C---------------AGAG--CAA-T-A 2781
GRV.ET.x.GRI_677          CG--A---CAT--TGAGACTAGA----ATCA-C-GG-----CATAG-A--A-GATG---C-AGAAAAA--T--AGAA--C----------G---GTC--A-G------G----GGGAA--------G---A--T-G-----GA-------------CA-TA-CAAA----AAT-A 3460
SAB.SN.x.SAB1C            CCAG---TAGGCC--AG-CAG-A---CT-GT--TGG--C--C--T-A---G-CT----AGAA-CT----GT---AA--C--T-A--G-AG---GTT---A--G--T-A-----A---C---------T-GT----AA-G-----CA-C-G---GAG--CACC--A 3685
TAN.UG.x.TAN1             ----A--T--GCC-GAGAAAG-A---GATGA---GG-----A---G-A---A-G-T----C-ATTAA-A--CT-CGAG--G-----T--G--GGTG-------AG---------ATGA------------A-CA-GT-GT-C---TC-C-C----GTC-CA-CAAA----A-A-A 3489
VER.DE.x.AGM3             -G---C-GGA---TGA-TACAC----GATCGGTTGG----AC-A--A---ATGA-AT-AAGTGCC------T-AGAA---------------G---GTA-------A----A--T-ATGAG--------A--CA--T-GTGG-AC-C--G----CAGA--AGCAGC---AAT-A 3012
VER.KE.x.9063             -G---C-AGA---TGA-TACACC---GAC-AGTT-G----AC-A----T---GTCA-G--AAAA-CA----T-AGAA--C-----------GGTG--A--GA-G--A---ATGAA--------A--CA--T-ATGG-AC-A------G----CA-C--A-TAG----AACT-A 3517
VER.KE.x.AGM155           ----AC-AGA---GGGTCCAA-----GATCAGCT-G--C-AACA--T--GA-TAGAT--CAAGAA------T-AGAA--------G------GGTG--A-G-----------C--TGAG--------A---A--T-ATGG---C--------G-----AGT----AAG-----AAT-A 3509
VER.KE.x.TYO1_patent      -T---C-AGA-A-TGA-CACACC---GAC-A-TT-G----AC--T-A---AC-A-AT-ACAAGCC-----CT-AGAA--C-----A--G---GGTG--A-G---------A---T-ATGAG--------A--CA----TTGG---C-C------GA-C---AGCAGA---CAA--- 3514
COL.CM.x.CGU1             ----A--TC-GCCTT---A-G---------GGGC-G--C-------C---G-TGCA--C---TAC-AT---GT-AGAG--------A--G---T-C---GC---T------A-AAG--------C-------T-A------A-GC------A-T----AAAGG-CAAAA-T 3059
```

```
H1B.FR.83.HXB2                CCAGAAAAAGACAGC.....TGGACTGTCAATGACATACAGAAGTTAGTGGGGAAATTGAATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAAACTCCTTAGAGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAGCT 3451
Pol                           _P__E__K__D__S__.__._W__T__V__N__D__I__Q__K__L__V__G__K__L__N__W__A__S__Q__I__Y__P__G__I__K__V__R__Q__L__C__K__L__L__R__G__T__K__A__L__T__E__V__I__P__L__T__E__E__A__E__L
H1A1.UG.85.U455               ---------------.....-------------T--------A-----A--A---C-A------------------TG----------------A----C-G--------T--C--G---G---------------TA--G--A-C--G--T-----------AT- 2897
H1B.US.90.WEAU160             -----------A--T.....---------------------------A-----------------------------G----------------A----C------------------G--G---------------A---------A-------------------- 3450
H1C.ET.86.ETH2220             -------G--T---.....-------------T-----A----------A-----A--C-------------------------------------C-G-----------------G---G---------------CA--G--A-------T-----------AT- 2843
H1D.CD.84.84ZR085             -----G-----A---.....------------T-----------------A-----A----------C--------T-----A----------A-------------------------G--------G---------G-------------------------AT- 2973
H1F1.BE.93.VI850              ---A-C--G-------.....----------------------A--A---C-A-------------------T-------------------------C-------------------G---G---G----------CA--G-G-----G--T-C---G-------T- 2788
H1G.SE.93.SE6165              -----C--G--A---.....-------------T-----A----------A---C-A-------------------T-------------------C---C----------------G--GG--------------CA--G--T----G----C---------AA- 2848
H1H.CF.90.056                 ---------------.....-------------T----------------A---C-A-------------------T---AAT-----------A----C------------------G--GG-------T-------CA--------------G---A----G-----AT- 2798
H1J.SE.93.SE7887              -----------AGA-.....-------------T-----A----------A---C-A-------------------T-----A----------A----C----------------A---GG-T-------------CA--G-----T-G--TAG---------AT- 2765
H1K.CM.96.MP535               -----C--G--T---.....-------------T----------------A-----A-------------A--A--T----------------AA----G-------------------G---GT-------------CA--G-----T---T-C-----------T- 2647
H101_AE.TH.90.CM240           ---------------.....-------------T--------A-------A---C-A-------------A-----TG---------G----A----C-G---------C--G---G-T--------------CA--G----------G--T-------------AT- 3025
H102_AG.NG.x.IBNG             ---------------.....-------------T--------A-------A---C-A-------------------TG---A------A---A----C-G---G------C--G---G-----------------TA--G--G----G--T--G--------AT- 2976
H103_AB.RU.97.KAL153_2        ---------------.....----------------------C-------A-------------------------TG-----------------------------------------G---G--------------------------C-------------- 2677
H104_cpx.CY.94.CY032          G-------G--T---.....-----------C--T--C------------A---C-A-------------------T-----------------A---------------T----G---G-T----C--------CA--G-----------TAC---G-------T- 2817
H1N.CM.95.YBF30               --------G--TGTA.....-------------T--------A-----A--A--G--A-------------C--T-----A--C-G----AA--GC-C------T-AA-C-------------TT-G----------G-CAACT-T--------------AT- 3045
H1O.BE.87.ANT70               --TA-C--G--TGTG.....-----A--A-----T-----A--AC--A-A--A--GC-A-------------A--C--T-A---A----G---G--AG----G-----GT-AA-------C-----GT--T-------------G-----TT---GTAG---G--------- 3506
H1O.CM.91.MVP5180             --T--C-----AGTG.....-----A--A-----T-----A--A------A--A----A-------------A--C--T-A---A----G-----AAG----G--C--GT-AA-C-------------T--T-G------G----G-----TT---GTA----G------A-- 3481
H1P.FR.06.RBF168              --------G--AGCT.....-----A--A--------C------C-----A--A----A--------------------T-----A--A-GG----AAG-------C------AA-------------T---T-CT---------G-TG-T--T--WAG---------AT- 2912
CPZ.CM.01.SIVcpzCAM13         --------G--GGTA.....-----A---------TC-T-----AC-CA-A--A----A--C------------------T-----A--A---A---AA--G--G--C---T-AA-A--------A---AGGT-------T---G-T--T--C---CC---------AT- 3031
CPZ.CM.05.SIVcpzEK505         -----G--------CA.....-------------T------AC-------A--A----A--C-----T--C-----C--------------------AA---C-T--C---T-AA------------G-A------T-----------G-CA-----T---AG---------AT- 3023
CPZ.CM.05.SIVcpzLB7           --------G--A--T.....--------------------A--G-----A-------A----------------A--A---------C--C--GA---AA--GA----C-GG--T-----G---GTA--GAATTC----------G-CA-CT-T--TAG---------AT- 3006
CPZ.CM.05.SIVcpzMB66          -----------A---.....-------------T--C-------A---A-A--A----A----------T------A--T-----A--C-G----AA---C-G--C--G---A-----------A------TT-----------TG--A-CT-T---AC---------AT- 2997
CPZ.CM.05.SIVcpzMT145         ------C-G--AGTG.....--------T-----T--T-------------A--C-----A-------------A--A--T-----A-------A-T-A---GC-T-------T-AA----G---G-G--G-GT--G----------G--G-T--C--CCC------T--A-- 2993
CPZ.CM.-.SIVcpzMB897          -----M---------.....---------------------A------A--A--G--A-------------A--A--T-----A------G----AA-----G-----G--TA-----------A-----TT-------G----G-TR-C--C---C-G---------AT- 3017
CPZ.CM.-.SIVcpzDP943          ---C-------G-C-.....-------------T--C--A---C-G--A--A----A-------------A--A--T----------A--------AA--GC----C--G----A-A-A-----G-A------TT----T-------G-GG-CA-G---C-G---------AA- 3037
CPZ.CM.98.CAM3                ---C-G-----G-C-.....-------------T--T-----AC-------A---C-A-------------A--A--T-----A--A-------AA--G-----C--G--GA-A-A---G--A------T----------CA--G-GA-CA-G---C-G---------AA- 2862
CPZ.CM.98.CAM5                ---C-G-----G-T-.....-------------T--T-----A------A--A----A-------------A-----TT----------A-------AA--G-------C---------A-A-A---G--A------CT-G--T-------G--G-CA-G---C-G---------AA- 3153
CPZ.GA.88.GAB1                -----------GGTA.....-------------T-----A--AC-GA-A--A--G--A-----------------------T-----A--A--A--A---------------G-GA-A----------A--GAA---G------T----G-T--T--C----CC---------AT- 3512
CPZ.GA.88.GAB2                --TC-G-----G-AT.....----------A-------T-------AC-G--A--T------A---------------C--T-----A------GA---AA------G------T-AA-C--------G-T---G--TT-----------A--G-CGAGA-G--TAG---------AT- 2844
CPZ.US.85.CPZUS               ---C-------A-TT.....-----A---------T--T-------A-----A------C----C--------------A--T-----A--A--A---A---G------C---T-AA-A-A----G-T------T------T-------TG--AATT-T---C-T--G------AA- 3511
CPZ.CD.90.ANT                 -----GCC---TGAT.....-----A--T------------C-----AC-----A--A------A------------------C----------A--C---ACT-A---G--G-----G----A-C------GT---GAGT---------TAG-G---A-A-G--TAGG---------AT- 2890
CPZ.TZ.00.TAN1                ------CC---GCAG.....-----A--G---C-T--C---------G--A--C---C-T---------C------A-----T-AT--A--A---ACT-AAG--C-----C---T-GA---------GTA---G-T----T--GCC-G--GA-A-G--CAGG---------AT- 3095
CPZ.TZ.2001.TAN2              ------CC---ACAG.....-----A--G---C----Y-----A---------C--G--A----------C------A-----T-AT--M------ACT-AAG--C-G--C--G--AA-Y-------GTAR-G-G-T-G--T--CCC-G--GA---G--CAG---------AT- 3100
CPZ.TZ.2002.TAN3_1            ------CC---ACAG.....-----A--G---C----C-----A---------C--G--A----------C------A-----T-AT--A--A---ACT-AAG-GC-G--C--G--AA----------GTA----G-T-G--T--CCC-G--GA-A-G--CAGG---------AT- 3100
GOR.CM.2007.SIVgor2139_287    --------G--A--T.....-----A---------T--C----GA--G--A--T---C-A----------T------A--A--T---------A-G-----AAG--C-C--C---T-AA----G--T--T---T----T---------G-GA-C--T-GTA----------AT- 2988
GOR.CM.2007.SIVgorCP2135con   --------G--A--T.....-----A---------T--C-----AC-G--A-------A----------C------A--A--T-----A--A-G-----AAG--C-C--C---T-AA----G--T--T---T----T---------G-GA-C--T-GTA-----G-----AT- 2987
GOR.CM.2004.SIVgorCP684con    --------G--A--T.....-----A---------T--C-----A--G--A--C------A----------T------A--A--T-----A--A-G-----AAG--C-C--C---T-AA-C--G--T--T---T----T---------G-GA-C--T-GTA----------A-- 2907
H2A.DE.x.BEN                  --CC-G-------TA.....-----A-----------C--A---C------AGTC--A--------GGCA--A--C--TT-----A--A---ACC-AA--C---------G---AA-----------AA-TGA----C---------A-G-G-AGTGG-------CT------AGC 3836
H2A.GW.x.ALI                  --CC-G-----AGTA.....-----A-----------C--A--ACAT-----TGTCC-A----------GCA--A--C----------A--A---GACC-AA--C---------G---AA-----------AA-TGA----C---------G-G-G-AGTGG-------CT------AGC 3829
H2A.SN.x.ST                   --CC----G--AGTA.....-----A-----------C--A--AC-G-----TGTCC-A----------GCA--A--C-------------A--GACC--AA-C-------GGT-AA-C---------AA-TGA----C---------AGG---AGTGG-------TT------AGC 3281
H2B.CI.x.EHO                  -----------GGTT.....-----A--A---------T--------------AGT---A----------GC---AC---T---G--------GACC------TA-----------AA-A--G----AG-TGA-C------------AGG---A-TGG---T---TTG------GC 3807
H2B.GH.86.D205                -----------AGTT.....-----A--G-----CA--T--A--AC-G--A--AGT---A--C-------GC---AC-C-TT--T--A------GAC-------CA-----C------AA------G----AG-TGA-C------------A-G---AGTGG-------CT------AGC 3809
H2G.CI.92.ABT96               --M--------G-CT.....-----A--A------T--T--A--AC-----A---GT-C-----Y-----GC-----C-A--T-----A--C---AC--AAA-CC-G--Y---T-AA--------T-AR-TGA-CT-G--G----A-G---AGTGG-------TTG-----AGC 3163
H2U.FR.96.12034               --------G---TGTT.....-----A--A------------T-----------A--AGT-C-A--------GCA--AG----T-----A--A-GAC--AA--TC-T--------T-AA---AG--G-AA-TGA----C---------AGG---AGTGG-------ATG------AGC 3318
SMM.SL.92.SL92B               --T-----G--GC-A.....-----A--A--------------A--AC---------AGTTC-A---------GCA------A--T-------------AC--AAA-CC-G---A-GA-C-----G-AA-TGA-T--G---------A-G-T-AGTGG-------TT------AGC 3184
SMM.US.x.H9                   --GC---G---A-C-.....-----A--A---A-T------A--A-----A--AGTGC-A---------GCA--A-----T-----A--A--GACT-AA--TC-T--C-G-T-AA-C---------AA-TGA-TT----------AGG-T-AGTGG--T--GATG-------GC 3226
SMM.US.x.PGM53                ---C---G---A-C-.....-----A--A--------------A--A-----A--AGT-C-A---------GCA--A-----T-----A--A--GACT-AA--TC-T--C---T-AA-C---------AA-TGA-TT----------AGG-T-AGTGG--T--GATG------AGC 3670
SMM.US.x.SIVsmH635F_L3        ---C---G---G-C-.....-----A--A--------------A--A-----A--AGTGC-A---------GCA--A-----T-----A--A--GACT-AA--TC-T--C-G---AA-C---------AA-TGA-TT----------AGG-T-AGTGG--T--GATG-------GC 3756

MAC.US.x.239                  ---C---G---G-C-.....-----A--G-----T-----------------A--AGT---A--------GC---A-----T-----T--A---ACC-AA--TC-C----GGT-AA----------AA-TGA-T---------G-A-G-T-AGTGG--T--GATG-----AGC 3753
Pol                           _L__P__Q__R__E__.__._T__W__T__V__N__D__I__Q__K__L__V__G__V__L__N__W__A__A__Q__I__Y__P__G__I__K__T__K__H__L__C__R__L__I__R__G__K__M__T__L__T__E__E__V__Q__W__T__E__M__A__E
MAC.US.x.251_1A11             ---C---G---G-C-.....-----A--G-----T-----------------A--AGT---A--------GC---A-----T-----T--A---ACC-AA--TC-C----GG--AA----------AA-TGA-T---------G-A-G-T-AGTGG--T--GATG-----AGC 3752
MAC.US.x.251_BK28             ---C---G---G-C-.....-----A--G-----T-----------------A--AGT---A--------GC---A-----T-----T--A---ACC-AA--TC-C----GGT-AA----------AA-TGA-T---------G-A-G-T-AGTGG--T--GATG-----AGC 3729
MAC.US.x.EMBL_3               ---C---G---G-C-.....-----A--G-----T-----------------A--AGT---A--------GC---A-----T-----T--A---ACC-AA--TC-C----GGT-AA----------AA-TGA-T---------G-C-G-T-AGTGG--T--GATG-----AGC 3227
STM.US.x.STM                  --TC-G-G----GTT.....-----A--G-----T--C--A--AC-G--A---GTC--A--C-------GCA--A-----T-----A--------ACT-AA--TC-G--C---T-AA-C-------T-AG-TG-----G--T--G-AGG---AGTGG-------ATG-------GC 3400
MNE.US.x.MNE027               ---C-------G-C-.....-----A--G-----T-----------------A--AGT---A--------GC---A-----T-----T--A---ACC-AA--TC-C----GGT-AA----------AA-TGA-T---------G-A-G-T-AGTGG--T--GATG-----AGC 3221
DEB.CM.99.CM40                -----TG----AGAA.....-----A--GC--A-A--T-------------A--AC-G--A--------------C--A-----TT------A---AC--AA--C--------GGGCAA-C-----GGT-CCG-G-T--------CCCGG-TA-C--CT-T--------TC-AGC 3260
DEB.CM.99.CM5                 ------G-G--GGAG.....-----A---C-CA-A--C-----AC-G--A--AC-G--A-------------------A---T----A--A--GACC-AA--C--G--C--GGCAA-C-----GGTTCC--GG---------TAA-G--AGC--TT-----G-------C-AGC 3254
DEN.CD.x.CD1                  ---TTGCC---GGATGGGCAG-----G--G-----CA--T-----C-----A--AGT-------------------T-G---T----A--A--GAC--AAG-----------GAG-A-------GTA----ACT--GG----ACC---TT----C--G--G-G----------GC 3344
LST.CD.88.447                 -----C-T----CTTCAGAAGACA--A--T--CA-----------C-----A--AGTC-------------GCA--AC----TGAG--A----G-ACT-AAG--C-T-------T-AA----G---GTA-----T---GAT------TG--A-CTGG------T-C----TTAGA 2725
LST.CD.88.485                 -----C-T----CTTCAGAAGACA--A--A--CA-----------C-----A--AGTC-------------GCA--AC----TGAG--A----G-ACT-AAG--C-T-------T-AA----G--GGTA-----T---GA-------TG--A-CTGG------T-C----TTAGA 2725
LST.CD.88.524                 -----G-T-----CAAGGAAAACT--A--A-----A--T-----AC-G--A--AGTT--A-----------GCA--AC-G--TGAT--A--C-G-AC--AAG-GC-C------GT-GA-----------GTA--GC-CT-GGA-C--C--G-GA-TTGG-------G-------CTAGA 2722
LST.KE.x.lho7                 -----G-TG--TCTAACAAAAACT--A--A---C-A---------AC-G-----AGT-C-T----------GC---AT-G--TGAT--T----GGAC--AAG--C-C--C---T-AA-A--G---GTA---C-CT-GGA----A-C----AACTGG-------G------CTT-GA 3809
SUN.GA.98.L14                 --CCCTCT---AGAAGAG...CCA--A--G-------------AA-T--A--AGTTA-C-----------TTCA--AT-A--TGA----C-C-G-AC--AAG--------------AA--------G-AT---C--T--G----GAA-G---AAGA-G-----------TAG-GA 3811
MND-1.GA.x.MNDGB1             ------TT--CAGAACAG...CCA--A--A-----A---------A--G--A--T-----A---------TGCA---T-A--T--T------C---AC--AA----C-G--C--G--AA-A------GGACT-AACA-------GAA-G-CA--A-G--------------AGA-- 3141
OLC.CI.97.97CI12              -----------AGAA.....-----A---------T---------A-----A-----C----A--------TTCA--A--C--TT----A--A---ACG-AAGCT--------------GA---AG----C-T-----GGT-GCT---TAA-G-GGAGTGG-----------G---A--GA 3325
WRC.CI.97.97CI14              ------G---TA-AG.........--A--A-----TG-------A--G--A--A-----A--C-------TGC-----C-A--TGA---A--AC-GAC--AA---------C-G-T-GA--------G-AT---C--T-G---------A-G---GA-TGG-------------TTTAGA 3748
WRC.CI.98.98CI04              -----GG---TA--A.........--A--A-----TG----A--A--G----A--------------------GCA-----------TGAG--A---C--AC--A--------------C---T-AA-----------A-----C-C--C---------A-G--A--TGG-------G------TTAGA 3763
WRC.GM.05.Pbt_05GM_X02        ------G--ATA--G.........--A--A-----TG------AC-GA-A--A------A-----------GCA--A--A--TGAG--T--A---AC--AAG----G-------T-AA-----G----AT----C-T-------C--A-G--GA-TGG--CATG---------AGC 2677
MND-2.CM.98.CM16              -----G-----AGAA.....-----A--A---------C--A-----G--A------C-T--------------------TT-C--A-------AC--AAG-GC-C-----G---AA--------GG-A---C-C---GAT---AA-G--GA-TGG--TAG----------AT- 3452
MND-2.GA.x.M14                -----------G-CA.....-----A--A-----T--T-----AC-G--A--A---C-C----------------A--C--TT-C--A-------AC--AAG-GC-C--C---T-GA-C------GG-A---C-T---GAT--------G--GA-TGG--TAGG---------AT- 3379
MND-2.x.x.5440                -----------A---.....-----A--A-----T--T-----A-----A-------C-T----------------C--A--C--TT-C--A-------AC--AAG-GC-C-------T-GA-A--G--GG-A-G-C-C---GATA--A--G-GGA-TGG--TAG---------A-- 3012
DRL.x.x.FAO                   --T--T-----A-TT.....----------T-----T--T-----A------A--AGTGC-T--------T-----A------TT-C--A-------AC--A-G-GC-C-------T-AA-----G---G------C-C----GAT------A--G--GA-TGG--TAGG---------A-- 3018
MON.CM.99.L1                  -----CC-G--ACAA.....-----------------G--C--A-----GA-A--A--GC-T--------GCA--A--C--TT-----A--A----ACC-AAA-CC-G--C---AGTA-C-------GG-ACGGAGC--G-----TAC-G-GGTC--CT-----TTT---TC-AGC 3157
MON.NG.x.NG1                  ------C-----CAG.....-----A----------GG--C--A--A-----A------------A--C--------GCG-------C----------A--C--GAC--A---TC----C-R-TGTA-CC-R--GRTA-------C-----G--CA--G-GGAG----T-----CTT----CC-AGC 1750
GSN.CM.99.CN166               -----CC-G--T-C-.....-----A--G---------C------A--G--A---------A--------T-TCA--A--AC----T--T--A-GGAC--AA--G--G--C---TGTA----G---G-T-----GG---------CCAGG--GAG---C--T--G--G---C--A- 3179
GSN.CM.99.CN71                -----TC------T-.....-----A--A---------C--A--A--G--A-----GC-C-------T-TCG--A--AC-T--T--C--A-GGAC--AA--G-----C-G-TGTA-------G--------GG----------CAGG-GGAG--C--T--G--------C-AA- 3161
TAL.CM.00.266                 -----T---CT--CA.....-----A---------GT-----------A-A--A---C-T---------T-C--A--A--TT-T--T-------AC--A-GCCC-C--C-GGTA----C--G----GTA----G---TCT------A-G----AGT---T----G----------GC 3420
TAL.CM.01.8023                -----------C-G-CA.....-----A--A-----TT---------C-G--A--A---C-A--------------------A--A--TT-------C---AC--AAGCCC----C---TG------G--TGTA----GCT-GTT------AGG-TA--T-G--------------AGC 2918
MUS-1.CM.01.1085              ---C--C----AGTT.....-----A--T-----TG-G--A---C-G--A--------A----------GCA--AGCA--------T--C--GAC--A-----------C---TG-A-C--G---GTA---T-T--G--T-----AGG-C-AGT-G------G-------C--GC 3141
MUS-1.CM.01.CM1239            ---C-GC----A-A-.....-----A--A-----T-------------G--A--C-----------------GCG--AGCC--------AG------ACC-AA--GC-----C---TGTA-A--G---GTT---T-C--G---------A-G-T-A----------G------TC-AT- 3131
MUS-2.CM.01.CM1246            -----CC-G----C-.....-----A--G-----A--T------C-G--A------G--A--C------GCC--AG-C--TT-------C---ACT-AA---T------C-G-TGTA-C--G----TA--G-GG---------GC---G-GGAG--C-GT--GC-----CC-AT- 3239
MUS-2.CM.01.CM2500            --G--CC----A-C-.....-----A--A-----G--T-----AC-----A--A-----------------GCA-----A----T----A--C--GACC-AA--TC-G--C----TGTA-C-------GTA------C--T-------AC-G--GA---T--C--GC-G--TC--T- 3198
SYK.KE.x.KE51                 -------TTACAGAATCA...CCA--A--T-----GT-------AA--A-A--AGTCC-A-----------------A--A--T--T--T--C---TGT-AAGC---GA----GTG-A-A-----GG-A--------T--C--CA-TG--GA---G-----------G-----AGC 3306
SYK.KE.x.SYK173               --TA-T-T---TGA-GAA...ATC--A--G---C-GT-------A--GA-A--AGT-C-T-----------C----------TT----A-----GAC--AAG-GC-C-----GTGTA-C--------T---CA-T-------------C-CA-----T--------C---T---GC 3648
RCM.GA.x.GAB1                 -----G-----AGAA.....-----A--T--C--T--T------A------A--A--G--A--C--------------A------AGT---C-A-GGACT-A-G----G--C--GT-GA-A--G---G-A-------T--GAT---AA-G--GA-A-G---A------------AA- 2930
RCM.NG.x.NG411                -T---------AGAA.....-----A--A-----T---------A--G--A--------A-------------------A--TT----A--A--GAC--A-G-G------C------AA-A--G---GTT--G---T-GGAT--CAAGG--GA-T-T---AG----------AT- 2945
GRV.ET.x.GRI_677              --CCCCTT---AG-A...GAA-----A--A--CA-A-----------G----A--AGTTC-A------------------A-----T-----A-------ACC-AA--TACC---GCCA-GT-G-----G-AA--GAAC--CCT------A-----GT-TGG----------G-------GC 3627
SAB.SN.x.SAB1C                -------GG--AGAA.....-----A--G-----T--T--A--A-----A--A--G--A------------------A-------A-------A--GACT-AA---T--G----G----AA------------G-T-G-C-TT-G------GA--G-C-A-TGG----------G------A-- 3849
TAN.UG.x.TAN1                 -----GC----ACA-.....-----A--G-----T--T--A--A-----A--AGTG--A--------T------AT-A--T------A------GAC--AAGCCC-----C---T-A--AC------G-TC---ATT-------CT----G-G-A-TGG--G--GC-G--------G- 3653
VER.DE.x.AGM3                 GA---C-----AGAA.....---------A-------T------A-GAC-----------C-A----------GCA---C------------AC-C-G-ACT-AAA-C--G-------T-AA-C--------AA--GAACT--CT------AC-G--A-CTGG-------G-------AGC 3176
VER.KE.x.9063                 GA------G--ACAA.....-----------T------T--C--A--AC----A-------------C-------GC----C---------C---T-A-G-AC--AA--GA-----C-GGT-AA-C--G------AG---AAT----TT------G-A-G-G-A-TGG--T-------G--G--AGC 3681
VER.KE.x.AGM155               GAGA-G-----ACAA.....-----A--G-----TC-T-----A--G--A----------A----------GCA--AT-A--T-----AT-G-G-AC--AAA-TA-C-----G---A----------AG---AATT--TT-----C--GG---GA-TGG--CCC----G-------AGC 3673
VER.KE.x.TYO1_patent          GAG--------TGAA.....-------------------C----------------T-------C-A--------GCA--AT-G--T------TC---GGACC-A-A-TA----C--GT-AA-----------AG---AAT--GTT----GC--G-GA-TTGG---CCT--G-------AGC 3678
COL.CM.x.CGU1                 ------C-G--T-AG.....-----A--A-----T-----A--A--G--A--AGCC--A-------TT---C--AT-A--------A--C-GGAC--A-G-GA--A------TG-AC--A---T-AG--GCATT-GTTG--G-A-G-G-AG--T-GT-C----------AGA 3223
```

```
H1B.FR.83.HXB2          AGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACCAGTACATGGAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATATCAAATTTATCAAGAGCCA...TTTAAAAATCTGAAAACAGGAAAATATGCAAGAA 3618
Pol                     _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__E__P__.__F__K__N__L__K__T__G__K__Y__A__R__
H1A1.UG.85.U455         ----T-------G-----G--------------C--T--G--------A-----------------------C--G----------------A--A-------A----------------------...--------------------------------- 3064
H1B.US.90.WEAU160       ---G--------------G--A--------------G-------------------------------------GC--------------------------G-------------------...-------------------G---------G 3617
H1C.ET.86.ETH2220       ----T-------G-----G--A--------G---------------A-T------------------------T-----------A-------A-T-A----------T----T-----C-----A---...--C-------------------G--G-T-----A-- 3010
H1D.CD.84.84ZR085       ------------------G-----------G------A-G------------------------------------T--------A--A-------T---------------------------A---...--------------------------G---------- 3140
H1F1.BE.93.VI850        ----T-----A----T--G-----------G--------------G--A-----------------------------------A--A--AG-C--G----------------------GA-C---...-------------A-----------G---------A-G 2955
H1G.SE.93.SE6165        G--GT-------G-----G--A--------G------T-----------C-------------------A-----------G----A--A--A----T--A------------------------...-AC-------------------------G-------A-- 3015
H1H.CF.90.056           G---T-------------G---------G-G-------A-----------A--------T------------------------G-----A--------------------------G-------...----------------G----------------A-- 2965
H1J.SE.93.SE7887        G----------------A-------------------------G--A----------T--G-------A-----------G-G-----A--A----TG-A-------------------G-------...--------C-----------G-----------A-- 2932
H1K.CM.96.MP535         ----T-A-----G---------------------G------G-----G--A-----------------------------------A--A---A-T-A--------------------A--------...CA------------------------G--G---------- 2814
H101_AE.TH.90.CM240     ---GT-------G-----G--------------ACC--T--G--------A-----------------------G----------G-------A--A------G-A------------------------...-----------A------------------------C---- 3192
H102_AG.NG.x.IBNG       ----T-------G-----G--A---------------T------------G--A-----------A----------G--------T-------A--A------A------------------------...---------T-A----------------------A-- 3143
H103_AB.RU.97.KAL153_2  ------------------G---------------------------------------------------------------G---------------------A-------------------...----------------------------------C 2844
H104_cpx.CY.94.CY032    ----T-A-----G-----G-----------------------G-CA--------------------------------A--------T---------------------A----------------...CA-------------------------G--G---------- 2984
H1N.CM.95.YBF30         ------A-----------G-----AT-----------CC-G---------C-----------GG------A---G---------T--A------A--A-----T--G--------------G---------G---TT-...CA----------T-A---------------G-------A-- 3212
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--AGA-...CA---G--C--C-----------G------A-T--GC 3673
H1O.CM.91.MVP5180       ----T-A-A-------------A-AG---------G----------------A------CC-G--TGAC--------GTGG-TTAGT--T----------T--AG----G---------T--C--GG-A----G--TGA-...CA---G--C--T----------------T--GC 3648
H1P.FR.06.RBF168        ------A-AG-----T-A----A---T-----------------G------T------CC-A---GA---G--A------T---C---------A------AG-G--A----------T----G--A--C--G--AGA-...CA--------CT-----------R-----------C--RC 3079
CPZ.CM.01.SIVcpzCAM13   ----T-A-----------------A--AG-T-GTAC-------------G--T---------A-T---GAT--G--GC-C----------G-----A--A-----AG----G--G---------G--C-T---G--A-A-...CA--G-----T-A--G----------------C---C 3198
CPZ.CM.05.SIVcpzEK505   ---GT-A--T--G--T--G-----AT-------------T--------G-CC-------A-----GAC------A--------------G--C--------------A-------T--------------G-----------CTT-...CA--------T-A--G-------------------A-- 3190
CPZ.CM.05.SIVcpzLB7     ----T-A-------T--G--A--AT------A-C--------------G--C-------------GG-T------C-G----------G----A--A--A--A--GA-T-----------C--------C-----G--A-A--...------------------------------------C---C 3173
CPZ.CM.05.SIVcpzMB66    ----T----------T-----A--AT----------------------G-CC-------------C-----------A-------------------A--T-----A--G-----T---------TC--G--A-AG...-A---G--CT--------------G-----------T--G- 3164
CPZ.CM.05.SIVcpzMT145   ---------------TCA------A--------------------------CC--C---------TGA---G--A--------------------A--A--G--T--G-------------A-T---G---------...CA-----------C--A--G--------------GC-A-C 3160
CPZ.CM.-.SIVcpzMB897    ---GT-A-------------G-----AT---G---C--T--------G-CC--C--------C--------A--G----------A--A--T---------------T-------G---------G--A-AG...CA------------------------------------- 3184
CPZ.CM.-.SIVcpzDP943    G---T---A-------T-------A---T-------C------G-------------C---------GAG--G--G-----------G----A--A-------AA-CA-T--G------T------------T--G--A-AG...CA---G--CT----------G--G------T--GC 3204
CPZ.CM.98.CAM3          G---T-A-A---G-----R-------T---RG--T-----G------G-----C-------------GA---G--A-----------N--------A-----AA-TA-T--G-----T------------TC--G--A-AG...CA---G--C--A---------C-------T--GC 3029
CPZ.CM.98.CAM5          G---T-A-A-------T-------A---T----G--T--T--G----------T--C------T---GA---G--G-----------GG----A--A-----AA-TA-T--G-----T--C--G----TC--G--A-AG...CA---G--CT-A---------G-------T--GC 3320
CPZ.GA.88.GAB1          ----T-A-------T--G-----AG----GCAC----------------G--A--C------T---GAC------GC-T-------------------A-------CA-CT----G-----T--------G--A-T---G--A---...CA---G-----T----G------G--------C---C 3679
CPZ.GA.88.GAB2          ---GT-A-A-------T--G--A---T-------------G-------CACC--C-----------GA-------AC-----------G-T--A-------AATA-T--G---------T--------G-----T--G--AA--...GGA-GG-----A----------------------C---C 3011
CPZ.US.85.CPZUS         G--GT-A-A-------------A-----------G------------G--C-----------------GA------A---G-----------G-----A--A--A--AAGGA-T--G-----------C-------T-----A-GG...CA---G-----A--G----------------C---C 3678
CPZ.CD.90.ANT           ----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-TGA-...GGA---CTGT-A---G--------------T--GC 3057
CPZ.TZ.00.TAN1          G--GT-A-A-------T-AGC---------------AAG--T--A----CA--C------T--TAA-TT-CCTC-GCA-----C-----------------------A--G-----------G--A-----G--AGA-...GGG--------T-A-------------------A-T 3262
CPZ.TZ.2001.TAN2        ---G--A-A---G--T-A-C-A-------------GAAT--G--A----CA--C------T--TAAGTT-CCTC-GCAG----C-G-------------A--G--A-------------C-----A--------AGA-...GGA--G----T-A--------C-----------A-T 3267
CPZ.TZ.2002.TAN3_1      ---G--A-A-------T-A-C-A-------------GAG--G--A----CA--C------T--TAAGCT-CCTC-GCAG----C-G-------------A--G--A-------------C-----A---------GA-...GGA--------T-A--------C-----------A-T 3267
GOR.CM.2007.SIVgor2139_287      ------A-A---G-----G--A--AT-----------C------------------CC-G--TGA---G--A------TGA-TG-G--A--A--A--AG----A--G------T------G--C-TC--G--TGA-...-AC--G-----A--G---------G-----T---C 3155
GOR.CM.2007.SIVgorCP2135con     ------A-A---G-----G--A--AT-----------C--G---------------CC-G--TGA---G--A------TGA-TG-G--A--A--A--AG----A--G------T------G--C-TC--G--TGA-...-AC--G--CT-A--G---------G-----T--GC 3154
GOR.CM.2004.SIVgorCP684con      ----T-A-A---G-----G--A--CT------------------------------CC-G--TGA---G--A------TGA-TG-G--A--A--A--AAG---G--G------T------G--C-TC--G--TGA-...CAC--G--CT-A--G-------------------T--GC 3074
H2A.DE.x.BEN            ---G--A-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--AGAG.........TC--A---GT-------G---------AG- 4000
H2A.GW.x.ALI            ------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-GAGA-.........T---A---GTG-A---G-------A-- 3993
H2A.SN.x.ST             G-----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-GAGG-.........T---A---GT-----------------AGG 3445
H2B.CI.x.EHO            ----T-CCAG-------A-ATC--C---G--C--GAGCAGG-A---TCC------CA-AGA-GGGGT-CCT---GA-----AC-G-G------AA-TCTAGC-AAT--G----------CA-G---C-----G-GAGAT.........TC--A---GT-----------------AGG 3971
H2B.GH.86.D205          ---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-GAAAT.......GTC--A---GT-----------------AGG 3973
H2G.CI.92.ABT96         ----T--CA---------AGATT--C---R--C-GGARCAGG-A--C-CC-----CA-AGA-GR-G-GCCT--RGA----CACTG--------AA-TTTAG-CAAT--G----------CA---G---C------GAAAC.......G--TC--A---GT---T-----------AG- 3327
H2U.FR.96.12034         ----T-T------G--T-AGATT---T-G-GTC-GGA-CA-G-G--G-CT-----CAGGGA-GA-G--A-TC--GA-----AC-G-G-TT--AA-TCAAG-TAAT--------C---CA-G---CC-----G-GTGAC.......GG-CAT-A---GT----------T---C-AGG 3482
SMM.SL.92.SL92B         ----T--------G--T-----ATT----------TC--GAGCAGG-A---AGA-----CAGAGA-GATG--CCGC--GA-----ACGG-T-TA---A-TCAAG-CAAT--G----GT---A-G---C-------GTGAC.......GG-T---C--GGT---G--G-T-----A-- 3348
SMM.US.x.H9             ----TAT-A-------AGATA----CC-GTC--GA-CC-G-G---TGT--C---C-GGAGGG---GCCAC---GAG----CC-G---ATA---AGTCA-G-TAAT------T-------A-----C-C---------GAC.........TA--A---GT---C-----T------AG- 3390
SMM.US.x.PGM53          ----TAT-A-------AGATA-----C-GTC--GA-CAGG-G---TGT--C--CC-AGAGGG---GCCA----GAG----AC-G---ATA---AGTCA-G-TAAT------T-------A-----CC-C-------AGAC.......G-TA-------GT---C-----T------AG- 3834
SMM.US.x.SIVsmH635F_L3  ----TAT-A-------AGATA-----C-GTC--GA-CA-G-A---TGT--C--CC-AGAGGG----CCAA---GAG----AC-G---ATA---AGTCA-G-TAAT--G---T-------A-----C-C-------AGAC.......GTA--A---GTG--C-----T------AGG 3920

MAC.US.x.239            ----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-------AGAC.........TA------GT---------T------AG- 3917
Pol                     _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__Q__W__S__Y__K__I__H__Q__E__.__.__D__K__I__L__K__V__G__K__F__A__
MAC.US.x.251_1A11       ----TAT-AG------T-A-ATA-----C-GTC-GGA-CA-G-A---TGT------CC-AGA-GGC--GCCA---GA---CACGG--ATA---A-TCA-G-CAAT--G---T-T---A-----C-C-------AGAC.........TA------GT---------T------AG- 3916
MAC.US.x.251_BK28       ----TAT-AG------T-AGATA-----C-GTC-GGA-CA-G-A---TGT------CC-AGA-GGC--GCCA---GA---CACGG--ATA---AGTCA-G-CAAT--G---T-T---A-----C-C-------AGAC.........TA------GT---------T------AG- 3893
MAC.US.x.EMBL_3         ----TAT-AG------T-ATATA-----C-GTC-GGA-CA-G-A---TGT------CC-AGA-GGC--GCCA---GA---CACGG--ATA---AGTCA-G-CAAT--G-----T---A-----C-C-------AGAC.........TA------GT-A------T------AG- 3391
STM.US.x.STM            ----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-------AGAC.........TAT-----GT---C--G-T------A-- 3564
MNE.US.x.MNE027         ----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-T---A-G---C-C-------AGAC.........TA------GT---------T------AG- 3385
DEB.CM.99.CM40          ------TAG---G-----AG--A--C--G--GC--GA---GTCA------T--C---A-AGAGCAGG--CCA--------------T--TCC---ATTA---A-----A------GG----GTG--CAGA---CCTAAG.......GG--TAT-A--G-----------TCT-T---G 3424
DEB.CM.99.CM5           -------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---ACAAAG.......GG-TTG--A---------G-----TCT---AGG 3418
DEN.CD.x.CD1            ---G--TAA----GCACA---A--AT---------AGTA-TTCA---AGT------AGG---GAG----CCT--G---------G--CTCT---ATTA--AG--AAT--G---GG----AC----CA-A--G--AAA-.......C-TCTA--C-----------C------CCAT-AGG 3508
LST.CD.88.447           ----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----GAAAT.......TCT-T------G-G------GG---CA----C 2889
LST.CD.88.485           ----TAT-A-C------CA-C----A--------AA-C-G--G----CT-----------------TA-T-------A--GG----T-AG-G------C--A-TAAAA-G--GGTCATT--C-TC--GTGGAGG----GAAAT.......TCT-T------G-G------GG---CA----GC 2889
LST.CD.88.524           ----TAT-AGACC---T-AGACA-----------AA--------G--G-CT--C------T----AGT------AC-TG----T-AGGG-----AC--A-CAA-A-----GGTTATA--T-TC---TGGAGG----GAAAT.......GCC-T----A-G-G------G----CA----C 2886
LST.KE.x.lho7           ----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----GAAAT.........C-TCT-A-G-G-T--G-GG---CA----C 3973
SUN.GA.98.L14           ----TACCAGAGT--TCAG---G-AT--C-G--GT------TCA---CAGT--C-------A----GAT------A------TTGCAG-G-G--A---GTAAA-------AAT--TA--C-T-----TGG-TG--G-GAAA-.......C--GTAT-A-GGGT-----G-----CA--AG- 3975
MND-1.GA.x.MNDGB1       G---TAT-A-C-----T-A-------CT-GGCT---GA-CA-G-A--GTCT--------T--TAAT--G--A---TAT-T-AG-T-T------AACAACAGG---AG-TATAT---T-----TGGA-G-----GA......AA---GGT-T-A-G-G-----G--------GG-A-C 3305
OLC.CI.97.97CI12        ----T-A-A---G--------------T----G---GA-CA-G-A-------ATG---A-AGAGGAGG-GCCTC-CCAGTGTACT...TTATGCT-TACAA-G--AACGGTAG-----AG----A-GG--GA---AT......GG--CCT-A--G--------GG-T-CA---G- 3486
WRC.CI.97.97CI14        ---GT-A-A-ATC---T-AGC-----AT---------GA-CA-TCA-----CC--C---A-AGG-AA-G-CTTAA--GC--A-TT-----A-A-CAGG-CCAGG-CAATGGGCCTATACAAT--GAG-AAT--G-GA-...AAG---CCA-------GT---C--------GG---C 3915
WRC.CI.98.98CI04        ---GT-A-A-AT----T-A-AGC-----------GGAGCA-ATG--G-CT--------ACTGGGAA-G--TTAG-GGC--A-TT-----A---CAGGATCAGG-CAATGGGCCTATAG-ATT-GAG-AGC--GAGA-...AAG---CCAT-A--GGT---G-----------GG-GC 3930
WRC.GM.05.Pbt_05GM_X02  ------A-A-ATC---T-----CT----------G---T--CA---A---AGT-------GAGG-GG-G-CTTAA--GCT-AG-T-----A-A-CAGGAACAGGGCAATGGGCCTACACAAT--AAG-AGC--GAGAG...ACA---CCAT-A---GT----------T-A------ 2844
MND-2.CM.98.CM16        ----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---GAT...AA---GCCG------GGT---G----------C----- 3619
MND-2.GA.x.M14          ----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--AGAG...AA---GCC----T--GGT---G--G-------T--G- 3546
MND-2.x.x.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--AGAC...AAC---CC----C--GGT-------G-------C--G- 3179
DRL.x.x.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-----TGAT...AAC---GCCA--A------------------T-AGG 3185
MON.CM.99.L1            T--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------AAA-.......TC-GTAT-----G------------T----C-A-- 3321
MON.NG.x.NG1            ----T-A----------T--G--A--------GTC----AGTCGAC----G-GA------C------CCAGCTC------G-AG-T---G--TACCTCC-T--YAG-----A--G---GGG--CRGGTA---TC--G--ARAT.......GGGT---A--G------C-----TC----AG- 1914
GSN.CM.99.CN166         ----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-----AGGC.......G-TGT-A---T---CC-----TC----AG- 3343
GSN.CM.99.CN71          ---GT-A---------------------------------C--T-CTC-G-G--CAGC--C--C---G--GA---GCCAC-TG-G-T------TACCTCC-TA--AG--CAG-------GG---CATGT-CAG---G--CAGC.......G-TG--A-G-T-----C-----T-----A-- 3325
TAL.CM.00.266           ----T-A--------------G--A--CT--GC-----GA-ACT--C--TACA-----CC-GGA-GG-G-GCC-C-GGAG-------C-GACC---ATT--CAG------T-----GGC---ATG---AA-A-----GCAA-.......CGGCTCT-A----------C--G-T----C-A-C 3584
TAL.CM.01.8023          ----T----------------CA-------GC---GGAGACC------TTCA------CA-GGA-GG-G--CCA---GA-------GC--ACC---ATT--CAG------------GGC---CATGG-AA-A--G-GAAA-.......CGGCTG--A--G------G--G-T-----AGC 3082
MUS-1.CM.01.1085        ----T-----------TCAG------C--TC--C-GA-----C--G--G-CA--C-----------GA-C-GCCA---G-G---------TGTCTCA-TA--AG-CTCA--G---GG-----A-CT-CACC--GAGCAAG.......GGGGTAT-A-GG------------T----C-A-G 3305
MUS-1.CM.01.CM1239      ----T-----------CAG---A-----T--GC-GG-----C--G--G-CA--C-T---T---GA-C---CC-C-GG-G--T-------TGTCTC-TTA--AG-CTCA------GGG----A-CT-CCTC----A-TAAG.......GG--TAT-A---T-T----------T----C-A-G 3295
MUS-2.CM.01.CM1246      ------A---------T--G--A--C--G--GC---GAG-CCGG---G-CC------------------GA------CC-C--G-CTT---G----GT-TCCTTA----G--CAA-------GGG--CACCT--ACC------CAG-........C-TG--A-G------------G-T----C-AG- 3403
MUS-2.CM.01.CM2500      ---GT------------G--------CA-G---C--GAG---CA-A------CA--C------T----GA------CCA------TCTG-------GT-TCC-TC---G--CAA-------GGG---CAC-T-CACC-------AAG........C-TGT-A---G------------T----C-A-G 3362
SYK.KE.x.KE51           ----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-----G-----CAAGAA---CC-------TC-----------ACA-G--A-G 3476
SYK.KE.x.SYK173         ----T-A-A---G--T---C-A--CT-----G---GA-CAGGCA--GTCT-----------------TAA-----CCT---GA-----C-C--CACT--A-TA---TCCCAG--------GG----CATG----A-G--------A-AAAGGCCC-CCC--A-T-------------GACA----A-- 3818
RCM.GA.x.GAB1           ----TAT-A-----------AGAT----------------AAGT-G------G-------------TGA-AAG----CC-----G-G-----A-C--T-------TTA-AAGGT--A--G---T-C--------AG-A--G---T--...GGA---CCA------------C-------T-AGC 3097
RCM.NG.x.NG411          ----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-GAGAC...CA---CC---A--G-----G--G-------A-C 3112
GRV.ET.x.GRI_677        ----TATAAGA-C--TCA--G---AG-GC-G---A--CA-G-A---ACA--C--------TCTC-------A---------AC-G-T--A------A--AG----G--------------C---T-CACC-----AGGG.......GC-GTAT-A--GGTG----G-------C-AGC 3791
SAB.SN.x.SAB1C          ----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--GCCAGA-...AA------TA------GGTT--C------------A-- 4016
TAN.UG.x.TAN1           ---GT-CAG-C-----TCA-------CT---------GAGCA---G----CT--C-----------AGT----TCT--G-G---CACT--TACC-----T--ATG-T--G--------GG-------T-CA-A--G---GG........GTA--A-----------------A---AG- 3817
VER.DE.x.AGM3           ----TAT-------------A-------CT--------ACGGA-CAGG-A--GACC--C--CA-A---GG--G-CC-A-C-G------C-G-G--A--A-TA-AAGG---T--------GT-----C---T-CA-G------GG.......C--GTAT-A---GT----T------CA---AGC 3340
VER.KE.x.9063           ---GTAT------G-----A---A---T-G---AC-GA-CA-G-A---ACT-------CA----GG----GCCTC-T-G-----C-G-C------ATTA-AAG-T--G-------GT------GT--A-A--G--AGG-.........TCT------GT-------G---------A-C 3845
VER.KE.x.AGM155         ---GTAC-A---------AG-------C-------AC-GAGCA-G-A--TACT--------CA-----GG----CC-C-T-GG-----C-G--------ATTA--AG-T--G--------T-----C---T-CA-G--G--AGG-.........TCT-A--GGT---G--G-TC--C-A-C 3837
VER.KE.x.TYO1_patent    T---TAT---------TGC----------T---AC-GA-CAGG-A---ACC------CA-A---GG--T-CCTA-T-GG-----C-G--------ATT---AAGG---A--G---GT---C---T-CA-A-----AGG-.......C--GTCT------GT----------CA-C-AGC 3842
COL.CM.x.CGU1           G--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-A-AT.......GG--CG--A-G--G--------C--AAC-C-G 3387
```

```
H1B.FR.83.HXB2          TGAGGGGTGCCCACACTAATGATGTAAAACAATTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGGAAAGACTCCT...AAATTTAAACTGCCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCT 3785
Pol                     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  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
H1A1.UG.85.U455         AA---TC---T---------------------------A-TG--------G-GT------------------------------TC---..-------G---A-----------------------G--------TG----------G--T---------- 3231
H1B.US.90.WEAU160       --------------------------------------------------G--------------------------------------..------------A------------A---------------------------------------- 3784
H1C.ET.86.ETH2220       GAG--AC--------------------G--G-------C--T---------G---G--CTG----------------------------..-------G-T-A-----C--G--A-----------G--------------C----------------- 3177
H1D.CD.84.84ZR085       ---------------------------G-----------------------G---T--------------G------------------..-------G---A-------------------------------T---C----------------- 3307
H1F1.BE.93.VI850        -----TCG------------------------------A-----A-----G---G-TTT-------------------G----GAT--CCT--G--------A-------TG--A--G--------T------------T----------------- 3125
H1G.SE.93.SE6165        G-G--TC--------------------G-----------A-T-----------G-------G-------------------------C..------------A-------G---A-----------T-----------C--------G--------- 3182
H1H.CF.90.056           ----AAC-----------------A-------------A-----------G---T-T---------------------A-T----..-------G---A--T--------A-----------G--C------------------------A--------- 3132
H1J.SE.93.SE7887        G----A--------------------------G----A-TG-----------G--TTG---GC----------------A------..-------G---A----------GA----------G-------------C--C-----G------------ 3099
H1K.CM.96.MP535         ----ATC-------------------G------------A------------G--------G-----------------A------..-------G---A------------A-----------G-------------A----------------C--- 2981
H101_AE.TH.90.CM240     GAG--TC---T----------------G------------A-TG---------G--G-------------------------C---..-------G---A----------GA-------------------TG----------G--T---------- 3359
H102_AG.NG.x.IBNG       A----TC--------------------G--------------T----------G--G-T-TG--------------------C---..-------G---A----------GA--G----------------TG----------G--T---------- 3310
H103_AB.RU.97.KAL153_2  --------------------C--------G-------------------------G----T-----------------------------..------------A------------A---------------------------------------- 3011
H104_cpx.CY.94.CY032    CC--ATC-------------------T-G-----------A-----------G---G---TG---T-----------------------..--G----G-T-A------------------------C------------A-------G---------C--- 3151
H1N.CM.95.YBF30         ----ATC------T---------A-------G---GTT--A-TG--AAGG---G-GG-A---------T--------T-------------..-------G-T-A--AG-----------GTG-----GG----------C--TC-----------A--T---------- 3379
H1O.BE.87.ANT70         AA-A--CCT-------A------A---G-------G----A-T-A-C--G--GG-GT-TCA----TCT---A-T--C--------ATTG---..---------G-----AG-CACT-GA------T----------------G-G--C------------------------A 3840
H1O.CM.91.MVP5180       AA-A--CCT-------A------A---G------GG----A-T---C--G--GG-GT-TCA----GCT-----T-----------G--ATTA---..-------C-GG-----AG-TACT-GA------T-------T-------G----A-----------G---------------- 3815
H1P.FR.06.RBF168        AA-A--CCA-T-----C------A---GG---C-GG----A-T-A-A--G--G-C-T-ACA----------------G------------ATTA--C..--G-----G---A--AG--A-T-GAA-TGTG-----G-AT------T----C---------------------------A--- 3246
CPZ.CM.01.SIVcpzCAM13   A---ATCAA-A--T--A------A-T-G------GG----A-TCA-A---------TG-TGTG-----T-----C--T------------A--A---..--G-----G-T-A--AG----------A----GT-----G---------G--------C------G-----------C--C 3365
CPZ.CM.05.SIVcpzEK505   -----TC-A-T--T--------CA----G---------T--A-TG--A------GG-GG-ACT------T-----T--C--------A------..-------C-G-T-A--AG-G------A----GT-----G--C-----------------------------G--A--T--------- 3357
CPZ.CM.05.SIVcpzLB7     AAG--TC---T--T--C----------------G----G-------T-A--------G---G---T-------------C---------A-TA--A..--G-----G-T---------T-------A----G---------G-------------T-------------------T------C--- 3340
CPZ.CM.05.SIVcpzMB66    ----ATC---A--T----------------G-------T-----T---A-------G---G--CT-------------C--C----------AGTA---..--G---C-GG--A--T-----------------------------GG-G-------------C-----------A--A------C--- 3331
CPZ.CM.05.SIVcpzMT145   AA--A-C---A--T----------A-T-GG--GC-TG-T--AA---A----------G-T-----------T---------------A--A--A..---------G---C--TG-T-----A----TG------G-C-------G----T--C-------G--T----------------- 3327
CPZ.CM.-.SIVcpzMB897    ----A-C--------TT-A------------------G--G--T--A-T----A--G---G---G---CTG-----T--------C--C--------A--A--C..--G---G--------------G--A--------GG-T---------------T--C------------A--T---------C--- 3351
CPZ.CM.-.SIVcpzDP943    AA--ATCA--A--T---------CA---G--------G-------T----A-------G--G-A--T-------T-----------C-------A--A--A..-----C-G-T-A--TG-T---------GGTC----------C--------T-----A--C------------------C--A 3371
CPZ.CM.98.CAM3          AA--ATCA--A--T---------CA---G--------G-------T----A---------G-T--T-------T-----------C-------A--A--A..-----C-G-T-A--TG-T---------GGTC----------C--------T------------------------C--A 3196
CPZ.CM.98.CAM5          A---ATCA--A--T---------CA---G--------G-------TG-------------G----T-------T-----------C-------A--A--A..-----C-G-T-A--TG-T----------GTC----------C--------T-----A---------------------A--A 3487
CPZ.GA.88.GAB1          AA---TCA--A-----A-----CA-C-GG---C-GG-G--A------A---------TG-T--T---------------------T------A--A---..-----------GGT-A--AG----G--A----GT------G---------G------A--C-------G--A------------C--- 3846
CPZ.GA.88.GAB2          AA--A-CA--T-CT---------CA----G----C-G-T--A-TG----C-----G-TT--C-T-------T----A-CC-T----------CTGT----..-----------GT-A--AG-G-----G-----GT-----------G-C--------------T-----------------------C--- 3178
CPZ.US.85.CPZUS         AA--ATCA--A--T------------A-C-G---GC--GTT--A-TG-----------------G-T--T--G-----T--C--T----------GTA--A..-----------T---TAG-G--G-------GTC--------------------T-------A---------G----------------A 3845
CPZ.CD.90.ANT           CT-CA--AA-T-----------G--T-GG-------G-T-GA-T--------------GGATT--------T------T--C-------G--GTA--A..-------C--T-A-------ACT-G--------T------TG-C------T------C---------------G-------------A--A 3224
CPZ.TZ.00.TAN1          CACCA---A-------C-----GA---G-------G---GACTGA-A--G------GG--AT--G------A----T------T-TTGTG---..-------TT-T-A--TG--TCC--A--G------AGCCAG--------T---T--C--------TT------G-A--- 3429
CPZ.TZ.2001.TAN2        -ACCA---A-------C-----GA---G-------G---GTTT-A-A---------GG--AT--G------T-----C-----T-TTGTA--C..-------CT-T-A--TG-GTCT--A--G--C--AGCCAG------------T--C--------G-T---A--G-A--- 3434
CPZ.TZ.2002.TAN3_1      -ACCA---A-------C-----GA---GG------G---GTCT-A-A--G------GG--AT--G-----CA-----------T-TTGTA---..-------CTGT-A--GG-GTCT-----G--C--AGCCAG------------T--C--------G-T---A--G-A--- 3434
GOR.CM.2007.SIVgor2139_287 AA-A--CCA-T--T--------CA-T-G---G--GG----A-TGA-A-----GG-CT-ACA-------------T--C---------ACTG---..--G--C-GG--A------TGTCG-AGT--C-----CG-T-TT-----T--C--------G--TT-------C--- 3322
GOR.CM.2007.SIVgorCP2135con AA-A--CCA-T--T--------CA-C-G---G--GG----A-T-A-A-----GG-CT-ACAG------------T--C---------ATTG---..--G--C-GG--A------TGTCGAAGT--C-----CG-T-TT-----T--C--------G--TT-T---------- 3321
GOR.CM.2004.SIVgorCP684con AA----CCA-T--T--------CA-C-G------GG----A-TGA-A-----GG-CT-ACAG------------T--C---------ACTG---..------C-GG--A------TGCCGAAGC--C-----CG-T-TT-----T--C--------G--GT------------ 3241
H2A.DE.x.BEN            -A-AAAA-A----T--C----GG--C-G-TT-C--G--C---T---T--G------GGA-A---GGCAC----C--T------CG--TA--A..-------C-C------AG-GG-G-GA--G--C-------GCAG-------GATA-C--C---------TG--A------C--A 4167
H2A.GW.x.ALI            ---AAAA-A----T--C--C-GG--C-G-TTG---G--C---T---T--------GGA-A----GCAC-G--C--T------CGA-TA--A..--G-----C-C--A--AG--G---GA-----C-------CAG-------GAT--C--C---------TG--A------C--A 4160
H2A.SN.x.ST             -A-AAAA-A-------C--C-GA--C-G--TCC--G--C-A-T---T--------GGA-A----GCAC----C--T------CGA-TA--A..-------C-C--A--AG--G---GA--T--C-------CAG-------GATA-C--C---------TG--A------C--A 3612
H2B.CI.x.EHO            -T-AAAACA-T-----C-----GA----G--T---GG-TC-T-T---C--------GGA-AG---GCAT-G--C--C---------G---TA--A..-TG--CC-T--A--AG--G---GA--G--------TCAG------------T--C---------TA--------C--A 4138
H2B.GH.86.D205          -T-AAAACA-G-----C--C-GG---G--T-C-GG--C-T-T---T--G-------GG--A----GC-C----C--C---------G---TA--A..GTG--CC-T-----AG--G---GA--G--------CCAG------------T--C---------TA--------C--A 4140
H2G.CI.92.ABT96         -A-AAAA-A----T--A-----GA-----G-TT---GG-TC-T-TG--C-----R---GGA-AG---GCAC------------------ATTA--C..TTC----C-T-----AG--G---G---C--C-------CA-----------------C--------TA--T------C--A 3494
H2U.FR.96.12034         -A-A-AA-A-T--T--A-----GA--TCGC-TGC--G--C-T-TG--A--G--G---GGA-A----GC--------------------A-T---A..-TG--CC-T-----AG--G---GA--GGTC------CA----------T--C----------G-T-------G-C-TA 3649
SMM.SL.92.SL92B         -A-AAAA-A-A--T--A-----GAA---GGTT-C-GG-TA-T-T---T------G---GGA-A-------TC------G-------------A--A..TTT----C-C--T--AG--G---G---GGTT------TCA---------------T--------------T--T--------A 3515
SMM.US.x.H9             -T-AAAA-A-A--T--A-----GA--C-GCTT-----G--C-T-T-------G------GGA-A-----GCA---A----T-------C--GTG--A..--G-----C-TT----AG--G-G-RA-----TT-------CA---------------T------------TA--------A--A 3557
SMM.US.x.PGM53          -T-AAAA-A-A--T--A-----GA--C-G-TT-----G--C-T-T-----CG-G-----GGG-A-----GCA---------T-------C--GTG--A..-------C-TT----AG--G-G-G-----TT-------CA---------------T------------TA--------A--A 4001
SMM.US.x.SIVsmH635F_L3  -T-AAAA-A-A--T--A-----GA--C-G-TT-C--G--C-T-T-------G------GGA-AG---GGCAC--------T-------G--GTG--A..-------CC-TT----AG--G---GA-----TT-------CA---------------T------------TA--------A--A 4087

MAC.US.x.239            -A-A-AA-A-A--T--C-----GA--G-G--T----G--C-T-T-A-A--G------GGA-AG---GCA-------G--C-------C--GTC--A..-------CC-CT-A--AG-TG-G------TGT-------CAG--------------C-----------G-TA---------A--G 4084
Pol                     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  . .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
MAC.US.x.251_1A11       -A-A-AA-A-A--T--C-----GA--T-G--T----G--C-T-T-A-A--G------GGA-AG---GCA-------G--C-------C--GTC--A..-------CC-CT-A--AG-TG-G------TGT-------CAG--------------C-----------G-TA---------A--A 4083
MAC.US.x.251_BK28       -A-A-AA-A-A--T--C-----GA--T-G--T----G--C-T-T-A-A--G------GGA-AG---GCA-------G--C-------C--GTC--A..-------CC-CT-A--AG-TG-G-G----TGT-------CAG--------------C-----------G-TA---------A--G 4060
MAC.US.x.EMBL_3         -A-A-AA-A-A--T--C-----GA--T-G--T----G--C-T-T-A-A--G------GGA-AG---GCA-------G--CGTA---C--GTC--A..-------CC-CT-A--AG-TG-G-G----TGT-------CAG--------------C-----------G-TA---------A--G 3558
STM.US.x.STM            -A-AAAACA-A--T--A-----GA--T-G-TT----GG--C-T-TGA-T--G--G---GGG-A----GCA-------G--C-------C---TC--A..-------CC-CT----AG--G-G-GA--GGTC------CA-------G----T-----------G-TA--A------A--A 3731
MNE.US.x.MNE027         -A-AAAA-A-A--T--C-----GA--T-G--T----G--C-T-T-A-A--G------GGA-AG---GCA-------G--C-------C--GTC--A..-------CC-CT-A--AG-CG-G------TGT-------CAG--------------C-----------G-TA---------A--A 3552
DEB.CM.99.CM40          AT-AA--A--T---TA-------T-CC-T-------G-TA----TA--T-T-------GGA--------------TT-T--------GA-TA--A..C-G--C-GG--C--AG-GGT---A---GA------C-AT-------CATA-TC-C-----------G--------C--A 3591
DEB.CM.99.CM5           AC-A---AA-T---TTA-------CT-TC-T---C--G-TA-----A--T-T-------GG-------G--T-------CTAT----------AGT---C..-----------GGT----AG--GT-------GA-------T-AC--------CATA-CC-C-------G------T-------C--- 3585
DEN.CD.x.CD1            AC-AA--AAG---TTA---CCCATATC---------G-CA-A-TGA--GC--------GGA-G----GC-T--------C------C-GACT-----..-CT--C-G----C--AG--AGC----T-GAT------TG--------T----A-------------TT-AT-----A--- 3675
LST.CD.88.447           AA-AATCA--T------A---CCCC--C--A----GG-----------A---------GGA-A----G--T------C--T------TC-TGTG--C..--GA--C-GG-A--GG--AC--GA---GTG-----TCA---------GT--TC----------G-T------------A--- 3056
LST.CD.88.485           AA-AA-CA--T------A---CCCC--C--A-G--GG-----------A---------GGA-A----G--T------C--T------TC-TGTG---..--GA-CC-GG-A--AG--AC--GA---GTG-----TCA---------GT--TC----------G-T------------A--- 3056
LST.CD.88.524           AA-A-A----T----TG---CCCC-TC-G------G----A--T--A--G-------GGA-A-------------------C--T------TTTTGTA--A..--GA-AC--G-A--TG-CACT-GA------------------AT--------T-A-C-T----------TA--T------A--A 3053
LST.KE.x.lho7           A--A--CA--A------A---CCCC--C-GA-----GT----A--TA-T--G--G---GGA-A--------------C--C------CTTTGTG--A..---A--C--G-C--AG-GACT-GA---GTC------GCAC-------GC--CC-C-------G-TT--A-----------A 4140
SUN.GA.98.L14           GAG---CA--A--TGAG---CC-TGTC----G----G---CT--CC-A------G----GGA-G-----------T---------------TTTGTG--A..--GA-AC-GG-A-----------GA--G-T----AGTCAG-------G-----C--C------GTGT--A-------C--- 4142
MND-1.GA.x.MNDGB1       A--AAACA--A--T-G------CC-C-TGA-----GG---GT--TAC---G--GG-GGA-G-------T---------C--------TTTTGTA--A..--A-GC-GA-A-----CTAC--G---G-T--------GAT--------CAT--------------GTGT--A-------A--A 3472
OLC.CI.97.97CI12        A--T-AC---T--T--A---A-GT-T-G---GA-G-----GA--CA-AG----G----GGA-A-------T---------G--T-----------TTA--G..-TT---C-GG-A--T-CTAC--GAA--GAT-------G------GCATTT--TG-----------TACAT------A--A 3653
WRC.CI.97.97CI14        AA--ACA-A-T--T-G------GT---G-TC-A-GG-TC-T--CA----G--G-----GGG---------T-------------C-------G-GTA--C..-------C-G-A-A-------G---------T-T------TT-T------T-----C-------------TGT-A------C--A 4082
WRC.CI.98.98CI04        AA--ACA-A-A-----C-----GC-T-G-AGC---GG-TC-T--CA----G-------GGG---------T-------------C----------GTA--C..-----------GA-A-------G---------T-TC------CT----------T-----TC-C--------G-T-T-A-----------A 4097
WRC.GM.05.Pbt_05GM_X02  -C---CA-A-T--T--A-----GC---G-AC-----G-TC-TTGTA-A--G--------GGA-----GGCT---------T------G--ACT---A..---------G-A-C--AG--G---------T-T-------C-TG--------T-----A--C--------G--------A------C--A 3011
MND-2.CM.98.CM16        CA-A-AA---T--T--A-----GC-T-G-GTGC-TG---GACT---A--G--G--TG-A-A----GCTT-----G--C---------ACTC---..--G---T-T------A---G---GA---GTC-----CCAG------C-----A-----------------------A--------C 3786
MND-2.GA.x.M14          CA-AACA---------A-----GT---GGGTGC-TG---GATT---A-----------G---A----GT--------G--T-------G-ATTG--A..------CT-T--C---C-GG-G-G---GT------TCAG-------CAT--T------------G--A--A------C--A 3713
MND-2.x.x.5440          CA-A-AA---------A-----AT-G-GGGT-C-TG---G-TTG--A-----GG--G-T-AG--GT-TC--------C-----------CTA--G..------CT-T--C--GT--G-G-G---GGT------CCA--------CAT--C--------------TA--A------C--A 3346
DRL.x.x.FAO             -A-A-AA-AGT-----C-------A-G-G-ATG---G---GACT---A------G--G-A-A----GGCAC-G-------C---------G-TTG--C..-TT--CT-CT-A--AG--G---G----GTT------CA----------CAG------C-------G-TA--A-------C--- 3352
MON.CM.99.L1            -T--AA--A-----T-C---ACCTATC-G--G---G----T--GC-AGCC---T-GGGA-A----GGC-------C-C----------CGACT---C..-TT---CGC--A--AG-GGTG--A--GCA-------CG-----------G-----TA----------AT--AT----G-A--A 3488
MON.NG.x.NG1            -T-AAAA-------T-C---CACATACC-G------GG----T--T--AACC---C---GGG-AG---GGC-T-G--C--T-----------CTA--A..-T-----G-T-A--AG--A-C--A---CAG-------C------------G-C--CA-----------T---T------A--A 2081
GSN.CM.99.CN166         CC--AAA---A----TA--CTCCTATC-G------TG----T--TT--ACC---G--GGT-GG--G--TCAG--T-CC-----------GTA--AGAC-------GGA-C--AG--GTC-------CA-------TG-----------TGA-T---------G--A--A------A--C 3513
GSN.CM.99.CN71          CC--AAC---A--T-TA--CTC-TATC-G------GG----T--TA--ACC---G--GGT-GG-----TCAG--T-CT----------AGTA---GAC-----C-GGA-C--AG--GTC------GCA--------T--G---------TGA-C---------G--A--A----------C 3495
TAL.CM.00.266           A---AACA--A----G---CCCCTATC----G-----T--GA--TA----G-------GG--AG---GTC-------CT-T------GC--GTA--A..GT---C-GGA-C--AG-GGTC------GGA---------CAG-----------AC-C--------TG-T-A------A--C 3751
TAL.CM.01.8023          A---AACA--A----G---CCCCTATC-----C--GTG-GA--CA-------------GGA-A----GGC-------CT-T-------C--GTG--C..-T-----GGG-A--AG-GGTC-----GGAG------GCA--------T----------------TG-T-A----C-G--C 3249
MUS-1.CM.01.1085        -A---TCA-TA---------AGCTATC-G------GG----C---A--GCC-------GGA-G--------C------C--T------G---C-C---GAC--GG-C-G-A-C--AG-GATC--A---CAG------CCAG---------T--CAC-C------G-T-T-G------A--- 3475
MUS-1.CM.01.CM1239      -C--ATCA-T----T--C--CAG-TATC-G--G---T-T--T--CA-AGCC-G-----GGA-GG---TCAT-----C--T-------G---C-C--CGAA--GG-C-GGA-T--AG-GATC-----GCAG------TCA-----------------AC--------G-TAT-A------A--- 3465
MUS-2.CM.01.CM1246      -C--AACA-----T-G---CCCCTACC-----GC-GG----A--CC-ATCT-G-GCT-GT-AG--GGC-C-----CTGT-------G------C--AGAC-----GC-GGA-A--GG--GTC-------CAG------C-AT---------G-----C-CA-------A-A--T----------C 3573
MUS-2.CM.01.CM2500      -T--AACA-----T-TC---AC-TACC-G---C--G---------T-AGC---GGCTG-A-G------TC-T-----CTGT-------------C---GAC-----G--G-A-A--GG-GGTC------GCAG------CCAC---------G-----A-CA-------GA-T---A---G----A 3532
SYK.KE.x.KE51           CATAT-CGA----TCAC-------TACC-GGC----GG-TC--CTGA--A-T-------GGA-T-C--GC-C--TGG---T----------A-TC---..G----CC-C--C--AG--A---G----GA-----------A-----------------TC-C----------A--A---T-G--A 3643
SYK.KE.x.SYK173         CCTTT-CA--A--T-G------CTATC--TCTC-TG-TC-ATTGT-AA-T-------GGG-TCC--TCAC--TGGTAT-----------GTA--A..-C---CC-T--C--AG--A---GA--GGA-------A--------------------C---------------------A---G-A--A 3985
RCM.GA.x.GAB1           A--AAACA--A-----C------AA-T-G-ATGC-GG---G-TT---A----------TG-A-A----GC-------T--T--------G-CTG--A..-C---C-G------A---G-G-GA--GGT------CTA-----C--T-CC-A--C-------G-TT---------A--A 3264
RCM.NG.x.NG411          AA-A-ACA--A--T--A------AA-T-G-ATGC-GG---GCTT---A--G--------TG--A----GCT---------T-----------CT--A..-------G-G------A---G---GA---GT-------CAG------T-G--T--C-------G-TTT----------- 3279
GRV.ET.x.GRI_677        AA--A-AAA-T--T----------C---GGACTC--G--C-CCTT--C----------CTGT-AG---GCAC-TACC--T--------GACT---A..CG-G-AC-----C--AG--G-C---A-------------T-TG-------CAG--C-------------TAT--------A--A 3958
SAB.SN.x.SAB1C          -C-AAACA--A-----C------AT---G-ATG---G---GCCT---A--G--G----GGA-AG----------------C--C-------C----TA---..-TTA-GG---------AG--G----GA--GCTC-------CA-----------T-T--C------------G-T----------C--G 4183
TAN.UG.x.TAN1           GC-AAAA-A-T--T--A------GT-T-G-GT----GG---GAATG--C------GGG-GTG--AG---GCG--CACCT-T---------ATTA--C..-TCA-GG--G-A--AG--G---GA---GT------CA-------------GT--C--------------T---A---C-A--A 3984
VER.DE.x.AGM3           A--AAAACA-T--T--C------GT-CCGTGT---GG---GATT---A---------C-TTGT-A----TCTT-----T-----------G--TTG--A..GTCC---G----C--A---G-G-G---GT-------CA-----------G-T--T--C--------G-TA-GT---------A 3507
VER.KE.x.9063           AA-A-AACA----T--A------AC-TCGCAC---GG-C-GCCT---A---------CTGT-A----GGCAC-G--C--T--------G-ACTA--A..-TT---G----C--AG--G---G---GT--------CA----------G-----C--C----------TG-G---------- 4012
VER.KE.x.AGM155         A--AA-C-A-T------C-----GT-GCGTGT-C--G---GA-T---A--G-------GGG-A----GGC-C-----------T--------C-ATTA--C..-CT---G----C--AG-GG-G-G---C---------------CA----------G-----C---------------T--GT-----A--C 4004
VER.KE.x.TYO1_patent    AA-A-AACA----T--A------AC-TCGCAC----G-T-GTTT-------G--G--TTG--A----GCTC-----T--T-------G-TATTA--A..GTTC-AG-----C--G---G---GA--GGT-------CA----------G-G--T--C---------G-TA-G----------C 4009
COL.CM.x.CGU1           GC-AAAA-CAG---TT----AG-A-GC-GG-G---G-TTCA-TTA-A------G---GGT-GG---GC--------C-CC-----------GTG--C..---A-G---G-C--AGC-A-G-GA--GGAT-------CAG--------GT--C-----------AGTG-T------C--A 3554
```

p51 end and p66 RT continue Pol p15 RNAse H start

```
H1B.FR.83.HXB2                 GAGTGGGAGTTTGTTAATACCCCTCCCTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAAGACAAAAAGTTGTCACCCT 3955
Pol                            _E__W__E__F__V__N__T__P__P__L__V__K__L__W__Y__Q__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
H1A1.UG.85.U455                --A-----------C-----------C----A-------------------------C-------C---------G--A-----------------------T----A-----GC------------G-----C---G-C-------------G-----TT---- 3401
H1B.US.90.WEAU160              --------------C-----------------------------T----------------------------T-------------------------A------------------------------C---------------G------T---- 3954
H1C.ET.86.ETH2220              --A-----------------------C----A---------------C-----------------C----T-----T-----------A---------T----A------A-----------G-----------G----------G--G---A----TT-T-- 3347
H1D.CD.84.84ZR085              --------A-----C----------T-----A-------------------------G--------A-------T-----------------T-AA--------------------------G-C--------------G----------C--T- 3477
H1F1.BE.93.VI850               --A-----------C--------C--TC-------C-------------A-C---G-------C--------C-------------------T----T--A----C---AA--------------------G---A-----A-------------T---- 3295
H1G.SE.93.SE6165               --------------C-----------TC----A--G---------T-G------A-C----------CC-----------T-A------------------T---------A-----------G-----------G-C-A-----AG------A--A-T----- 3352
H1H.CF.90.056                  --A-------------C-------ATC----A---------------T-------A-C---G-------C----------T-A----A---------------T-----A-------------------------C---G-----------AG-----------T---- 3302
H1J.SE.93.SE7887               --------------C-----------C--TC----A--G-----------A-----A--G----------A-G-------T-----------------T-----------------AC------G-----G---------G-C-A---------G-----A--T---- 3269
H1K.CM.96.MP535                --A-----------C-----T-------C----A---C------------A-C--------------G---------T-----C--------------C-T-----A------AAG----G------------G-C-------------------G------T--A- 3151
H101_AE.TH.90.CM240            --A-----------------------TC----A-----------A-----A-----C---------------G--T--------------------GT-----------GC-------------G-----C---G-C-------------------G--A--TT---- 3529
H102_AG.NG.x.IBNG              --A-----------C-----------TC----A--------------------A-----C-----------------------T----------A---------T--A--------GA--------------G-----C---G-C-------------------G-----TT---- 3480
H103_AB.RU.97.KAL153_2         --------A-----C------------------A--------------------------------------------T----------A--------T-------------C---------------------G-C------------------G-----T---- 3181
H104_cpx.CY.94.CY032           --A-----A-----C-----------TC----A---------------A-C---C-------C-----------T--------------A---------T--A--A--A---CAG--------------------G---------C----------------T---- 3321
H1N.CM.95.YBF30                --------A-----C--C--T-------C-T--A------------T-------A-C---G--A--CAGT--G---------T-------------A---------T-----A--A-------G-------T-T---G--AG----G----------G-----G---T-TA- 3549
H1O.BE.87.ANT70                --A-----A-----C-GC--A--C--A--GA-C-------------AG-C----A-GT-----T--TA-G--G----------AT-------------A---------T--A------A----C--------G--------------AG-ACA---G-A---G--GA-AA-A-AAT- 4010
H1O.CM.91.MVP5180              --A-----A-----C-GC--A--C--A--GA-C-------------------A-C------T--T-----G-------------T-------------A---------T---A-T--A----C--------G--G-----------AG-ACA-----A---G--CA-AA-A-AGT- 3985
H1P.FR.06.RBF168               --A-----A-------G---A--C--TC-TA-T--GC-C-----T---C-----A-----C-----CC----AT---------T-------------------A-----A---------------T------T--------A--AG-----G--T--G--G--GA-AA---AA-- 3416
CPZ.CM.01.SIVcpzCAM13          -----------CA-----C-----C--A------A-----------TAGT-------C----C--T---CC-ACG-----C--A-A-------------A------A--T------A--A--GAA------------------CA-A--AG-C-AG-----A--------A-AA---GT-- 3535
CPZ.CM.05.SIVcpzEK505          --T-------------C--C---------AC-----A-----------T------------C-------A--TAGC--G----------AT-------------A---------T-----A------------G-----------T---G--AG----G--T--G------G--GAC-T-TA- 3527
CPZ.CM.05.SIVcpzLB7            --A-----A-A----------------T--CAT--C--G--T--------A------C-------T---CC---G---------A--T---------------------------------A-A------G------------A--AG-----G--G-AG--G---A-AA--C--T- 3510
CPZ.CM.05.SIVcpzMB66           --A-------A-------C-------------A---C-T---------A-----C-----C-----CC---------------A--T------------------------------A--A--C-----------G------------A--AG-C-A-------------GA-CA--T-TT- 3501
CPZ.CM.05.SIVcpzMT145          --------A--CA-A--C-----A---C----A--G---------------------C---G--T----A-------------A-A--------------------C------A--A--A---AG--------C---------G--A--C--G--G------------GA--CAT-GT-- 3497
CPZ.CM.-.SIVcpzMB897           --A-----A-A-------C-----------C-G--A--GC-G------------A-C--------------G--G--------------T--------------------A--T--A--A--C-------------------A---G-------G-----------GA--A--T--- 3521
CPZ.CM.-.SIVcpzDP943           --C-------------A-----------C-G--A------------R-----A-C--------------CC---G-------A---------------------------C------A--C--A--GAA----------T------C--GG-------T------G----A-A----AT-- 3541
CPZ.CM.98.CAM3                 --C-----N--C--G-------------C-T--A-----------A--G--A-C---------------CC-----------G--------G--------------------A-----A--GAA----------C--C--C---G--------C------G----A-A--A-A-T- 3366
CPZ.CM.98.CAM5                 --C-----A-----A-------------C-T--A-----------A--G--ACC---------------CC---G--G----A---------------------------------A--A--A--GAA----------T--CA----AG-----G--C-----G----A-AA-A-ATT- 3657
CPZ.GA.88.GAB1                 --A-----A---A-C---------A---------C-----------AGT---------C-------T---CC-ACCA----T--T-AT-----------A-----A--T------A--A----ACT--G--------------A--AG-C-A----G-A--------A-CA-T-G-T- 4016
CPZ.GA.88.GAB2                 --A-----A--C--C-GA------A---C-T--A--G----------T--A--G--A-----------T---CA------------------AT--------T--------------C--T--------A-G-AG-----------------------T------AT----A---G--G--CA-TT--T- 3348
CPZ.US.85.CPZUS                --T-----A-----C-----------------C-C--A---C-T-------TA-----------C----GG--------AG--G--------A--T----------G----A------CC-AC-A--A--C--GAA------G------------G-----G--T---------A--A-A--TT- 4015
CPZ.CD.90.ANT                  --A----------------A--C--A---A-T-G-C-C--------TA-TC--TT-GC----C--T--TCC--AG--T----------T------T------------T--AA-CT---C-----G--G--G-------C--------G--AG-C--------AGT-GG--AAAACA--- 3394
CPZ.TZ.00.TAN1                 --------A---A----C------A--AC--A-C-GGC----------A-TC-GTT-TCT--C------CCC--A---------------T------------------------A-------A--C-G----AAG----G--------------A--A--C--------C---T-C-GGTCAAAGGA-T- 3599
CPZ.TZ.2001.TAN2               --------A--CA-----C------A--A---A-T-GG----------A-CC-GCTA-CT--C------TCC--A------------T--T------------------------A--------A-GC---AAG----G-------G-------A--AG-C--G--T---T-T-GGACAAAGGA-T- 3604
CPZ.TZ.2002.TAN3_1             --------A---A----C------A--T---A-C-G-----------A-TC-GTT--CT--T------TCC--A------------T--T------------------------A-------A--A-GC---AAG----G-------G-------G--AG----G--T---T-T--GACAAGGGATT- 3604
GOR.CM.2007.SIVgor2139_287     --------A---A----G---A-------C-CA-C--GC-C-------T--A--G--A--G-------------CC--------------------T--------------A----------T--A--A-G----CA------------------A--AG----AG--T-------G---A-AAGA-AA-- 3492
GOR.CM.2007.SIVgorCP2135con    --------A--------G---A--C---C-CA-T--GC-C-------T--A--G--A------------------CC--------------------T-----C-----A----------T--A--A-GC---CA----------------C--A--AG-----G--T------G---A-AAGG-AA-- 3491
GOR.CM.2004.SIVgorCP684con     --------A---A----G---A--C--TC-CA-T--GC-C-------T--A--G--A------------------AC--------------------T-----------A----------T--A--A-GC---CA-----------------------A--AG-----G--T------G--GA-AAGA-AA-- 3411
H2A.DE.x.BEN                   --------C-----ATC------A--AC-G--C-GG---ACA-TTA-CC----TAGG---T--T---CC---C-----G---------CAC-------AT--TGC---T--AC--T-A---GAG-------------------A--AG------------A-G-C-------AAAAGTAT- 4337
H2A.GW.x.ALI                   --C-----C-----ATC------A--AC-G--C-GGC--GCA-TTA-CC----TA-----T--T----C----C-----G---------CAC---C---C-CTG---T---C-AT-A---GA-----------------A-A--AG-------G----G-C--G--GAAGGTA-- 4330
H2A.SN.x.ST                    --C-----C--CA-ATC------G--AC-G--C-G----GTA-TTA-CC-G-T-------T------C----C--------------CAC-------AT-CTGC--T-A-C-AT-A-G-GA-----------------CA-A--AG---------------G-C--G--GAGGCTAT- 3782
H2B.CI.x.EHO                   --A-----T-----CTCA-----A--A---A-A-GG----GCC---TA-CC-G-TC------C---C----A-----T-------T-A---CAC-------AT-CTG------AA-CCT-A---GA---G--------------C--AG-C--G----AGG-T-------AAAC-AT- 4308
H2B.GH.86.D205                 --------C-----CTCG-----A--A---A-A-G-C--GCC----A-CC----TC------C---C---A----GAG----------A---CAC--------T-CTGC--T--AACCT-A--GGA-----------------C----G-C--G-----A-G-T--G---AAAGTGT- 4310
H2G.CI.92.ABT96                --T-----C-----GTC---A--G-ATC----A-GG----GTC---A-CC----TC---------TC---AGCAGA-----G----AT---AC---C---T-CTGC--T-AAACCT-C--GGAR-------------C--C--C--AG-C---------A-G-C--G--AAARGT-T- 3664
H2U.FR.96.12034                --C-----T-----ATCC--A--A---C-G----G-C-GGCA--TA-TC----TA---------AT--C--...-A---G--A-AT---AC-------CT-TTG----------CTT-A---CAG-G---------C----------G--AG------------A-G-C-G---AAAAC-TT- 3816
SMM.SL.92.SL92B                --------C--CA-CTCC--A--A--G---A-C-GG---GT--TTA-T-----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- 3685
SMM.US.x.H9                    --A-----C-----GTCA--A-------------C-G-C--GTT-T-A-CC----TA------G--T---CAG----------A--T-----------A-AT-CTG---T----C--T-A-G-GA----------C------G--AG---------C---GGC--GACAAAACTG-- 3727
SMM.US.x.PGM53                 --A-----C-----GTCA--A-------------C-G----GTC-T-A-CC----TA---------T---CAG---------G--A--T-----------AT-CTG---T---C-AT-A-G-GA------G------C------A--AG---------C---G-C--G-CAAAAC-TT- 4171
SMM.US.x.SIVsmH635F_L3         --T-----C-----GTCA--A-------------C-G----GTC-T-A-CC----TA------G--T---CAG-------------A--T--C---------AT-CTG---T---C--T-A-G-GA----------C------G--GG----G--C----G-C----CAAAACTTT- 4257
```

Pol RT end Pol p15 RNAse H start

```
MAC.US.x.239                   --A-----T---A-CTCA--A--A--GC----A-G----GTC-T-A-TC---T---G--C--T---AG----A--------AT---AC------AT--TG---T-AAC--T-A---GA---G---------------A-C--AG----G--C-A-G-C-----AAAAGTGT- 4254
Pol                            _P__E__W__D__F__I__S__T__P__P__L__V__R__L__V__F__N__L__V__K__D__P__I__E__G__E__E__T__Y__Y__T__D__G__S__C__N__K__Q__S__K__E__G__K__A__G__Y__I__T__D__R__G__K__D__K__V__K__V
MAC.US.x.251_1A11              --A---A-T---A-CTCA--A--A--AC------G----GTC-T-A-TC---T---G--C--T---AG----A--------AT---AC------AT--TG---T-AAC--T-A---GA---G---------------A-C--AG----G--C-A-G-C-----AAAAGTGT- 4253
MAC.US.x.251_BK28              --------T---A-CTCA--G--A--AC----A-G----GTC-T-A-TC---T---G--C--T---AG----A--------AT---AC------AT--TG---T-AAC--T-A---GA---G---------------A-C--AG----G--C-A-G-C-----AAAAGTGT- 4230
MAC.US.x.EMBL_3                --------T---A-CTCA--G--A--AC----A-G----GTC-T-A-TC---T---G--C--T---AG-T--A--------AT---AC------AT--TG---T-AAC--T-A---GA---G---------------A-C--AG----G--C-A-G-C-T---AAAAGTGT- 3728
STM.US.x.STM                   --A-----T-----ATCC--A-----T-----A-G----GTC-TTA-T----TA---------T-G-A----A--------A--T-----G-----CT--TG---------C-AT-----GA---G----------C--CA-A---G-------------A-A-C------AAAAG-TT- 3901
MNE.US.x.MNE027                --A-----T---A-CTCA--G--G--AC----A-G----GTC-T-A-TC---T---G--C--T---AG----A--------AT------------T--TG---T-AAC--T-A---GA---G---------------A-C--AG----G--C-A-G-C-------GAGAGT-T- 3722
DEB.CM.99.CM40                 --C------GCAA-AC-------A-ATC-G--C--GC-C-----TG--C-G-T-TC-----------CC--AT-----C----A-----------------------------T-C----------G-----T--C-----C--AG-GT-G----A-----GT--AAAATG-T- 3761
DEB.CM.99.CM5                  --C------GCG--GC-------A-AAC----C-GGC-------TG--C-T-TATC---G--A---CC--AT--T--C----A-------------A---------T-----AT-A----A-G--------C--CA-AT-AG-AT-G----A---GGG---AAAA-GT-- 3755
DEN.CD.x.CD1                   --C-----AA-G--ATCA--T--A---C-TC-A-G-AAT-TT--TA-T--G-TATC---G-------ACAA---GCCC----AT---AC------A-----A--T--AC---------CA-----G-------------GG-C-CC-AT--C-AG-----G--A--A---T- 3845
LST.CD.88.447                  --A-T---A---A-CTCA-----A-AG----AAC-GGA------TAGC-----AGC-------T----A---G-T---T----AT-------------A--T--AG-A-AA-TAGGG---AC-------G-----T------C--AC-ATC-----AGG-G-------AAAGAAT- 3226
LST.CD.88.485                  --A-T-------A-CTCA-----A-AG----AAC-GGA------TAGC-----AGC---------------A---G-T---T----AT-------------A--T--AG-A-AA-TAGGG---AC----------------T------C--AC-ATC-----AGG-G-------AAAGAAT- 3226
LST.CD.88.524                  --C-TA--A--CA-CTCC--A----AA----AAC-GGA------TAGC--G--AGC-----------A---TA---T----A------G-----A--T--AG-A--A-TAGG---AC-----------------G---A-A---C-ATC-----A-G-G---A-AAAAGAAT- 3223
LST.KE.x.lho7                  --C-TA--A--CA--TC------G-AA----AAC--GAG------ATT-GG---GC-----------A-----G-T---C----A-----------------A------AG-A-A--TAGGA---AC-----------------A-A--AC-ATC----G-A-G-G-----AAAGGAGT- 4310
SUN.GA.98.L14                  ----TA--A-----C-GC--T---AAG--G-A-C--GAG------TACTC----CC-C-------AG-GCC-----...--C----A------G-----A-------G-A-A-CTAGAA--GAG---G-------------A-A--CC-GG---------TCC-G-----AAG-AAT- 4309
MND-1.GA.x.MNDGB1              --AGTA--A---A-C-GC--A---ATG-----AA-GGGA------TAGC--GTCCCC-------TC----AG--G-T------A-AT------T-----A---------------C-G-----A-G------------A-----AG--------TTTT----GG--A-AAGAATA 3642
OLC.CI.97.97CI12               --CATA--------ATC--ATT--GAAAATCAA-GGAC--TC-GGG------T-CC--C-------G...-AG-T-----------------G-----A---TA---T---A-A-GC-GGA-------G-----------C-G-TCAG-G----AG--T---A-A--ACA-A- 3820
WRC.CI.97.97CI14               --CACA--ACA---GTC---G---TT-C-T-CT-GGGA-------TA-------T-CT-------TC--CC--A--T-------T-AT---A-------AT-T--A------A--C-G-----GA---G---------------...TGG-CC-A-TC--A-G-----------A-GTG- 4249
WRC.CI.98.98CI04               --TACA--ACA---ATCA------TA---G-CA-GGGA-------A-AA---TACTG--G--T--------AG-TG------A-AT---CA--------AT-T--A--T--A--C-G----GGA---G---------C-TC...TGG-CA-ACTCT-A-G-------AA-T-ATT- 4264
WRC.GM.05.Pbt_05GM_X02         ---GTA--ACA---GTC---T---TTTC--AAA-GGGA-------ACA-----TATC-------T----A---G-T-----GAAA-G---A-------T-T-----T-AA----GC----GA---G---------------...TGG-C--AGACT--GA-C--GA-ACAG-AAA- 3178
MND-2.CM.98.CM16               --A-----A-----ATCC--T--A-A-C-TA-CGGG--G-----TA-CC-GTT--G-------TG-GCC-----A---TGTG-AT-----G-----A-------------AA-CT-C---GA---G-------------------A--AGC----AAT-A-TCC-G---AA-AG---- 3956
MND-2.GA.x.M14                 --A-----A--CA-CTCA--A--A--AC--A-A-GG--G-----TA-CC-GTTA---------A---CC---G-AG--TGT--AT-----G-------------A--T--AA-CT----GGA---G---------------CTA---AGC---G----A-AGT------AA-AG--T- 3883
MND-2.x.x.5440                 --------C--CA--TCA--A--A--A---A-A-GG---------A-CC-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- 3516
DRL.x.x.FAO                    --T-----A---A-ATCA--A--C-ATC-CA-C--G-----------TA-C--GCT--GT--C--A---CC---G-A----GTT-AT-----G-----A-------------A-TT-G----GA----------T--C-----------CTCC--G------A-G--------AA-AG--A- 3522
MON.CM.99.L1                   --CATA---GCG--GT--T----A-A-C-GC-C-G-CAG------TACTC-G-TAC----C--------A---G--TCC---A-A---C--G--------C--C--T---ACAT-A--GGA---G-------T------A--C--CTTT--GGC-AG---G-CC------AT- 3658
MON.NG.x.NG1                   ---ATA---CCA--AC-CT-T--A-A-C-GC-T-G-CA-------ACC----T-C--R-T-----T-A-CAC--TCCC------A------------------G-----------A-CAT-A---GA---G--G-------------C--C-G--C---CCA-AG---G--AA-TC-AT- 2251
GSN.CM.99.CN166                ACCATA--AGCAA-AC-C--A--A-A-C--C---G-CA-------ACCC---TCTCG------C---A--AT----TT--T-AT-------------A-----A--T---ACCT-C---C---------------T--T-----G--C----C-----A---T-GG-CCA-AGAAT- 3683
GSN.CM.99.CN71                 ACCATA--AGCA--AC-C--A--A-AT--GT---G-CAG------ACCC---TCTCG------C---A-AAT-----TC----AT---------C-------A--T---ACCT-----------------------T--C------G--C----C-----A-T-T-G--CCA-AGAA-- 3665
TAL.CM.00.266                  ACCATC-T-CCCA--C-C--A--A---A-----T-GGC-------TA-CC---T-C-G--G-------A----------C--G-AT--C-----C--T-----CC--C-A----T-A---GA---C--G----------C--G---GCA-C-----A-G--C-T--CA---AGT- 3921
TAL.CM.01.8023                 ACCATT-TCCCCA-AC-----------------C-GGC-------TA-CC---TATC---G-------A------------A-A------G-----T--T--CC----A--AT-A--GGA---C--------C------G--C-CGTC-----A-G--C-T--A--AGGA-- 3419
MUS-1.CM.01.1085               ACCATA--AGCA--GC-----A-C-A-C-GT-A---CAG-----T-ACA--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--- 3645
MUS-1.CM.01.CM1239             --CATA---GCA--AC-----A-C-A-C-GC--CG-CAG-----T-ACA--G-T-CC---G--AT-GTC--A---TCCC--A-AT-----------------A-----AC-T-AA-TAT----G-----G--G----------------------C-C-------A-G-----G--G--T-G-T- 3635
MUS-2.CM.01.CM1246             --CATA--AGCA--AC-C-----ATA-C-TC---G-CAG-----TTACC--G-TCCC---G-------A-AC----CCC----A------G-----A-----A--T---A-CT-----C------G---T--------A--CG----G---CAGG-G-G--C-A-A-ATT- 3743
MUS-2.CM.01.CM2500             ---ATA---GCG--GC-C--A----A-C-GT---G-CA------T-ACTC-G-T-CC---G--TC---AG-CG---CCC----AT-----T---------C-----T--AA-CT-C---------T-GG-----G-----C--AG----------CAGG----G-CCA-AGA--- 3702
SYK.KE.x.KE51                  --AGTAA-A-G---GC-C------CATA------T-G--GG-AT---A-T--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- 3810
SYK.KE.x.SYK173                --AGTCA-A--CA-A-G------G--A--G--T-GG-GG-AC--TA-C--G-TACC-----------CC--A-----T---------C-------A--------T--A--C-GC--GAC------T-----------AG-A-G-GAT--G-C---G-G---ACAGTAT-- 4155
RCM.GA.x.GAB1                  --T-----A-------GC--A--A--TC-TA-T-G-C-CG------A-CC---TT--G--T------CC-----AG---GT--A------G---------------------AA-T-G----A------G--G-----------A-----------A-G----G---AAAAGAAT- 3434
RCM.NG.x.NG411                 --A-----A-------G---T--A--AC--A-A-GGC---------A-TC---TA--G--T--T---CC-----A----GT--AT-----------------A--T----A-T-G---G--G-----G-------------A--AG-----G----A-G----G---AAAAGAAT- 3449
GRV.ET.x.GRI_677               --A-------------GC--A--A-T-C----A---C-G-----TTCC----TA---------A--CAA-----A---TGTT-AT-----G-----------ATC--AA-T---C--------T-G----------C-GT-AG-G--------A-AGT-G-A--AGGGAAT- 4128
SAB.SN.x.SAB1C                 --A-----AA-G--C-G---T-----AA--GA-T-G---G-----A-A----TA-----T-----CCC-----A--C-GT--AT----------------------T--AA-TT-C---GA-----------------CT-A--AG--------GGAT------G--G--AG-AT- 4353
TAN.UG.x.TAN1                  ---ATA--A--CA-CTCC--A--G------A-C-----------TA-C--G-TA-----T------CC-AA----A-GGTG-AT---TGC---------TGC--T--AA-TT-C-----------G------C-TCC-A-----C-C----GA-TC-C-G--CA-AGAAT- 4154
VER.DE.x.AGM3                  --C-----A-----C-G------A---C----A-----------TACCC-GACA------------CC-AAG-A---TGT--A--------G------A--TTG---T--AA-TT-A-GGGA-----------------A-C--AC-ATAT--G-A------GG--G-AA-AAT- 3677
VER.KE.x.9063                  --A-----A-----G-G-GTA--A--TC-----------G------ACTC-GACC--G-----T---CC-----AG--TGTA-AT-------------CTG---T--AA-CT-A---GA----------------CA----AC-ATAT-----A---G-G----A-AG-AAT- 4182
VER.KE.x.AGM155                --A-----C-----C-G-GTT--G--------A-CT--G-----TACAC-GACT--G-------CCCG-----AG--TGT--A-------------A--CTG---T--AC--T-G---GAG-------------C--CA-A--CC-ACA---C-A-----G---ACAACAG-- 4174
VER.KE.x.TYO1_patent           --A-----T-----C-GC-----A--T--GC-C---C-------ACA---ACA-------------CCCAAG-AG--CGTT-A-------------A---TGC------AA-TT-A---GA----------------CA-CT-AC-ATAC-----A---G-G----A-AA--AT- 4179
COL.CM.x.CGU1                  ---ATA---------CTCAT--T--TATG-----A------GTA-GGA-C--G-TTTT---G--TC-GAAG-AG---CC---T-AT-GGAC------A-G-TGCTC-----A-ATTAGG-GCG------G--------GGA-A-A-TCA-A----GA-G--G-----A-----AT-- 3724
```

```
H1B.FR.83.HXB2            AACTGACACAACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGACTCACAATATGCATTAGGAATCATTCAAGCACAACCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGAGCAGTTAA 4125
Pol                       __T__D__T__T__N__Q__K__T__E__L__Q__A__I__Y__L__A__L__Q__D__S__G__L__E__V__N__I__V__T__D__S__Q__Y__A__L__G__I__I__Q__A__Q__P__D__Q__S__E__S__E__L__V__N__Q__I__I__E__Q__L__
H1A1.UG.85.U455           ------G----------A--------A-----T-----CC-------C----------A---C---------T-------------G-----------G--------G-----------CAGG-----------AA----------------------A--C--- 3571
H1B.US.90.WEAU160         -------------------------------------T---C---------------------------G-------C--------------------------------------------G-------------------G------------A-------- 4124
H1C.ET.86.ETH2220         ------A-----------------A--------G--CC-G-----G--A--------A---C--------------------------G-----------G-----C-TG-----------A-G-----------A----------------------A-------- 3517
H1D.CD.84.84ZR085         T---------------------------------------A-------------------------------------------T-----------C-----------------------A-G---------A-------G------------A-------- 3647
H1F1.BE.93.VI850          ---G--G--C-----------G------------------T--------------A---C-----G-----G------------------------------------G--G---------A-G---------AA-----------------------A---- 3465
H1G.SE.93.SE6165          ------A--------C--A---G---A-----G------C-G------------A--C--AA---C--------------------G-------G------C-----------------------AGG------G----A--------------------A---C--- 3522
H1H.CF.90.056             ---G--A-----------------A-----------------------------A----A--G---------G---------G--------G----T-----G------C----------C---A-G-------------------T-------------G-A---- 3472
H1J.SE.93.SE7887          ------------------------AC----C--C------T-------A-G------A--GC-----------------------T--------------G--G--T--A--------------A-G---------A----------------------G------ 3439
H1K.CM.96.MP535           ------A------------AG---A-----------C-G-T-------A--------A---C-------------------------G-----------------------G-----A-------------T-----T--------------------A---- 3321
H101_AE.TH.90.CM240       ------G----------A--------A-----T--G--CC--T----C----------A---C--------T-----------------------------G-----------CAGG-----------AG-------C------------G--C--- 3699
H102_AG.NG.x.IBNG         ------G----------A--------A-----T------C--T----C--A-------A---C--------T-------------G----------------G-----------CAGG------------------------------------A--C--- 3650
H103_AB.RU.97.KAL153_2    ---------------------------------------C-----------------------------------------------------------------------------A-G-----------------G-------------------G- 3351
H104_cpx.CY.94.CY032      -T----A-----------------A--------------CT---------------A---C----------------------------G-----A--------------------------AG-----------T----T----------------------- 3491
H1N.CM.95.YBF30           TG-A-----C--C-----A---G-------------T--CCT-A-G--C--A--A--G--A---CGG--T-----------C--T-----T--G-----TA-G-----A-----TT----G------A-------------A--G--G-GC------------AG--C-C- 3719
H1O.BE.87.ANT70           -GA---G--C--C-----A--AG----A---ATG--G--A-TA-----C--A--------CAA-GA-AC------T-----------T-----------------G--CG----CTCCT-C-----TACA--G--------CCCTA----TC-G-------------G-AC--- 4180
H1O.CM.91.MVP5180         -GAA--G-----C-----A---G----A---ATG--TG-A-TAA----C-----------CAAGGAGC---------------------------T---G--C-----ATCCT-C------ACA--G-----C--CCCTA----TC-G--G----------G-AC--- 4155
H1P.FR.06.RBF168          -GAA-GA-----C-----A--AG-A--A--RG-----G-GCTGT----C--AA-A--A--A---AA-C-G-CT----------------C-------TG-----G-----CTC----ACT----------------C--CCCCC----GC-AA-----------AG-AA-G- 3586
CPZ.CM.01.SIVcpzCAM13     -GAGA----C--T-----AC--G-A------GGG--CT-G-----------A--A--C--T-A-CA-C----G-----------T-----T-------TG-----G--T-----GT--------------C-----------A-----A-----------------AG-A---- 3705
CPZ.CM.05.SIVcpzEK505     CT-A--A--T--C--C--AC-AG----A-----G--TG-CCTCA-G--C--A--A---G-A---CA----------T-----C--T-----T--G---TCC--------A-----CT-------------A-------------A-----A-----G--------AG-AC-T- 3697
CPZ.CM.05.SIVcpzLB7       ---A--A-----C---------G------------C---C-AT----C--A--A--C--A---CCT--G-----T-----T--T--T-------------------C--------C----G-----G-----T-----A--G--G--------C-----AA-TC--- 3680
CPZ.CM.05.SIVcpzMB66      ---A--A--C--T-------AG-A--A--------T--AC-AT-G-----A--A-----A-A-G----------T-----C--T--T------------C-G-----A-------GT-----T--AC-------------AA----A-----G---------G-AC-G- 3671
CPZ.CM.05.SIVcpzMT145     TGAA--T--C--T-----A--AG-A--A---------G---TAA-G--C--A-----C--A---CAG-----C--TG-G-----C--TG-G--C--T-----C--------TGC-G------T-A---T-----C-----C--C---------TA----A-----G--------AG------ 3667
CPZ.CM.-.SIVcpzMB897      ---A--A-----T-----A-GAG-A--AC-G-----T--A--C-----CC-A---------A-A-G----G-----------------G-----------T-----T-----------------G----------T----G--T--CATT-----------CA----A-----G-------------C--G 3691
CPZ.CM.-.SIVcpzDP943      -GAA--T--C--T---------AG-A---C-GAC---TG----------CC-A--A---G-A-A-CATAT---T-----------C--T--T--C--------TG--------A-----CT-T--G----------------G-----A--G--T---TT----------AG------ 3711
CPZ.CM.98.CAM3            -GAAA-T--C--T---------AG-A--AC-GAC---TG----------A--A--A---G-A-A-CATAT---T--T-----C--T--T--C--------TG-----G--A-----CT-C--G----------------G-----A--G--T----T----------AG------ 3536
CPZ.CM.98.CAM5            -GAAA-T--T--T---------AG-A--AC--AC---TG------G--A----------G-A-A-CATAT---C--T-----T--T--T--T-------TGC-------A-----CT-T--G-------------------------A-----AGGGTCT--------AG------ 3827
CPZ.GA.88.GAB1            -GAAA-T--C--T-------C-AG-A--A---A-G--TT-GCT---G--C-------------A-ATCA-C-G--T-----G--T--------------TG-----G--T-----GT-----G----------C-----------A-----------G--------AG------ 4186
CPZ.GA.88.GAB2            GGAA--G--C--T--C--A---G-A--A---G-G---C-CAGAA---CC-A-----C--A--GCA---------TG-G-----------C--------TG--G---------AA-TT-C--------C--------G-----AA----G----------------AG--C--- 3518
CPZ.US.85.CPZUS           -GAAA-T--T--T-----A---G-A---C-CAC---GG-A--CT----A--AA-A-----A-AGAATAC---T--TG---C--T-G---C-------TC------------C--CT-T--G-----C-----------------A--G--------G----------AG-A---- 4185
CPZ.CD.90.ANT             -CAAA-G--C--C-----AC-AG-A--A-----G--G---CT-A-G---C-AG-------AAC-GGCCC---C--T-----C-----T----------------G---G--T-G---GTACC-----------------------CCCC----GG-AG-------CC--A-A---- 3564
CPZ.TZ.00.TAN1            -GAGA----C--T-----AC-AG-A--A---TGG---G-AG-------C--AA-A--C--A---GC-C-G-----T--C-G-----T-----C-------TTA-G---G-TT-A--G-G-TT-----------------C--CCCCA----AG-G-----T--TC-AA------ 3769
CPZ.TZ.2001.TAN2          -GAAA----T-----C--AC-AG-A--AC--TGG-----AG--T-G-----AAGA-----A---GC-C-G--G------C-G--G--------T-----T--T-------TTA-----G--C-A----G-TTG--T-----G-----C--CCCCA----AG-G-----T--TC-A--A--G- 3774
CPZ.TZ.2002.TAN3_1        -GAAA----T--------AC-AG-A--AC--TGG-----AG-C--G--A--AAGA-----A---GC-C-G--G------C-G--G--------C-----T--T-------TCA-G---G--T-A----G--TG--T--C---------C--CCCCA----AG-G-----T--TC-A--A--G- 3774
GOR.CM.2007.SIVgor2139_287 -GAG-G------T--------AG-A--AC--G-G---G-CCTCT-G--C--AG-A--A--A---AA-A-T--T-----A-T--T---------G--------C--G----T-C----G--T---GC----AGT--T-----------C--CCCC-----GC-GA-----------G-AA-G- 3662
GOR.CM.2007.SIVgorCP2135con -GAG-G------T--------AG-A--AC--G-G---G-CCTCT-G--C---G-A--A--A--GAA-A-T--T--T--A-T--T--T-----G-----T--C--G----T-C----G--T---GC----AGT--T-----------C--CCCC-----GC-GA-----------G-AA-G- 3661
GOR.CM.2004.SIVgorCP684con -GAG-G------T-----A--AG-A--AC--G-G---G-CCTCT-G--C---G-A--A--A--GAA-A-T--T--T--A-T--T--T-----G--------C--G----TGC-G--G--T---GC----AGT--T-----------C--CCCC-----AC-GA-----------G-AA-G- 3581
H2A.DE.x.BEN              -GAAC-A--T--C------C--G-A--A---G---TCT--CGGA-G--AC--GCA--C--A--CCC-A-G--T--T--A-G--T--T--CA--GT---T-----G----T-A-G--G--AG-AGC--GC--G---ACAG-GTCA---AATAGAA----G--C--G--C-----AG-AA-G- 4507
H2A.GW.x.ALI              -GAAC-A--T--C------C-AG-A--A---G----CT-CGCG--G--AG-AACA--C--A--TCC-A---CC--T--A---CC--T--TA--GT---T-----G----T-A-G-----AG-AGC--GC--G---ACAG--TCA---AATAGAA----A---------C-----AG-AA-G- 4500
H2A.SN.x.ST               -GAGC-A--C--C------C-AG-A--A---G----CT--GCGA-G--AG-AACA--C--A--TCC-A-G-CC-----A-G-CC-----TA--GT---------------T-A-G-----AG-AGC--GC------ACAG-GTCA--GAGTA-AA----A---------C-----AG-AA-G- 3952
H2B.CI.x.EHO              -GAAC-A------------C-AG-A---C-TG-----T--GCA-----AC-A-----C--A---CC-C-G--C--T--C-G--C--T--CA--GT---T----------TCA-G-----AG-AGCT---------ACAG---CA------CC-A----A-GAG-------T--AG-AA-G- 4478
H2B.GH.86.D205            -GAAC-G---------C--AC-AG-A--AC-TG-----T--GCAT----A--AACA--C--A-A-CC-C----T------C----T-----CA--GT---T----------TCA-G-----A--AGCT----G---ACAG---CA------CCAA----AGCAA-------T--AG-AA-G- 4480
H2G.CI.92.ABT96           -GAGC-G---------C---C--G-R--AC-TG-G---T--GCAA-G--C---------C--A--TCC-A----C--T--A----C--T--CA--GTR--T--C--G----TCA-G--R-----AGC--GR------ACGG---CA-----CCC-------A----A--------T--AG-AA-R- 3834
H2U.FR.96.12034           -GAAC-G--C--C--C--AC-AG-A--A---G----CT--GC-A-G--C--AG-A-----A---CC-A------------A-----------T--G-----T----------AT-A-G-----A--AGC--G---G---ACAG---CA--G--CCCAC----A-G-A-----C-----AG-AC-T- 3986
SMM.SL.92.SL92B           -GAAC-GG-T--C------C-AG-A---C-G--G--CC-G-TAT-G--C---A-A----TCC-AGTA-G--C--TG-TA-G--C--TG-------G---------G-----TC--GAAC--A--AAC--G---G---TCAG--TCG--T-----TA----GGCA-----T--T--A--AC--G 3855
SMM.US.x.H9               GGAAC-G--T--C-----AC-AG-A-----GG----CT----------C--AGCA-----A---CC-A---C---T--A---C---T--TA--GT---T--C-------TTA-G--C--AG-AGC--GT-----CAC-G--TCA---AGTAGA-----A--C--G---------A-AA-G- 3897
SMM.US.x.PGM53            GGAAC-G--T--C-----AC-AG-A-----GG----CT-C--------C--AGCA-----A---CC-A-G-CC--T--A-G-CC--T--CA--GT---T--C-------TTA-G--C--A--AGC--GT-----CAC-G--TCA---AGTAGA-----A--C--G---------G-AA-G- 4341
SMM.US.x.SIVsmH635F_L3    -GAAC-G--T--C--C--AC-AG-A-----GG----CT-C--------C--AGCA---------CC-A---C---T--A---C---T--TA--GT---T--C-------TTA-G--C--AG-AGC--GT--G--CAC-G--TCA---AGTAG------A--C--G--------AG--A-G- 4427

MAC.US.x.239              -GAAC-G--T--T-----AC-AG-A--A--GG-----T--CTCA-G--A---ACA--C--A--GCC-A-G-C---T--A-G-C---T--TA--GT---T-----------TTA-G-----A--AAC--G-TGC--TACAG--TCA--GAGCAG-C----T---------------AG-AA-G- 4424
Pol                       __L__E__Q__T__T__N__Q__Q__A__E__L__E__A__F__L__M__A__L__T__D__S__G__P__K__A__N__I__I__V__D__S__Q__Y__V__M__G__I__I__T__G__C__P__T__E__S__E__S__R__L__V__N__Q__I__I__E__E__
MAC.US.x.251_1A11         -GAAC-G--T--T-----AC--G-A--A--GG-----T-CCTCA-G--A--AACA--C--A--GCC-A-G-C---T--A-G-C---T--TA--GT---T-----------TTA-G-----A--AAC--G-TGC--TACAG--TCA--GAGCAGAC---T--C---------A-AG-AA-G- 4423
MAC.US.x.251_BK28         -GAAC-G--T--T-----AC-AG-A--A---G-----T--CTCA-G--A---ACA--C--A--GCC-A-GAC---T--A-GAC---T--TA--GT---T-----------TTA-G-----A--AAC--G-TGC--TACAG--TCA--GAGCAG-C----T--C-----------AG-AA-G- 4400
MAC.US.x.EMBL_3           G...ACA--T--T-----AC-AG-A--A--GG-----------AT-GCA-TG-A--C--A--GCC-A-GAG---T--A-GAG---T--TA--GT---A-T----GTATGT-AT--G-AT-A-A-CAGGTTC--TACAG--TCA--GAGCAG-C----T--C-----------AG-AA-G- 3895
STM.US.x.STM              -GAAC-G--T--T-----AC-AG-A--A---G-----T--GCCA-G--C---GC---C--A--TCC-A---C---T--A---C---T--T--GGT-------------TGA-G---------AAC--G---G--CACAG--TCG--GAGCA--------A-----G--------AG-AA-G- 4071
MNE.US.x.MNE027           -GAAC-G--T--T-----AC-AG-A--A--GG-----T--CTCA-G--A---ACA--C--A--GCC-A-G-C---T--A-G-C---T--TA--GT---T-----------TTA-G-----A--AAC--G-TGC--TACAG--TCA--GAGCAG-C----T--C-----------AG-AA-G- 3892
DEB.CM.99.CM40            -GAGA-T--C--T--------AG-A--A---G----C--A-TGT-G--C---G-A--AGGCCC--C-A--A-G--T--A--A-G--T--------------C--G------C-G--G--A---TT--A---C---TC-G-G-CA---CATA--A----AG-AA-GG-----C--GCC--GC 3931
DEB.CM.99.CM5             GGAAA-T--T--C-----A---G-A--A---G----C--CCTA--G--C---G---GGGACCC-CCA-GA-G-------------------C--G-----T---------AT-GTT--A---C--CACGG---CA---CAT---A----AG-AA-G--C----C-AGCC---C 3925
DEN.CD.x.CD1              GGAA--A-----C-----A---G-A--A---AT--G-G-C-TAA-G--C--A--A---AGTCCCAACA-G--C--------C--T--T------------CC----C-----AGC--GTAC---CAC-AC-TC----CATC--A----GG-A--G--C----C-AGCTA-GC 4015
LST.CD.88.447             ---A-----C--C-----AC-AG-A--A---G----TG-C-TAT-G--A--AAGA--CAGTAAGAGCAGG--C-------------------------TGA--AA-G-TT-ATCTCA--GG--TACAG---CA---CATCCTG----A--A--G--T------G-A-GT- 3396
LST.CD.88.485             ---A--T--C--C-----AC-AG-A--A---G-G--TG-C-TAT-G--A--AA-A--CAGTAA-AGCAGG--C-------------------------TGA-GAA-G--T-ATCTCA--GG--TACAG---CA---CACCCTG----G--G--G--T------G-A-GT- 3396
LST.CD.88.524             -CAA--T--T--T------C--G-A--A---G-G---G-A-TGT-------A--A---AGCAA--C-AC---T----------T-----C-------T-A-GAAGG--T-AAGTCAGAG---TACAG---C---GCATCCTA----G--GA-C--C---A--G--C-GC 3393
LST.KE.x.lho7             ---------C---------C-AG-A--AC--G-G---G--CTAA-G--A--A--A---AGCAATAGTA-------T-----------T-----------T-A-GAA---AT-GTC-CA-AG---ACAG---CA---CACCCCA----A--AG-C--C-----A--A-GC- 4480
SUN.GA.98.L14             GGAAA----C--C------C-AG-A--AC--G-G--T--AA-AA-G--C---G-----AGTA-GAGCTC---T--T----------G-----C---A--C-GC-CAGTAAGAG---CACAG---CA--T------C----A--GG-G--CG------TCA--- 4479
MND-1.GA.x.MNDGB1         TCTAA-T--C--C------C-----A--A-----T---G-AA-A------C-AG-A---AGT---AGTT-T--T--T--T-------------T----------T-G----T--AC-AGC-AGCAG---TAC-G---CA--TCACCCCA----A--GG-----------ATTAA-G- 3812
OLC.CI.97.97CI12          -CAG--G-GT-GT--------AG-A--AA--AG---T---ATAGA---CC-TAG---C--AAA-AGT--G--C--T--G--C--T--T--------T---TTG-----CA-G-----AG-GTT--AGA-------AGTGTC----GA-CCCC-CA-AC-AG--T-----A--G----GG 3990
WRC.CI.97.97CI14          GCAAA-T-----T------C-A--A--A--GAT----T-AA-A-----AC-AG-A---GG---TCA-A--A-G--T--A--A-G--T---A-C------AGT-------TCA-G--G--AG-AA---AT--T--CACAATG-----G--TCCT-----AC-GG---T--T--ATT------ 4419
WRC.CI.98.98CI04          -TA-A----C--T------C-AG-A--A--GCG--TT-GA-GT----CC-AG-A---GG---T-C-A-GA-C--T--A-GA-C--T--TA-T----TAGT-------TTA-G-----AG-AA---AC--C--TAC-ATT-----G----CTC----AC-AG---T-----T-TTC--- 4434
WRC.GM.05.Pbt_05GM_X02    C--AA-T--T---------C--------C-GATT---T-AAGAA-G--A--AG-A--CAGTTC-GATA-G-----T--TA-G-----T---A-C------AGT-------TT-------------AA---AT--C---AC-------------TCCC---A-GC-AG---------C-TTC--- 3348
MND-2.CM.98.CM16          -GAGA-T--C-----C--A--AG-A--A--GG-G-----AA-AA-G--A--A--A--C--A---CC-A--------T--A-------T-----G-----T-----G-----TA-G--T---C-ATCG---GC---C--C---TCA--CAACCCCA----A-GAG----T-----A-TCA-G- 4126
MND-2.GA.x.M14            -GAAA-------T--------AG-A--A---G-G-----ACTGT----C---A-A-----A--TCCTAGG-C---T--TAGG-C---T-----C-----------------------C--AC--TC---CACT-------G-TCA--CAACCCTA----T-GAG-------CA-C--AA-G- 4053
MND-2.x.x.5440            GGAA--A--G--C-----A---G-A--A--GAC---T--AA-GA-G--------A-----A---CCCAGG--G--T--CAGG--G--T-----C--T--T--C--G------C----C--AC--TC---GCC--------GTCA--CAACCCCA----T-GGG----------A-TCA-G- 3686
DRL.x.x.FAO               -GAA--A--C--C---------G---A-----G--C---CTGT----C---A---C--T--TCC-A----C--T--A----C--T--G--G--T-----G--C--CA-G--T--A--ATCCT-CTCC-----AATC-----CAATCCTA----G---------T-----A--AA-G- 3692
MON.CM.99.L1              -GAAC-G--C--C--C--A---G-A---C--G-G--CG-G-TA-----CC-CAGA---GGACC-CCTC-GA-G------A-T-----T--T--G----T-C-G--C--TC-AGCCT-CTGC--G--GATCTCA---AGTCCCA----GG-GGCC--C---C--G-AC-CT 3828
MON.NG.x.NG1              -GAGC-G--T--C--C------G---AC--G-----G-GCT-T----A-----AA--GG-CCCGCTAGTA-G--T--------------G----TCA-G------C-AGC-T--A-G-----GGTCTCG--G--CCCAA----AG-GA----C--CC------C-CC 2421
GSN.CM.99.CN166           -GAA--A--T---------C-AG-A--AC-C--T---G--CT-T-G--A--AA-A--AGG-CC-CC-A--A-G--TC-G--C-----T--C--G----T-A-G--G---T-G---AGT--G-----GGT-TCCACC--T-C-C----GG-A-----C---C-AG-AC-GC 3853
GSN.CM.99.CN71            -GAA-GA--T--------AC-AG-A--AC-C--G---G--CT-T-G--A--AA-A--AGGACC-CC-AG-A-G---T-G--C-----T--C-----C-T-A-G--G---T-A---AGC---------AGT-TCCACC--TCCC-----GG-A-----C---C-AG-AC-GT 3835
TAL.CM.00.266             -GAAA-A--T--C--C--A---G-A--A--GG-----G-ACTAT-G--AC-CA-A-----A--GCCCA-G-----T-----T-----C---------T--G---TC-AT--C--AT-GGC--G-AGC--CACAG---CA--TAACA--A--A-AG-AG-C--T--CC---TAC--C 4091
TAL.CM.01.8023            T-GCA-T--T--T---------G-A--A---G-----G-GCTAT-G--A--AA-AC-C--T--TCC-A-G----------------C---------T-------TC-AT--C---T-AGT--G-A-T--CACAG-G-CA--CAACA--A----GG-AG-C--C---C-----C--T 3589
MUS-1.CM.01.1085          -GAAAGT--T--C--C-----AG-A--A--GG-G---G-C-TAT-G--C--AA-A--AGG-CCTCCTAGGA-G--TT---------C-----T--T--G----TC-----G--AG-AGCTT-T--G---C-AG--TCCACT---CCA-----AG-G---------C--G--C--C 3815
MUS-1.CM.01.CM1239        -GAAA-T--T--C-----A---G-A--AC--G----TG-C-TGT-G--C---A-A--AGGCCCTCCCAGCA-G--T--------C-----T--C-------TGC-G--G--AG-GGCTT-T-----CC-AG--TCCACC--GCCC--G--AG-AG-------TC-A--A---C 3805
MUS-2.CM.01.CM1246        GGAAA-------C--C---C-AG-A---C-C-----C---CTCT-G---C-AAT----GG-CC-CC----A-G---C---------T--T-------------G-----A--CAC----GC------GGT-TCA--G--CCCTC-G--GG-G--G------C----AA-GC 3913
MUS-2.CM.01.CM2500        -GAAA-T-----C--C--A---G-A--A--G--G--T--A-TAA-G--CC--ACA---GGACC-CC-A--A-G--------G-----T--C--G-----CC-G--T-----CAC----GC------GAC-TCA--G--TCCCC-G--AG-GA-G--C---C-A-----GC 3872
SYK.KE.x.KE51             GGAAA-------C-----A---G--------G-----G-A-TAA-G--A--A--A--C--A---CC------T-----------------T--T--G---------------AT-AATGAACTGT--TACATCCTCG---CACCCAG-----G-A--G--T--GC-AA--GCT- 3980
SYK.KE.x.SYK173           GGAGC-A--------C--AC--G-A--A---G-G-GCT-GCTCA-G--A--------C--AAA-GACA-------T-----GT---T--------CT-T-AT--C--TC--ATGA--TGC--CAC-A-C-CA---CACCCAA----GG-G--G------C-AG--GCC- 4325
RCM.GA.x.GAB1             -GAA--A--T--T-----A--AG-A--A---G-----G-A-TA--G--A--AA-A-----A--GCCTA--------------C-----T-----G----TT-AT-----TC-AG---------------ACT-----C----G------G-CAG----T---A-T---A-G- 3604
RCM.NG.x.NG411            -GAAA-T--T--C-----A--AG-A------GG----G-A-TA--G--C--AA-A--C--A---AA-AGG--------------------T--T-G-C---TG--T--C--AT-AGC--G-AGT-----CAC------GAGCCCCC----AC-A--G--T-----A--AA-G- 3619
GRV.ET.x.GRI_677          -GAAA----C--T--C--AC-AG-A--A---AC----G--A-GA-G--A---G---CAGT---GA-A-T-----T-----T-----C-----T--T-------T-A-GAAC---T-GAC----TGT---C-GG-----A-C---CCC-----GG-A--G------C-AGCCC--- 4298
SAB.SN.x.SABIC            -GAGA-T--T--C--C-----AG-A------G-G--C---CTGT----C--AAGA--C--T---AGCA-------T---A-------------------TA-G--T--T--AGC--GGG---TACAG--TCA--TAATA-CA---AC-A-----T-------G-AC--- 4523
TAN.UG.x.TAN1             -GAG--T--C--G--C--AC-AG-A--A---------G-CCT------A--AA-AC--AGTAAGAGCA---T--T--------G--T-----G----TCA-G------C-GAC----------TACA---------C--C-T-C----A-----------T--AG-AA-G- 4324
VER.DE.x.AGM3             -GAAA-T---------C---C-AG-A--A---ATG--C--AA-AA-G--AC-AG-A---AGT--GCCTA-T-----------------T-------------A-G-----AT-AACT--C-----CACA--G-----C---CCC---A-AG-A-----T----CA-TAA-GG 3847
VER.KE.x.9063             -GAGC-G--T-----C---C-AG-A--A--GGT---C--AA-AA-G--C---G-----AGT---GA-A-----------------------T-----------CA-G-----AT-GAC------G--CACA---------C--CCCA-----AG-A--G--C-----CATTAA-GG 4352
VER.KE.x.AGM155           -GAAA-------------AC-AG---AC-GAC---C--AA-AA-G--C---G-----AGC--CCCTA----C--T--C--T-----------T-----------GA-G--C--AT-GAC------G--CACA--G-----C--CCCAC----AG-A-----------CA---A-GG 4344
VER.KE.x.TYO1_patent      -GAAA----T--C------C-AG-A--A---AC---T--AA-AA-G------G-A--CAGT--GCCTA-T--G----TA-T--G----------------T-----------A-G-----TT-GAC---------CACA---------T---CCA-----AG-G-----T-----CCTTAA-G- 4349
COL.CM.x.CGU1             TCA---AGGGT--------C-AG-A---C-GAC--G-G---TG------C-AA-AC--GG-CC-AA-AG-A-G--TT-G---------AGT-TT-----------G--A---ACT-GG--G---T--G-TCAGCTTC--TCTAGTAGAGGGG-G--T--TC--GCAGGT- 3894
```

```
                                                                                                     Pol p66 RT, Pol p15 Rnase H end Pol p31 Integrase start
H1B.FR.83.HXB2                TAAAAAAGGAAAAGGTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGCTGGA...ATCAGG......AAAGTACTATTTTTAGATGGAATAGAT...AAGGCCCAAGATGAACATGAGAAA...TATCACAGT 4280
Pol                           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  .  .  K  V  L  F  L  D  G  I  D  .  K  A  Q  D  E  H  E  K  .  Y  H  S
H1A1.UG.85.U455               --G----------A-----C---T-------------G-------G--------------G---------------------T----..------.....--G--G-------------G------..-----T-----G--C-----A---..------T-C 3726
H1B.US.90.WEAU160             ------------------C--A----------------------------------------------------T-A----..------.....---------------------------------..-----------A-------------..--------- 4279
H1C.ET.86.ETH2220             ---GC-------G------C---T---------------T---------------------------------A---AG---..------.....------G--G---C--------------G------..-----T-----A--G-----A---..---------C 3672
H1D.CD.84.84ZR085             ------------------C----------------T---------------------------------------AG----..------.....----------------------G------------..-----T-----G-------------..---------C 3802
H1F1.BE.93.VI850              --C---------G------C---T------------------G-------------------------------------..G-----.....---A----G------------G------..-----A-----A-------A---..--------AC 3620
H1G.SE.93.SE6165              ------------------C---T------------------G------------------C-----------AG----..------.....----------C------C--------..--A-------A--G-------G-..--------AC 3677
H1H.CF.90.056                 ----G-------------C---T-----------------------------------------T---T-----..G----A......-----G------C------G-------..--A--T-----A-------A-GG..------T-AC 3627
H1J.SE.93.SE7887              ----G-------------C---T-G----------------------G------------------T-----..------.....-----G--G---C------G-------..--A--T-----A--T-----A---..------T--C 3594
H1K.CM.96.MP535               -----------G-A----C---AT--------------------------------------------------..------.....-----G-----C--------------..------T-----A-------A---..------T-AC 3476
H101_AE.TH.90.CM240           ----------G--A-----C---T-------------G--G------------------------------T-A----..------.....--G--G-------------G-------..------T-----A-------A-G-..---------C 3854
H102_AG.NG.x.IBNG             ----------C--A-----C---T------------------G---------------------------AA----..------.....--G-------------------C--------..--A-------A--G-----A-G-..---------C 3805
H103_AB.RU.97.KAL153_2        ------------------C----------------------------------T-----------------------..------.....G-------------------------..-----A-----A-------------..------G-- 3506
H104_cpx.CY.94.CY032          --CGG-----C-------C---T------------------G-------------------------------CAA----..------A.......--G--G------------------..-----T-----A-------------..---------A- 3646
H1N.CM.95.YBF30               -----------GA--T-----CT-T--------T------T-----T-----------G--G-----C--------T--CT-A---..--T--A.......--A--T----CC-------T-----A..--A-------A--T-----C-G-..---------C 3874
H1O.BE.87.ANT70               C----------C----G-----TA-------T--T--C--T-----C--A------------A--A-------------A--CAAG-AT..--T--A......-G---C--G--CC----A-------C..C----A-------A--T-----A---..-----T--- 4335
H1O.CM.91.MVP5180             C----------CGA--G-----TA-------T--T--T-------C--A------------A--A-------------A--CAAA-AC..--T--A......-G---C--G--CC-G--A----------..C----A-------A--T-----A---..-----T--- 4310
H1P.FR.06.RBF168              C-------------A--A--C--AT-------G-----------C--A--G--T------A-TA---------C--------AAA-AC..--T--A......-G---GT-----C------A--C-----C..C-A--A--G--G--T-----A---..--------- 3741
CPZ.CM.01.SIVcpzCAM13         -T---------C-----T--C--TT-C--------T-----T--G--C--A-----------------G----------T----..------.....-----------CC-G-------------..--A--A--G--A-------A-G-..------T-C 3860
CPZ.CM.05.SIVcpzEK505         -------------GAA-A-----TT-C--------T-----T------T-----------GA-----C--------CA-A---..------.....------CT----CC----------T----..------T-----A-------A-G-..------T--- 3852
CPZ.CM.05.SIVcpzLB7           -------------A--T--CT-AA-----------T-----T--G-----A--------------C--------T-A---..--T---.....------C--T---C-G-----------.CGA-----G--A--G-----A---..----C--T-A- 3835
CPZ.CM.05.SIVcpzMB66          -C------A--G--A--G-----CT-----------T-----------------------A---------------CT-A---..--T---.....------C--C--CC-G----G-----------..-----G--A--G-----A---..----C--T-A- 3826
CPZ.CM.05.SIVcpzMT145         -------------AA-T------AT----------------G---T--A--------------G---------C----T---T-A--G..-------A......-----T--T--C-----------A..--A--T-----A-------A-G-..----C---TC- 3822
CPZ.CM.-.SIVcpzMB897          -G-------------A-T-----TT------------------G-------------------A----C-------T-----A---..------.....------C--T--CC-G-------------..-----------A-------------..----C---AC 3846
CPZ.CM.-.SIVcpzDP943          -------------AAG---C--CT-------G------------------------------G------------CT-----..--T---.....-G---T--C--C-----------T-------..------A-----A-------A---..------T-A- 3866
CPZ.CM.98.CAM3                ----G----------AAGT--C--TT-------G--------T-----G--------------------------T----T----G..A----A.......------T--C--------------T-------..T---CA-A-G--A---------G--A..-----T-AC 3694
CPZ.CM.98.CAM5                --------------AAGT--C--CT-------G-----------------------------------C----T----T-A--C..--T--A.......------C--C--C------------T--C-------..-----A--G--A---------A--G..------T-AC 3982
CPZ.GA.88.GAB1                -T-----A-----AA----C--CT-C-----------------T--A-----------G-----G-----------------A---..------.....------G-----CC------------C..-G-----------A-------A-GG..------T--- 4341
CPZ.GA.88.GAB2                ----G--A--G---------T--A--------------------A-------------------G--------G-------AG---C..-------A......-G--G--------------T-------..------A-----A--C-------G-..------T--C 3673
CPZ.US.85.CPZUS               ----G--A-----AAGT---A-CT-------------T---------------------------A--------C--------T-A--G..-------A......------T--T--CC--------C--------..--A--A--G--A--C-------C---..------T--- 4340
CPZ.CD.90.ANT                 -------GA---C--A-T--C--CT-C----------T----C--A------------------G------C---------CAA--G..-------A......C-----C--C---C-G--A--C----------..--A--T-----G--C-----T---..------T-C 3719
CPZ.TZ.00.TAN1                C-C------AC-GCAA-T-----A----------------C--T------T--A--G--T------G-------C-------G--T---AAAAAT..--T--A.......----A--T-----CC-G----------TA--..G-A--A--G--A--C------T---..---------- 3924
CPZ.TZ.2001.TAN2              CCC-G---AC-GCAA-A-----------------------C--T------C--A--G--C-----GG-------C-----------T---AAAAAT..--T--R.......----A--T-----CC----------TA--..G---------G--G--C-------C---..------T-C 3929
CPZ.TZ.2002.TAN3_1            CTC-G--A-C-GCAA-A-----------------------C--T------C--A--G--C------G-------C-----C-----T---AAAAAT..--T--A.......G---------------CC----------TA--..G---------G--A--C-------C---..---------C 3929
GOR.CM.2007.SIVgor2139_287    C--CG------------A------T-A--T--------------T-------T--A--G---------G--A-----C-----------------AAA-AT..-------A......-G---CT-----CC----A--C-------..C-A--G--G--A--T-----A---..--------- 3817
GOR.CM.2007.SIVgorCP2135con   C--C---A----------------T-A--T--------------T-------T--A--G---------G--A-----C-----------------AAA-AT..-------A......-G---CT-----CC----A--C-------..C-A--A--G--A--T-----A---..--------- 3816
GOR.CM.2004.SIVgorCP684con    C--C---A-------A-----T-A--C-------------T-------T--A--G---------GG--A-----C-----------------AAA-AT..-------A......-G---CT-----CC----A--C-------..C-A--A--G--A--T-----A---..--------- 3736
H2A.DE.x.BEN                  ----G------GCA------G-T--------C------C--T------C--A---------CC-GG--------CC-T-----A----CAA--C..-------A......C------T-----CC-G--AAAG-----G..CCC---T-----G-------A---..------T--C 4662
H2A.GW.x.ALI                  -----------GCCA-----G-T--G-----C------C--------C--A--G--------C-GG--------CC-T-----A----CAG--C..-------A......C------T-G---C----AAA-----G..CCA---T--G--A-------A--G..------T--C 4655
H2A.SN.x.ST                   -----------GCAA-----G-T--------C------C--T------C--A-----------C-GG-G------C--C-----A----CAG--C..-------A......C------T-----CC----GAA-----A..CCC---T--G--G-------A---..------T--C 4107
H2B.CI.x.EHO                  -C------------AA-A---G-A-G-----------T------G---C-G--T--T---C-GG--------CC-CC------G--CCAA---..--T--A......C--A-C------C-----AAA-----A..CCA---T-----A-------A---..------T-A- 4633
H2B.GH.86.D205                -C-------A--GGCA--A---G-A-G-----------T------G---C-G--T--T---C-GG--------CC-CC------A---CAA---..-------A......C-G--CT-G--CC-----AAA-----A..CCA------G--A--G-----A---..-------TG-C 4635
H2G.CI.92.ABT96               -T-----A---RCTC-A---G-A-G-----------T--T-----------A-----T---C--G-G------CCRTC-----A--CCAG---..--T--A......C----------------C--GR-AAA-----A..CCT---T-----A---------A--R..-T---T--- 3989
H2U.FR.96.12034               ---GG--A---GCA--G---G-T-GC--------T--T--T--G--T------T--C---C--G----------C-T--G--G---CAG---..--T-----.......C-----------G---C-----AAA-----A..CCA---T--G--A----------------..---C--T-C 4141
SMM.SL.92.SL92B               -TC-G--A---GCA--A--CA-A-G------T------T--T------C--A------G--CA-TG-G------CCGT------T---CAA---..--A--A......C-G--C--G--CC-----AA-T-----G..CCA---A-----A--C-------C---..------T--C 4010
SMM.US.x.H9                   -T-----------GCAA-T---G-A-----------T------T---RR---A-----A-----C--G-----------C-TC------T---CAA-A-..--T--A......C------C------C-----AAA-----A..CCA--G--G--A--G-----A---..----C--T--C 4052
SMM.US.x.PGM53                -T-----------GCAA-T---G--------------T--T--T-----------A--------------C--G----------C-CC-G--T---CAA---..--T--A......C------C------C-----AAA-----A..CCA---A-----A--G-----A--G..----C--T--C 4496
SMM.US.x.SIVsmH635F_L3        -T-----------GCAA-T---G-A-----------T------T-----------A--------------C--G----------C-CC-G--T--CCAG---..--T--A......C------C------C-----AAA-----A..CCA---A-----A--G-----A--G..----C--T--C 4582
                                                                                                          Pol RNAse H end Pol p31 Integrase start
MAC.US.x.239                  -T------TC-G-AA-T---G-A---------------------T--A---------CC--G--A----CC-CC-----T---CAA--G..--T--A......C-----T--C--C--G--AAAG-----G..CCA--A-----A---------T---..----C--T--- 4579
Pol                           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  Q  V  L  F  L  E  K  I  .  E  P  A  Q  E  E  H  D  .  K  Y  H
MAC.US.x.251_1A11             -T------TC-G-AA-T---G-A---------------------T--A---------CC--G--------CC-CC-----T---CAG--G..--T--A......C-----T--C--C--G--AAAG-----G..CCA--A-----A--G-----T---..----C--T--- 4578
MAC.US.x.251_BK28             -T------TC-G-AA-T---G-A---------------------T--A---------CC--G--A----CC-CC-----T---CAG--G..--T--A......C-----T--C--C--G--AAAG-----G..CCA--A-----A---------T---..----C--T--- 4555
MAC.US.x.EMBL_3               -T---...-TC-G---T---G-A----------------TTAG----T--A---------CC--G--A---G-CC-C----T---CAG--G...T-T--A......C-----T--C--C--G--AAAG-----G..CCA--A-----A---------T---..----C--T--- 4047
STM.US.x.STM                  -T----------GCAA-T---G-A-----------------T------------A---------C-GG-G-------C-T--G--A---CAG---..--T--A......C-G--C------CC-----AAA-----A..CCA---T-----A--G-----A---..------T--C 4226
MNE.US.x.MNE027               -T-------AC-G-AA-T---G-A---------G-------------T--A---------CC--G--A----CC-CC-----T---CAA--G..--T--A......C-----C--C--C--G--AAAG-----G..CCA--A-----A---------T---..----C--T--- 4047
DEB.CM.99.CM40                A-GGG--A---C-AA-T-----A----------------C-----G-----A--------------C--G--------------------AAA---..--A---......GGA-C--C--C--G--AA-GG----G..G-A----------G--C-----A---..----C--TG-G 4086
DEB.CM.99.CM5                 A-G--------C----A-----TAAT--------------C--------T--A------G-----T--G-----C-------G--A---AGA---..--A---......G-A-C------C--G--AA--G----A..G----T-----G--T---------------..--------TG-- 4080
DEN.CD.x.CD1                  A-GGC------G-AA-A--CA-CA-C-----T------C-------------A-----C---AC-G--A-----C--GC-----T---AAA--G..-------A......C-A-C------C------ACA---TCCA..C-A--T-----A---------T---..----C--T--C 4170
LST.CD.88.447                 A---G--A--CC-A--G-----A-G-----------------T-------T--A--G-------C--G---------CC-CC----------AAA---..--T--ACAAAAAC-------G---C---GAAAG-----G..CCC--A--G--A--GT--AGC---..-T---T-A- 3557
LST.CD.88.485                 AG--G--A--CC-A--A-----A-G-----------------T-------T--A--G-------C--G---------CC-CC----------AAA---..--T--ACAAATAC-------G---C---AAA-------G..CCA--A--G--A--G---AGC---..-T---T-A- 3557
LST.CD.88.524                 AG--G--A--TGCA--T--C--A-G-----------------T--T-----T--A--G-----CC-GG-GA-----C-------T-T-AAA---..--T---ACAAAAC-G--CA-G---C-G--GAAG-----A..CCA--TGT---A--G-----GA--G..-T-----A- 3554
LST.KE.x.lho7                 AGC-G--A--TC-A--T-----A-G------G--T--T--T---------A-----T---C--G-G------CC-CC-----A---AAA--T..--A--ACAGAAAC-G--CA-G---C-----AAAG-----A..CCT--AGT---A--------AGC---..--TC--T-AC 4641
SUN.GA.98.L14                 GGC--------CC----A--C--A-GT--------G--------------------A----------C--G--A------C--C-----T--CCAG---..-------AAAGAGAC-G---A-G---A----AAAG-----A..CCA--TGT---A---------G---G..--TC--T-AC 4640
MND-1.GA.x.MNDGB1             A-GG---A-----AA-T---T-AAGT---C-------------------G--A-----G------G--A-------GC-----A---T-A---..-------A.............CT-----CC--C-AAAT-----A..CCA--A--G--A--------------------..------T--C 3967
OLC.CI.97.97CI12              AG--G--A----GAA-A---A-T-G------T---------T--G--CT-A--G--C---C--G---------C------------------AAA---..--A---ACTAAGC----GA-G------G-CA-AC-----G..CCA--A------GAGC--CA-T--G..------TCAA 4151
WRC.CI.97.97CI14              A-GGG--A---G-AA-A--CA-A-AT-----------T--T-------G--A--G--C---AC-G--A--------G--------G------A--G..--T--ATCA...---A-TA-G---C-----AAAT---ACA..CCA--AAT---AAGT---ACT---..----C--TCAA 4577
WRC.CI.98.98CI04              A-C-G--A---G-AA-T---A-T-AG-----G-----------T-------G--A-----T---AC-G--A-T------------A-----A---..--T--ATCA.........---A-G---C-----AAAT---AC-..CCA--GAT---AAGC---ACT---..--------TCAG 4592
WRC.GM.05.Pbt_05GM_X02        A-GGG--A---G-AA-A---A-CA-------G-----------T-------G--A-----------AC-G--A-----GC-------A-----A--C..--T--ATCT...-G---A-G---C-----AAAT---ACC..CCA--AAT---GAGT---ACA---..--------TCAA 3506
MND-2.CM.98.CM16              -CC-C------GCA--A--C--A--T-----------------------C--A-----G-----G----------------C----G--CAGA---..G-A---......C--------G--------A--C--------..--A--A-----A--G-----T---..----------A- 4281
MND-2.GA.x.M14                --GCC--A---GCT--A--CT-A-----------------T--T------C--A-----C-----------A----C---------A--CAAA---..-------A......C----------------CC-G--A---------..--A--T--G--G-------C---..----C--T-A- 4208
MND-2.x.x.5440                --GGC------G--A-T-----CT-------C------C-----G--T--A-----T-----------GA--------------A---AAG---..--A--A......C-G----------------------G------------..--A-------G--A--------C---..----------A- 3841
DRL.x.x.FAO                   -C-GT--A---GCA--T--CT-AAAT-----------T------------A------T-----GG-G-AG--C---C-----T---AGA---..--A--A......C-----GT-----CC------AAC--G--A..--A--A--G--A--------TTT-..----C--T-A- 3847
MON.CM.99.L1                  -----------CC----A-TCT--T---------------T--------------A-----------------G--------CGCC-----GG-CCAG---..--T------......C------C--G---A-G--AAATT-G--A..CCC--AAGG--A--T-----A--G..------T--- 3983
MON.NG.x.NG1                  --GG---AA-GC----A---T--T--------C--G--C--------C--A--G--------------G----R--C----------A--YCAA--G..-------A......C-G--C--C--CA-G--GAAC--T--A..CCR----G--G--T-----A---..-----------C 2576
GSN.CM.99.CN166               ------------GCCA----CT-AAGT-----T--G-----------G-----A---------CAC-G-G--G--C-------------CAAG---..--C-C......C-----G--T--CC-G--AAAC-----G..CCT--AGT---A--C-----A--G..------T--- 4008
GSN.CM.99.CN71                --GGG--A--GGCC-----CT--AGT-----T--------------G--A---------CAC-G-G------C--------------CAGA--G..--C-C......C----------C--CA-G--GAAC-----G..CCC--AGT---G--C-----A---..------T--- 3990
TAL.CM.00.266                 ---GC------GGCA-----CT----T---C-G--G--C-------C--A---------------------G--G-----G-----A--CCAA---..-------.....---GA-C--T---G----ACAG--CCCA..G-A--A-----A-----------G-..----C--T-A- 4246
TAL.CM.01.8023                --GGC------GCT--G--C--TA-----C-------C--T---------A-------------------------C--GC-G-G---CAA---..-------.....---G-A-T--C---G----ACA---CCCA..G-A--A--G--A--G-----A-GG..----C--T-AC 3744
MUS-1.CM.01.1085              -GTC---A---GCCA-A--CT-AT-------G-----C--T--G--T--A----------GG----G--C---------A-A-AG---..-------.....C-----C------CA-G--ACAG-----A..CCT--AA-G--A--T-----A---..--------- 3970
MUS-1.CM.01.CM1239            -G-CC-----GGC---A--C--AT-G-----G-----C--T------T--A----------GG-T--G--C---C-G--A--CCA---..--T---.....C-G--G------CA-G--GCA------G..CCT--AA---G--C-----A--G..----C--T--C 3960
MUS-2.CM.01.CM1246            -G-GC--AA-TGCCA---T-A-AT-CC-A-------C-----G--C--A-----G-----GG-------CC-CC-----T--CAGG---..--C-C......C-G--------------A-G--AAAT-----C..CC---AGTT--A--C---------G..----C--TTC- 4068
MUS-2.CM.01.CM2500            -C-CC--A---GCAA-A--C---T-C-------------C--T--------A--G-----CC-GG-------CC-CC-----T---AAA--G..--C-C......C-G--T--G---A-G--AAAC--T--G..CCA---AT---A--C-----T---..-----------TCC 4027
SYK.KE.x.KE51                 -GG-G--A-------A----CA-AA-C-----T-----C--T-----------A-----------------C--G-------C--------A---AAG---..--T-----.....----A----G---C-G--AAC--CCCA..--A--A--G--A----------------..----------C- 4135
SYK.KE.x.SYK173               ------------GCCA-A--G--A------C-----C--T-----------A-------------GCT--T------C---------AAA---..--A--A......----A-C-----CC-----AA----TCCC..C-A--A--G--A--T-----A-GG..--C---TCA 4480
RCM.GA.x.GAB1                 --GGG--A---GCA--G--C--TT-C-----G--C-----T--G--C--C------------GG-----------------T---AAA---..-------A......C-G-----G--CC-------G-----A..--A--A-----A---------A--G..------T-A- 3759
RCM.NG.x.NG411                --GGG------G-A--G--C--CT-C--------G-----T------C-----G---------G--G-------------G---CAG---..--T--A......C-G-------------CC----C--G----GA..--A--A--G--A----------A---..------T--C 3774
GRV.ET.x.GRI_677              -G------AGGC--------CT-ACA-------------T--T--G--G--A------C---AC-G--A-----------------A--CAAA---..--A--A......C-GA-C--C--C-------A--------A..G-A--A---------C-----CA--G..----C--T-AC 4453
SAB.SN.x.SAB1C                ----G------GGCA--G---A-A------------T--C--T------G-A--T--C-----GG--A-T------C-----T---CAA---..--A--A......C-------------CC-G--CA----T--A..G-A--A-----A--G------T------..----C--TGCC 4678
TAN.UG.x.TAN1                 -T--T--A--TGCTA-----T-AA-C-----------------T------G--A--G-------C--G--A--------GC----A--CCAG--G..--T-----.....C-----G-------C------CA----G--A..G-A------G--ATC------T------..-------T-C- 4479
VER.DE.x.AGM3                 --C----AC-TC--A-A--CT--CA---------------G--------G--A--------C--------G-GA--------C------A---CAA--G..--G---.......--A-TT-----------AAA-------A..G-A------G--G--------A-GG..----C--T-A- 4002
VER.KE.x.9063                 ----G--AAGGGCAA-A--C--ACA-------G--T-----------T--A--G--------------G--A-----C-----------A---CAA---..--A--A......-G---T------A---GAA--------A..G-A---T-----A-----------C-G-..------T--- 4507
VER.KE.x.AGM155               --C-G--A---GCCA-------------CA--------T--T--T-----T--A--G--C-----------G--A-----C------------A--CAAG---..G-T--A......-G-A--T-G--CA-T-GCA-G------A..G-A--A-----A---------T-GG..--------- 4499
VER.KE.x.TYO1_patent          --C-----C--C-AA-A---T--CAG-------------------T--------A-----------------GG-GA----------------G---AAA--C..--T--A......-G---TT-----C-------AAA-------A..G-A--T-----A--G------A-G-..------T-A- 4504
COL.CM.x.CGU1                 -GGCT--A---GCCA-TC--G-A--C---TGT--------T-----G--A--------------TGA-----C--AAG--A..........--.G-A--A......C-G---A-G-GGA-----AA------A..GCA--AG----A--C---C-A--G...T---T--- 4043
```

```
H1B.FR.83.HXB2              AATTGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAA...GGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGTTAT 4447
Pol                          N  W  R  A  M  A  S  D  F  N  L  P  P  V  V  A  K  E  I  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
H1A1.UG.85.U455             --C---------------------T---------G----G--G-----------------A---------A-----...-G---------------------------------G-------T------C-----C-----------A-- 3893
H1B.US.90.WEAU160           -------------------------------------------------------------------------------...----------------------------------------------------------------G------A--- 4446
H1C.ET.86.ETH2220           ------------------A---A-----TA-C-----C------C-C------------TT-------------------...-G-------A-----------A-T----------G---------T------------C--------G---A-C-- 3839
H1D.CD.84.84ZR085           -----------------------------------G------------------T--------------------------...--------A-------------------------------------T---C------------------------ 3969
H1F1.BE.93.VI850            -----------G----------------T---------A-----------------T---------G-------------...-G-----------C----------------------G-------T------C--------------C--G----- 3787
H1G.SE.93.SE6165            --C------------C------------TT--------A-------------------G-----------------------...-G------------------------------------------T-------------------------A---- 3844
H1H.CF.90.056               -----------G----------------T--A------A-----------------T--------------------------...-G-------------------------------C--------T------C--------G------C-------- 3794
H1J.SE.93.SE7887            ----------------------------T----------------------G--------T---------------------G...-G--------------------------------G-------T-------------C---------------- 3761
H1K.CM.96.MP535             --C-------------------------T---------A-------C-----------T-------C-----------------...-G-------A----------------------------------T-------------------------A---- 3643
H101_AE.TH.90.CM240         ---------A------------------TT--------A---------G---------A---A-------------A-----...-G----T----------------------------G-------T------C-----C----------------C-- 4021
H102_AG.NG.x.IBNG           -------AG-------------------T---------A-----------------G------------------A-G---...-G-------------------------G--------------T----------------------G---A---- 3972
H103_AB.RU.97.KAL153_2      -----------------------------------G------------------------------------AT-----...---------------C-------------------------------------------------------A---- 3673
H104_cpx.CY.94.CY032        --C-------------------------T------T-A--G------------G-------T------A-------------...-G-----------------G----------G-----GT-------------------T-------------- 3813
H1N.CM.95.YBF30             -------A---------C------------T-A--C--CA----G---------------------C-----C---------...-G------------------G--CA-T---------G-G-------T----------C-----G------A-C-- 4041
H1O.BE.87.ANT70             -------A----C-A--------A---GGA--A-----A--G--G--C--G-----CA-T--T------CC---------TA-----...-G-----A--T-----T--G--------C-------A-G---------A-------C-------A-G-----C---A-C-- 4502
H1O.CM.91.MVP5180           ------------T-A---------C---GGAT-A-----AA-------C--G-----CA-T--T--T---CC-------C--TA-----...-G-----A-C-------T----------AC--C----AG---------A-G---------------------C---A-C-- 4477
H1P.FR.06.RBF168            --C---------CC-A---------------GG---T-----AA--------------G--TA-CAA---T--CCCC------------G-----...-G------------------G-------C-----------T-----------G--C-----C--CAC----------A---- 3908
CPZ.CM.01.SIVcpzCAM13       -------A---TT----GG----A------T-A-----CA-------------G-----G--T-A-----------------G-----...----G--TC-------G--G--T--T-----C--------C------GA-G-----C-----------------C---A-C-- 4027
CPZ.CM.05.SIVcpzEK505       -------AG---------C------------T--A--C--CA-------T----------------T-----------G-----AT-G---...-----------A--------G---A-CA-------------GG-------T-----------C--------G--G------A---- 4019
CPZ.CM.05.SIVcpzLB7         ------------T-----C--------C--T----------CA-T-------------G-----------T--T--------------------G...----G--A----------------G---------------C-----------GT---------------CC----------------C-- 4002
CPZ.CM.05.SIVcpzMB66        ----------G--C-----C-----------T---------CA----------------G--------A-AT----------G-----AT-----...----G------A--------G-----------------------------T-----C--C-----C--G--------GA-C-- 3993
CPZ.CM.05.SIVcpzMT145       -------AG-----------CAG--C-------T-A-----AA-----G--T--------------CAA----------------------...------------A--------------T-----------C----------GA-T------C--C--------------------A-A-- 3989
CPZ.CM.-.SIVcpzMB897        -------A---T-----AGA---A------TT--------CA----------C----------------------------C--AT-----...-----------------------G---------------------------------C--C---------------------GA-C-- 4013
CPZ.CM.-.SIVcpzDP943        ------------T------A------C--C------C-----------G--C--------------------T-----C--G--C--AT-----...----------------------G---------------C--C--------G------T------C--C--------C--------G-----C-- 4033
CPZ.CM.98.CAM3              ------------T------A----------C--TA-A----A-A-----G--C-------C------------C-T-------------C--A------...----G----------------G-------------C--C------C-------------T------C--C-------CC------------------ 3861
CPZ.CM.98.CAM5              --C----A----T-------C----------C--T--A-----------G--C-----------------------T-----C--G--C---T------...-G----------------------------------------------C--T--------T------C--C----------C----------A-C-- 4149
CPZ.GA.88.GAB1              -------A----T------------------TT-A------CA-----------------------G-----CAT-----------G--C--G------...------------------G--G---------------C-----G--T-------G-----------C--CC--------C------G-- 4508
CPZ.GA.88.GAB2              --C---------C------GA---A--CC-A--C-----G---------C------------------CA--------------------T-----...---------A--A--------G--T--C---------G--C-----G--G--C--C--C--C---------G---A-A-- 3840
CPZ.US.85.CPZUS             -------C----T------C------------------A--G--C--T--G--G------G-----------------------C------...-----G-------A-----G--G------T--C--------T--C------G--------------------G---A-C-- 4507
CPZ.CD.90.ANT               ---------T--T-A--AGA---A-AC--T--T--C---A-T--G--T--------TA-----ACAG---------------CG-----G...----------G-----------------------C-------------G---------C--C----------T------A-- 3886
CPZ.TZ.00.TAN1              -------A----TT-A---GA----A-A----T------C--A--T--G--T--------TA-T--TCAG---CCA------C--TA------...-----G--T--A--------G--G-----AC-------A----C--------A-----C-----C--CC----------G--C-- 4091
CPZ.TZ.2001.TAN2            -------AG--CT-A---GA---G-A-C-T--C--C--AA-T-----C--------TA-T--TCAA--CCC-------C--CA------...------------A-------G----------A-----------AG--------A------C--C--C--CC----G---------C-- 4096
CPZ.TZ.2002.TAN3_1          --C-----G--CT-A---GA---G-A----T--C--T--AA-T-----C--------TA-T---CAA--CCC------------CA------...-------G------A-----G-----G----A---------A---------GA-----C--C--C--CC------------------C-- 4096
GOR.CM.2007.SIVgor2139_287  ---------A--C-A--C-----C---GGT-----G--A--G-----C--------CA-CAAT-----CCC---G--C--TG------...-G-----T-G------------------T--C--C-----T--C-----GT-G------C--T--C----------G---A-C-- 3984
GOR.CM.2007.SIVgorCP2135con -------------C-A--C-----C---GGT-----G--A--G-----C--------CA-CAAT-----CCC---G--C--TG------...-G-----T-G------------------T--C--C-----T--C-----GT-G--C--C--C--C----------G---A-C-- 3983
GOR.CM.2004.SIVgorCP684con  -------------C-A--C-----C---GGK--------A--G-----C--------CA-CAAT-----CCC---G-----TG------...-G-----T-G------------------T--C--C-----T--C-----GT-G------C--T--C----------G---A-C-- 3903
H2A.DE.x.BEN                -T-ATA-A--A-C-AA-CCA-A-A---GGAA-A--C-T-C---------G-C-G------AA-TCA----CCC-----C--A-AG---...-----------A-------G------A-TGCAGAAAT---CG-T-------A-G--C-AC-------C--------------A-C-- 4829
H2A.GW.x.ALI                ---GT--A--A-C-AT-CCA-A-A---GGAT-A--CAA-C-G--G----G-C-G-------AA--CA-----CCC-----------AG--G...----G--T--A-------G------GA-TGCAGAA-T---C-CT-------A-G--C--C-------C------------------C-- 4822
H2A.SN.x.ST                 ---GTA-A--A-C-AT-CCA-A-A---GGA------CAAAT----G----G-C--------AA--CA--CACCC----------AG-----...-G--G--T--A-------G------A-TGCAGAATT---C-CT-------A-G--C--C-------C----------------A-C-- 4274
H2B.CI.x.EHO                ---GTA-A--A-C-A-TCCA-A-A---GGGA-T-----AAT----G----G-C--------AATTC-------------C--A-A-----...-G------T--T-----------G----A-T-CAGAA-T---G-C-------A-G--C----------------G--------G------ 4800
H2B.GH.86.D205              ---GTA-A--A-C---TCCA-A-A---CGGAA-T----AAT----G---------C-G--------AATTC---------------C--A-A-----...-G------T--T-----------G----A-TGCAGAC-T---G-C-------GA-G--C--------------------A------ 4802
H2G.CI.92.ABT96             ---GTA-A--AGT-AA---CA-A-G---GG-A-T--T-AAT--------------C-G-------AATTCA-ACC-C-GT--C--A-AG---...------------T-------------G---A-TGCAGAATTG--T-CT-------A-G--C--C--C-----------------------A-- 4156
H2U.FR.96.12034             ---GTT-A--AGC-ACAGCA-A-A---GGTA-A-----AAT--------C---C--------AAT-CT----C--GTT---C-----AG-----...-----------------------G----A-TACAGAAGT---G-C-------A-G--C------C--C------G------------G- 4308
SMM.SL.92.SL92B             ---GTA-AG-AGT-A---CAGA-A-A----A-C----AAT--------T---C-G------AATGCT---A-C------C--A-AG--G...---------G--A-----T---GAC-A-TGCAGAGGT---C-C-------GA-G--C--C--C--C------------G------ 4177
SMM.US.x.H9                 ---GTA-A-AA-T---TATTCA-A---RGTT-A--TAGGC----------------C-----------A--CA-----------C--C--G---...-----------Y--A-------G-------A-TGCRGRA-T---G-CT-------A-G--C----------CC-----G--C---A-A-- 4219
SMM.US.x.PGM53              ---GTA-A--A-T---TATTCA-A---GGTA-A--TAGGC----------------C-G---------A--CA---C-------C--C-AG---...------------A-------G-------A-TGCAGAA-T---G-CT-------A-G--C----------CC-----G--C---A-A-- 4663
SMM.US.x.SIVsmH635F_L3      ---GTA-A--A-T---TATTCA-A---GGTA-A--TAGGC----------------C-G---------A--CA-------G---C--C-AG---...------------A-------G-------A-TGCAGAA-T---G-CT-------A-G--C------G--CC-------------C---A-A-- 4749

MAC.US.x.239                ---GTA-A--A-T---TATTCA-A---GGAT-A--CAGAA----G--C-G-C-G-------A--C-------------------T-AG---...-----G--T--A-----G--G-C-A-T-CAGA--T---G-CT-------A-G----------C----C-----G------A-A-- 4746
Pol                          S  N  V  K  E  L  V  F  K  F  G  L  P  R  I  V  A  R  Q  I  V  D  T  C  D  K  C  H  Q  .  K  G  E  A  I  H  G  Q  A  N  S  D  L  G  T  W  Q  M  D  C  T  H  L  E  G  K  I
MAC.US.x.251_1A11           ---GTA-A--A-T---TATTCA-A---GGAT-A--CAGAA----G--C-G-C-G-------A--C-------------------T-AG---...----------T--A-----G--G---A-T-CAGA--T---G-CT-------A-G----------C----C-----G------A-A-- 4745
MAC.US.x.251_BK28           ---GTA-A--A-T---TATTCA-A---GGAT-A--CAGAA----G--C-G-C-G-------A--C-------------------T-AG---...----------T--A-----G--G---A-T-CAGA--T---G-CT-------A-G--C--------C----C-----------------A-AG- 4722
MAC.US.x.EMBL_3             ---GTA-A--A-T---TATTCA-A---GGAT-A--CAGAA----G--C-G-C-G-------A--C-------------------T-AG---...----------T--A-----G--G---A-T-CAGA--T---G-CT-------A-G--C--------C----C-----------------A-AG- 4214
STM.US.x.STM                ---GT--A--AGC-A-TCTTCA-G---GG-A-A---AGGC-----------GC--------G-AT-CA-----C--G--C--C-AG---...----------T--A-----G-------A-TGCAGAGTT---G-CT-------GA-G--C---------------C--------C--GA-A-- 4393
MNE.US.x.MNE027             ---GTA-A--A-T---TATTCA-A---GGAT-A--CAGAC----G--C---C-G-------A--CA------------------T-AG---...----------T--A-----G--G---A-T-CAGA--T---G-CT-------A-G----------C----C-----G------A-A-- 4214
DEB.CM.99.CM40              -------A-CA-T-AAGGGA---A-ACC-A--A--TA-AC-CA----T--GC----TA--A-ACAA--CCCC----------AC--G...--G--GC---CA-----T--G--G--TGCCTCAATG--C--C-------GA-G--C--C--C---C--G---AAC----A-A-- 4253
DEB.CM.99.CM5               -------A-CA-T--AGAGA---A---C-GT-A--TA-CT--A----C--GC-G--TA-T--GCAG----CCA--G--C----A----G...------GC-T--------T--------TGCCTCCATG--GG---------A--------C--C--CC------AAC-----C-- 4247
DEN.CD.x.CD1                -------A--AG--CAGAGA-A-G-A-GG---A---A--A-G-----T--G-----CA-TAATGT----CCA---------AC-C-T...--G---C-A--CAGA-----G--G--TGCCTC--T-----C-------GA-G--C--C--C---C------GAAT-----A-- 4337
LST.CD.88.447               ---GCT-A--ATT-A-AAGAAA-A---C--A-A--C---A-G-----T--GC-------GAAT--T---A------------AC---G...------------AACA-----------TAC-TCA-----T--T-------A-----C--C--T--CA-G--G---C-G--A-- 3724
LST.CD.88.485               ---GCT-A--ATC-A-AAGAAA-A---C----A--C---A-G-----T--GC-------GAA---T---A------------AC---G...------------AACA-----------TAC-TCA-----T--T-------A-----C--C--T--CA-G--G---C-G--A-- 3724
LST.CD.88.524               ---GCA-A--ACT-A-AAGAAA-G-A-GGGA-A--C--CA-G-----T--GC---------AATGAG-----C---G---------AG--G...------------AACA--T--------TGTATCAGTG--T--T----------G--------T--CC---------GC-A-C-- 3721
LST.KE.x.lho7               ---GCA-AG-ATC-A-AAGAAA-A---------A--C--CA-G-----C----C----T--CAATGA-----CA--C------AAAG-----...------------AACA---------G--TGTCTCAGTG--T--T-----G------C------T--C---------C-G-C-- 4808
SUN.GA.98.L14               ---GCAGCCT-TT--CAGGAGATG---G--A-A--C-TA--G---------GC-G-----CAATGAA-----CAC---------A-AG---...--G------------AACA-----------TGCCTCAGTG--C--C------GA--------C--C--CA-G----A-------C-- 4807
MND-1.GA.x.MNDGB1           ---GAAGC-CA-T-AAGAGAGA-A--CC--T-A---G--C--------C----C-G--T--GCAA--T--CAG---G--CTGT-ATC-T...-----GC-----AA-G-----GAC---TGC-TCA-TT---G-C------GA-----------C-------C-G---AATC--A---- 4134
OLC.CI.97.97CI12            ---GCACA-AGTT--CAGCAA--G--CC-AA-A-----TAA-----G------AGG---------ATGAA------AG------A-AGGG---...--G-TTC-T-G--------AT---G-CAGTGG-GGTTTAT-GG------GA-----C--C--T------------------G--G--G-- 4318
WRC.CI.97.97CI14            GG-----AG--CT-ACAAGAA--A---C-T--T--T-AAT-GA--------AG-----A---AGGAA---AG--CC--C--AGGG---GGA--G-C-C-A---------TATG-----T-A-GAGGTT--G-C-------A-G--C---------T-----------------G---A---- 4747
WRC.CI.98.98CI04            GGC-----A-----T-ACAAGAA--A---C-TA-T------AAT--A-T-------AG---------AAGAA---AG--CC-----AGG----GGA-----CTC-AC-A-------TATG---------A-GAAGTG---C-------A-G---------------T------------------A-C-- 4762
WRC.GM.05.Pbt_05GM_X02      GGG---G-G----T-A-AGGAA--A---C-T------CA-CA-TC-T--------AG---------AAGAA-----C--CC------GG-----GGG---C-CC-T--T------CATG---------A-GAGAT---G-C-------GA-G--C--C--C--C--G-----------A-C-- 3676
MND-2.CM.98.CM16            --------G---T-A--ACAA--C--CTGTA-A---AACA----G---------------------CAG---CCA--G-----AAC-----...------GC-A--A-----C--G-----TACATCC---------CC-------A-G--C--C---------A-G--------------A-- 4448
MND-2.GA.x.M14              ---------T--T-AAG-CAG--A----G-A-A--C---A-----G------------------T--ACAG--CCCA--G-----G------...------C-T--C---------G--G--TGCGGCC--------C-------A-G-----C--C-----C--------------G------ 4375
MND-2.x.x.5440              -----------GT--T--CAG--C---------A--C--CA----G---------G------T--ACAA----CCA----------AC---G...------C-TG-T-----G--G---A-TGCAGA------G-C------GA-G--C--C--C----------------A-C-- 4008
DRL.x.x.FAO                 ---------T-CT-A--ACAA--A---GGTT-A---GGGA----G--------------------G--ACAG---CC----------AA--C-T...------GC-A--A-------C-------TGCA--C-------CT-------A-G--C--C-------------------GA-A-- 4014
MON.CM.99.L1                -------A-TACT-AAGAGA-ACC-ACC--A-T--TA--C-GT-G--C-----------CAA-CAT---C-C----------AAACC-T...------C---AATCA-----G---A--GCAGAAGT---G-------GA-G--C--C--C----C----G----GCA-A-- 4150
MON.NG.x.NG1                -------A-TATC-CAGGGA-AC--A-GG-A-T---A-AC-CC----C-----G-----CAA-CAT--CCCC--G-----CAC-C-------...------C----AA--------G------Y-CAGA-AT---G-----------A-G------C--T----AT------GTA-C-- 2743
GSN.CM.99.CN166             -------AGTACC-AAGAGA-C-A-AC--AA-C--CG-AC--T-G--T-----------GAA--AA---TCC--G-------G--C-T...--G----C-A-AA------G--G---A-TGCAGAATT---C--G-------GA-G--C--C--G------G---------G--C-- 4175
GSN.CM.99.CN71              -------AGTATT-AAGGGA-C-A-AC--AA-C--GG-AC--T-G--T-------------AA--AG--CTCC--------------A--C-T...--G----C-A-AA-------G-------A-TGCAGAATT---C--------------A-G--C--C-------C-G-------C--------- 4157
TAL.CM.00.266               --------G-ATC-TAAGGCAAGA-----GT-A--CA-CA-------C--G-CT--TA---AAGCA--CCCC---------AG--C------...------C-T-A-ACA-----GAACA-TGCAGC-GT---G-C-------A-G-----C--C------G--G---C--A-C-- 4413
TAL.CM.01.8023              -------AG-ACC-AAAGGCCAGG-----A--A--CA-CA-------C-----CC--CA---AGGCA--CCCC----------G-CC-----...------C-A-A-ACA-----GACCA--GCAGCAGT---G-C-------GA-G-----C--C--CC-G------C-G--A-- 3911
MUS-1.CM.01.1085            -------AGTACC-CAGGGA-A-A-ACC-AA-C---G--T-GT----C-------------AA-CAT---CCA-----C---AC-C-T...------C-A-AAACA--G-----CA-TGCAGAGAT---GG--------A-G------C--C--CC-G--G-------A-C-- 4137
MUS-1.CM.01.CM1239          -------AGTATC-CAGGGA-A-G-AC---A-C---G--T--C-G--T--G-----------AA-CTT---CCC----------AC-C-T...--G--GC-A-AAACA-----G--CA-TGCAGAATTG---G-G-------A-G-------T--C------------------A-C-- 4127
MUS-2.CM.01.CM1246          -------AGTATC-CAGGGA-A-A-AC---A-A--CA-CA-CT----C--G-----------AA--AA--CTCAGC-------AAC-C-T...--G---C-A-AA------G-----CA-TGCAGACTTG--GG--------A-G--------C--------C-G------C--G--C-- 4235
MUS-2.CM.01.CM2500          -------A-TATT--AGGGACA-A-A---TA-A--CA-AC-TC-T--C--G-----------AA---A--CTCAGC---C--AAC-C-T...------C-A-AA-----T-----CA-TGCAGACAT---GG--------A-G--C--C--C----------------------A-C-- 4194
SYK.KE.x.KE51               ---GTAGA-TAC--AAGGCAA--A---C-T--A--CAGACAG-----C-----CT-------AA-TG---CCC-----C--AA-C-G-...------C-A-AA-------G---------ACAGA--T---T-C-------A-G---------C--C------------------A-- 4302
SYK.KE.x.SYK173             ---AT-GAGTATC-TAGGCAG--G---C-T--A--C-GACAG-----------C---TA--CAACAA---CCA---------AAAT-G-...------C-A-AA-----G---------TGTAGA-AT-TAT-AT----GA-G--C-----T---GAG-----C------A-- 4647
RCM.GA.x.GAB1               -------------T------AGAA------C-AA-T----AAA----G-------------------ACAG---CCA---------AG-----...--G-------A--C-------G-----G--TGCA-----------G-CT-------A-G--C--C------------G---------A-A-- 3926
RCM.NG.x.NG411              -----------------AGAA--A---C-AA-A--T-AAA--------------------------CAA--CCCC--G-------G------...-------------AG-A----------G------TGC----------T-CT-------A-G-----C-------C----------------A-C-- 3941
GRV.ET.x.GRI_677            ---------AGT-----TACAG--A---GGAT-A--TAA-A--------------G-----------GGCA---CCC------C--AA---G-...------C-T-A---------G-------GCCTCCATT-A--CT-----GA-G--C--C--C--------------------------- 4620
SAB.SN.x.SAB1C              ---------GAGT---CAACAA--A---GG-T-A--TG--A----------------------GGCA---CC-------C---A-------...--G--GT-TG-A-----G---------TGC-----------G-G-------A-G-------------CC------------A-C-- 4845
TAN.UG.x.TAN1               ------CA-TTT--AAGAGA--CA--CGGTA-A---G-AC-G-----------------------TGCT--CCCA-----------AA-TCG-...------C-G--A-----G---------TGCC---GT---GG-C-------A-G--C---------------G-----------A-C-- 4646
VER.DE.x.AGM3               --------GAACT-A--AGACAC--A-GGG--A----AAA-T--G-------------------TG---CCA--------------A-----...--G---C-AG-C-----G---------TGCCTCG-----GG-------A-G--C--------------------C--G--A-- 4169
VER.KE.x.9063               -------A-AATT-A--GGACACA---GGAT-A----AGA----T--C-----G----------A-TG---CCA--TGCCA-AA-------...--G---C-A--A------G----------GCCTCC-----GG---------A-G--C--C--G--G-------------A-AG- 4674
VER.KE.x.AGM155             --C------AATC-A--AGACACA---GGAT------AAA--------T---------T-----A-TG---CCCA----------AG-----...--G---C-A--A------G----------TGC-TCA------G-G-------GA-G--C--C-----A-----------A-AG- 4666
VER.KE.x.TYO1_patent        -------A-AACC-A--AGA-ACA-A-GGG--T----AAA-------------G-----G----TG---CCA----------A----G...------C-AG------------G--TGCCTCA--T----C-------GA-G---------T---C------------------AG- 4671
COL.CM.x.CGU1               ---GT-CAGTATC-AAAAGAAC-A---GGGT-A--CA-------------------TTGG-AA--A---TCAG-G------AAC---...---C----TG-T-----C--GC----T-A---CTAT--CT----------T----------T---GAG--------G----- 4210
```

```
H1B.FR.83.HXB2          CCTGGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTTAAAATTAGCAGGAAGATGGCCAGTAAAAACAATA...CATACTGACAATGGCAGCAATTTCACCGGTGCTACGGTTAGGGCCG 4614
Pol                      L  V  A  V  H  V  A  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  A  T  V  R  A
H1A1.UG.85.U455         ---A--------C-------------C--C-----------------C---------------A----G---------C---A-AC--------------------------------GT----..--C--A-----------------A-C---G-A---AA--A- 4060
H1B.US.90.WEAU160       --------------------------C--------------------------G---------A--------------C---A-C------C----------------------------...------A------------------A--A---------A----- 4613
H1C.ET.86.ETH2220       ------------C-------------C--C-----G-----G---------------------A--A-----------C------AC--------------G---------C-GGGT----..----A--T--------T--C------A--AA-G-A----AA--A- 4006
H1D.CD.84.84ZR085       --------------------------C----------------------------------G--------C--------C-----------------------------------GT-G-G..----A------------------A--------A----A----- 4136
H1F1.BE.93.VI850        ------------C--------T----AG--C--------------------------------A-----------C--C--CA-A------G------------------------T----..----A------------------A----CG------A--AT 3954
H1G.SE.93.SE6165        -A-A----------------------C------------------------------------A--------------C---A-A-----------------G---------C-GT----..----A--T-----------------A-----G-A--A-A---A- 4011
H1H.CF.90.056           T-----------C-------------C----------------C-C-----------------AA-------------C--C--G--G---C-----A-C---------------GT----..----A--------------------GA-----G------A---A- 3961
H1J.SE.93.SE7887        --------------------------C------------------C-----------------A------G------T----A-A-----------------C----------GT----..----A----------------C-----A---G-G-T--G-A---A- 3928
H1K.CM.96.MP535         ------------C-------T--C--C------------------C-----------G--A--A------C--C---A-AC-------------------------------GT----..----A-------------C---------A-CA--GT-----A---A- 3810
H101_AE.TH.90.CM240     ------------C--C--G-------------------------C------------------A------G--------C------GC------C-------------------GT----..--C--A-----C--T------------A-C---G-A----AA--A- 4188
H102_AG.NG.x.IBNG       -T-A--------C--------------------------------C-----------------A-----G--------C---A-A--------------------------G---GT----..--C--A--------------------A-------G-A--A-A---G- 4139
H103_AB.RU.97.KAL153_2  ---A-----------------------------------------------------------A--------------C---G-C-----------------------------T----..----A------------------A---A--G-------A---T- 3840
H104_cpx.CY.94.CY032    -A--------------------G--T--------C-------------C--------------A-----------C--C--CA-AC---------------------------G---TG---..--G-A-----C---CC---------A----G-------A---A- 3980
H1N.CM.95.YBF30         ---T--G--G--C------G--------C--CT-----------------T------G-----A-------------------A-----G-----T------------------GTT---..--C-----T-----ATC---------TA-----C--T--A-AA--A- 4208
H1O.BE.87.ANT70         AA-A--T--T--C------G--A-----G-TC--------------G--A-------------A--A----T--C--C--C--G------C-G--T-C-----------T--T---GT----..----A-----C--GCCT-----T--AA---A-A--TA-G-A---T- 4669
H1O.CM.91.MVP5180       AA-A--T--T--C---------A-----AC-T--------------G-G-A------------A------T--C----C--C--G--------------C---------T--C--GT----..----A----------ACCT-----T--AA----AG-CA-G-AA--T- 4644
H1P.FR.06.RBF168        ---T--------C---------------C-TCC-G-----------A----------------A--A----T--------G-GC-C-----G------C------------T--GT----..--C--A-----------GCCT--C--T--TA-----G-A--A-A--A- 4075
CPZ.CM.01.SIVcpzCAM13   -A-A-------------A----------C------------------C---------------A--A----T-----C--CA-A-------------A-C----------------G---GTCT---..--C--A--T----G-CA-----T--AA---AA-G-A---AA--A- 4194
CPZ.CM.05.SIVcpzEK505   ---T--------C------G--------C--CT-------------C-----T----------A-------T-----------A-CC-----G-----T--------------------GTC---..-------------------------------A-----C--T--A-AA--A- 4186
CPZ.CM.05.SIVcpzLB7     -T-A--------G--------T------C------G-----------C--A--T---------A--------------T------A-C------------TCG----------------TT---..--------T-----T-------T--TA-CA--G-A--C-A---T- 4169
CPZ.CM.05.SIVcpzMB66    ---A--G-----A---------------G------G-----------A--A-T--T-------A--A-----------C---A-C--------G------C--------------------GTC---..--------T-----A--T--C--T---A--AG---T--C-A---A- 4160
CPZ.CM.05.SIVcpzMT145   TA-A--G-----C------G--------C------------------C--C--T--T------A----------------G-CC-T--G----------------------------G--T----..--C--A-------GCCT-------TA---AA-G-A--A-A---A- 4156
CPZ.CM.-.SIVcpzMB897    -T-A--------T--A------------C------G-----------C--A--T--T----------A--G----T-----G-GC-G-------------------------------GGTT---..--C-----T-------------C--T--TA-----G-T--C-A---T- 4180
CPZ.CM.-.SIVcpzDP943    ---A--G-----A---------------CC-----G-----G--C--C--T------T--A-----------T-----A--C-------------------------------------GT---T..--C-----T-----ATCT-----T---A--AA---A----A---A- 4200
CPZ.CM.98.CAM3          ------G-----A---------------G--CC-----G----G--C-----T-----G--T--A--------T-----A--C------------------------------T------GT---C..--C-----T-----GCCT--C--T---A--AA---A----A---A- 4028
CPZ.CM.98.CAM5          ---A--G-----A------G--------CC-----G-----G--C-----T--------A--A-----------T-----A--C------------------------------T------GT---..--C-----T-----GCCT--C--T--TA--AAC--A--C-AA--A- 4316
CPZ.GA.88.GAB1          -A-A--G---------C-----------C--------------------C-----T--G----A--A-----------T-----C--G------------------T--------------T--T..--C--A--T-----GCCA-----T--AA------G-A--C-A---T- 4675
CPZ.GA.88.GAB2          ------G--------G--A---------CT------------------A--------T--G--T--A---------T--------A-C---------------------G------------G-AG--..--C--A--T-----A-CA----------A------AG-A--A-AA--A- 4007
CPZ.US.85.CPZUS         TA-A--G-----A-----G---------CC-------------G--C-----T----------A--A--G-----C-------A-C------------------------------T------GTG--T..--------T-----ATCT--C--T--AA---AG-----A-----A--A- 4674
CPZ.CD.90.ANT           -A-A--------G------T-T------C-TC--------T-----A--GG---AT--G------A-AAGT-------C--C--G-------------CA-C---------------------------..--C----------AGCT-------AA-----AG-A--A-AA--G- 4053
CPZ.TZ.00.TAN1          -A-A--------G--------T------T-TC-----------C--A----A--------AAGA------C--T--C--CA-CC-------G-----------------T------G-A----..----A--T-----ACCA-----T--TA---A-AG-A--G-A---A- 4258
CPZ.TZ.2001.TAN2        -A-A------------------------T-TCC---------G--C--A----A--------AAG----------T------CA-C---------G--T---------------T------G-A----..----A--T-----GCCA--C--T--TA-----AG-A--A-AA--A- 4263
CPZ.TZ.2002.TAN3_1      -A-A--------CA--------------T-TCC---------G--C--A----A--------AA-----------T------CA-CC------C-G-------------------------G-A----..----A----------GCCA-----T--TA-----AG-A--A-AA--A- 4263
GOR.CM.2007.SIVgor2139_287 TT-A--------C-----G-----C--C-TC--------T-------------------------------C--C--C--GC-C--G--------CT--G--------------------T----..--C--A-----------GCCT-----T--TA-----C--A--C-A---T- 4151
GOR.CM.2007.SIVgorCP2135con TT-A--G-----C-----G--T--C--C-TC--------T-------------------------------C--C--C--GC-C--G--------CT--------------------------T----..----A-----C--GCCT-----T--TA-----C--A--C-A---T- 4150
GOR.CM.2004.SIVgorCP684con T--A--------C--------T------C-TC--------T-------------------------------T--C--C--GC-C--G--------CC--G--------------------T----..--C--A--T-----GCCT--C--T--TA-----C--A---A---A- 4070
H2A.DE.x.BEN            TA-A--------A-----T--A------TC-------------C--C---CAG---T----AAG-C-G-------CTC---C--A-------C-G--CA-T--G------A---CGCACT-G...--C--A-----------CC----C-----TTCACAGGAA--G-A-ATG- 4996
H2A.GW.x.ALI            -A-AA-------G------T--------TC-------------G--C--C--GCAG---T----AAGAC----------CTC---C--A--------C-G--TA-T----------A---C-CACT-G...--C--A--T-G---TGT----C-----TCACAGGAA--A-A-ATG- 4989
H2A.SN.x.ST             TA-A--------A-----T--A------TC-------------C--C---CAG---T----AAG-C----G---CTC---C--AC------C-G--CA-T--G------A---C-CATT-G...--C--A-----------TGC----C-----TTCACAGGAA--A-A-ATG- 4441
H2B.CI.x.EHO            AA-A--G-----------------------TC--------------A--A--CCA---------AAGAC-G-----TCTC---C--G-----GC-G--CA-C--------TA-C-C-CACC-G...--C--A-----------TGC----C-----TTCACAGAT--G-AAATG- 4967
H2B.GH.86.D205          AA-A--G-----C---------------G-T-----------G--A--A--CCA---G-----AAGAC-G-----TCTC---C--AC-----G--G--CA-C--------TA-C-C-CACC-..--C--A-----C--TGC----C-----TCAC-A-GT--A-A-ATG- 4969
H2G.CI.92.ABT96         -A-A----------------G-------C-T-G---------A--C---CA-----------AAG-C------------TA-C--G-------------------G-------TA-C-C-CACC-..----A-----------TGC----C-----TCGCAAGA---G-A-ATG- 4323
H2U.FR.96.12034         -A-C--G-----A-----G---------C-T-------------G--C---GCAG-----T--TAG-C--------CCTG--C--AC----GC-----------G-----TA-C-CTCATT-G..--C--A--T-----TGC----C-----TTCACAGGAA--A-A-ATG- 4475
SMM.SL.92.SL92B         -A-A--------A-----G--T------C-T-------------G--A--A--TAGG-------AAGAC-G-------TG--C--C------A-----TC----------TA----GCATC-G..--------A-----C--TGC----------TCCCAGGA---G-A-ATG- 4344
SMM.US.x.H9             TA-A--T-----A-----G-GT------TC------G---------A--C--GCRG-------AAGAC----------CTG-----GC----------CA-C--------CA-C-C-CATC-G..--------T-----TGC---------ATCACAAGAA--G-AAATG- 4386
SMM.US.x.PGM53          -A-A--------A-----G--T------TC------G---------A--C--GCAG-------AAG-C----------CTG-----GC----------TA-C--------CA-C-C-CATC-G..--------T-----TGC---------ATCACAAGAA--G-AAATG- 4830
SMM.US.x.SIVsmH635F_L3  TA-A--T-----A-----G--T------T-------G---------A--C--GCAG-------AAGAC----------CTG-----G------GC---T--C--------CA-C-C-CATC-G..--------T-----TGC---------ATCACAAGAA--G-AAATG- 4916

MAC.US.x.239            -A-A--T-----A-------T-------TC----------G--A-----CA---G-----AAGAC-G-----CTA----G--------G-----C--------TA-T-C-CATC--..--C--A--T-----TGCT--C--TG-TTCGCAAGAA--A-A-ATG- 4913
Pol                      I  I  V  A  V  H  V  A  S  G  F  I  E  A  E  V  I  P  Q  E  T  G  R  Q  T  A  L  F  L  L  K  L  A  G  R  W  P  I  T  H  .  L  H  T  D  N  G  A  N  F  A  S  Q  E  V  K
MAC.US.x.251_1A11       -A-A--T-----A-------T-------TC----------G--A-----CA---G-----AAGAC-G-----CTA----G--------G-----C--------TA-T-C-CATC--..--C--A--T-----TGCT--C--TG-TTCGCAAGAA--A-A-ATG- 4912
MAC.US.x.251_BK28       -A-A--T-----A-------T-------TC-----------------A-----CA---G-----AAGAC-G-----CTA----G--------G-----C--------TA-T-C-CATC--..--C--A--T-----TGCT--C--TG-TCGCAAGAA--A-A-ATG- 4889
MAC.US.x.EMBL_3         -A-A--T-----A-------T-------TC-----------------A-----CA---G-----AAGAC-GCACTAT...--C--G--------G-----C-------TTATTT-CAC--CTACACACACAGT-----TGCT--C--TG-ATCGCAAGAA--A-A-ATG- 4381
STM.US.x.STM            TA-A--------A-------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---------TCGCAGGA---A-A-ATG- 4560
MNE.US.x.MNE027         -A-A--T-----A-------T-------TC-----------------A-----CA---G-----AAGAC-G-----CTA----G--------G---A-C--------TA-T-CGCATC--..--C--A--T-----TGC---C--T--TTCGCAAGAA--A-A-ATG- 4381
DEB.CM.99.CM40          AA-A--G-----C-----G--T---------------GA-CA-GA-CC-A---------------AA-A--G-----TCTG--C--G---C-G---G--C-------------CTCAG---..--C--A--T-----ACC---C--T-T-A--CAAC-A--C-A---A- 4420
DEB.CM.99.CM5           -A-A--------G-------T-------T---G--GA-TA-G--CT-AT-----------T--AA-----------C-TA--CT-A---C-G--G-GG-C---------CA-T---CAG---..--C--A--T-----ACCA--------TAA--CAGC-A--C-A---A- 4414
DEN.CD.x.CD1            AA-A--------AA---C----------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--CAAG-CA-GCA---A- 4504
LST.CD.88.447           -A-TAAT-----------G---TC-----T---G-T-----------A-------AT-----------A-AAC-------C-C--C--G-------C-GTGCA-C-------------C-GCA-G-C..----A----------ACCA--C--TGT-A-CAAGGA---ACAA---- 3891
LST.CD.88.485           -A-TAAT---A-C-----G---T-----T---G-T-----------A-------AT-----------A-AAC-------A-C--C--G-------C-GTGCA-C-------------C-GCA-G-C..----A----------ACCA--C--TGT-A-CAAGGA---ACAA--A- 3891
LST.CD.88.524           -A--AAT-----C-----A-G--ATCA--C-TC--G-TT--------G--------AT-----------A-AACT------A------T-G------G---TGTA-C-------T-------CA---T..--C--A--T-----TCCA-----TGTGA--AAAGAT--ACA---A- 3888
LST.KE.x.lho7           TA-AAAT-----C-----G---TCA--C-TC--G-TT----------A----AT--G-----AA-AAC---GT--C----C--G---------TGCA-T-------T------CA-----..----A--T-----TCCA-----TGTAA--AAGGA---ACA---A- 4975
SUN.GA.98.L14           AA-A-----G--G-----G--TTCA--G-----G-T-----------AC-A--CAAT--GCA----A--AC-------CAC--GG--G------GC-CTGT-CC-TG----------CAG---..----A--T-----GCCA-------T-A--AAAGAT--GGA---A- 4974
MND-1.GA.x.MNDGB1       TA-A--------G-------TTCA--C-TC--GA-G-----------AA-----------T--AA-AA-G-------G-G------G-----G-----------C-CA-------TA-T-GT-A-C--..--C--A--T-----GCCT--C--T--TA--CAGGAA--AGAAA--A 4301
OLC.CI.97.97CI12        -TGTT-G----CA---AC--GGTCA--G-TCC--T-T-G----A-CC--AG-------GAT--AG-AA-T----TGA-A---G-GC-T-G-C----TA-TCA-------T----------T---T..--C-----------G-CA-----TGTGAA-CAG-A-A-GCA---A- 4485
WRC.CI.97.97CI14        TT-AAAT-----C-----------TCA-----CC---------A--A--C--AT---------GAT-CCT-TC-G-------CTA---ACC--GC-GC-G---A-T----------CA-C--G-A----..----------------G-CA-------AGTC-CAG-AAA-GCA---TA 4914
WRC.CI.98.98CI04        T--AAAT-----C-------ATCA--G---C-----------A--A-C----------GAT-CTT-TC-------TCTA--C--CC--C--A-C------T-------TA-G-----AG--T..-----G--T-----A-CA------T-AGTCACAG-AAA-GCA-A-TA 4929
WRC.GM.05.Pbt_05GM_X02  AT-AAAT-----C-----G--ATCA--C--CT-------GA--A-A------A-----GAT-CCT-TC-------CT---CACCA--C-------------------------AG--T..-------T-----A-CA------AATCACAG-AAA-GGA----A 3843
MND-2.CM.98.CM16        -A-A-C------C---------A---A-G---C-----------A--A--CA-----------AA-A---------C-C--C--G-------------------------CATC--..--C--A--T-----CCT---------A---AA-AA--AGCCA-A- 4615
MND-2.GA.x.M14          -A-A--------C-----G---------------C-------------G--A-------------AA-A---------C----C--GC-----C------------------G---CATT---..--------------TCCA--C--TGT-A---AG-A---AGCCA-A- 4542
MND-2.x.x.5440          -A-A-----------------------T-----------G--A--A--------G----AA-A--------GC-C--C--G----GC--------T--G----------CATC--..--C--------------TCCT--C--TGT-A---AA-A---GGCCA-A- 4175
DRL.x.x.FAO             -A-A--G-----C--C-----A-----C-C-----------------A--A-------------AG--A-----T-------GC---------T--C----------TC-CACT-G..--C--A--T-----ACCA---------A---AA-A---AGCAA-A- 4181
MON.CM.99.L1            TT----------A-----------------T---G--TGG--CA-GA-CC-A---AGG--------AA-ATGT-----GGATAGCC--G--GG-GC-G--C-CT-TG------G--C-CTCAG---..--C--A--T-----GCCA---------A-CCAAGAAT-CGAA--A- 4317
MON.NG.x.NG1            ---A-----C---------------T-----C--CG-TTGG--CA--A--C----CAG-----------A-A------GCAT-GCC--G--GG--C-T----CG-C-------A--C-TCATG---..--C-----------ACCA---------A---CAGGAAT--GCA--Y- 2910
GSN.CM.99.CN166         -T-A--G--G--C--C-----A-----G---G-CTGG--TAGGA-A--A--CCAG--------CAGAC----T----TGAAG--C--G-------C--C-------T---CCCACC---..----A--T-------CC-------GT-TCCAAGGAAT-GGAA--A- 4342
GSN.CM.99.CN71          -T-A--G-----C--C-----------G---G-TTGG--TAGGA-A--A--CCA---------CAGAC----T----TAAAG--A---G---------C--C---------CTCACC---..----A--T-------CC-------T-TCAAGGAAC-GGA---A- 4324
TAL.CM.00.266           -TGT--------A--C--G-----------G--C-----GA-TA-GA--C-G---AGG--G--G-----G---G--C--CCTC-----C---C--G-G--CA-C-------TA---GCCACC-G...----A----------GCCA--C--TGT-TCA--AGAAA-GCA----A 4580
TAL.CM.01.8023          ATGT--------A-----G--A------T--C--T--GA--A-GA-CT-G--CAG-----C--A-G------C--CCTG-----C--GC-GG-G--CA-C-----------A--G--CATC--..--C--A----------GCCG--C--TGT-TCG--AGA-A-GCA----A 4078
MUS-1.CM.01.1085        T--A--------C-----G--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----AGAAT-CGAA--A- 4304
MUS-1.CM.01.CM1239      -T-A-----C--C--C--G------------C-CGTGG---A--A-AT-A--CAGG--------AAG-C-------CCTAGGC--G---G----------CTCTT------T-----C-CA----..--C--A--T-----AGC---C-----TTA---AAGAAT-CG-A--T- 4294
MUS-2.CM.01.CM1246      -T----G-----A------------A-T-TC---TGG--CAGGA-A---T-CCA------T--AAG-CT-------CCTAGAG--G-C---C--G------CC-C--------A-T-GTCAG---..--CC-A----------T-C---C--T--TA-CAAGGAAT-CCAA--T- 4402
MUS-2.CM.01.CM2500      T--A--------A------------T-T----TGG--TAG-A-----T-CCA-----AT--CAG-CT--------TAGAG--G----C------C--C-----CA---C-CAG--C..--C--A----------T-C---C-------A---AAGGAAT-CCAA--A- 4361
SYK.KE.x.KE51           TTGT--------AA-CAC---------G----C----A-TA--A-AT-GAA-AG--------------A--------G-CTA--C-----GC--A-------C---------A-T---CAT---..----A-----------TCCA--C-----T-A---A-G-CT--GC--A- 4469
SYK.KE.x.SYK173         -TGT--------AA-CAC---T-------------------A-CA-GA-CT-AAA-AGG-----G---G-T----------TG-----CA-GC--A----CA-T--------A-----CAG---..--C-----T-----ACCA--C--TGT-A---AC-A-T-C-AA--A- 4814
RCM.GA.x.GAB1           -A-A--------C-----G--A--------------G---------A--A-------------AA-A--G-------C----C--G--G---C------C-------------GG-AGC--..--C--A--T-----AGCA---------AA----AG-A--ACA---G- 4093
RCM.NG.x.NG411          -A-A--------C----------------------C-----------A--A---------G----AA-A--G-----TC-----C--G------------CC-----------G---CATT---..--C--A--T-----A-CA--C--T--AA---AG-A--GCAA--A- 4108
GRV.ET.x.GRI_677        AA-A--------A-----------------TC-----------G--G--C----AG------T---A----G------C-C-----GC-G---C-GTT--C-----------G---CATC---..--C-----T------CCA--C--T---TC-CAG-AT--GGCA--G- 4787
SAB.SN.x.SAB1C          -A-A--G--T--------T--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-CCAGCAA--AGCA--TA 5012
TAN.UG.x.TAN1           TA-A--------C-----G-----C-TCT-------------C-AG-TAGG-------AA-A----------A-A--C--G-------A-CATTA-T--G-----TA-T-C--A-T-..---------------GCCA-----TGTTA--CAGGA---ACAAA-TA 4813
VER.DE.x.AGM3           -A-A--------C-----------------TC--------------A--TAG----------AA-A----------A-A---T-A-----GA--CT-A-T-------CA---CCCA-C-G..----A----------ACC------T--GTC-CAAGAA--AGCA--AA 4336
VER.KE.x.9063           -A-A--------------G-----C---TC-------------C--A--TAGG-------AAGA--G------CA-A--C---C-----A--TTGA-T--------A--GTCCAGT---..--C--A--T-----ACCA-----T--TTCCCAAGA---AGC---AA 4841
VER.KE.x.AGM155         GA-A-----G--C-------------G-T-----------G-----C--TAGG-------AA-A--G------A-G--CT-G------A--AT-----------CA-C-CTCACC-C..----A--T-----ACCA---------TTC-CAGGAA--AGCT--TA 4833
VER.KE.x.TYO1_patent    -A-A--T--G--C-----------------TC-------------C--A--TAGG-------AA-A-----G---A-G-----A------A--CTGA-T--------TA---C-CAGT---..--C--A----------GCCT--C--T---TCCCAAGAA--GGCA--AA 4838
COL.CM.x.CGU1           -T-A--------A--C--CTGT-C-CTG-TCTGCTGG---ACCA-AT-AAAGAG-----------G-A-------GTAGGGCC--CA-------G--TA-TCAG---GAG--T-G-CA-G-T..--C--A--T-----GCCA-----TGTAA--CAGCACT---A---A- 4377
```

```
H1B.FR.83.HXB2                CCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCATCCACAATTTT 4784
Pol                           A__C__W__W__A__G__I__K__Q__E__F__G__I__P__Y__N__P__Q__S__Q__G__V__V__E__S__M__N__K__E__L__K__K__I__I__G__Q__V__R__D__Q__A__E__H__L__K__T__A__V__Q__M__A__V__F__I__H__N__F_
H1A1.UG.85.U455               T------------AAAT---C-A-----------G--C-----------------------------G-----C-----C--G--------------C-----G-----------G--A--------C--------------------------------T--------- 4230
H1B.US.90.WEAU160             ----------------G-----------------C-----------------------------GA----------------------------------------------------A--------------------------------------------------- 4783
H1C.ET.86.ETH2220             -------------A--T--TC-A--------------------------------------------------------------------------C-----G-----------A--A-----G--C--------------------------------T--------- 4176
H1D.CD.84.84ZR085             ----C--------A--T-----------------------------------------------------------------------------------------------------A-----------------------------------------T--------- 4306
H1F1.BE.93.VI850              -------------A--T---C-A--------------C--------------------------------------A--------G-----A--G--C--------------------A--------------------------------------------------- 4124
H1G.SE.93.SE6165              -A-----------AAAT----CA--------------------------------C--------------------------G--------------C--C--G-----C--G-----A-------------------------G---------------T--------- 4181
H1H.CF.90.056                 -------------A-AT---C-A-----------G-----------------------G------------------------------------G-C-----G-----------C--A--A-----C--------------------------------T--------- 4131
H1J.SE.93.SE7887              -------------A-AT-------------------------------------------------------------------------------C-----------------A--A--------C--------------------------------A--------- 4098
H1K.CM.96.MP535               -------------A--TG----------------G--------------------C-----------------------------------------C--------------G-----------------C--A------------------------------------ 3980
H101_AE.TH.90.CM240           -------------CAATG--C-A-----------G--C--------------------------------------------G--------------C-----G-----------G--A--------C----A---------------------------T--------- 4358
H102_AG.NG.x.IBNG             -A-----------AAATG---CA--A-----------------------------C-----------G-----------------------------------------C--------A--------C----------------G---------------T--------- 4309
H103_AB.RU.97.KAL153_2        -------------A--G-----------------C------------------------------------------------C--------C---C---------------------A--------------------------------------------------- 4010
H104_cpx.CY.94.CY032          -------------A-AT-----C--------------------------------C-----------G-----------------------------C-----G-----C--G-----A--------C----------------G---------------T--------- 4150
H1N.CM.95.YBF30               -------------AAAT-----A-----------G--A-----------T--------G----C------G--C--------------------------------AA-C--------A--A--------A-----------G--------G--T-----T--------- 4378
H1O.BE.87.ANT70               -A-----------CAAC--AC-A--T--G--------A--A--T-----A---------------------G-C--------G--------ATC-------CAG-----G--G--C--A--A-----CT-A-GA-----------------------TG-T--------- 4839
H1O.CM.91.MVP5180             -A---------A-A--C--AC-A--T--G-----G--A--A--T-----A---------------------G-C-----------------ATCT------CAG-----G--G--C--A--A--G---T-A--A--------------------C--TG-T--------- 4814
H1P.FR.06.RBF168              -T---------CTTAAT--A-CT--T--G-----G--A--------------G-----G--------G-----A-----------------------C---CTT--A---C-------A--------CT-A-----------T--------------TG-T--------- 4245
CPZ.CM.01.SIVcpzCAM13         -A-----------CAAC-----A--AA-G-----------------C--A--------------G--------C-----C-----------------C--T-A---A--T-----------A--------A--A--------------------G----GT--------- 4364
CPZ.CM.05.SIVcpzEK505         -------------AAAT---C-A-----G-----------A--------T--------G--T-C------------------------------------------AA-T-----A--A-----T-A---------------------------C-----T--------- 4356
CPZ.CM.05.SIVcpzLB7           -A------------CAG---C-------------------------------G-----------G--------A-----C-----C-T---------A-----G-----------------A-G--C--A--A---------------------C-----T--------- 4339
CPZ.CM.05.SIVcpzMB66          -T--------------T--C-A--A------------------------T-----------G-----------------------------------A-----T--A-----------A--A-----------A--------C-----------------T--------- 4330
CPZ.CM.05.SIVcpzMT145         -A-----------A--T-T-GA------------------A-----C--A-----------------G-----------------------A--G--C---AG--A--C------T-A--A---------------------------------C-----T--------- 4326
CPZ.CM.-.SIVcpzMB897          -A-----------AAAT---C-A--------------------------T-G---------------------A-----C-----G-----------C-----G-----------A--A--T-A------------------G-----------C-----T--------- 4350
CPZ.CM.-.SIVcpzDP943          -------------A--C---C-A-----------------T--T--C--A-----C-----------------A-----C---------------G-C--------AA-T--A--A--A--G---T-A--------------T-----------G-----T--------- 4370
CPZ.CM.98.CAM3                ----C--------A--C---C-A--------------------T-----A--------------G--------A--------G--------A--G--C--------AA-T--A--A--A-----CT-G-G----C--G----------------G-----T--------- 4198
CPZ.CM.98.CAM5                -------------T--C---C-A-----G--------C-----------A-----------------G-----C-----------------------C---------A-T--A--A--A--G--G--A------T--C----------------C-----T--------- 4486
CPZ.GA.88.GAB1                -------------A-AC--------------------A-----T-----A--G-----------G--------CT-A--------GC----------A-----T--G--------A--A-----T-A--A------------------------G-----T--------- 4845
CPZ.GA.88.GAB2                -------------A--C--A--A--A-----------A--T--T-----A-----------G-----------C-----C-----------A--G--------G-A--T------C--A--A--CT-A-----------------------T--------T--------- 4177
CPZ.US.85.CPZUS               -T--C--------A--C--AC-A--A--G-----------A--------A--------------G--------C-----------------------C--------AA-C--G-----A--A--G--T-A-----T------G-----------------T--------- 4844
CPZ.CD.90.ANT                 -A-----------TAAT---C-A--------------A--A--------A-----------------G-----C---------C---G--C---------------AA-T-----C--A--A--AT-A------------GT--------GCA--T--------- 4223
CPZ.TZ.00.TAN1                ----C--------ACA---TC-A--T-------G-----A--T-----T-------------------------------C--------C--------AG--A--C--G--C--A--A--G--A--G-G----------ATC---------G-AT----------- 4428
CPZ.TZ.2001.TAN2              -------------ACA----C-A--T-------G-----A--T-----T-----------------------C-------C--------C-----------A---AA-T--G--A--A--A--G--C--G-GR--------GRYT--------G-AT----------- 4433
CPZ.TZ.2002.TAN3_1            --------------CA---TC-A--T--G----G-----A--T-----T-----------------------C-------C--------C---C-----------AA-T--G--A--A--A--G--CT-G-G---------ACT--------G-AT----------- 4433
GOR.CM.2007.SIVgor2139_287    -----------CTCAATG-A-CA--T--G--------------T-----------------------G-----------C-----------A--G-----CA---A--G--------A--A-----T-A----------G--G-------------TG----------- 4321
GOR.CM.2007.SIVgorCP2135con   -----------CTCAATG-A-CA--T--G--------------T-----------------------G-----------C-----------A--G-----CA---A--G--------A--A-----T-A----------G--G-------------TG----------- 4320
GOR.CM.2004.SIVgorCP684con    -----------CTCAATG-A-CA--T--G--------------T-----------------------G-----------C-----------A--G-----CA---A--G--------A--A-----T-A----------G--G-------------TG----------- 4240
H2A.DE.x.BEN                  TGGCA-------TA--T--AG-A--ATCC------G-A--T-----C--A-----C--G------------G-A------C-CC-CC------TCAG---A-TAGAA-T-----A-----AA-TACAA-AG-A---AT-----TG--------TCATTG-ATG------ 5166
H2A.GW.x.ALI                  TAGCA-------TT--T--AG----ATCC------G-A--T--------A-----C---------------G-A------C-CC-CC----A--TCAG---A-CAGAA-T-----G-----AA-TACAG-GG-A---AT-----T---------TCATTG-ATG------ 5159
H2A.SN.x.ST                   TGGCA------ATA--T--AG-A--ATCC--C---G-A--T--------A-----C-----------G---G-A------C-CC-CC----A--TCAG---A-CAGAA-T-----G-----AA-CACAG-AG-A---AT-----T---------TCATTG-ATG------ 4611
H2B.CI.x.EHO                  -AGCC------ATA--G--AG-A--AAC---C---G-G------T-----AG--------G----------G-A------CC-TC-TC-G--A--TCAG---ACAGAA-T-----------A-T-TCAA-AG------TT--GTT-------ACTCA-TG-ATG------ 5137
H2B.GH.86.D205                TAGCC-------TA-----AG-A--AACT------G-A------T--C--A--------------------G--G-A-----CC-TC-CC-G--A--TCAA----ACAGAC-C-----C--A--A-T-TCAA-AG------TT-----T-------ACTCA-TG-ATG------ 5139
H2G.CI.92.ABT96               TAGCC-------TA--C---G----A-C-------G-A--A-----C-----G---------G------------A------C-TC-TC-------CAA----ATA-AA-T---------A--AA-TTCAA-AG-A--CAT---GTT-----------CATTG-ATG------ 4493
H2U.FR.96.12034               TAGCA-------TA--T--AG----TACC------G-A------T--C--G-----C--------G-------G-C------C-CC-TC-------CCAA----AGAGAA-C-----A-----AA-TACAA-GG---CAT---C-T----------CTCA-TG-ATG------ 4645
SMM.SL.92.SL92B               TAGCA------TTA---G-GG-A---TCC------G-A--G-----C--A----------G------------G-C---G-CTT-C-TC-T-------AC----ATA-AA-----G--A--A--A--GTCAG-AG---CCT------T---------C-CATTG-ATG------ 4514
SMM.US.x.H9                   TTGC---------A--R--TG-A---ACC---R-GG-A--T--T-----A--G--C---R-G------G---G-G------CC-TC-TC----A-CCCAA----ATAGAA-T-----A-----AA-TTCAA-AG-A--TAT------T-------A-TCATTG-AT------- 4556
SMM.US.x.PGM53                TTGCC--------A-----TG-----ACC--C--GG-A--T--T-----A--G--C---------------G---G-A------CC-TC-TC----A-CCCAG----ATAGAA-T-----A-----AA-TTCAA-AG---TAT------T---------TCATTG-ATG------ 5000
SMM.US.x.SIVsmH635F_L3        TTGCC--------A--G--TG-A---ACC------G-G--T--T-----A--G--C---------------G---G-A------CC-TC-T-----A-CCCAG----ATAGAA-T-----A--A--AA-CTCAG-AG---TAT------T---------TCATTG-ATG------ 5086

MAC.US.x.239                  TTGCA--------A--G--AG---CACC-----GG-A--A---------A--G-----G---------G---G-A------C-CC-CC-G--A--TCAA----ATAGAA-C--G--A--A--AA-TTCAG-AG-A--CAT----TT----------TCATTG-ATG------ 5083
Pol                           M__V__A__W__W__A__G__I__E__H__T__F__G__V__P__Y__N__P__Q__S__Q__G__V__V__E__A__M__N__H__H__L__K__N__Q__I__D__R__I__R__E__Q__A__N__S__V__E__T__I__V__L__M__A__V__H__C__M__N_
MAC.US.x.251_1A11             TTGCA--------A--G--AG---CACC-----GG-A--A---------A--G-----G---------G---G-A------C-CC-CC-G--A--TCAA----ATAGAA-C--G--A--A--AA-TTCAG-AG-A--CAT----TT----------TCATTCGATG------ 5082
MAC.US.x.251_BK28             TTGCA--------A--G--AG---CACC-----GG-A--A---------A--G-----G---------G---G-A------C-CC-CC-G--A--TCAA----ATAGAA-C--G--A--A--AA-TTCAG-AG-A--CAT----TT----------TCATTG-ATG------ 5059
MAC.US.x.EMBL_3               TTACA--------A--G--AG--GCAC-CC--T-GG-A--A---------A--G-----G---------G---G-A------C-CC-CC-G--A--TCAA----ATAGAA-C--G--A--A--AA-TTCAG-AG-A--CAT----TT----------TCATTG-ATG------ 4551
STM.US.x.STM                  TAGCC----------------AG-A--AACC------G-A--T--------A--G--C-----------G---G-C------C-TC-T-------CCCAG----ACAGAA-T----------A--AA-CACAG-AG-A--T-T-----TC---------TCA-TG-ATG------ 4730
MNE.US.x.MNE027               TTGCA--------A--G--AG---CACC-----GG-A--A---------A--G-----G---------G---G-A------C-CC-TC----A--TCAA----ATAGAA-C--G--A--A--AA-TTCAA-GG-A--CAT----TT----------TCATTG-ATG------ 4551
DEB.CM.99.CM40                -T--C------CTA--T--TG----CACCACA--GG-A--A--T--C--A------C--G------G-------AG--A---C-GGATTC----AG-T-CC--CAA---AA-C-------G-T--A--GAGAT-GG-A-----------GC-------TACC-A---AATT------ 4590
DEB.CM.99.CM5                 -A----------TT------AG-T--TACTACA--GG-A--A--T--C--A-----------------G----------AGC-AA---GG-TCC----AG---C---CAA---AA-T-------AG-T--A---AGA--AG-A-----------GC-------CACC------TATT------ 4584
DEN.CD.x.CD1                  -A-------------AAAT-----A--CACCACA---G----A--T-----A-------------G--TA-T-------AA--A---C--GC-----------G--G-------C--TA------G--AG-T--A---AGAT-AG-A------C--GC--------ACCC-----T-TT------ 4674
LST.CD.88.447                 TAAT-------CTA--T--TG-A--TAC-ACA--------T--T--C--T--G-----G-----------T---G-C-AA-------C-C-----AC-----------A-AG---C-----AG-T---C-----AT-A---------T--C-T--------AC-CA---T--T------ 4061
LST.CD.88.485                 TAACC------CTA--T--TG-A--CTC-ACA--G-----T---------------G-----G-----------T---G-C-AA-------TGC-----C-----C-----A-AG---T-----AG-T---C-----AT-A---------T--C-T-----------CA---T--T------ 4061
LST.CD.88.524                 TGACC------TTAA-T--AG----CAC-ACA--G--A--A--T--C--T-----------------------------GG-A-AA-------TT------C-----C-----A-AGA--------AG-T--AC-G--G--A-----------GTT------------CCA---T--T------ 4058
LST.KE.x.lho7                 TAACC------ATA--G--AG-A--TAC-AC---G--A--T--T--C--T--G----------G--T------G-A-AA--C-----TT------A-GT---------A-AGA--------AG-T--AC-G--A--A--A-------------T------------CCA---T--T------ 5145
SUN.GA.98.L14                 TA---------TTA--C--AC----TAC-ACA--G--A--T------C--T--G------------------G--G-A-A---C--G-T-----AC-------CA-CAGAA-T-----AG-T--AC-GG-AT-G--A--T-----------TC-------TC-GCA---T--T------ 5144
MND-1.GA.x.MNDGB1             TG---------TTA--G--AG-A--CAC-----------A------T--C--A----------G--G-------G---AA--AA-----GT-TC-----AG--T-G--T-AGA-AA--------AG-TTGCA--G-AT-A--A-----------GCC-------CAC-------T--T------ 4471
OLC.CI.97.97CI12              TA--C------TTA--G--TG----T---ACG--C------G-GG-----T-------------A-CA-----G-A-AA-----G--TG-----G-G--AT----GAGAA-C------A------AAC-----T-GG-A-------T-GGCT------TA-C-------T--T------ 4655
WRC.CI.97.97CI14              T---C------TTA-----AG-A--A---------G--C------T-----T----------------------A-----AG--A---C-GGAGG-----AG-C-------CACAGTA-T-----AG-T--A---A-AT-AG-A------T--GCG-------T------------T--------- 5084
WRC.CI.98.98CI04              T---C------TTA-----AG-A--A---------G--C--T--T-----T----------------------A-----AGC-AA----G-AGGC-----AG----A---CACAGAA-TC----AG-C--A---A-GT-AG-A------T--GC--------C------------T--------- 5099
WRC.GM.05.Pbt_05GM_X02        TA--C------TTA--C--AG-A------------G------------------T--G--------------A-T--GAG--AA-----G--T-------AC----A---CA-AGAA-------AG-T--A---A-AT-GG-A------T--GCT-------A-C-A---T--------- 4013
MND-2.CM.98.CM16              T------------TCA---AG----CACTACA--------------------T--G--------------A----GG-A-AA--CC-TC-TC-C--AC----C-----G--A--T------C--A--A--GA-A--AG-A----------T--------------A---T--------- 4785
MND-2.GA.x.M14                TG-----------TCAG--AG----CACCACA-----A------T--C-------------------T--G----G-A-AA----C-TC-TC-T---GT---------AG--A--T--G-----A--A--GAGA--GG------------------------C-----T--------- 4712
MND-2.x.x.5440                T---C--------TCA---AG----CACCACA---G-A--T------C--A--G-----------------------G-A-AA----C-TC-TC-T-----G--C-----A---A--T--G-----A-------AGG--AG-A--------G-----------GC-A--T--------- 4345
DRL.x.x.FAO                   TA--C--------TAAG--AG----TACCACA--G-------T--C-----G---------------G---------A-----C--TC-GC----------------------AA-T-------------A---A-A--AG----------T---------------C-G--T--------- 4351
MON.CM.99.L1                  -TGCA--------AAAT--TC-T--CACCACA--GG-A---------C--A------C-----G--T-----GAA--------G-C--C-----AG-G-CC--TAAGA-AA-C--G---G----T----T-C--CCCA------T--GC-CAA---C-G-------TTA------ 4487
MON.NG.x.NG1                  -TGCA--------AAAC--AG-A--CAGTAC----G-G---------C--A---------------T--------AAC--------GC--A-C--A-CC-CAT----GA-AA-C--G---G-A-T-C-GTTC--CTCT-------C-GGC-CAAA-GC-C-A---T-TA------ 3080
GSN.CM.99.CN166               ----C--------A-AC--AC-A--CACCACA--GG-G--A-----C--G-------------G--------AAC-----C---C--C-C---G---CA--CACGA-AA-C-----AG---TCAC-T-CT-AG-A----------GC-CA----TGC-A---T--T------ 4512
GSN.CM.99.CN71                -T-----------AAAC--AC-A--CAGTACA--GG-G--A-----C--G-------------G--------AAC-----C---C--C-C---G---CA--CCA-A--A-T------G---TCAC-T-CT-AG-A----------GC-CAA---TGC-A---T--T------ 4494
TAL.CM.00.266                 TGGTA------CTCAA---AG----CTCCGCA---G----A--T-----T---TCC--G----C-------AAC-AA------C-GC-C-----G-CA--CAC---AA-T--G---G-A-TAC--T-C--GTCC---------GGC-CA---GACT------TTTG----A- 4750
TAL.CM.01.8023                -AGCG------CT-AA---AG----CACCACA--GG-A--A--T-----A---TCC-----GTC------GAA--A---C---C-G--G-----G-CA--TCA---AA-T--G---G-A-TGC-GT-C--GTCC------G--GGC-CA---TACC------T-TT------ 4248
MUS-1.CM.01.1085              -A--C--------CA-C--AG-A--TACCACA--GG-A--A--T--C--T--G------------G------AAC-AA--C---C-G-----AG--GCC---CACA-AA-C--G--AG---TCACTT--T-GG-A----------GC-CAA-----C------TTTA------ 4474
MUS-1.CM.01.CM1239            -A-----------TA-C--AG-T--CACCACA--GG-A--A-----C--T---------G------------AAC-AA--C---C-G-----AG---C----CAGA-AA-C-----AG-A-TCACCT--T-AG-A----------GC-CA------C-------ATG------ 4464
MUS-2.CM.01.CM1246            TTGCA--------TAAC--AC----CACCACA--TG-C-------------A--G--------------------AA-GCA--C---C-G------G---CC--TCACA-A-----G--AG-A-TAAC-T--T-AG-A----------GC-CAA-----C------T-TA---AA- 4572
MUS-2.CM.01.CM2500            TTGCA--------A--C--AG-A--CAGTACA--TG----T--T--C--A--G--C-----G---------AA-GCT-----GC---------G-G-CC--TAACA-A--C-----AG-A-TCC-GT-C--AG-----------GC-CAA-----C------TTTA---AA- 4531
SYK.KE.x.KE51                 -------------A---G--G-A--CACCACA-----A--A--------A-----C--G---A--------AAC-AA----G-C-------AG---CC--CAAT--AA-T--G--AG-A-TA---AGAT-AG-------G--GGC-------CACG------TTTA------ 4639
SYK.KE.x.SYK173               -A-----------TGT--C--AG-A--CACTACA--T------A------C--A------C--G---A--------G--C-AA----GGT-T------AG--GC---CTC---AA-T-----CG-C-TCAC----T-GC-A-----------GGC--------ACC--------TTA------ 4984
RCM.GA.x.GAB1                 T---C--------TCAG--AG----C-CC------G-A--T--------T------------------G---AGC----------C---------A-T---C-----A---A---------A--A--A---A-AT-AG-A----------C-------------T--GG-T--------- 4263
RCM.NG.x.NG411                A------------TCA---AG----CACC------G-A-----------T------C---------G----------A------C-TC----G----T---A-----G--AA-T-------------A---A-GT-AG-------------T-----------GC----T--------- 4278
GRV.ET.x.GRI_677              TG--C--------GTAAT--AG----CACCAC------A--T--T--C--A--GTCA--G--TAGT-------AGC------C-G-C-GC-C---G-----C--CTCT--AA-----------G-TTG---GAGAT-GG-----------G---------TACGCAT------------ 4957
SAB.SN.x.SABIC                TT--C--------GAAA---AG----CAC------GG-A---------C-------------G-------------------------GC-G------AG----C----------AA-T------G-T--A---AGAT-GG-A----------CAT----------GCA---T--------- 5182
TAN.UG.x.TAN1                 TT-----------GACA-G-AG----CAC-ACA--G--C-----T--C--A---TCA------AGT-------AG---------G-C--------AG----C------A-A-AA-T-----AG-CTG--C-TT-T-GG-A--G------C-TC------CTGCCA---T--------- 4983
VER.DE.x.AGM3                 TG-----------GAAA---AG-A--CACCACA--TG-A------T--C--T---TCA-----CTCTA----GAG-----------C-G------AG----A--T--GA-AA-----------G-CTG-C--T-CACAG-A-------------TT------TGCCA------------ 4506
VER.KE.x.9063                 T------------GAAAG--AG----CACCACA---G-G--A-----------A--GTCA--G---TC--------AGC------C---C--C-C--AG----A--T---A-AA----------G-TTG-C-GT--ACAG-A-------------TT------TGCCA------------ 5011
VER.KE.x.AGM155               TG--C--------GAAAGG-AG-A--CAC-ACG--GG-A--A--T-----A---GTCC--G---TCTA-----AG-------C---C-----G--AG-G--A--T---A-AA----------G-CTG-C--T--ACAG-A-------------TT------CTGCCA---T--------- 5003
VER.KE.x.TYO1_patent          TA-----------GAAA---TG-A--TAC-ACA--T--A--A--T--C-------TC--------TC-A-----AGC-----C---C--------AG-G--A--T--GA-AA----------G-TTGCC--T--ACAG--------------TG------TTGCCAT--T--------- 5008
COL.CM.x.CGU1                 -AGTG------CTA-----AGCA--CACCACA---CAC-----T-----T---TCA---------T-----GCAG-GA--C--G--TG-T--A-G--AG--TAAGA-AA-G-A-------A--A---ACAT-AG-A-GCAA----GC-----------C-ATGCG-T-------- 4547
```

```
H1B.FR.83.HXB2            AAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGA...ATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATT...ACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAA 4948
Pol                       _K__R__K__G__G__I__G__G__Y__S__A__G__E__R__.__I__V__D__I__I__A__T__D__I__Q__T__K__E__L__Q__K__Q__I__.__T__K__I__Q__N__F__R__V__Y__Y__R__D__S__R__N__P__L__W__K__G__P__A__K_
H1A1.UG.85.U455           ----------------------A-------------------...---A--------------------------------------------------...T-----------------------------------------G-C--CA------------------- 4394
H1B.US.90.WEAU160         ------------------------------------------...---A--------------------G---C--C------------------------...-------------------------C---------------G------------------------- 4947
H1C.ET.86.ETH2220         -------G----------------------------------...---A----T--------T---------G---------C-C-----C----------...TT-------------------------------A---------G-C--TA---------------C-- 4340
H1D.CD.84.84ZR085         --G---------------------------------------...---A------------T----T---------G-----------------C-------...------------------------------------------G----A------------------ 4470
H1F1.BE.93.VI850          ------------------------------------------...---A---------T----------------G---------------------------...-G--------------C---------------------------C--G------------------ 4288
H1G.SE.93.SE6165          ----------------------------A-------------...---A------------T----T--------------C--------------------...------------------------------------------G-C--G---------------C-- 4345
H1H.CF.90.056             ------------------------------------------...---A--------------------------------------------------...T----C--------A---------------------------G-C--A------------------ 4295
H1J.SE.93.SE7887          ----------------------A-------------------...---A-----------------------G-----------------------------...------------------------------------------G-C--A------------------ 4262
H1K.CM.96.MP535           ------------------------------------------...--------T-----------------------------------------------...TT---C--------A---------------------------G-A--A------------------ 4144
H101_AE.TH.90.CM240       ------------------------------------------...---A--------------------------------------------------...------------------------------------------G-C--A------------------ 4522
H102_AG.NG.x.IBNG         ------------------------------------------...---A----T--------T----T--------------C---------G-------...-T----------------------------------------G-C--CA----------------- 4473
H103_AB.RU.97.KAL153_2    ------------------------------------------...---A--------------------------------------------------...-T-------------------------------A---------G----A------------------ 4174
H104_cpx.CY.94.CY032      ------------------------------------------...---A------------T----T--------------C--------------------...------------------------------------------G-A--A------G----------- 4314
H1N.CM.95.YBF30           --------------------------C---------------...---A--------------------G--A-C-A-T-------C--------------...TT---G---------------------------A---------G----CA---------------C-- 4542
H1O.BE.87.ANT70           --------------------------C------A----G---...---A-------T-----T--C-A--------A-C----------------------...TT----N----C--A---------C--------A---------G-C--TA-C-----------G---C- 5003
H1O.CM.91.MVP5180         --------------------------C---------G-----...C--A-------C-----T--C-A--------A-C----C-----------------...TT------CA-C------------C--------A--T------G-C--TA--------------G---C- 4978
H1P.FR.06.RBF168          --------------------------C---T--A--C--G--...--CA----T--T--G--T---C-A--------A-C----------------------...TT-----------------------C--------A---------G-C--TA-----------------G-C 4409
CPZ.CM.01.SIVcpzCAM13     --G-----------------------C---------------...---A----------------------------A-----C-----------------...TT--------------A----------------A---------G-C--GA-------------G--C-C 4528
CPZ.CM.05.SIVcpzEK505     --G-------------------A----C--------------...---A------------------T--------A-C-A-GC------C-----------...TT---G-------------------------A---------G-G--CAC-------------C-- 4520
CPZ.CM.05.SIVcpzLB7       ----------------------A----------A--------...---A------------------T--------C------------------------...T----G--------A------------C------G----TA----------------C-- 4503
CPZ.CM.05.SIVcpzMB66      ---------------------------------A--------...---A--------C---------T--------C-A-------C--------------...T----T---------A------------------------G-C--TA----------------T-- 4494
CPZ.CM.05.SIVcpzMT145     ---------------------------------A--G-----...---A-------C-------------------A-------C----------------...TT-----------C-A-------C-------A------A-G-----A------------------G-- 4490
CPZ.CM.-.SIVcpzMB897      ------------------------------------------...--CA----T----------------T--------G----C-----------------...T----G--------A---------------------------G-A--TA-C--------------C-- 4514
CPZ.CM.-.SIVcpzDP943      --------------------------C------A----G---...--TA-------T--------------------A-CTA-T-----------------...TT---G------------------------------------G-----A-------------T--C-G 4534
CPZ.CM.98.CAM3            --------------------------C------A--------...--TA-------C--------------------A-CTA-TC----------------...TT---G-------------------C------------------G----CA-------------T--C-G 4362
CPZ.CM.98.CAM5            ------------------------T-C------A----G---...--TA-------C--------------------A-CTA-T-----------------...TT---G------------------------------------G-----A-------------T--C-G 4650
CPZ.GA.88.GAB1            --------------------------C---------------...---A-------------------------A-GC-----------------------...TT---G--------A------------------A---------G-C--A--------------G----C 5009
CPZ.GA.88.GAB2            --------------------------C------A--------...-----------GC--------T--------AC----C--------------------...CT--------------A----------------A---------G-C--CA--------------CTC 4341
CPZ.US.85.CPZUS           ------------------------T-C------A--------...--CA-------C------T---AC--------AG-CTT-------------------...TT---G------------------C----------------G-----A-----------------C-- 5008
CPZ.CD.90.ANT             --------------------------CAC-T--AC-G-----...---T-------C----------------G--A-CTC---------T----------...TT-----------C-A---------C-------T------G-C--TG-G--------------C- 4387
CPZ.TZ.00.TAN1            --------------------A-----C---------------...C--T-------C--A-T---A-T----G--A---C----------------------...TT---G---------------------T------GC----G-----A-------G---------GCG 4592
CPZ.TZ.2001.TAN2          --------------------A-----C---------------...C--T-G-----C--A-T---A-T----G--AC--C----------------------...TT---G---------------------T------GC----G-----A-------G---------GCG 4597
CPZ.TZ.2002.TAN3_1        --------------------A-----C---------------...C--T-G-----C--A-T---A-T----G--AC--C----------------------...TT---G---------------------T------GC----G-----A-------G---------GCG 4597
GOR.CM.2007.SIVgor2139_287 -----------------------A----C-------A--C--G...---A------GT----C---C-A--------A-C----------------------...TT-----------------------C--------A---------G-C--TA-----------------G-C 4485
GOR.CM.2007.SIVgorCP2135con -----------------------A----C-------A--C--G...---A------GT----C---C-A--------A-C----------------------...TT-----------------------C--------A---------G-C--TA-----------------G-C 4484
GOR.CM.2004.SIVgorCP684con -----------------------A----C-------A--C--G...---A------GT----C---C-A--------A-C----------------------...TT-----------------------C--------A---------G-C--TA-----------------G-C 4404
H2A.DE.x.BEN              -------GG--A--A--A---ATATG-CCC---CA-------...C--A-CA---G--TA-C-----ACA-G---TAC--TTCC-C----G-A---A-...T----TT--A-------C-A---C---------A--AG-----G---AG--G------------T-GTG- 5330
H2A.GW.x.ALI              -------GG--A--A--A---ATATG-CCC---CA-------...C-CA-CA-T--G--CT-C-----ACA-G---TAC--TTCC-C-----C-A---A-...TTG---T--A-------C-C---C-------------AG-----G---AG--G------------T-GGG- 5323
H2A.SN.x.ST               -----G-GG--A--A--A---ATATG-CCC---CA-------...C--A-CA-T--GG-CA-TG----ACAGG---TAC--TTCC-C---GC-A---A-...T-----T-A---------------C----T--A--AG-----G---AG--G------------T-GGG- 4775
H2B.CI.x.EHO              -------GG--A--A--A---ATATG-CCC-T-CA-------...------CA----G---A-T-----ACA-G---TAC--TTCC-C-----C-A---A-...TT----T-C---------C-----C---------A--AG-----G---A---C-----G-----T-GTG- 5301
H2B.GH.86.D205            -------GG--A--A--A---ATATG-CCC-T-CA-------...C----TA----G---A-C-----GCA-G---TAC-GTTC--C---GC-A---A-...TT----T------------C-A---C---------A--AG-----G---A---C-----G-----T-GTG- 5303
H2G.CI.92.ABT96           -------GG--A--A--A---ATATG-CCC-T-CA-------...---A-TA----G--TA-T-----ACA-G---TAC--TTTCA-----C-A---A-...T-----T--A--------------C---------A--AG-----G-C-AG--C------------T-GTG- 4657
H2U.FR.96.12034           -------GG--A--A--A---ATATG-C-C-T-CA-------...--CA-TA----G---T-C-----ACA-G---TAC--TTCC-C---TC-A---A-...T-----T--A--------------C---------A--AG-----G-C-A-T-G---------C-------G- 4809
SMM.SL.92.SL92B           -------GG--A--A--A---ATATG-C-C-T-CA-------...---------A-T--G--CA-C-----AC--G----CC--T-C---A-TTC-----A-...T-----T--------------------------A--AG-A---G-C-A---G-----G---------G- 4678
SMM.US.x.H9               -------GG--A--A--A---ATATG-C-C---CA-------...T----CA-T--G--TA-C-----ACA-G---TAC--TTCCA----TC-A---A-...T-----T--A--------------C---------A--AG-----G-C-AG--G------------C-GTG- 4720
SMM.US.x.PGM53            -------GG--A--A--A---ATATG-C-C---CA-------...T----CA-T--G--CA-C-----ACA-G---TAC--TTCCA----TC-A---A-...T-----T--A--------------C---------A--AG-----G-C-AG--G------------C-GTG- 5164
SMM.US.x.SIVsmH635F_L3    -------GG--A--A--A---ATATG-C-C---CA-------...T----CA-T--G--CA-C-----ACA-G---TAC--TTCCA----TC-A---A-...T-----T--A--------------C---------A--AG-----G-C-AG--G----G------C-GTG- 5250

MAC.US.x.239              -------GG--A--A--A---ATATG-C-C---CA-------...T--A-TA----G--CA-T-----ACA-G-G-TAC--TTTCA----TC-A---AC...T-----T--A--------------C---------A--AG-----G---A---G-----G-----C-GTG- 5247
Pol                       __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__I__Q__F__Q__Q__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_
MAC.US.x.251_1A11         -------GG--A--A-CA---ATATG-C-C---CA-------...T--A-TA----G--CA-T-----ACA-G---TAC--TTTCA----TC-A---AC...T-----T--A--------------C---------A--AG-----G---A---G-----G-----C-GTG- 5246
MAC.US.x.251_BK28         -------GG--A--A--A---ATATG-C-C---CA-------...T--A-TA----G--CA-T-----ACA-G---TAC--TTTCA----TC-A---AC...T-----T--A--------------C---------A--AG-----G---A---G-----G-----C-GTG- 5223
MAC.US.x.EMBL_3           -------GG--A--A--A---ATATG-C-C---CA-------...T--A-TA----G--CA-T-----ACA-G---TAC--TTTCA----TC-A---AC...T-----T--A--------------C---------A--AG-----G---A---G-----G-----C-GTG- 4715
STM.US.x.STM              --G-A--GG--A--AC-A---ATATG-CCC---CA-------...T----TA-T--G--CA-C-----ACA-G---TAC--TTTCA----TC-A---A-...T-----T-CA-------------C-----------AG-----G---A---A-----G-----C-GTG- 4894
MNE.US.x.MNE027           -------GG--A--A--A---ATATG-C-C---CA-----G-...T--C-TA----G--CA-T-----ACA-G---TAC--TTCCA----TC-A---AC...T-----T--A--------------C---------A--AG-----G---A---G-----G-----C-GTG- 4715
DEB.CM.99.CM40            -------GG--A--A--A---ATATGTC-C-GATA-------...C-----A----G---A-------A---G----A---ACT-------C--A-------...T-----C--TT-GGC--CAA---C---------A--AG-TGCTG-----AC----------G-------C 4754
DEB.CM.99.CM5             -------GG--A--A--A---ATATGGCA---ATA--G----...C-----A----G---A-T------T-GG----AC--ACC-------C--A-------...T-----C--TT-GG--ACAA---A---------A--AG-AGCTG-C---AC-------G-----C--T-C 4748
DEN.CD.x.CD1              -------GG--A--AC-A---ATATG-CA----CA------TAC--TAACATGC---ATA--CATTTAGA-AT-CAAC-T......ACC----------......C-----T--TCG-------AA---C---------A--GG-AGCT-----G-AA---C-----------TG- 4832
LST.CD.88.447             ---CA--GG--A--AC-A------CT-TCAC---CA-----G...T-TA-TA-T--G--TAATG-----T--G----AC--T-T----------T--A-...T--------TT---A---AA------C-----AC-AG-AC--G-----C-AG-----G-----T---C- 4225
LST.CD.88.485             ---CA--GG--A--AC-A------CTTTCAC---CA-----G...T-TA-TA-T--G--TAATG-----T--G----AC--T-T----------T--A-...T--------TT---A---AA------C-----AC-AG-AC--G-----T-AG-----G-----T---C- 4225
LST.CD.88.524             ---CA--GG--A--AC-G--------TATCAC---CA--G---...TACA-CA-T--G--CAATG----TT--G-----C--T-CA---------TT--A-...T--------TT---A---AA------C-----AC-AG-A---G---------AG----G------T--TC- 4222
LST.KE.x.lho7             ---CA--GG--A--AC-A------CTT-CAC---CA--G--G...T-TA-TA-T--G--TAATG-----AC--G----AC--T-TC---------TT--A-...T--------TT---A---AA------C-----AC-AG-A---G-----T-AG----------------GC- 5309
SUN.GA.98.L14             ---CA--GG--A--AC-A-----AATG-CA----CA--G--G...T-TA-CA----G--TAATG----TC--G----AC--T-C-----G---AT--A-...T--------TT---A---AA---C--C-----A--AG-A---G-----T-AG--------------T--T-- 5308
MND-1.GA.x.MNDGB1         ---CA--GG--A--AC-A------ATG-CA----A--G----...-------A-T--G--CAAT-----AT--G--TA-C--T-TCA------T--------...T--------A-TT-------AA------C-T---A--AG-A---G---AG--G------------T-GT-T 4635
OLC.CI.97.97CI12          ----AC--G--A--ACAA-----AGGT-CA-TG-CA--GGA-...T------A--CA----TAC-----T-C-...CAA-T------ACAC-C--A-T-AA...TTT----AATTTTCAGGATATCGAGTG--TTATA-A-A--A-G-AGGC-AG---TGT--C-------G- 4816
WRC.CI.97.97CI14          ---GA---G--A--AT-A--------...TCCC-T-CA------...T------A-T--G---A-CT-----TC--G----AC--C--ACT---C------AA...TT----T--A------------A---G-----------AACAG-AGCC-----AG-AG----C-----------GT-- 5245
WRC.CI.98.98CI04          ---GA---G--A--AT-G------A...TCCC---CA-------...T------A-T--G--CA--G-----AC--G----C--TC--AC----C------AA...TT----T--A------------A----------------AACAG-AGCC-----AG-AG----C-----T----GT-- 5260
WRC.GM.05.Pbt_05GM_X02    ---GA---G--A--AT-G------A...TCCC---CA-------...-------A-T--G---CA-T-----T---G-------TC--AC----C------AA...TTC---T--A------------A-------C-------ACAG-AGCCTCG-AG-AG----C-----------GT-- 4174
MND-2.CM.98.CM16          -----------------A----AA-------C-----A------ATAG---ACATT---GC-A--GACCT-C--AC--------TT-CA---C--...T----...CA---------------------------------A--AG-A--GG-C-A--AG---------------G- 4949
MND-2.GA.x.M14            -----------------A----A--------C-----A----...------G--T------A--------T--T-T-----CCA--------C--A-T----...T------------------------------------A--AG-A--GG-G-AG--C---------------G- 4876
MND-2.x.x.5440            -----------------A----A--------C-T---A----...----------------A-C--------C-T---A-CCA-------G-C--A-T----...T-------------------A-------------A--AG-A--GG-C-A--AG---------------G- 4509
DRL.x.x.FAO               -----------------A----A-------------CA----...-----------T------CT-T---C-CTT---A-GCA--C------C--A-T----...CTC-------------------------------A--GG-A---G---AGT-G---------------G- 4515
MON.CM.99.L1              --------G--A--A--A---ATATGGCCC-T-TA--C----...T-TA--A-----G---CAC-----AT--G--CTAC--ACC-C-A-C--C---------...-------T--TC---A--------------T----ACAG-TGCCG-C--T-AG--------------GCG 4651
MON.NG.x.NG1              -------RG--A--A--A---ATAT----C--ATA--G--G...T-----CA----G--TCAT-----TT--G--CTAC--C-C-C-A-C--C---------...-------T--TCG--A--------------T----AACAG--GCTG----TTCG------G-----------C- 3244
GSN.CM.99.CN166           --------G--A--A--A---ATATGT-CC-TACA-------...---------A-T--G--TCAC-----AT--G----AC--C-C---A-C-C------A-...T-C---T--------A------------------C-AG-AGCT--C--T-------C--------------GT 4676
GSN.CM.99.CN71            --------G--A--A--A---ATATGT-CC-TACA--G--G...T----GA----G--CCAC-----AT--G----AC--C-C---A-C-C------G-...T-C---T--------A------------------C-AG-AGCT-----T-A----C--------------GT 4658
TAL.CM.00.266             -------G------AT-A---ATATGT-CC---CA---GC-...----A-TA----G--CTAC-----AC---------C-C--C--------C---------...CA---TT--TCGG-----AA---C----------AA-GG-TGCT------T-------C-G---------TC- 4914
TAL.CM.01.8023            -------G------AT-A---ATATGT--C---CA---GC-...C-CA--A----G--CTAC-----AC--------A-C-ACCC--------C---------...CAC--TT--TC-G------AA---G--C--T--AA-GG-TGCT------T-------C-G---------TC- 4412
MUS-1.CM.01.1085          ----AG--G--A--AG-G---ACCTATCAAAT-CT--G--G...---A-CA----GC--CAC-----AC--G--CTAC--C-C-------C------AA...TAC---T---------------------------T---ACAG-TTCTG----GTC-------G------C--CGT 4638
MUS-1.CM.01.CM1239        ----AG--G--A--A--A---ACCTAGCA--T-CT--G----...---A-TA-T--GT--CAT-----AC--G--CTAC--C-C-------TC-----AA...T-C---T---------------------------T---ACAG--TCTG-C---TC-------G-----T--TGC 4628
MUS-2.CM.01.CM1246        --------G--A-----A---ATATG-CCC-TACA-----G...T----TA-T--GC--TAC-----AC--G---TAC--C--C--------T--C-A-...-------T--TC---A-----------------AAGAG-AGC---C--TTA-----CTGA--------TCG 4736
MUS-2.CM.01.CM2500        --------G--A--A--A---ATATG-C-C-CACA-----G...T----CA----GC--TAT-----A---G----AC--C--A--------CAC--AA...........T--TC---A-----------------AAGAG-AGCG-------AG-----C--------T-TC- 4689
SYK.KE.x.KE51             --------G--A--A--A------ACT-CAC---CA--T--G...TATA-TA----GC--TAC-----AC------CTAC--A-TAC---C-C-..........CA----T--TCG-------AAA--A-----T--AC-TG-T-CT...T-TGA-----C-G----------TTC 4794
SYK.KE.x.SYK173           --------G--A--A--A------ATATCAC----A-----G...TACA-TA-T--GC--TAT-----AC-----TTAC--C--AATAC--C-TC-......C-----T--TCG-------A-------C---C-CC-GG-G-AG...AACGAG------G----------TCG 5142
RCM.GA.x.GAB1             ---------------------------------A--------...---A----T------------------T--GC---C--TA-----------T-------...T----------------------------------------A--AG-A--GG---A---G-----G---T------T-- 4427
RCM.NG.x.NG411            ---------------------------------A-----G--...---A----------------T------T-------A-CCA-----------C--------...T----------------------------------T---A--AG-A--GG---AG-AG------G--------T-- 4442
GRV.ET.x.GRI_677          --------G--A--A--A-----TAT-TC-AGT-CA------...T-G--TA-T--GC--A-----C-AC--G--CTA--TACTC--------C-----C...CA--------TTG-------AA---C--C-----A--AG-T---G------G-G---------------GCG 5121
SAB.SN.x.SABIC            ---------------------------TC------CA-----...C--A-CA-T-------CAT-----AC--G-----C---AC-C-T---C--A-------...T---------------------------C-----------AG-----G-C--TG-G---------T--G--T-- 5346
TAN.UG.x.TAN1             --------G--A--A--A-----TATG-CCC---CA--G--G...C--A--A----G---A-C---C-AC--G---TAC--C-CA------C----CAA...CA-------TC--------AA---C--C-----A--AG-A---G-G---G-G---------------C-C 5147
VER.DE.x.AGM3             --------G--A--A--A--------TA-CAC-G-CA--G--...T--A-CA-T--G--TA-T---C-AT--G--TTAC--C-CC-------CCA-------...CA--------TT--------A-A--C--C-----A--AG-G---G----TG-C---------------G-C- 4670
VER.KE.x.9063             --------G--A--A--A--------TA-CAC---CA-----...C--A-TA-T--G---A-----C-AT--G--CTAC--AC--------CCA-------...CA--------TT--------A-A--C--C-----A--AG-G---G------G-G-----G-----------C- 5175
VER.KE.x.AGM155           --------G--A--A--A------CTA-CA--T-CA--G---...C--A--A-T--G---A-----C-AT--G---TC--CACTC-------CCA----C...CA--------TTG-------A-A--C--C-----A--AG-----G------G-G-----G-----T--TCG 5167
VER.KE.x.TYO1_patent      --------G--A--A--A-------AC-G-C-T---CA--G--...C--A-TA-T------A-----C-AT--G---TAC--C-T--------CCA-------...CA--------TT--------A-A--C--C-----A--AG-G---G-C--TG-G----------------C- 5172
COL.CM.x.CGU1             --------G--A--AT-A-------A-G---C--T-------...CA------A-G-GC-AT--T---AT--G-T--AC--A-CC--AC----TT-CAAAATCA----T--A----------AA--CC--C-GG-AA--GCA-...-CAGG-GAG---C---------G-G- 4711
```


PLV Complete
Genomes

```
                                                                                                               Vif start                              Pol p31 integrase end
H1B.FR.83.HXB2        GCTCCTCTGGAAAGGTGAAGGGGCAGTAGTAATACAAGATAATAGT...GACATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAG...................... 5096
Vif                                                                                                             M  E  N  R  W  Q  V  M  I  V  W  Q  V  D  R  M  R  I  R##__.__.__.__.__.__
Pol                   __L__L__W__K__G__E__G__A__V__V__I__Q__D__N__S__.__D__I__K__V__V__P__R__R__K__A__K__I__I__R__D__Y__G__K__Q__M__A__G__D__D__C__V__A__S__R__Q__D__E__D__*_
H1A1.UG.85.U455       A--A----------------------------C-------...--T-----G-----A-------------------------------------------------------A-----G------------A-----....................... 4542
H1B.US.90.WEAU160     ---T------------------------------------...--------------------------------------------------------------------------------------------------....................... 5095
H1C.ET.86.ETH2220     A--A------------------------------------...--------G-----A-----G--G-------A--------------------------C-----------G------------A-----....................... 4488
H1D.CD.84.84ZR085     ---T----------------------------C-------...--------G-----A-----------------------------------------------------------------------------A...................... 4618
H1F1.BE.93.VI850      ---A------------------C-------C-------...--A-----GA----A-----G--G-------------------------T--------------G--------------------A...................... 4436
H1G.SE.93.SE6165      A--A----------------------------C----AC...--A-----G-----A--------------------------------------------------------G------------------------....................... 4493
H1H.CF.90.056         A-------------------------------C-------...--A-----------A--------G-G-----A-----------------------------------------------------------A...................... 4443
H1J.SE.93.SE7887      A--A-C------------------------G--C-------...--A-----G-----A-------------------------A----------------------------G--------------------A...................... 4410
H1K.CM.96.MP535       ---A----------------------------C--C----...--A-----G-----A-------------------T-------------------------------------G--------------------A...................... 4292
H101_AE.TH.90.CM240   A--A----------------------------C-------...--T-----------A--------------------------------------------------------G---------------------------....................... 4670
H102_AG.NG.x.IBNG     A--A--------------------------G--C-------...--T-----G-----A--------------------A---------------------------------G---------------------------....................... 4621
H103_AB.RU.97.KAL153_2 A--A-----------------G-------G--C----AC...--T-----------A----------------------------------------------------------------------------------....................... 4322
H104_cpx.CY.94.CY032  A--A--------------------------G--C--C---...--T--C--------A-------------------T--------C-----C----------------A-------------G------------A-----....................... 4462
H1N.CM.95.YBF30       A--T--G--------A----------G----T---------CG-G...--T--------C---C-T--G--------A--A---------------------------G---------------G----------AA--C--GAAATGGAATAG....... 4702
H1O.BE.87.ANT70       ------G-----------G--A---------C-----------GG-A...-----T--G-----A---------G--G------A--A--C--A--G--------------------AC---A--A---------G-----ACA--AAG-G-AAGCGTGGAACAGCCTAGTG 5170
H1O.CM.91.MVP5180     A-----G-----------G-----------C-----------AG-A...-----T-----G--A-----------G------A--A--C--A------------------------AC---A--A-----A-------ACA--AAG-G-AAGCATGGAACAGCCTGGTG 5145
H1P.FR.06.RBF168      ---T--G-----------------------C--------C--AG-A...--T--T--------C--T--G--------A--A--C--A---------------------------------GA--TC---AGA--AAG-G-....................... 4556
CPZ.CM.01.SIVcpzCAM13 C--G--G-----------------------T--G---C-AG-A...--A-----------A--T-----G------A--A-----A------------A-----------------------------------------AAGCATGGAATAG....... 4688
CPZ.CM.05.SIVcpzEK505 A--A--G---------A----------G----T---------CG-G...--T-----------C---C-T-----G--C--A--A-----A-----------------------G-------------G----------AG-C--GACATGGAATAG....... 4680
CPZ.CM.05.SIVcpzLB7   A--A--G-----G--------A--------GC-G-------C-GGAA...--A-----G-----C---C-T--------T--A--T----------------A-----C-A-----C-----GA---------------A---AAGCATGGAACAG....... 4663
CPZ.CM.05.SIVcpzMB66  A--A-------------G--A---------GC-C-------CC-AGAA...--A-----G-----C--TC-----G-----A--------------------------------------------------------A--C-AAACATGGAATAG....... 4654
CPZ.CM.05.SIVcpzMT145 A-----G-------------------G----T-------C-GGAC...--A--------G-----T-----G-----T-----A---------------------------CA---A--------------------AG---AAACCTGGAATAG....... 4650
CPZ.CM.-.SIVcpzMB897  ---T--------------G--A--G-----GC-T---------GGAA...--T-----------C---C-C-----G--T--A--A--C-----C-----------G---A-----A-----C-----GA-------------AG---AAGCATGGAA......... 4671
CPZ.CM.-.SIVcpzDP943  A--T--G---------------------CA---------AGAG...--AG-T-----------C--G--G--------A--A--C--A--------------------------G-A--A--------G------------AG-C-AAACCTGGAATAGCCTAGTA 4701
CPZ.CM.98.CAM3        A--T--G-------------------G-----TA-----G-GAGAG...--AG-T-------A-A--C-----G--------A--A-------------------T--------A--A-----G--G------------AG-C-AGGCCTGGAATAG....... 4522
CPZ.CM.98.CAM5        A--T--G-----------------------G--TA-----G--CGAG...--AG-T-----------A--C---------G--T-----A--C--A-----------------------G-A--A-----G--------A---------C-AAACCTGGAATAG....... 4810
CPZ.GA.88.GAB1        CT-A--G-----------------------G--C--G---C-AG-G...--AC----G-----A---------G-----------G-----------------A--------------------------A-----------AGACTTGGAATAG....... 5169
CPZ.GA.88.GAB2        T-----G-----------G--T--------T--T--G--C--CG-A...--T--------------A--C-----------------T--A--T----AAC-----------G---G---------------C----TG--G--------------A--C--AAC....................... 4492
CPZ.US.85.CPZUS       A--T--G-----------------------C--A---G--C-AG-A...--GG-T-----------A--C-----------------A--T--A-AA--C-----------------------C-----A--A---------G----------AG---AGGCTTGGAATAG....... 5168
CPZ.CD.90.ANT         A--T--G-----------------------C--CA-----CC-AGAG...--A--T--G-----C--TC---G---------T--A-A-----G---A-----A-----A-A--A-AGA---GA-T--------G--------A----------ATATCTGGAAATC....... 4547
CPZ.TZ.00.TAN1        A--A--G------------------------A-----AGGAGAA...--T--------A--C--G----------A-----A-AA--G---------------------CA-G-G--A---ATGA------A-----AC---AACATGGACAAG....... 4752
CPZ.TZ.2001.TAN2      ---G--G------------------------A-----AGGGGAA...--T--------A--C---------------A-----A-AA--A------------------------------CA-G-G--A---ATGA------A-----AC---AACATGGAATAGCCTAGTT 4764
CPZ.TZ.2002.TAN3_1    ---G--G------------------------A-----AGGGGAA...--T--T--------A--C-----G--------A-----A-AA--A--------------G--------CAA--A--A---ATGA------A-----AC-C--AACATGGAATAGCCTAGTT 4764
GOR.CM.2007.SIVgor2139_287 ------G-----------------T-----C--T--G--C--AG-G...--T--------C--T--G-----G-----A--A-----------------------------------------------------GA--CC----GA---AG-G-AAGCCTGGAACAGTCTAGTT 4652
GOR.CM.2007.SIVgorCP2135con ------G-----------------T-----C--T--G--C--AG-G...--T--------T--T--G-----G-----A--A-----------------------------------------------------GA--CC----GA---AG-G-AAGCCTGGAACAGTCTAGTT 4651
GOR.CM.2004.SIVgorCP684con ------G-----------------T-----C--T--G--C--AG-G...--T--------T--T--G-----G-----A--A-----------------------------------------------------GA--CC----GA---AG-G-AAGCCTGGAACAGTCTAGTT 4571
H2A.DE.x.BEN          A--A--G-----G--A-----A-----CA--G-CA-G-TAGGG-CA...----------------A--------G--G--C-----T--C-----C------GG-AGACA--A-CTG---AG-A--CCCCACCTGGAGGGT-CCAG---GG-TGGAGAAATGGCATGC.... 5493
H2A.GW.x.ALI          A--A--G-------G--C--A-----CA--G-TA-G-TAGGG-CA...----------------A--------G--G--C--------C-AA--C------GG-AG-CAA-AGCTG---AG-G--CCCCACTTGGAGGGT-CCAG---GG-TGGAGAAGTGGCATAG.... 5486
H2A.SN.x.ST           A--A--G-----G--C--A-----CA--G-CA-G-TAGGGGC-...------------A-A-A--------G-----T----------C-AA--C------GG-AG-CAA-AGATG---AGCG--TCCAACTTGGAGGGT-CCAG---GG-TGGAGAGGTGGCATAG.... 4938
H2B.CI.x.EHO          T--AT-G--------G-----A-----CA-C---A-G-TAGGG-CA...--A--C------A-A--C--------------------A---A-C------GG-GGA-AA-A-TTG---TGCA---CC-ACGTGGAGG-TACCAT-C-GGCTAGAGAGGTGGCACAG.... 5464
H2B.GH.86.D205        A--AT-G--------G-----A-----CA-C---A-G-TAGGG-CA...--A--C---------A--C--G----------A--T--A---C-C------GG-GGA-AA-G-TTG---TG-A---CC-AC-TGGAGG-TACCAG-C-GGCTAGAGAGATGGCACAG.... 5466
H2G.CI.92.ABT96       ---TT-G--------G-----A--T--CA-C---A-G-TAGGG-CA...--A--C-----T--A--C-------------------A-----A-AA---------GG-GGA-AA-AG-TG---AGCA---ACCAAC-TGGAGG-TACCAG-C-GACTGGAGAAGTGGCATAG.... 4820
H2U.FR.96.12034       CT-A--G-----G--A-----A--G--CC-T--TA-G-TAGGG-CA...------T--G---A-A--C--G-------------------C-A-----------GG-GGA-AA-A-TTG---AG-A--TCCAACTTGGAGGGT-CC---A-GGTTAGAGAGATGGCACTC.... 4972
SMM.SL.92.SL92B       ----------------------T--T-----TA-G-TAGGG-CA...--A------G--A-----------------------------C-AA----------GG-AG-CAA-AGATG-G-AGCA---CCAGT-TGGAGG-TC-GT--A-GA-TAGTGGAGCGCTGGCATAG. 4844
SMM.US.x.H9           ---AT-G--------G-----A-----CA-CC--A-G-TAGGG-CA...--G--C--G-----A-----G--G-----T--A--T--C-AA--C------GG-GGA-AA-A-TTG---AG-G--TCCCACTTGGAGG-TACC-G---GGCTAGAGAAGTGGCATAG.... 4883
SMM.US.x.PGM53        ---AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA...--G--C--G-----A-----G--G-----T--A--T--C-AA--C------GG-GGA-AA-A-TTG---AG-G--CCCCACTTGGA-G-TACC-G---GGCTAGAGAAGTGGCATAG.... 5327
SMM.US.x.SIVsmH635F_L3 ---AT-G--------G-----A-----CA-CC--A-G-TAGGG-CA...--G--T--G-----A-----G--G-----T--A--T--C-AA--C------GG-GGA-AA-A-TTG---AG-G--TCCCACTTGGAGG-TACC-GA---GGCTAGAGAGGTGGCATAG.... 5413
                                                                                                               Vif start                              Pol, Gag-Pol, and p31 Integrase end
MAC.US.x.239          ---AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA...-----T--G-----A--C--------G--T--A--T--C-AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA---GGCTAGAGAGGTGGCATAG.... 5410
Pol                   __E__L__L__W__K__G__E__G__A__V__I__L__K__V__G__.__T__D__I__K__V__V__P__R__R__K__A__K__I__I__K__D__Y__G__G__G__K__E__V__D__S__S__S__H__M__E__D__T__G__E__A__R__E__V__A__*_
Vif                                                                                                             M__E__E__E__K__R__W__I__A__V__P__T__W__R__I__P__E__R__L__E__R__W__H__S##_
MAC.US.x.251_1A11     ---AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA...-----T--G-----A--C--------G--T--A--T--C-AA---------GG-GGA-AA-AGATG---AGCA--TCCCAC-TGGAGG-TACC-GA---GGCTAGAGAGGTGGCATAG.... 5409
MAC.US.x.251_BK28     ---AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA...-----T--G-----A--C--------G--T--A--T--C-AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA---GGCTAGAGAGGTGGCATAG.... 5386
MAC.US.x.EMBL_3       ---AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA...-----T--G-----A--C--------G--T--A--T--C-AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA---GGCTAGAGAGGTGGCATAG.... 4878
STM.US.x.STM          ---AT-G--------G-----A-----CA-CG--A-G-TAGGG-CA...-----T--G-----A--C--------G--T--A--T--C-AA---------GG-GGA-AA-AG-TG---AG-G--TCCCACCTGGAGG-TACCAG---GGCTGGAGAGGTGGCATAG.... 5057
MNE.US.x.MNE027       ---AT-G--------G-----A-----CA-CT--A-G-TAGGG-CA...-----T--G-----A--C--------G--T--A--T--C-AA---------GG-GGA-AA-AG-TG---AGCA--TCCCAC-TGGAGG-TACC-GA---GGCTAGAGAGGTGGCATAG.... 4878
DEB.CM.99.CM40        C--G--G--------A-----T-----G--GTGCA-GACAG-GGTAGGA--T--C-----G--C--T-----G---------AG-A--C-AA--A---------G-C----ATA--A-G-T-AACACACA--A-TCCA-T--AT-A............... 4899
DEB.CM.99.CM5         CT-A--G--------A-----A--G--G--GTGTA--ACAG-GGTAGGA---------------------------------C---G-G--A-AA--------GGAG---ATA--A-GAC-GACACACA--A-CCCAGT--AT-A............... 4893
DEN.CD.x.CD1          ---G--T--------------A-----T--C--CA-GACAC-GGCAGGA-----CTTCCCG--A--T--------G-----A-----C-A-CCA---AATGC-G-A...T-TA-GA-G-GAGAACA-AAC-AAG--AT----AGT--GACTGACATATAA.......... 4989
LST.CD.88.447         ---T--G--------A-----T--T-----GG--A----AGGAGAC...AC---CTTCTC---C--C-------------T---C-AG-A-AA---------G-GGGACCAAAG-A-AG---GG---GCCT-CA--ACA-CT-G.......................... 4364
LST.CD.88.485         AA-T--G--------A-----T--T-----GG--A----AGGAGA-...AC---CTTCTC---C--C-------------T---C-AG-A-AA---------G-GGGACCAAAG-A-AG---GG--AGCCTGGA--ACA-CT-G.......................... 4364
LST.CD.88.524         A--T--G--------A-----T--T-----GG--A-G--GGGAGAG...A----TTTCTCTA-C--C--G--G-----------AT-AG-G-AA---------G-GGGACCAAAG-A-AG-C-GA--AGCATGGA--AC--AT-G.......................... 4361
LST.KE.x.lho7         A--T--G--------A-----T--T-----GG--A----AGGGGAG...A----CTTCTC---C--C------------------AC-AG-A-AA---------G-G...GGAC--AAG---AG-GAGAGTAGCCTGGATA-CA-CT--......................... 5448
SUN.GA.98.L14         A--T--G--------G-----A--------C--CA--C-G...G--GAGA----CCTG------A--T--G--------------AC-AG-A-A---C------GG-G-ATCCT-CTCA-TG--GATG----G-T-A...........................AAGAGAATGGGAAG 5452
MND-1.GA.x.MNDGB1     C--TT-G--------A-------------TT--A--T--C-AGAA...--G-----GA---A--T--------GTGT--A--A--A-AA---------G-G...--GT-G-AAGA--AG-CAG--TAACTTGGA-AGT-TCT--......................... 4774
OLC.CI.97.97CI12      A-----G--------A-----A-----CC-GC-TA-----AGGAGA...A-GTATTTCTCT-----TCTG--G--------A--AG-G-A-CC------GTGG-T-GCAGT-CCA---TA..............T--............................ 4937
WRC.CI.97.97CI14      -----CT--------A-----A---T--------G--ACACCGGAAGGT...----T-AC---T--G------------------C-T--A-AA-T--GGAATGG-G-AGGAATG-A-AGGAG--CCTCT-AT-AG-ATT--........................... 5381
WRC.CI.98.98CI04      ------T--------A-----A--CT-----G-GG--ACCCCAGAAGGT...----T-AC---T--T-A---G-----------A--A--A--A--TC-GGAATGG-G-AGGAATG-A-AGGAG--CCTCTAAT-AAGATT--........................... 5396
WRC.GM.05.Pbt_05GM_X02 AT-GG--------------G--A--TT----TG-G--GAC-CC-GAAGGA--TT--CTTAC---T--C---------------------A-A-AC--GG-ATGG-A-AGGAATG-A-AGCAG-ACCACTAAT......T-A............................ 4307
MND-2.CM.98.CM16      A---A-T--------A-----A--G--G-----CA----GGGG-CA...--T-G--G-----A-----G--------------A-----C--A-------------ACAG---ATA-----CCCAAC----A-GC-T--T--.......................... 5085
MND-2.GA.x.M14        A---A-T--------A-----A--G-----G-----CA----AGGA-CA...--T--------T-----C--------------T--A--A--C--------------ACAG---ATA--A--CCCCACA---AGGA---------GACA-C--GA................ 5024
MND-2.x.x.5440        ----A-T--------A-----C--T--G-----CA----AGGAGT...--TT----G--G--T--T-------------A--A-----A-----C-------------ACAG---ATA-CA--CCCCACA---G---ATAG........................... 4642
DRL.x.x.FAO           A--TG----------G-----A--G--G--C--CA-G--AGG--C-...---T-G-----A-A-T--------------------------AA-------------A-TG---ATA--C--ACCAACA---AG---T--.............................. 4648
MON.CM.99.L1          C--A--------------T--C------C--A--AC-G-GGAGGGC--T--CCTGAC---C--T--G-----G--------A--C-AACCC-----C-CCAC-GAAAAT-TG-G--G-GACACCAACCAATATA-TCTCAGAA-GC-AGATGGATTGGCTAATTAG. 4820
MON.NG.x.NG1          T--T--------G--G-----A--TT-G--TA-GACAG--CAAGGT--GG---T-AC---A--C-C--G--G-----------C-CCC------CTCCA--GAA-AT-TG-G-AG-AAGCCCAGTGCACAT---ATCAGA--GGT-GATGGAATGGCTGATTAG. 3413
GSN.CM.99.CN166       A------------------A-------G-C---AC-C-AGCAGGT--G--T-TTAC---C--T-----G--------------C-A-CCA------G-GGCA-A-A---AA----TGG---GT----CCCATACAA--A-T--C-GAGAGGAAGGTAGACTGGTTG 4846
GSN.CM.99.CN71        A--T-----------G-----A-------GG-C---ACCC-AGCAGGT--G--C-TCAC---C--T-----G--------------C-A-CCA------G-GGCA-A-----AA----TGG---AGT---GCCCATACAAG-A-T--CAGAGAGGAAGGTAGAATGGCTG 4828
TAL.CM.00.266         C---A-T--------A-----A--G--T--CT-G-ACAG-CGAAGGA--GG---TCAC---C--T-----G--G-----A--A-----A-CCC------C--GCTT---GTAACA-GAT-GACC---A-----TCCA----AC----GCAGAGATGGGTAGGGACAATT 5084
TAL.CM.01.8023        C--G-----------A-----A-----C--CC--AGGACAG-CGAAGGA--GG---TTAC---T--C-----G--G---T--A--A--A-AACCC------C--GCT-----GTAACA-GAC-GACT---A-G--TCCA----AC----GCAGAGATGGGTAGGGACAATT 4582
MUS-1.CM.01.1085      CT-A--G--------A-----A-----C--CA--ACAG-CTCAGGA--GG---TCAC---T--T--------G-----A-----A--C-AACCC-----CGCCA--GAA-AT-TG-G-AG-AAGTCC-AT-CAG--G-CCTCAGAA-GG-AGATAGATTGGATAATTAG. 5097
MUS-1.CM.01.CM1239    CT-GT-G--------A-----A-------CG-CA--ACAG-ACAAGGAC-GG---TCACT--A--T--------G-----A--A--A-CA-CCA------CGCCA--GAAAAT-TG-G-AG-AAGTCCAAT-CAG--G-CC-CAGAA-GG-AGATGGATTGGATAATTAG. 4797
MUS-2.CM.01.CM1246    A--T--T--------A-----A--TC----G--CA--ACA--GGAAGGT--G--CGT-AC---C--------------------A--A--C-A-AACCC------C-CCAG-CAAAAT-TG---AG-GAGCCC-ATTCAGT---T-TCAGAA-GG-AGATGGATTGGCTGATTAG. 4905
MUS-2.CM.01.CM2500    A--------------G-----A--A-------CA--ACAG-CCAAGGT--G--T-T-AC---A-----G--G--G--------T--C-AA--C-----C-A--GAAAAT-TG---AG-GACCCCAATCCAGTTT-T--CAGAA-GG--GATGGACTGGCTGATTAG. 4858
SYK.KE.x.KE51         T--A--G--------G--A----A-----C--G--GACACCAGA-CAAC-GG---T--C---A--T-----G---------A--A--C--CAG-G-----G--AGAG---ATA---G-AC-CACC---A--T--C-TCCA-ATCAA---GATAA.................. 4948
SYK.KE.x.SYK173       T--A--T--------------T--------GG-G--ACAG-GGAGGGA----TTTT-C---A-----------G-----A--A--CA--CC------G--AGA----ATA---G-TCCCAC----AG-A-GAC-CCA-GACA----GATAG.................. 5296
RCM.GA.x.GAB1         ---GA----------A-----A--------C--C--G--GG-G-C-GGA--T-G--G-----C--C--G--------------A-----A-AA--A-----C-G-A-AGAT-TG-A-AG---GGCCAATTTGGCGG-TAGAC-G-----AA-TTGA.................. 4581
RCM.NG.x.NG411        A---A-----G----A--------T--C--G--C--G--C-AGAG...--TC-----G-----C--T------GTGT--------A-AA---------C-G-A-AGA--TG-A-AGA--CCCAATATGGAAG-TAGAC-G-----AAGTTGA........GCAGTGGC 4601
GRV.ET.x.GRI_677      A---A----------A-----C--G--G-----TA----GGGGGAA--------C--G-----C--C--G-----G--T-----T--C-AA---------G-GAGA-AAA--ATG---AG-GAG-GTAGT-TGGAGGGT-TCAGA--GGCAAATAAGCAGATGGAGGGGG 5288
SAB.SN.x.SAB1C        ----A----------A-----A---------T------AC-GG-...--AT-G---AC-A-C----------------C--A--T-A-AA-------------TTG---AG-CAG-CCC-TCTGGA-GGTA-C-G-CGGACAGCAGGAGAGGTGGACTAG. 5506
TAN.UG.x.TAN1         T-------------G--A-----A--G-----C---A----AGG-GAG...--GT-G-------A-----G-----G-----A--T--C-AA--A------G-GAGA-AAA-T-TG-G---CAAGACTAAC-TGGAGGGTT-G.......................... 5286
VER.DE.x.AGM3         -T-AA-T--------G-----T-----G--C--CA--...GGAG--GTG--AT------A-TAC---------G-------A--T--A-A--------A-CC-AGA-AAAG-ATG-G---GAGAGTAACTTGGA-GGT-CC-GA-GAGCTGATAACTAAATGGCAAGG. 4836
VER.KE.x.9063         AT-AA----------G-----T-----G--C--CA--...G-AG-AGAA--GC-G---A-T--C-------------------------C-AA-------A-CC-AGA-AAAG-ATG-G---GAG-GTAAC-TGGA-GGTCTCAG--GACCTGATAACCAAATGGCAAGG. 5341
VER.KE.x.AGM155       C--GA----------A-----C--G-----TC-CA-G...G-AG--GAA--AC-G--G-----T--G-----G-----A-----A-AA-C-----AGCC-AGA-AAA--TTG-G---GAGACTCACCTGGA-GGT-CA-GA-GAAGTGATCACCAAATGGCAGGG. 5333
VER.KE.x.TYO1_patent  AT-AA----------G-----A-----G--CC-CA-G...G-CG-AAGT--C-----G--T--A-----------G-----T--A--T-----A---------A-CCCA-ACAAAG--TG-G-A--GAG-GT-ACGTGGA-GGTACCAG--GA-CTGATAACTAAATGGCAGGG. 5338
COL.CM.x.CGU1         A--GG-G--------------A--G-----C--CAG-A--TC-CAAGGTAC-T-GTTT--GAAA--T---------T--------CC--ACTCC-ATATGG-G-AGAT-TG--CAG---GAA-T--TT-TCC-ATGG-C-AA----AGCCGAAACCGTTAAGGGAATGG 4881
```

```
H1B.FR.83.HXB2               ...................................................AACATGGAAAAGTTTAGTAAAACACCATATGTATGTTTCAGGGAAAGCTAGGGGATGGTTTTATAGACATCACTATGAAAGCCCTCATCCAAGAATA...AGTTCAGAAGTACACATC 5211
Vif                          .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__T__W__K__S__L__V__K__H__H__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_

H1A1.UG.85.U455              ...................................................---------C--------------T--C--------C--CAA-------CAA--T------------------------TAGA---T-----G--...----------------- 4657
H1B.US.90.WEAU160            ...................................................---------------------------------A-----AA--------A---------C---------------C-----A----------------...-------------T--- 5210
H1C.ET.86.ETH2220            ...................................................---------T-----------G---------C--A-----A---G-----AT------G-------------T-----C---AGA-------AGG--...----------------- 4603
H1D.CD.84.84ZR085            ...................................................C--------------------T--------C--A-----AA--------AA---------------C--------C--------C---A----...----------------- 4733
H1F1.BE.93.VI850             ...................................................C--------------------T-T--------------AA------C-AA-------C---------------T-C-----AGG---------G--...----------------- 4551
H1G.SE.93.SE6165             ...................................................-------C-C-----G--------T------------C---AA---------A--C-------------C--------C---AGG---------G--...----------------- 4608
H1H.CF.90.056                ...................................................C-----------C--------GT--------C--A-----A----------A-------------------T-T-------A----------G---...----------------- 4558
H1J.SE.93.SE7887             ...................................................C--------------------GT--------A--------AA----------AAA----C-G-----G-----------C---AA--------A---...----------------- 4525
H1K.CM.96.MP535              ...................................................T--T-------------------T--------C--A-----AA--------ACC------A-------------------------AGG-------A---...----------------- 4407
H101_AE.TH.90.CM240          ...................................................---------C--------------T---------A-C---AA--------AAAAG----------------T----------AG-------AGG--...--C-----------T--- 4785
H102_AG.NG.x.IBNG            ...................................................---------T--------------T----------C---AA--------A---T------------------T------TAGG-------A-G--...T---------------- 4736
H103_AB.RU.97.KAL153_2       ...................................................---------C--------------------A---A-----AA--------------------G------A---------------AGGA--------...----------------- 4437
H104_cpx.CY.94.CY032         ...................................................---------C---C----G-----T--------------AA------------AA-----A--C------------------TAGG--C----A-G--...-------------T--- 4577
H1N.CM.95.YBF30              ...............................................................C-----------T----------G---AAA--G--A-AA-------A------------T-------CA-A---C----A---...-------------T--- 4805
H1O.BE.87.ANT70              AAATACCATAA..........................................................---CAGG--TA-A--GA-CGA-AAC----GG--C---------T------TC-AGAA--------G-T...------AGT---T-T--T 5266
H1O.CM.91.MVP5180            AAATACCATAA.........................................................A--CA-G--TAA---G--CGC-AAC---CG------G-----T------TC-AGGA------A-G-C...-----G-CG--GT-T--T 5241
H1P.FR.06.RBF168             ..................................................A-G-C-----T---C-G--C---T-T--C-AA---AGG--TAAA--GA-A-A-AAT-----------G-----T-----G-CRGTC-----T--G--T...--------G---TT----T 4672
CPZ.CM.01.SIVcpzCAM13        ................................................................CC----T---T-T---------AAAA--AA------A-AA-AT-----C--C-----C----------CAT---A---------G-T...TC-------G--------T 4791
CPZ.CM.05.SIVcpzEK505        ................................................................---G-----G--T--C--------G---AA---G--AC-A--T-----C------------------T-------CAGA-A------A------...----------A-------- 4783
CPZ.CM.05.SIVcpzLB7          ................................................................---G-------T-T-----C------A--CA-------A--A--C-------------------T-----G--TGAAA------A-G--...----------A----T--- 4766
CPZ.CM.05.SIVcpzMB66         ................................................................CC-----------T--C---------G--TA-------C--AAAT------C-------------C---------AA-A-------A----...----------------- 4757
CPZ.CM.05.SIVcpzMT145        ................................................................-------G---T-T--------CAGA---AA-------A-AAAAT-----C-----------------TG--A-CA-C--T--G--G...GC----C-GA--T-T--T 4753
CPZ.CM.-.SIVcpzMB897         ...............................................................C-----G--------T--C--------G--TAAA--G--AC-A--T---------C----------T------TGAGA-------A-G-T...-----------A-T-----T 4777
CPZ.CM.-.SIVcpzDP943         AA..................................................................-T----------CAGAA--AA------A--AAA-------------G--C--T--------TT--A-C--C---G---...-----------A----T--- 4797
CPZ.CM.98.CAM3               ................................................................CC---------TGT---------AAAA--AA-------A-A---AC-------------G--C--T--------TT-------T-A-G---...-----------A-------- 4625
CPZ.CM.98.CAM5               ................................................................CC---------AC----------AGAA--AA---G--A-AAAAG-----------G------T-----G--TGA-A-----T-A-G-G...-------------T--- 4913
CPZ.GA.88.GAB1               ................................................................-------T---T-T-----T---AGG---AAA--G--A--A--T-----------C--------T-----TCA----A-------A-G-T...GC---T----A-------A 5272
CPZ.GA.88.GAB2               ..........................................................------T------------GTTT---------CAAG---AAA--G--C--A--------------------G--C--T------T--T--CA-----T-A-G---...-----------C-------A 4604
CPZ.US.85.CPZUS              ................................................................-C-G-------T-T--C-------AGA---AA----------AA---------C--C--G-----T-----G---A---A-----C---G---...-A------------T--- 5271
CPZ.CD.90.ANT                ................................................................C-----G-------C--A-GG-AGA-TAAAGT-CT--AACC----AAG------------T-------ATGA---C-----AG-A-...G-AGA----------T--A 4650
CPZ.TZ.00.TAN1               ................................................................CC----T------T------C-T-ACAA-CAAATGCTG--AA-AT---AAG------------T-------CTGA-ACA----A--G-...GCAGGG---A-------A 4855
CPZ.TZ.2001.TAN2             AA.....................................................................---T-----C-T-ACAA-TA-ATGTTG---A-A----AAG------------T-----G-CTGA-GTA----A--G-...GCAGG----A-------A 4860
CPZ.TZ.2002.TAN3_1           AA.....................................................................---T--C--C---ACAA-TAAATGTTG--AA-A----AAG------------T-----G-CTGA-GTA----A--G-...GCAGG----A-------A 4860
GOR.CM.2007.SIVgor2139_287   AA.....................................................................-T------AA---AGAA-TAAAC----A-AACAG-----C-----G------------GCAGTCA----T--GG--...--------GG---T----A 4748
GOR.CM.2007.SIVgorCP2135con  AA.....................................................................-T------AA---AGAA-TAAAC----A-AACAG-----C-----G------------GCAGTCA----T--GG--...--------GG--GT----A 4747
GOR.CM.2004.SIVgorCP684con   AA.....................................................................-T------AA---AGAA-TAAAC----A-AACAG-----C-----G------------GCAATCA----T--GG--...--------GG---T----A 4667
H2A.DE.x.BEN                 ............................................................CCT-G-CAAGT-C-TGA-ATACAGAACAAA--ACCT--AGGA--TGC-C-A-GT-CCC--C---A-G-TGG-ATGGGCATGGT-G-CT...T-CAGCAGG---ATAT-- 5599
H2A.GW.x.ALI                 ............................................................CCT-G-CAAGT-C-TAA-GTACAGAACAAA--ATCT--AA-A--TG--C-A-GT-CCC--C--TA-G-TGG-ATGGGCATGGT-G-CT...T-CAGCAGG---ATAT-- 5592
H2A.SN.x.ST                  ............................................................CCT-A-CAAGT-T-TAA-ATACAGAACAGG--ATCT--AG-A--TG--C-A-GT-CCC--C--TA-G-TGG-ATGGGCGTGGT-G-CT...T-CAGCAGG---ATAT-- 5044
H2B.CI.x.EHO                 ............................................................TCTAA-TAA-T-C-TGA-ATATAGGACAAA--ACTTGCAACA--TC-CT-A-GT-CCT--C--TA-G-T-G-ATGGGC-TGGT-G-CT...T-CAGTAG----ATAT-T 5570
H2B.GH.86.D205               ............................................................TCTGA-TAAGT-T-TTA-GTATAGAACAGG--A-TTGCAACA--TC-CT-A-GTCCCT--C---A-G-T-G-ATGGGC-TGGT-G-CT...T-CAGTAG-A--ATAT-T 5572
H2G.CI.92.ABT96              ............................................................TCTGA--AA-TTCTTAA-GT-TA-AACAAA--ACCT-CAG-A--CGGTT-A-GT-CCT------A-A-TTG-ATGGGC-TGGT-G-CT...T-CAGCAGG---ATAT-- 4926
H2U.FR.96.12034              ............................................................CCTGA-CAA-C-C-TGA.GTATA-CACAAA--A-CT-CAG-AA-TG--T-A-GTGCCC------A-G-T-G-ATGGGC-TGGT-G-CT...T-CAGCAGG---ATAT-T 5077
SMM.SL.92.SL92B              ............................................................CT-CA-TAA-TTC--TA-GTATA-AACAAG--A-CT--AG-AA-CC--C-A-GTGCC-------A-A-T-G-ATGGGCATGGTAC-C-...GCC--TAGG--GATTT-- 4950
SMM.US.x.H9                  ............................................................CCTCA-CAA-C-C-TGA-ATATA-CAC-AA--ACCT-CAG-A--CT--T-A-GTACCC-----TA-A-TTG-ATGGGCATGGT-G-CT...T-CAGCAG----AT-T-- 4989
SMM.US.x.PGM53               ............................................................CCTCA-CAA-T-CTTGA-AT-TAGAAC-AAG-ACCT-CAG-A--CT--T-A-GTACCC-----TA-A-TTG-ATGGGCATGGT-G-CT...T-CAGCAG----AT-T-- 5433
SMM.US.x.SIVsmH635F_L3       ............................................................CCTCA-CAA-C-C-TGA-ATATA-CAC-AAG-ACCT-CAG-T--CT--C-A-GTGCCC-----TA-A-TTG-ATGGGCATGGT-G-CT...T-CAGCAG----ATTT-- 5519

MAC.US.x.239                 ............................................................CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC---TTTA-G-TCG-ATGGGCATGGT-G-CC...T-CAGCAG----AT-T-- 5516
Vif                          .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__L__I__K__Y__L__K__Y__K__T__K__D__L__Q__K__V__C__Y__V__P__H__F__K__V__G__W__A__W__W__T__.__C__S__R__V__I__F_

MAC.US.x.251_1A11            ............................................................CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC---TTTA-G-TCG-ATGGGCATGGT-G-CC...T-CAGCAG----AT-T-- 5515
MAC.US.x.251_BK28            ............................................................CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC---TTTA-G-TCG-ATGGGCATGGT-G-CC...T-CAGCAG----AT-T-- 5492
MAC.US.x.EMBL_3              ............................................................CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC---TTTA-G-TCG-ATGGGCATGGT-G-CC...T-CAGCAG----AT-T-- 4984
STM.US.x.STM                 ............................................................CCTCA--AA-C-TTTAA-GTATA-CAC-AA--A-CT-AGC-A--C---C-A-GTGCCC--C--TA-G-TTG-GTGGGCATGGT-G-CT...T-CAGTAG---GATTT-T 5163
MNE.US.x.MNE027              ............................................................CCTCA--AA-T-T-TGA-ATATA-AAC-AA--ATCT-CAA-A--TT--C-A-GTGCCC-----TA-G-TCG-ATGGGCATGGT-G-CC...T-CAGCAG----AT-T-- 4984
DEB.CM.99.CM40               ..........................................ATTAGATAGGTGG--C-GTGC-A--AA--G-G-AA-ATACA-GACCAA--AATT--A--CT-T-CAA-GGATACAC---TTTG-ATT---GAACTCATATTAT-C-...CAAA--A--A--GTAT-- 5023
DEB.CM.99.CM5                ..........................................ATTAGATAAGTGG--T-GTGCCA--AA--G-GGAA-GTATA-GACCAA--AATT-CAAGCAAT-AA--GGATACAC---T-TG-ATT---AAGGGCATTCTAC-CT...CAGA--A--A--GT-T-- 5017
DEN.CD.x.CD1                 ..........................GATTGGGGAAACCAAATTAGAACATTGGAT--GC-C---GAA-TCTATAATATACA-AACAAAT-A-TT-AAAG--AC-GT-A--ATACC----TTTAGAATC-CATGGG-ATGGTAT-CC...-TGA--A--A-------- 5128
LST.CD.88.447                .............................................GAGCCT-GT-GA---GTTCC---AT-TCAGGAGAGCAC--AAAGAGACTC--AA-GATT-GGTAAGC-TACATGGAACAGGG-T-G--TGGG-ATTTTAT-CT...TA-AATA-TA--GTG--A 4485
LST.CD.88.485                .............................................GAACCTG-T-GA----TTCC---AT-TTAGGAGAGCAC--AAAGA-ACTC-G-T-GATT-GGTAAGC-TACATGGAACAGGG-T-G--TGGG-ATTTTAT-CT...TA-AATA--A--GTA--A 4485
LST.CD.88.524                .............................................GAGCTT-GT-GA----TT-C---AT-TCA-AAGG-ATC--TCAGA-A-T---AA-GATT-GGTTAG--TACATGG-ACAGGA-T-G-ATGGG-GTTTTAT-CT...TACAATA--T---TA--T 4482
LST.KE.x.lho7                .............................................ATCCTT--T-GAA--GTT-----AC-T-GT-AGGT-AA-CACAGAG-CA----AA-ATT-GGTAGGC-TACATGGAACAGG--T-G-TTGGG-ATTCTAT-CC...TA-CATA--T--GTG--A 5569
SUN.GA.98.L14                ATAGAAGACAGTATAAGATAGTTAGAATAGTATGGCTTATAGATAGAATTGCTGT-GA----TT-----ATTT-AGAAGA---C--AGGGA-ACA----A-GATT-GGTA-C-ATGTATGGAACAGG-ACTG-ATGGG-GTGGTAC-C-...TACAATA--A-TATT--T 5619
MND-1.GA.x.MNDGB1            ...................TCAGAGAATAGAAAAGTGGCACTGGTTAGTAAG--G-CA--TGGCA-GG-CC-CTGCAA---ATG-G-AAGG-T-TTGGTGGCT-TATCCTCA--T--TGGC-T-TA-----T-GTA-AC-TGC--T-A-...GTAGTGATTA--ATA-AT 4922
OLC.CI.97.97CI12             ..................................GCTAAATAATGGGCAGGT-GA-CATGTG-AA--CA-GGCTA-GTTAGCC--CAAA-TTT--GG-TACCATT-GATA-GGC-A-ATTG-T--A-GT-CCCAAGGG-ATGGT-G-C-...CA-GG-A--A-TT-T--T 5070
WRC.CI.97.97CI14             ..........................................TAAAGGTAA-TTCA---G-T-GACCTATT-TATAGCA-ATCTAA--AA--CAGGGCAGGCAAAT-TTC-A-GGG--TC-TGGC---C-GCAG-AG-ATGCTTT-CT...--GCATAT----A-A--A 5505
WRC.CI.98.98CI04             ..........................................TAAAGGAAGGTTTAAA-G-T-GGC-TATTGGATAGC-CATCTAA-GCA-AAT---AAGGCAA-T-TT--A-GGG--AC-TGGC--ACTCAGG-AG-ATGCTTT-CT...--A-ACCT-T-GAGA--A 5520
WRC.GM.05.Pbt_05GM_X02       .....................................AATCTCAAAAGGAAG-TTTA-A-G-T-GAGCTATTGGATAGCAGA-CTAA-ACA--AA---CAAGCAA---TCC-G-T-G-CTCATGGC---T-CATG-AG-ATGTTTC-CT...--ACATAT----AGA--T 4437
MND-2.CM.98.CM16             ..................................ACATGCAGTAGAAAGATGG-AT-CACTGGTAAA-TACC-CA-AT---AAGGA-AAAA-CATTTG-AC-AAT-GGAT--AA-GGC----TTTC-ATGCTCA...GGATGGT-G-C-...CAC---C--AA-T-T--T 5215
MND-2.GA.x.M14               .............................................AAAATGG-AT-CATT-GTCAA-TACC-CA-AT---A-G-A-AAAA-CACTT--A--A-T-GGAAC-A-TCCAC--CTTTC-ATGCTCA...GGATGGT-G-CG...CAC--TC--AAGATT--- 5142
MND-2.x.x.5440               ............................ACAGGACGTCAAGTTGAGAGATTGG-AT-CA-T-GTAAA-TATC-TA-GT---AAGGA-AGAA-CATTT--AA-ATT-GCAA--A--CCCC---TTTC-ATGCTCA...GGATGGT-G-C-...CAC--TC--AAGA-G--- 4778
DRL.x.x.FAO                  ...............................GAGGTACATGGTAGAAAGGTGG-AC-CCCTGT-CAA-TATC-TA-G----A-GGA-AAAAGTATTT--AA--AT-GCAT-A-GCACCT--CTT-C-ATGCTCA...GG-TGGT-G-CT...CA-AGTC-GTGGAC---- 4781
MON.CM.99.L1                 ..................................................................-TGC-CTA-ATGGCACATAC-CA----A-GT---ACCCTTTGTC-A-GTACAT--CT--C-GCTGCA-AA---GAG-TTCTC-...CAGAACA--A-TA-AC-- 4920
MON.NG.x.NG1                 ..................................................................--C--CCA-AT---ACATATA-A---------TGCCATTTGT--A-GTACAT--CT-TC-GCTT-CT-AC--G-GCTTTTCG...CAA-GCA-GA-TA-GC-T 3513
GSN.CM.99.CN166              CTTAG.............................................................--CT-CTA-A--CCACATATGG-----C---A-ACCATTTGTA-A-GTCCAT--CT--C-ATTGCAA-AC--AAG-TTT-C-...CAAAACA-GA-T-GAT-A 4951
GSN.CM.99.CN71               ATTAG.............................................................--C--CTA-A---CATATATGG-----A---A--CCATTTGT--A-ATCCAC--CT-TC-ACT-CAA-A---AAGGTTT-C-...CAAAACA--A-T-GAT-A 4933
TAL.CM.00.266                AG...............................................................GT-TC-CATA----AAACCAAG...TACTT-CAA-A-TTGATT-GGGT-CAT--CTTTC-ATTG--A-ACGCCT-CTAT-C-...CAAG--A-GTGGATAC-- 5183
TAL.CM.01.8023               AG...............................................................GTATC-CATGT-C-AAACCAAG...TACTT-CAA-A-TT-ATA-GGGTCCAT---TT-C-ATT---AATGGCCT-TTTC-CG...CAGG-CA--TGGAT-C-- 4681
MUS-1.CM.01.1085             .................................................................G-C--C-A-AT-CTACATCC---------G---AGCCTTTACA-A-GTGCAT--CT-TC-ATTGTA--A---GAGGTTCTCC...CAGAGCA--A-CAGA--A 4907
MUS-1.CM.01.CM1239           .................................................................G-CT-C-A-A---CATATCT--------A---A-ACCATTTGT--A-GTACAC--CT--C-G-T-CAA-AC--AAG-TTTTCC...CAAAACA-GA-CA-GC-A 4897
MUS-2.CM.01.CM1246           ..................................................................-TGC-CTA-A---TATATCT-C------------CCATTTACA-A-GTCCAC--CT-TC-ATTGCAG-A---GAG-TTC-CC...CAGAACA--A--A-A--A 5005
MUS-2.CM.01.CM2500           ..................................................................-TGT-CCA-A--CCACATCT-A-----------ACCATTC-CA-ACATCCAT---T--C-A-TGCAG-A---GAG-TTC-CT...CAAAATC-GA-TA-G--A 4958
SYK.KE.x.KE51                ....................................................--TGCA-C-T-TA---AAGGT----A-GTATA-AAGCAA-CAATT--AG-AA-CCACA-A-A-ACAT---T-TC-GATT-CATGGG-GTGGT-G-CT...C-AG--C--TGGG-A--A 5063
SYK.KE.x.SYK173              .....................................................-TTGCA-C-T-TAA--AA--C---TA-GTATA-AAGCAA--AATTG-AG-AA-CCACA-A-A-ACAT---CT--C-AAT-GAATGG--ATGGTAC-CT...TA--GCC--TGGACA--A 5411
RCM.GA.x.GAB1                ..........................................ACA..ATGGCAC-CTTTGGTAAA-TATC-T--GT-C-AAGGAAAGAA--CAGC-AAAGA-T-GGA--A-GTGCCT--CTTTA-A-T-CCATGGGGATGGT-GTCG...CA---------T-----A 4702
RCM.NG.x.NG411               ATTCCTTAGTAAAGTATCATCAGTAT.........................................-A-GGAAAGAAG-AAGCTAAAGATT-GGA--A-GTACCC---TTTA-A-T-CCTTGGGGGTGGT-GTCT...C---------C------ 4714
GRV.ET.x.GRI_677             ATAGTGACTTACAAGATCAGGAATAA.................................................AC--TTGCCTT-GGAA-ACAGACAT---TGGC-G-TGCAATGG--GTTTT-G-CC...TACAGCC-GT-CATT--- 5387
SAB.SN.x.SABIC               ..............................................................CCTAG-GAA-T-C---ATGCATGTGTCCAA-CA-TGT-T-CATT-GC---A-ACCCCT---ACTA-AAT---ATGGA--TGGTACTCC...TA-CA----TGGGT---A 5612
TAN.UG.x.TAN1                .......................................TGGAGAGCATAT-GACAAATGG-AGGGCA--GTGA-AT-CCACATGAGGAACC--TT-CAGGATT-GACA-A-CT--TG--CT--C-ATGTG-ATGGGC-TGGTAC-C-...T--AGTAG-T-C-T---- 5413
VER.DE.x.AGM3                ..GATAGTGAGGTACTGGATGAGGAC.....................................................TA-A---TTAGACT-GAAA-A-CGA-TG--CT--C-AATT-CATGGGCATGGTAC-C-...-TGAGTAG-TATG-G--A 4936
VER.KE.x.9063                ..AATAGTCAAATATTGGATGACTAA......................................................-A----CCT--AAT-GCAA-ACATG-TG--CT--C-AATT-ATTGGGC-TGGTAT-CT...-TG-GCAG-TACAT---- 5441
VER.KE.x.AGM155              ..GATAGTTAGATATTGGATGAATAA......................................................-A----TCTG-AAT-GGAA-ACA-A-TG---T-TC-AATC-CTTGGGCATGGTAC-CT...-TGAGCAG-TATGTA--A 5433
VER.KE.x.TYO1_patent         ..AATAGTCAGATATTGGATGAGACA......................................................-A-A--TTTG-AAT-GAAC-A-----TG---T--C-AATT-CATGGGC-TGGTAC-CC...-TGAGTAG-TATGTA--A 5438
COL.CM.x.CGU1                ATTAA.............................................................................-A-GCAG-TTCCAG--TGCCAATAT-AT-A--TAG--A-AACAG-GTC-CC-GAATGGTGG-C-GA-...G--AA-AT--AGTTTTA- 4971
```

```
H1B.FR.83.HXB2               CCACTAGGG......GATGCTAGATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCAT........................ 5280
Vif                          _P__L__G__.__.__D__A__R__L__V__I__T__T__Y__W__G__L__H__T__G__E__R__D__W__H__.__.__.__.__.__.__.__
H1A1.UG.85.U455              ---------......--A--------A---G---G------------------------A---------C.............................. 4726
H1B.US.90.WEAU160            ---------......--A-G--A---A---------------------------------------------............................ 5279
H1C.ET.86.ETH2220            ---T-----......--G--------AA------A------------T----A------------T------............................ 4672
H1D.CD.84.84ZR085            ---T----A......---------C-----G---------------------------------A------............................. 4802
H1F1.BE.93.VI850             -------A-......--A-T--A---A------------------G-----CC-----------A------............................. 4620
H1G.SE.93.SE6165             --------A......-------C-C-A---G--------------G---------------A---------A............................ 4677
H1H.CF.90.056                ---T----A......--A-----G--A--C-----C-----C---------A---------------A------.......................... 4627
H1J.SE.93.SE7887             --------A......--G----T---A--------------------------G--------------T-----C......................... 4594
H1K.CM.96.MP535              ---------......-----AGAG--A---G--------------------T-----------G---------........................... 4476
H101_AE.TH.90.CM240          --------A......--G--------A-------G---------------------A---------AG-------C........................ 4854
H102_AG.NG.x.IBNG            ---------......---------G--A---G---G------------------------C---------C............................. 4805
H103_AB.RU.97.KAL153_2       ---T-----......--------A----------A---------------------------------------.......................... 4506
H104_cpx.CY.94.CY032         ---------......--G--------A---G---G-------------------GC----G---CA---------C........................ 4646
H1N.CM.95.YBF30              ---G----T......C-G--A-----A--G-C-GTC--T--------G--AACA---------CAGTCT------......................... 4874
H1O.BE.87.ANT70              ---G-----......-TG--CCATG-A---G----C-----------AT--ATGC----G--------TGAA---......................... 5335
H1O.CM.91.MVP5180            ---G---CA......--A---GATA-A--GG-C--C-----------AT-AATGC----G--------GGAA--C......................... 5310
H1P.FR.06.RBF168             ---G----A......-CA--A-TCA-T--TG-G--T--T--------G--TATGC-T--G----A---AGAA--A......................... 4741
CPZ.CM.01.SIVcpzCAM13        --TT-TAAA...GATA--T-A-A---AA----C-----C-------C-T-AACC-------------CG------......................... 4863
CPZ.CM.05.SIVcpzEK505        ---T----A......--A--A-A---A------GTT--C--------AT-AATGC----------GCCA------......................... 4852
CPZ.CM.05.SIVcpzLB7          ---------......--G--C-A---A------GT---T--------A--TACAC--------C-G--------C......................... 4835
CPZ.CM.05.SIVcpzMB66         -----G--A......--A----A---A--------------------A--CACA-------------C------.......................... 4826
CPZ.CM.05.SIVcpzMT145        ---A--AAT......C-G----AGC-AA-TG-C--------------A--AATGC-------------T------......................... 4822
CPZ.CM.-.SIVcpzMB897         ---G----A......------C-A---A------GT---C-------A--CATGC--------------------......................... 4846
CPZ.CM.-.SIVcpzDP943         --TT---AA......A-A--A-A---A--GG-C--C--C--------AT-AATGC-------GCAT--A-----C......................... 4866
CPZ.CM.98.CAM3               --TA---AA......A-G--A-A---A--GG-------C--------AT-AATGC-----C-GCAT-C-------......................... 4694
CPZ.CM.98.CAM5               --TT---AT......ACA--A-A---A--GG-C--------------A---ACA-----GC--CAT-CT-----C......................... 4982
CPZ.GA.88.GAB1               --TT-CA-A...GATT--T---A---AA--G-G-----C-------CCT-AAGCC-----------G-C------......................... 5344
CPZ.GA.88.GAB2               --CA--A-A..........-----A-A-A-A--GG-CTGT-----C---AA-T-AATGC--------CA-CCA------..................... 4673
CPZ.US.85.CPZUS              --TA---AT......A-A--A-A-A-A--G--T-T---G--------GT--A-A---------CA--C------C......................... 5340
CPZ.CD.90.ANT                ------CCAACATTG---AAA-A-------G-G---GTG-TC-----AT-A--ATGT------C-CCCA-----C......................... 4725
CPZ.TZ.00.TAN1               --T---ACA...GAAAGAT-A-A---A--GG-TTT-CAT-----------AGCCTGT---------CCA------CTAGGTCATGGC............ 4939
CPZ.TZ.2001.TAN2             --TA--ACA...GGAA-G--A-A-A-A---G-CTT-CAT-----------AGC-TGT---------CCA------......................... 4932
CPZ.TZ.2002.TAN3_1           --T---ACA...GAAA-AT-A-A---A---G-CCT-CAT-----------AGC-TGT---------CCA------......................... 4932
GOR.CM.2007.SIVgor2139_287   ---G----C......CCC---GA-A-TA--G----C-----------A--CATGC---------AG--AGAA--A......................... 4817
GOR.CM.2007.SIVgorCP2135con  ---G----C......CCC---GA-A-TA--G-G--C-----------A--CATGC---------AG--AGAA--A......................... 4816
GOR.CM.2004.SIVgorCP684con   ---G----C......CCC---GA-A-TA--G-G--C-----------A--CATGC---------AG--AGAA--A......................... 4736
H2A.DE.x.BEN                 ------CAA...........G--AAAAGTCATCT-GA-ATACA--CATATTGG-ACCT-AC-CC---AAAAGGATGGCTCTCCTCTCATGCA....... 5680
H2A.GW.x.ALI                 ---T--CAA...........G--AGAAGTCATCT-GAGATACA--CATATTGG-ACCT-AC-CC---AAAAGGATGGCTCTCCTCTTATGCA....... 5673
H2A.SN.x.ST                  ---T--AAA...........G--GAAAGTCATCTGGAGATACA--CATACTGG-ACCT-AC-CC---AAAAGGATGGCTCTCCTCCTATTCA....... 5125
H2B.CI.x.EHO                 --C--GAAA...........GA-GGA-C-CATCT-GA-GTCCAA--ATACTGG-ACCTGACCCC---AA--GGATTCTTGAGTTCCTATGCT....... 5651
H2B.GH.86.D205               --C---AAC............-A-GGA-C-TGGCT-GA-GTCCAA--ATATTGG-ACCT-ACCCC---AA--GGATTCTTGAGCTCCTATGCT...... 5653
H2G.CI.92.ABT96              ---T--ACA............-A-GAA-C-CATTTGGA-ATACAA--GTATTGG-ACTTGACCCCT--AAA-GGATGGTTAAGCCAGTATGCA...... 5007
H2U.FR.96.12034              --T---AAA............-A-GGATC-CATTTGGAGGTGCAA-CCTATTGG-ATTTGAC-CC---GAA-GGATGGCTATCCTCTTATGCA...... 5158
SMM.SL.92.SL92B              --CT-G-A-...........GAGGGAAGTCACTTGGAGGTACAA-TGTACTGG-ATCTGACCCC---AAAAGGATGGTTATCAAGCTATGCA....... 5031
SMM.US.x.H9                  --GT--A--...........GATGA--CTCATTTGGA-GTACAA--ATATTGG-ATTTGAC-CC---GRAAGGGTGGCTCAGTACTTATGCA....... 5070
SMM.US.x.PGM53               ------A-A...........GATGAATCTCATTTGGA-GTACAA--ATATTGG-ATTTG-C-CC---AAAAGGATGGCTCAGTACTTATGCA....... 5514
SMM.US.x.SIVsmH635F_L3       ---T--A--...........GATAA-ACTCATTTGGA-GTACAA--ATATTGG-ATTTGAC-CC---AAAAGGATGGCTCAGTACTCATGCA....... 5600

MAC.US.x.239                 ------CA-...........GA-GGAAGCCATTT-GA-GTACAA--GTATTGGCATTTGAC-CC---AAAAGGGTGGCTCAGTACTTATGCA....... 5597
Vif                          _P__L__Q__.__.__.__.__E__G__S__H__L__E__V__Q__G__Y__W__H__L__T__P__E__K__G__W__L__S__T__Y__A__.__.__
MAC.US.x.251_1A11            ------CA-...........GA-GGAAGCCATTT-GA-GTACAA--GTATTGGCATTTGAC-CC---AAAAGGGTGGCTCAGTACTTATGCA....... 5596
MAC.US.x.251_BK28            --C---CA-...........GA-GGAAGCCATTT-GA-GTACAA--GTATTGGCATTTGAC-CC---AA-AGGGTGGCTCAGTACTTATGCA....... 5573
MAC.US.x.EMBL_3              --C---CA-...........GA-GGAAGCCATTT-GA-GTACAA--GTATTGGCATTTGAC-CC---AA-AGGGTGGCCCAGTACTTATGCA....... 5065
STM.US.x.STM                 --CT-GCAA...........G--GAA-C-CACCT-GA-GT-CAA--GTATTGG-ATTT-AC-CC---AAAAGGATGGCTCAGTGAATATGCT....... 5244
MNE.US.x.MNE027              ------CAA...........GA-GAAAGCCAGTT-GA-GTACAA--GTATTGG-ATTT-AC-CC---AA-AGGGTGGCTCAGTACTTATGCA....... 5065
DEB.CM.99.CM40               --CA-CACA............-A-GAA-GGGG-GT-GT-GTAGAATTA--ATGGTGCTTGACCCC---GAAAGGATGGCTACCCACTATGGCA...... 5104
DEB.CM.99.CM5                --CG-CACA...........GA-GAA-G--A-GT-GT-GTGGAAATAT-CTGG-GCCTG-CTCC---AAAAGGGTGGATGCCCTCTATGGCG....... 5098
DEN.CD.x.CD1                 -----GA-A...............GC--CTTAC-TGGA-GTAAA-CTCTATTGGCATTTGACCCC---AAAAGGATATCTGTCACAAGTAGCA...... 5206
LST.CD.88.447                --TT---A-...............ACA-G--CCTT-GT-GTGA-AT-CTATGCACACTTG-C-GC--GAA-AGGATGGATCTCTCAGTGGGCA...... 4563
LST.CD.88.485                --TT---A-...............A-A-G--CCTT-GT-GTAA-AT-CTATGCACACTTG-C-GC--GAA-AGGATGGATCTCTCAGTGGGCA...... 4563
LST.CD.88.524                ---T---AT...............ACA-G--CTTT-GT-GTAA-AAT-TATGC-CATCTG-C-GC--GGA-AGGATGGATTTCACAGTGGGCA...... 4560
LST.KE.x.lho7                --T----A-................-CA-G--CTTT-GT-GTCA-AATATATGCACATTTG-C-GC--GAA--GGATGGATCTCACAATGGGCA..... 5647
SUN.GA.98.L14                ---G--ACT................-AT-GG-C-GT-GT-GTAA--ATCTATGGACACCTCAC-CC--CAAAAGGGTTGATTAACCAGTGGGGA..... 5697
MND-1.GA.x.MNDGB1            AGGGAC....................A---G-TT--T-GT-A-AA-CTATTGGCATTTGC---T---GGTAGGATGCTTAAGTACTTATGCAGTA... 4997
OLC.CI.97.97CI12             ---G--CCA..............GG--GTGAG-T-GA-AT-ACACTCTACTGGCATATTAG-CAT-GGAAAGGATTACAAAATTCAGGTGGGGTG... 5151
WRC.CI.97.97CI14             ---A--CA-...........G--CACTGGGAG-TGCA-AT-A-AT-CTATTGG-ATTTGTC-CC---AAA-GGATGGATGGACACCATGGGATTT... 5589
WRC.CI.98.98CI04             --CT--CAA...........G-GCATTGGGA-TT-CA-ATCA--T-CTATTGG-ATCT--C-CCT--AAA-GGATGGTTAGACACCATGGGGTTT... 5604
WRC.GM.05.Pbt_05GM_X02       --T--TCAA...........G-GCATTGGGA--TGCAGATAA--T-CTATTGG-ATTT--CTCC---AATAGGATGGATAGATACTTTGGGATTC... 4521
MND-2.CM.98.CM16             --C----AA...........GA-GAT-A-CAG-TC-T-ATCACCATC---TGG-ACTT-ACCCC---AAAAGGATGGTTATCAACTTATGCC....... 5296
MND-2.GA.x.M14               --CT-CAAA...........GATGGCTC--AG-TT-TTAT-ACT-CC--ATGG-ACTTGACCCC---AAAAGGATGGTTATCTCAATATGCA....... 5223
MND-2.x.x.5440               --CT-TAA-...........GATGGTTC--AG-TT-TCATCACC-TCT--TGG-ACTTGACCCC---AAAAGGGTGGCTGTCTCAATACGCC....... 4859
DRL.x.x.FAO                  --CT-CAAA...........GACAAAAGT-A--T-GT-GTAACA-TC---TGG-ACCTCACCCC----AAAGGGTGGTTAAGTAAATATGCCATA.... 4865
MON.CM.99.L1                 --CT---ACCTAGGAATCA--CA-GAA-G-GAGT--TGGGCAACATA-T--G-G...AT-AC--TTT-T---G-CTTAACAAACGTGGGACCTAGTGCACTCTCACCAACAGTGTACAACAAACAGGCCTACACAATAGAATGGGTTTATTGGATGCAAAAGGTGGAACG 5087
MON.NG.x.NG1                 --TT---A-...ATG-G-ATC-C-CAARAGGRGGA-CA--GGGCAACATACATRGAGATCACG-TCCTT---G-CAYTACCAATGTAGGGCCCGCGTCACTATCTCCATCAGTCAAATGGAAGCAAGCATACATCATAGAATGGGTATATTGGAGGCAGCCCAGGGTGCC 3680
GSN.CM.99.CN166              G--A-G-ATACTAGGA-G......G-A-GGGA-GA-GT-G-AGCAACATATATAGAGAT-ACT-TCCTA---G-CACAACCAGTCATGGACCAGCTAGCTTATCCCGGTCCACCTATTACCAACAAGCAGTCATAATAGAATGGGTATACAAC............. 5098
GSN.CM.99.CN71               G--T---ATACTAGGC-G......G-A-G-GAGGA-GT-G-AGCCACATATAT-GAGAT-ACC-TCCTA---G-CACAACTAGTCATGGACCAGCCAGTCTATCCCGATCCACCTACTACCAACAGGCAGTCACAATAGAGTGGGTTTACAAC............. 5080
TAL.CM.00.266                --GT---AAGAAAAG--CA-C---G-A-A-TATTT---CATAACACAGTATTGGTGTTT--C-CC----CATGGATATACGTCCGATTGGGAC....... 5276
TAL.CM.01.8023               --C----AAGAAAAA--GAA----G-A-AGTAC-TT--CATCACTCAGTATTGGTGCTT--C-CC----AA-GGATACACATCAGACTGGGAT....... 4774
MUS-1.CM.01.1085             ---T-GACACAA...A--A-CC-GGAATC-GGTG--GT-G-AA-TAC-TACATAGA-CTGAC-GTCCTG---G--GTGACCAATGTGGGACCAGCAAGTTGCTCTAAAAGCACATGGTGGAAGCAATCCTTTATCCTAGAGTGGGTATATCTA............. 5057
MUS-1.CM.01.CM1239            G--T-GACACAA...A-CA-CTT-GAATC-GG-GAG---G-AGTCACCT-CATAGAGAT-AC-GTCCTA---G--GTGACTAATGTAGGACCAGCAAGCCTATCCAAAAGTACCTATTGGAAGCAGTCATACATCCTGGAATGGGTATACATG............. 5047
MUS-2.CM.01.CM1246            --TT-GA-CATA...A-CAGA-T-GAA-G-GGCC-----G-GACTACATACATAGAGAT-ACC-T-TTG-TAG--GTCACAAACGTAGGGCCAGCCAGTTTATCAAGGTCCACCTATTGGCAGCAATCCTATATCCTGAAGTGGAGATACATC............. 5155
MUS-2.CM.01.CM2500            --CT-GACAATA...A-GAAG---GA--A-GAC-A-GT-G-A-CTACATACATACA-AT-AC--TTCTA---G--GTTACCAATGTAGGGCCTGCCAGCCTCTCAAAGTCTACTTACTGGCAGCAGTCATATATATTGGAGTGGGTTTATGTT............. 5108
SYK.KE.x.KE51                ---G----A...............CAA-GTG--TT--TTATAAAATTCTAT----ATTTGAC-CC---GAAAGGCCTGTTAAGAACAGAGGGA....... 5141
SYK.KE.x.SYK173              ---G----A...............GAT-G--CT-T-TGGATCAC-TTCTAT--C-ATTTG-CCCC---AA-AGGATGGTTACATATGCAGGGA....... 5489
RCM.GA.x.GAB1                --TT---A-...........GA-GGATC--AGTT--A-ATAACCACCTATTGG-ATTTGAC-GT---AAA-GGATGGCTAGGGACATATGGG....... 4783
RCM.NG.x.NG411               --CT-G--A............-ATAACACT-AG-T--AGGT-ACAACATATTGG-GTCTCAC--T---AAA-GGATGGTTAGGAACCTATGGA...... 4795
GRV.ET.x.GRI_677             --CT--TCA............-A-GAT-ATTAC-T-GA-GTGAATAT-TAT--C-ACCTCACCCC---AA-AGGATGGCTCTCAAGTCATGGA...... 5468
SAB.SN.x.SAB1C               --C--GAA-...........GATGGA-CTC---T--A-GT-ACTAA-TATTGGCATTT-ACCCC---GAAAGGATGGTTAGAAACCTATGCA....... 5693
TAN.UG.x.TAN1                --CT----A...........G--GAA-G--A--T-GT-GTAGATT--TACTGGCACCT-ACGCC---ACA-GGATGGCTCTCTACCTATGCA....... 5494
VER.DE.x.AGM3                --C------...........CA-CAT-G--GT-T-CATGTAGATCTATATTGGCATCTGAC-CC---AAA-GGATGGCTATCAACATATGCT....... 5017
VER.KE.x.9063                ---A----A...........TT-GAA-G-GA--T-CAGATAGACT-CTATTGGCATTT-AC-CC---AA-AGGCTGGCTCTCTACTTACGCC....... 5522
VER.KE.x.AGM155              --C--CCCA...........G--AGT-G-GA--TCCATGTGGATATCTATTGGCATTT--CTCC-A-ACAAGGATGGCTCTCAACTTATGCA....... 5514
VER.KE.x.TYO1_patent         ---A----A............-A-CAT-GGGA--T-TGTGTAGACCTATATTGGCATTT-AC-CC---GCAAGGATGGCTATCCACATATGCA...... 5519
COL.CM.x.CGU1                TT-GGGTCA...............GCTA-T-AGTT--GCATAATATACT--GGC--TTT-ACCCC---TA-AT--CAT..................... 5034
```

```
H1B.FR.83.HXB2        ....................................TTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAAAGAGATATAGCACACAAGTAGACCCTGAACTAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCTTAT 5413
Vif                   __.__.__.__.__.__.__.__.__.__.__.__._L__G__Q__G__V__S__I__E__W__R__K__K__R__Y__S__T__Q__V__D__P__E__L__A__D__Q__L__I__H__L__Y__Y__F__D__C__F__S__D__S__A__I__R__K__A__L__
H1A1.UG.85.U455       ....................................-------T--G------------------CTG--A------------------T-----C----------T--------C---C---T------------------A-----C-----G-GA---A--- 4859
H1B.US.90.WEAU160     ....................................-------A-----------------------C-------------------------C---------------------C-T-----T------------G--C------------T---A--- 5412
H1C.ET.86.ETH2220     ....................................-------T------------------ATTG-GA--C----A-----------------GC--G--------C---------A--C----T-----T------G-----A-----C-----G--A---A--- 4805
H1D.CD.84.84ZR085     ....................................C------------------------------G-------------------------GC--G--------------------A-----------T------G----------------C-----A---A--- 4935
H1F1.BE.93.VI850      ....................................C-------------------------C-GGG--AG-----G------A----------GC--G--G--------G--C---A-A-----T-----T-------------A-----C-----G--A---A--- 4753
H1G.SE.93.SE6165      ....................................-------T--G-----------G------C-G-GA--------A-----------G-----C----------T--G-----------C----T-----T--------------------C-----G--A---A--- 4810
H1H.CF.90.056         ....................................--A--C---------------------ACTG--A--G-----------------G----GC--G--------------------A--C----T-----T-------------A-----C-----G--A---A--- 4760
H1J.SE.93.SE7887      ....................................--------------------------C-G-GA--G--C--A--------------GC--G--------------------A---G---------T--------------------C-----G------A--- 4727
H1K.CM.96.MP535       ....................................--------------------------CTG--A--G--C--A------------A-----C---------------------A-A-----T-----T-------------A-----AG----G------A--- 4609
H101_AE.TH.90.CM240   ....................................-------T------------------C-G-GA-A-------------A----T-----C----------A----G--------AC-A--T----G-------------------------C-----G--A---A--- 4987
H102_AG.NG.x.IBNG     ....................................-------T--G---------------A-C-G--A--------------A----------CT-----------------G--------C--------T---A-T----------------------C-----------A---A--- 4938
H103_AB.RU.97.KAL153_2 ...................................-----C-------C--------------A---G-----------------------A-C----------------------------------T-----T-------------A------------------T---A--- 4639
H104_cpx.CY.94.CY032  ....................................-------T-------------------CTC-GA----------------G--T-----C--G---------------------A--C---------T-----T-------------A-----C-----G--A---A--- 4779
H1N.CM.95.YBF30       ....................................C-A--A--T-----A------------ACT--GA-A---C-AG---------T--T--------A-G------A-G------A------TC----T-----T------A----C--------C-----GC-A--GG-C- 5007
H1O.BE.87.ANT70       ....................................-----A--T--G--TAGT---------C-TT-T----AG----AA-----GA-T-----------AC--------AGGA-G--A------C----T---AC--------A---CA--A--AG-C--G--A---A-CC 5468
H1O.CM.91.MVP5180     ....................................-----A--T--G--TAGT---------CAAT-C---GAG----AA-----GA-T--T--------AC--------AGGA-G--A------C----T--CACA--------A---A--A--A-C--G------A-TC 5443
H1P.FR.06.RBF168      ....................................--A--A--T----YAAGTG-G--G------AC--GGT-A---C-C---C-GA-----T--A---AC--------TAGG---------------CC-C-------C-A-----------------T--G---G-G--G-G---AA--C 4874
CPZ.CM.01.SIVcpzCAM13 ....................................--A-----T-------------C-G-------TGGGATC----GTT------G------T---T-TG--------TA--T-G--C--CTCTC-A--T------T------G-----------C--C----GA--TA--- 4996
CPZ.CM.05.SIVcpzEK505 ....................................--A--A--T-----A--------------AC--GGA-T---C--G-------A-----T--------T-G------A-G-----A--C--T-----T------T------A---C---------C--C--GC-A--GG-C- 4985
CPZ.CM.05.SIVcpzLB7   ....................................--A------T--G--A--T--C--G----AAC-G-----A------T-------A-----------GT-G------T--TA-G--------T-A-----T---A-T-------G-----A---------C--------G------A--- 4968
CPZ.CM.05.SIVcpzMB66  ....................................C-A--C--T--G--AAGT-----------AC-G-GA--T-----G------A------T---------T--G---------T-----C----T-----C--T-----T-------G-----A----------C--------G------A--- 4959
CPZ.CM.05.SIVcpzMT145 ....................................C----A--T-----AG-A-----------AG---GAG-T--C--T--C--G--------------AC--------TAGG-----A--C--TC-----T-----T------A----CAA-------C--C--G--C--TA-C- 4955
CPZ.CM.-.SIVcpzMB897  ....................................--A--A--T--G-----T------------AC-G--A-----------------A-----T------T-G-----------C------------T--C-----T-----T------G-----G---------C----------G--A---A--C 4979
CPZ.CM.-.SIVcpzDP943  ....................................C-A--A--T-----G---------------AT-TGG-TC---C--A--------A-----T------A-G------TAG-----------C--TC-----T-----T------G----CT--------C--C-------A--TG-T- 4999
CPZ.CM.98.CAM3        ....................................C-----G--T-----G---------------ATGTGG-TC------T--------------------GA-G-----TA-----------C--TC-A--T-----T------G----CT--------C--C--G--A--TG-T- 4827
CPZ.CM.98.CAM5        ....................................C-----A--T-----G---------------ATGTGG-TC---C-AA-------------T-------ACG------TAG-----------C--TC-----T-----T------G----CT--------C--C--G--A--TG-T- 5115
CPZ.GA.88.GAB1        ....................................--A--C--T-----------TC-----------CT-GGATCT---GT-------G--G--T--CTTCAC------TAG-T-G--C--CTCTC-A--T------T--C--G------AA-----C--C-----GA--TA--- 5477
CPZ.GA.88.GAB2        ....................................C-A--C--T-----G---------G------C-G-GA-C-------A--T--GA----T----C--AC-----T-GG--G--A--CTGT------T------T-------G----C-----------G-C-----GA---A-CC 4806
CPZ.US.85.CPZUS       ....................................--A--G--T-----A--A------------AC--GGA-CT--C-AG------A-----T-------GA-G------A-G------------------T------T------A----C------------C--C-------A--TG-T- 5473
CPZ.CD.90.ANT         ....................................--A--A--T--T--GAG--------------ATGTGGA-AG----T------------------A----AC---T--T----A-G--A--C-AA-----T--CCC--------------TCAA--AG-C----C-A--AA-CC 4858
CPZ.TZ.00.TAN1        ....................................A-A--ATTA.............--------AC--GGA-A---C--T------A-------------AC-----------T-G------CACTAGG--T---AC--------G-T-CAGGA--AG-TC-GC-A--AA--- 5060
CPZ.TZ.2001.TAN2      ....................................C-A-----T-----AGGGC----------AC--GGA-A---CT---------T--------------AC-------T--GT----------ACTAGG--T---ACA--------A-T-CAGGA--AG-TC-GC-A--AA--- 5065
CPZ.TZ.2002.TAN3_1    ....................................C-A--A--T-----AGGAT-G---------C--GGA-A---CT-----------------------AC-------T--GT-G--------ACTAGG--T---ACA--------G-T-CAGGA--AG-TC--C-A--AA--- 5065
GOR.CM.2007.SIVgor2139_287 ..............................--A--C--T--G--AGG-----------AC-GGGA-A---C-TA-----T-----T--A---AC------TAGGA----C------T---------C-A------A-----A------AG-G--G--A--AA-TG 4950
GOR.CM.2007.SIVgorCP2135con ..............................--A--C--T--G--AGG-----------AC-GGGA-A---C-TA-----T-----T--A---AC------TAGGA----C------C---------C-A------A-----A------AG-G--G--A--AA-TG 4949
GOR.CM.2004.SIVgorCP684con ...............................--A--CA-T--G--AGG-----------AC--GGA-A---C-TA-----T-----T--A---AC------TAGGA----C------C---------C-A------A-----A------AG-G--G--A--AA-TG 4869
H2A.DE.x.BEN          ....................................G-AA-ATTAACCTGG-AT............-C-G-A-AG-TCT-G---G-T--TAC---A--CTGT-------ATC------A---AGCAC---T--CTCT--C---A-G-CAGG--AAG-------GA---A-CA 5801
H2A.GW.x.ALI          ....................................G-AA-AATAACTTGG-AT............-C-G-A-AG-TCT-G---G-T--TAC---A--CTGT-------TCC------A---GGCAC---T--CTCT--C---A-G-CAGG--AAG-------GA---A-CA 5794
H2A.SN.x.ST           ....................................G-AA-A-TAACTTGG-AT............-C-G-A-A--TCT-G---G-T--TAC---A--CTGT--G---TCC------A---AGCAC---T--CTCT--C---A-G-CAGGC-AAG-------GA---A-CA 5246
H2B.CI.x.EHO          ....................................G-AA-A-TAAC-TGG-AT............G-G-G---C-T-TAT---G-T---ACT-----TG----------G-T--C-G---GG--C---T--CTCT-CG---A---CTAA--AAG----G-GA---A-CA 5772
H2B.GH.86.D205        ....................................G-AA-A-TAAC-TGG-AT............G-G-G--AC-T-TAT---G-T---ACA------TG-G-----------G---C-G---GG--C---T--CTCT--C---------C-AA--AAG----G-GA---A-CA 5774
H2G.CI.92.ABT96       ....................................G-AA-ATTAACCTGG-A-............-C--GA-A--T-TAT---G-T--GACA--A--GAC-------T---T--T-G---GGA-C---T-----T--C-----A---CTGG--AAG-------GA---A-CA 5128
H2U.FR.96.12034       ....................................G-AA-GTTAACTTGG-A-............-C--GAG-T-T-T-G---G-T---ACA------TTGC----------T--C-G---GG-AG---T--CTCT--C-----A---CGGG--AAG-G--G-GA---A-CA 5279
SMM.SL.92.SL92B       ....................................G-TA-GATTACTTGG-A-............TC-G-A-A--T-T-G---G-T---AC---A--TG--------T--GT--GCA---AGTAC---T--CCCT--C---G---CGCAC--GG-G---C-A---A-CA 5152
SMM.US.x.H9           ....................................G--A-AATAACCTGG-A-............TCC-GA-AT-TCT-G---G-T-C-ACA--A--TTAT-------ACTT--C-G---GGCAC---T--CCCT--C---------GGGA-AAG--C----G----A-CA 5191
SMM.US.x.PGM53        ....................................G-AA-AATAACCTGG-A-............TCC-GA-A--TCT-G---G-T---ACA--A--TTAT-------ACTT--C-G---GGCAC---T--CCCT--C---------GGGA-AAG--C----G----A-CA 5635
SMM.US.x.SIVsmH635F_L3 ...................................G-AA-AATAACCTGG-A-............TCC-GA-AT-TCT-G---G-T---ACA--A--TTGT-------ACTT--C-G---AGCAC---T--CCCT--C---------GGGA-AAG--CA---GA---A-CA 5721

MAC.US.x.239          ....................................G--A-GATAACCTGG-A-............TC-----AC-T-T-G---G-T---ACA--AA-CTAT-------ATTT--C-G---AGCAC---T--CCCT--C---A---CGGGA-AAG-G----G----A-CA 5718
Vif                   __.__.__.__.__.__.__.__.__.__.__.__._V__R__I__T__W__Y_.__.__.__.__._S__K__N__F__W__T__D__V__T__P__N__Y__A__D__I__L__L__H__S__T__Y__F__P__C__F__T__A__G__E__V__R__R__A__I_
MAC.US.x.251_1A11     ....................................G--A-GGTAACCTGG-A-............TC-----AC-T-T-G---G-T---ACA--AA-CTAT-------ATTT--C-G---AGCAC---T--CCCT--C---A---CGGGA-AAG-G----G----A-CA 5717
MAC.US.x.251_BK28     ....................................G--A-GATAACCTGG-A-............TC--G--AC-T-T-G---G-T---ACA--A--CTAT-------ATTT--C-G---AGCAC---T--CCCT--C---A---CGGGA-AAG-G----G----A-CA 5694
MAC.US.x.EMBL_3       ....................................G--A-GATAACCTGG-A-............TC--G-GACCT-TTGGACAG-TGTA--A-CAG----TGCAGG--TT-TTCCTGCA-AGCACTTA--TC-CTC---A---CGGGA-AAG-G----G----A-CA 5186
STM.US.x.STM          ....................................G--A-GATAACTTGG-AT............-CC-GA-AT-TCT-GT--G-T---ACA--A--CTGT-----T--G---C-G---GGCAC---T--CCCT--C---A---CGGGA-AAG-G----GA---A-CA 5365
MNE.US.x.MNE027       ....................................G--A-AATAACCTGG-A-............TC--G--AC-T-T-G---G-T---ACA--A--CTAT-------ATTT--C-G---AGCAC---T--CCCT--C---A---CGGGA-AAG-G----G----A-CA 5186
DEB.CM.99.CM40        ....................................G-A-CAATA-C-TGG...........GAA-C---A-C--GG-AG---G--C--ACA--A--TT-------T--C--T--A--C--AAGA-------CCA------C--CC-ACAG--AG-G---C-A---A--A 5225
DEB.CM.99.CM5         ....................................G-A--AATCTCTTGG...........GA-C-G-CAG-G-GGCAG---G--T--ACA-----C------T--T--TT-G--A--CT-AAGA-----CCCT------C--CC-ACAG--AG-C-C-C-A---C--A 5219
DEN.CD.x.CD1          ....................................A-CA-CATA-CTTAT...G--C-CCTC--C-C-GG-T-G...GCT---G-------------T-AC------TG-GA----A---AAAG-CC-T---CTCT--C---A-CAGACAA--AG-C----C-A------A 5333
LST.CD.88.447         ....................................G-ATCAATA-A-TGG............-TTT-TGG--AC---CAA---G--A----T--CATT-------A-T-----A-G--A--CTGTA-A--T---A----C------TACAAGG-AC--T--G-----T---- 4684
LST.CD.88.485         ....................................G-ATCAATA-A-TGG............-TTT-TGG--AC---CAA---G--A----T--CATT-----------T---A-G--A--CTGTA-A--T---A----C------TACAAGG-A---T--G-----T---- 4684
LST.CD.88.524         ....................................G-ATCAATA-A-TGG............-TCT-TGGA-AC---CAG---G-----G--T--CATT------T--T-----A-GC-----G--A-G--T-----------GGGTGC-TAGG-AC---GTT-G----A--A 4681
LST.KE.x.lho7         ....................................G-TTCCATA-A-TGG...........GTCT-T-GAGAC---CAA---G--A-------CAT----------------A-G--C--CTATA-A--T---A----C--CAT-C-AAGA-AC--T-----G---AC-TC 5768
SUN.GA.98.L14         ....................................-GTA--AT--A-TGG............-TTT-T--C--T---CAA---G--A-----T---CTTG-G------T--GA-G--C--CT-TA-G--T-----T----GGA-CTCTAGGTG---C-----GA--TA-G- 5818
MND-1.GA.x.MNDGB1     ............................AGCATAGAA-CAGTA-TTAGA...CCGCC-CCCTTTG-G--AGAG-GGT-T---G-GA--ACT--A--GG-------T--T------A---T-AC---TT-A------C--CATG---AG---AG-T-TG--A---A-CA 5133
OLC.CI.97.97CI12      ....................................AATTT-A-ATATA-AGTA............GGG-TTTC--GG-TG...T-TT-GACT--A--GTAT-T---TTTCA-GG-A--C--AGTG--------C------GGA-TC-A-----------G--A---TA-T- 5269
WRC.CI.97.97CI14      ....................................AACAT-AGAATG--G-AT............G----T----GGGAA---TTTA--ATG--ATC-GAG---A-TA--A-C--A--CAGTC-AC-T------A-----GGA--C-GAA--TGG-G------TA-AA-TA 5710
WRC.CI.98.98CI04      ....................................AACATAAGAATG--G-AT............G----T----GGGAA---TTTA-TATG--CTCTGA-------TGT-A-G-----CAGTC-AC-TG---------GGA--C-AAA--TAG-T----G-TA--A-TA 5725
WRC.GM.05.Pbt_05GM_X02 ...................................AACATAAGAATG--A-GG............G----T-A--GGGAA---TTTTA--ACTGTAA--T-G------TGTCT-G--A---TCTC-AC--------G----GGA-CC-AAG---AG---A---CA-AA--A 4642
MND-2.CM.98.CM16      ....................................A--ACAATA-A-TAT............-A-TTC-CA-AC--CTA---T--CA-------CC--AC-------AGGA-G--A---TG-A-A--TC-CCC-------A----ACAA--A-------C-A---C-GC 5417
MND-2.GA.x.M14        ....................................A-CACAGCA-A-TA-............-AC---GG-GAC---TAT---T--CA--------A-T-AC-------AG-A-G-----TG-G-A--T--CCCA-------A---CTG-C---G-G--------TGC-G- 5344
MND-2.x.x.5440        ....................................--TACAGTA-A-TAT............CAA--GG-A-AT---TT---TT-CA--------A-TGACG-----TAG-A----C--CGG-G-A--T--CCCA-------A----TCAA--A----------A-----G- 4980
DRL.x.x.FAO           ....................................ACAATAG-ACATA-T.................C-GG---AG-T-T-T--CTT-A--------AAC-AC--------AG-A----C--CT--G-A--T--CCCT-------A----TCAA--A------GGT-T----C 4983
MON.CM.99.L1          G.........GTCCCTTTAGATCCCGTGTGG......AATCCAA-CCC-AGAGAT-G---GGTC...-G--CAT-G---GCT--C--TC-GACG--G--C--G--GA-T--GA--G-A---ACAC----T----CA--C--CCA-C-G-TG-ACG-GC-C-GA--AA-CA 5239
MON.NG.x.NG1          ACAGCAAGAGGGAGTAGTAGATCGAGTGTGGTTTCCTAGTCCCAGAA--CCGGA-CGG---ATTGAT-G-GTCT-G---GCT--CT-CC--ACA--A--TG----T-------A----C--CACCCGC-------TCA------CCAG-C------ACG--C-T-GA--AA-TA 3850
GSN.CM.99.CN166       ....................................AGAA-CA-ACT--GAG-A-G---GG-AGAC-T--TTT-G--C------TA-TC-GAC---A-G-G-----ATG---A-T--A---GGAA-A-------CCA--C----CAGA--GAA-AC---C-G-G---AA--A 5231
GSN.CM.99.CN71        ....................................AGAA--A--CT---GAG-A-GG---ACAGAC-T--CAT-G--C--T--TA-TT-GAC---A--CG-----ATG---A-C--A--CACAAGG-------CCA-------CAGA--GAA-ACG--C-G-G---AA--A 5213
TAL.CM.00.266         ........................ACAATATCCATTGA-T-GTCA-TCTCGGAT----TCC-AGACG-----GTG--CGC---TG-----AC---AA-TG-G-----T--CT-G------ACCA-A--T--CCCA-------A---CG--A--AG-C----C-A---A-CA 5421
TAL.CM.01.8023        ........................ACAATATCCATTGA-T-GTCA-TCTCGGATG-GACC-T-GATGC----GTC---TC----G-GT--ACA--A--GG----------C--G------CACCA-G--T--CCCA-------A----C---A--AG-C----C-A---A-CA 4919
MUS-1.CM.01.1085      ....................................AG-A-A--TCC-CAGGA--GG---CA-GATT-C-TCT-G---TCT---T-CT-GAG---CAC---G--CATGAG-A-C--A---ACCC-G--T---TCT--C--C-ATAG-CAG-AC--CCC--G--T-A-CA 5190
MUS-1.CM.01.CM1239    ....................................AG-AAAGGTCC-AGAGAT-G----C---GACT-TGTTT-G---CA----T-CC-CACA-----GA----CATGA-GA-C--A--CACCC-CC-----CTCT------C-ATAG-CAG-A----CCCG-G--T-A-CA 5180
MUS-2.CM.01.CM1246    ....................................AG--CAGCCCC-CAAAAT-G---GATAGAC-TGGT-CAT--CGAG--CTTCC--------A----G-GA-CATG--GA-T--A--CACAC-C------CTCT------CCA-C-GAGA-AC--CCA--G---AA-CA 5288
MUS-2.CM.01.CM2500    ....................................AG--C-GCCCC-CAAGAT-GG---GGT-GAC-TGGTAT-G--C--T--CT-CC-T-----A----T-G--CATG---A-T------ACT----C-----CTC--------CAG--ACAG-AC--C-T--GA---C-CA 5241
SYK.KE.x.KE51         ....................................G-A--ATTAATTTGGG--...C-TCC-TCTGG-TG-.......--T---G-----ACA--AC--AC-------TGC-T--G-A---AGTC-A-----CCCT------------TAGG--AG-CCA-C-A--AC-TA 5265
SYK.KE.x.SYK173       ....................................A-TA-GATTCA-TA-CAA...TGGAACCA-TGG--C.........--TG-CT--ACA--A-C-G----------GC--T--A---AAT-T---T--CCCT--C---A----CAAGA--GG-C-ACC----AG-CA 5610
RCM.GA.x.GAB1         ....................................G-A--AATCTT-...................----TA---GGAGAT---GTA---G-T---TTT---TGGACT------TAGTT-----A----AAAAT---T----CCA--------A----TAGA--A--C----C-AAGTC--C 4904
RCM.NG.x.NG411        ....................................G-A--AATA-C-TA-ATAGATA-CA-AT-TG-TCCCT-T--CTT-----G-TA-T-----AAC-G-------TA--T--G-A---AATC-G--T---CCT--------A----TAAA--A--T----C-A---A-C- 4928
GRV.ET.x.GRI_677      ....................................G-A--GTTATCCTAT-A-......CATCAA--GGGA.......AG---G--------T--A-G-AC---------AG-A-G--A--C--A-----T---A-------------A----TAGA--C--CCA-C-----TA-CA 5592
SAB.SN.x.SAB1C        ....................................ACA--AATA---TA-............--T---GG-GA--GGTTT---G-GC-T--T---TGGAC---G------TA-T-------TG---CC--T--CCC--------A----TAGA---G-CCA-C-A---A-CA 5814
TAN.UG.x.TAN1         ....................................G-A-CAATAA-TT-T........--GAATT--C----C-CG--C-AA---G------ACT------TG-G--G------TA-G------TGCC-C------CCCT--C---A----TAGA--C--CCA-C-------A-CA 5621
VER.DE.x.AGM3         ....................................GA---GATACAGTATCTA...AGCAAT---G-TCCTT-G------G---G--T-G--T----C-AC-------TAGC--G--A---ACCC----T---ACT--------A----AAGG--C--C--G--A---C--- 5147
VER.KE.x.9063         ....................................A-A--AATACA-TA--TG...-G-GCAGTAGG-...-AT-TC--A---G--T--------A-C-AC-------TGG-A-G--A--C--AC-C--T---TCA--------A----AAGG--C--TCA---A--TA--A 5649
VER.KE.x.AGM155       ....................................G-A--AATACA-TATGTT...AGCCTAGTA--TG-T-A-------A---G--T-----T--CA-TAC--------TCCA-G--A---TGTC----T---AC---------A----TAGA--C--CCA-C-----AC--A 5644
VER.KE.x.TYO1_patent  ....................................G-A---ATACA-TATGTA...AGCAATTTAG--TCT-A-------A---G--T-----------CTAC--------TAGTA----A---GGTC-C--T---A-T--------AA---AAGA--C--CCA-C-A--TC-GA 5649
COL.CM.x.CGU1         ....................................AGA--A-CACC------TTAT----TACA--T-GGT-AG-GGGAA---G-TT--AG-TGGC---A------TAGGGTAC--G--CA-CATAAGGGGGA-G---G-ACGCATAGAG-AAGA-GC--GAAAAGCTA 5167
```

```
H1B.FR.83.HXB2          TAGGACACATA.....GTTAGCCCTAGGTGTGAATATCAAGCAGGACATAACAAG................GTAGGATCTCTACAATACTTGGCACTAGCAGCATTAATA..............ACACCAAAAAAGATA...AAGCCACCTTTGCCTAGTGTTACG 5541
Vif                     L  G  H  I  .  .  V  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  P  P  L  P  S  V  T
H1A1.UG.85.U455         -------A--.....-----A-------------------------------------................------------------T-----T--GAA------G--..............--C---C--G-GC-..----------------------A- 4987
H1B.US.90.WEAU160       -------TT--.....----TT--------------------G---------------................-------------------------A------------..............---------------..----------------------A- 5540
H1C.ET.86.ETH2220       -----T-T-G-.....--------------------C---------------T---................-------------------------GA------G---..............-AG---------GC-..---------C-----------GT 4933
H1D.CD.84.84ZR085       -------T---.....----T---------------------------------................-------CT---G--T--A------A------------..............G------------G-..----------------C------A- 5063
H1F1.BE.93.VI850        -------T-G-.....A-------------A-C---------------T---................-------C--------T--------A------------..............G-T--G-G---C-..---------GC----------CCA- 4881
H1G.SE.93.SE6165        -------A---.....------------------------------------TC--................-----------------T------T-GAA--T---G--..............---T------G--GT...-G------------C---------- 4938
H1H.CF.90.056           ------GTG--.....-----A---------A-C----C-----------AC--................------A----------------T--A--------G-G..............G----------------------------------------GA 4888
H1J.SE.93.SE7887        -------A---.....------------------C--C----------------................-----------------TC-------A------------..............-G----------G--G-..-----------------------CA- 4855
H1K.CM.96.MP535         -------T-G-.....-----T------------G--------------------................-------CT-------T--------CA--------G--..............G-T---G--G-CC-..-----------G-----------A- 4737
H101_AE.TH.90.CM240     -------AG--.....-----A-G-------------C-T--------------................-------C--------T--------GAA-------C-..............-----------G--..-G-------C----------A... 5112
H102_AG.NG.x.IBNG       -----G-A---.....-----A------------------------------T---................-------G--------T--------GAAT--T---G--..............G------C-----C-..-----------------------G- 5066
H103_AB.RU.97.KAL153_2  -------T-G-.....-----------T---------G-----------------................----------------G---------------------G-..............---------------..----------------------- 4767
H104_cpx.CY.94.CY032    -------T-G-.....-----T--------------------------------T---................-------CT-------C------------------..............T-C-----------C-..----------------------A- 4907
H1N.CM.95.YBF30         ----GAGACC-.......--ATTA--------------C--------G--C--AC--................-----CA-C--------TC-A------A----C-GGG-G..............GG-G----G----G-..--------C--A--------G--T 5135
H1O.BE.87.ANT70         -------G-G-.......--ACTGA-C-A---------C-CTA---------GTC--................-----GA-A-------CTGC-A--T---AG----G--G--..............-A-G---G--GC-GG..------T--CC-A--C-----CCA- 5596
H1O.CM.91.MVP5180       ----G--G-G-.......--GCTGA-C-A---------C-TG----------GTC--................-----GA-A--------T---A--CT-GAA----G--G-G..............-A-GT-----GA-AT..------T--CC-A--C-----CCA- 5571
H1P.FR.06.RBF168        ----GG---GG.......--ATTGAAC-A---------CTC-----------GTC--................-----C--CT-G--G-TT--A--CT--AA--TGG--G--..............GGGAA-GT---A-GG..----------CC-C--------CCA- 5002
CPZ.CM.01.SIVcpzCAM13   ----G--T--T.......--AGAA--CC------CT--CAG--A------C-GGC--................-----CA-C--------T-C-A--TT--AA---T---G--..............-A-GA-----G-C-T...-GA-----C--A--C-----AG-T 5124
CPZ.CM.05.SIVcpzEK505   ----GAGACC-.......--A-TA----A---------C-CT-----G------AC--................-----C--C--------TC-----T--AT---C-GGG-G..............GG-GT-C-G----G-...-GA-----C--A--------G--C 5113
CPZ.CM.05.SIVcpzLB7     ------GAG--.......-------------A-------T--C-------C--G---................------A-C--------T------A-CT-T--C---T--..............CA-AG-C-GCCA-C-...-G---G-------------C-GC 5096
CPZ.CM.05.SIVcpzMB66    -------A---.......--------------------C-A---------------................-----CAG-T-------T------T-GT--A--------..............-A-AGC---CCA-G-...-GA-----C------C-------GA 5087
CPZ.CM.05.SIVcpzMT145   ---A-A-GGC-.......--CCT-----A-------TC-C--A---G--C-G---................-----G---T---------G---T---AAGCAG--GG-T..............-A-A-C--G--A-AT...-GA------C----CTCA--AG-A 5083
CPZ.CM.-.SIVcpzMB897    -------A---.......------------A-------TC-C-------C-CA---................-----CA------------T-------A-CT-----C-GC--..............-A-AGG--GCCA-AG...-A------C-A-----------GT 5107
CPZ.CM.-.SIVcpzDP943    ----GACACA-.......--AGTT----A------G--C-C------G--CC-AC--................-----G--C---------T--A--CT--AAG--C-GGG--..............GG-GT-G------A-...-GA-----C-------------CG-C 5127
CPZ.CM.98.CAM3          ----GA-ACA-.......--ATAT----A---------C-C------G---C-AC--................-----G--AT--------T--A--C---AG---C-GGG--..............-G-GT-GG-----A-...-A-----C-------------C--T 4955
CPZ.CM.98.CAM5          ----GA-ACA-.......--ATAT----A---------C-C-A----G--CC-AC--................K---G--A---------T--A--C---AG---T-GGG-G..............GG-GT-C------A-...-CA-----C------C---A--GSC 5243
CPZ.GA.88.GAB1          ----G--GC-T.......--AGCG---C-T--------CA---A------C-GGC--................-------CT---G-T-----C---AA---T--------..............T--GAG-G--G-CAT...-GA--------A---------CG-C 5605
CPZ.GA.88.GAB2          ----GG-G--T.......--G-CAAAGGT--------TC-CTAGG--G---CCTC-A................--TCAG--A---------TC----CA---T-CA-G--G--..............-AGA----G--A-AG...---------CC-A--------C--T 4934
CPZ.US.85.CPZUS         ----GAGACC-.......--A-TT--C-AA--------C-CTA-----G--C--T--................-----C----T-G----T------TT--AAG--C---G--..............GG--AG-GC--A-G-...-GA-----C--A--------G--T 5601
CPZ.CD.90.ANT           -----G-A-G-.......A-ATTGA-ATAC--CC-C--CA-GAAG-------GTC--................-----GA-A--G--G--T--A--TT-TTGTAAGA-CC--.......GAATTCAGAGG-TACCCC---GGTCCA-GAAG--AG--C------C-GT-T 4998
CPZ.TZ.00.TAN1          -----G-A-G-.......A-ATTGA-ATTC--CC-C-T----T-------C-GAC--................-----GA----G---T---A--TT-CAG-AAGG--G-TGAGAGCCAA...GAT-A--AGCC----GG-CCA-G-AGG--C-----ATC------A 5203
CPZ.TZ.2001.TAN2        -----G-A-G-.......A-ACTGA-ATTC--CC-T-T----A-------C-GAC--................-----GA---------T---A--TT-CAA-AAGR--G-Y..............GAGATTC--T-T-AG..--------AGGGA--A--GAGGC-C 5193
CPZ.TZ.2002.TAN3_1      -----G-A-G-.......A-ACTGA-ATTC--CC-T-T----A-------C-GAC--................-----GA---------T---A--TT-CAA-AAGG--G-T..............GAGA-CC--T-T-AG..--------AAGGGA--A--AAGGC-C 5193
GOR.CM.2007.SIVgor2139_287 -G--GG-A-A-.......C-CCTGTAC-A----A-C--C-CG----------GTC--................-----GA-AT-G--------A---T--AG---T---G--..............GGGAA-GTC-G-G--...-GA--------T--A------CA- 5078
GOR.CM.2007.SIVgorCP2135con -G--GG-A-A-.......C-CCTGTGC-A----A-C--C-C-----------GTC--................-----GA-AT-G--------A-----AG---T---G--..............GGGAA-GTC-G-G--...-GA--------T--A------CA- 5077
GOR.CM.2004.SIVgorCP684con -G--RG-A-A-.......C-CCTGTGC-A----A-C--C-C-----------GTC--................-----GA-AT-G--------A---T--AG---T---G--..............GGGAG-GTC-G-G--...-GA-----C--T--A------CA- 4997
H2A.DE.x.BEN            G---GG-A-AG.....T-ATTGT-CT-C--CA-C----CCCA--CT-----AGCACAG................---CC---A--T------C-A--C----T--T-G--CA-CAAA........................................ 5897
H2A.GW.x.ALI            G---GG-A-AG.....T-ATTGT-CT-C--CA-T----CCCAG-CC-----GTCACAG................---CCG--A--C----TTC----CT---TG-T-G-GCA-CAAA........................................ 5890
H2A.SN.x.ST             G---GG-A-AG.....C-ATTAT-CT-C--CA-C--C-CCCA--CC-----GT-CCAG................---CCG--A--C--G-TTC----CT---TG-T-G-GCA-CAAA........................................ 5342
H2B.CI.x.EHO            GG---G-A-AG.....A-ATTGT-CCAC--CA-C--C-C-T---CT----CAGGCCAG................---CC-AG-T----G-TTC-A--C---AG--TTG--CA-GAAGGAAAAG.................................. 5874
H2B.GH.86.D205          GG---G-A-AG.....A-ATTGT-CTAC--CA-C----C-T---CT--CG-AGG-CAG................---CC-AGCT----G-TTC-A--C---AGG-TCG--CAGGAAGGAAAAA................................. 5876
H2G.CI.92.ABT96         GG---G-GCAG.....A-ATTGT-CT-C--CA-C----C-A---CT-----A-G-CAG................-GCCTAGCT-G----TTC-A--CT--CA--T-G-GCA-AAAGGAC..................................... 5227
H2U.FR.96.12034         G----G-ACAG.....T-ATTGT-CT-C--CA-C-T--C-T-G-CT----GAC--CAG................---CC----T------TTC-A-----CAG-----GCA-G........................................... 5372
SMM.SL.92.SL92B         G----G-GCAG.......--GTTAT-CTAC----GC--CGCG-T--CC---C--TCTAGT................--CCA-AGC-----GCTG-----C----AAG-T-G--T--CAAA........................................ 5248
SMM.US.x.H9             GG---G-G-A-.....T-GCTGT--T-C--CA-G-TC-CGAA--CT-----G--TCAG................---CC-AGC-----G--TC-A------A---T-G---GTC............................................ 5284
SMM.US.x.PGM53          GG---G-G-A-.....T-GCTGT--T-C--CA-G-TC-CGAA--CT-----G--CCAG................---CC-AGC-----G--TC-A------A---T-G-G-GTC........................................... 5728
SMM.US.x.SIVsmH635F_L3  GG---G-G-A-.....T-GCTGT--T-C--CA-G-TC-CGAA--CT-----A--TCAG................---CC-AGC-----G--TC-A------A---T-G---GTC............................................ 5814

MAC.US.x.239            GG---G-ACA-.....C-GCTGT--T-C--CAGG-TC-CGAG--CT-----GT-CCAG................---CC-AGC-----G-----A-----GAA--T-G---GCG............................................ 5811
Vif                     R  G  E  Q  .  .  L  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  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
MAC.US.x.251_1A11       GG---G-ACA-.....C-GCTGT--T-C--CAGG-TC-CGAG--CT-----GT-CCAG................---CC-AGC-----G-----A-----GAA--T-G---GCG............................................ 5810
MAC.US.x.251_BK28       GG---G-ACA-.....C-GCTGT--T-C--CA-G-TC-CGAG--CT----GGT-CCAG................---CC-AGC-----G-----A------AA--T-G---GCG............................................ 5787
MAC.US.x.EMBL_3         GG---G-ACA-.....C-GCTGT--T-C--CA-G-TC-CGAG--CT----GGT-CCAG................---CC-AGC-----G-----A------AA--T-G---GCG............................................ 5279
STM.US.x.STM            GG---G-A-AG.....T-ACTGT--T-C--CAGG-TCACGAA--CT-----G--CCAG................---CC-AGCT----G-----A--TT--AA--T-G-GGA-C............................................ 5458
MNE.US.x.MNE027         GG---G-ACA-.....C-GCTGT--T-C--CAGG-TC-CGAG--CT-----G--CCAG................---CC-AG-----G-----A-----GAG--T-G---GTT............................................ 5279
DEB.CM.99.CM40          GG---G-G--C.....C-ACTG-AACA----C-GC-GT-CCATA-GGGACCAG-AGGT.........AAAGGACC-CCT--A-----G----A-A-----TAG--T-GC-G--AAAC........................................... 5327
DEB.CM.99.CM5           AG---G-G-A-.....C-ACTA-AACA----C-GC-AT-CCATA-GGGACCAG-AGGT.........AAAGGACC-CCC--AT----G-----A------AGGA--GC-G--AAAT........................................... 5321
DEN.CD.x.CD1            GG---G-AG-G.......--CCTGAGC-CT--C-CCAGGA-GAAG-C----T--C-TCCA......AGGACAGGTCC---TAGGGA--CGCCA-CATT--AGTATAT-G--T--......CAAGCAGCC-TTAG-T-TGGA-AGGTT-CCA-CATCACA--C-CCAGC-TT 5485
LST.CD.88.447           GG---G-A-AG.....A-A-TGG-ATTC----CTATG-CCA-T--------A...............GGTTGT--TCTC--AT-G--G-T-C-CTGCT-GAG-CA-C--CA-C............................................. 4777
LST.CD.88.485           GG---G-A-AG.....A-A-TGG-A-TC--C-CTATG-CCA-T--------A...............GGTTGT--TCTC--AT-G--G-TTC-CTGCT-GAG-CA-C--CA-C............................................. 4777
LST.CD.88.524           GG---G-A-AG.....A--CTG-AC--C---A--ATG-CTA-T--------A...............GGTTGT--TCTC--AT-G--G---C-CTGCT-GAG-CA-C--CA-C............................................. 4774
LST.KE.x.lho7           -G---G-A-G-.....A-A-CTATGT-T--CC-G...-CTTT---------A...............GGTTGT--TCTC---T-G--G-TTA-CTGCT-GAG-CA-C--CA-C............................................. 5858
SUN.GA.98.L14           ----GG-A-A-.....A-ACTA-A-GA---CAG-A------T--CT--C--A...............GGCCTG--TCTC---T-G--G-TTC-CTGCT-GAG--TTC-CC............................................... 5908
MND-1.GA.x.MNDGB1       GG---G-AGA-.......--GTTAAAAGTT---AG--T--C---T--C------AGCA.............CAAGGT--TCTC---T-G--G-TTC-CTGCT-GAG--TCA-CT............................................... 5226
OLC.CI.97.97CI12        -G---ATGCC-.......-A-CAT-GATAC--CC-GCTGG-GA----T---CTT...............GGG...--CCTG--GT----G-TGC----TT-GAG----G-C-G-A............................................. 5359
WRC.CI.97.97CI14        GGAC-GGT-A-ATAAAC--G-TAGA-GCA---C-G-----CAA-------GCA...............GGA...--GCTTA-CT-G--GAC----TGCT--CAG-TC--GCA-CATGAGCAACTCCAG-G-G--.......................... 5826
WRC.CI.98.98CI04        G-TC-GGA-GGTTAGATA-A--AAACGCA---C-------CAA-------GCA...............GGA...--GCTTA-CT-G--GAC----TGCT--CAG-TC--GCA-CATGAGCAACTCCAGGG-A--.......................... 5841
WRC.GM.05.Pbt_05GM_X02  G-CAGGGTTATTTACATCC--TAGA---C---A-------TCA-------CTA...............GGA...--GCTTA-CT-G--GGC---ATGTT-GCAG--C--GCA-CATGAGCAACTCAGG-AGAG-.......................... 4758
MND-2.CM.98.CM16        -----A-G-G-.....C-A-CAGTG...--CT-C-TC--CTGG--------A...............GGGAAG-----GAGC--G--G--------TT-GCT-AGT-AT-C-GCCTACTGTAACA.................................... 5519
MND-2.GA.x.M14          AC---G-A-G-.....A-AGTAG-G...---T-CAGC-CCTGG--------A...............GGTCAG-----GA--------CT------TT--AG---C-AC--CAAGTTTTGCAGAA.................................... 5446
MND-2.x.x.5440          -T---G-A-G-.....C-GGTAG--...---T-C-TC-CCTGG-------GA...............GGTCAG-----GA-------G-TT-----CT--CA---C-ATC-CAGAGGTAGAAAGC.................................. 5082
DRL.x.x.FAO             -G---T-T-G-.....A--CAAG--...--CT-T-TC-CTAG---G--C--G...............GGTCAG-----GAG---T--G--TC-T--CT-GCA---CCAC---AAGTACCTTAAGC................................... 5085
MON.CM.99.L1            G----G-G-AG.....C-CTTGGGACA----C--C---T----AC---CT--...............CCGAAG-----GA-C-CGTTGAC---A-AGAA-TT---C--CC-TGCCCTAGTGCGATATGG-G---CCCCC-C-AAGTCAT-CG--CCATTGGACC-CC-- 5388
MON.NG.x.NG1            G---GG-AGCG.....M-CCTAGGG-AC---C-CC-CAT-CAGAC---CT--...............AATAAG------A---CT-C-AC-C-A-AGAA-TT---C--CC-TGCTTGCTTCAAGTATGG-AA-TGC-CC-CCGAGCCAT-CT-CACATTGGACC-CC-- 3999
GSN.CM.99.CN166         GG--TG-GCAG.....A-CCTAGGA-AT------C-C-CG-A--CT--CCTGC--....................------A-A-CT--GACGC-T-AG-AGTT---C--CT-TGCATATGTCAAACATT............................... 5336
GSN.CM.99.CN71          GG--TG-ACAG.....C-CTTAGGATAT-----CC-C-CG-A--CT--CCCAC--....................------A-A-CT--CACGC-T-A-AGGCT---C--CT-TGCGTACGTTAGATACT............................... 5318
TAL.CM.00.266           G---GG-GG-G.....C-ACTAG-AGTT----G-GTCAG-CACAA-GG--GG...............GACACGCC-CCG--------G-T-C-A--CT-GA-----GC-G-TAGAT........................................... 5517
TAL.CM.01.8023          G---GG-GG--.....C--CTTG-GGT-------GTCAG-CATAA-GG--GG...............GACTCGCC-CC---CT----G-T-C-A--CT--AAG-T-GC-G-TAGGC........................................... 5015
MUS-1.CM.01.1085        G------GCAC.....AC-CTAGGGGAT------C---C--AC-C---C-CA---...................---------C---CT-CT-CT--ACA-G-CCT---T---T--GCATTAAAGAAGC................................. 5292
MUS-1.CM.01.CM1239      G------ACC-.....CCCTTAGGGGAT------C-C-C--G--C-----CA...AAG..................------C-C-CT-CCACTC-ACAGG--TT----C--T--ACATTAAAGAAACATGG-AAG-GT-CC-CCCAGCCA-AGA-C-CC........... 5314
MUS-2.CM.01.CM1246      G----G-ACAG.....C-CCTAAGA-A------GC-CATCAAGAC---CT--CC-AAG..................------A---CT-T--CTC-A-A-AAGCT------CT-TGCATACATCAAACAT---AATGG-GCA-GCGCCTCCAAGT-ACCCT--CTC...... 5431
MUS-2.CM.01.CM2500      G---GG-A-GG.....A-CCTAGGG-TT---C--C-C--GA---C---CCT-CCCAAG..................------A-C-CT-CT-C-C-A-AGAAGTTG--C--CT-CGCATATATCAAACAGGG-GATGG-GCA-C-CCCTCCAAGT-ACCCT--CCC...... 5384
SYK.KE.x.KE51           G---GG-G-AG.....T-G-CTT-GCAC--CTGGA--TTCCATAA-G-AC-G...................---CTT--AT-G-----T-----TT--CA-AAG-ATT--AGCAAG............................................. 5357
SYK.KE.x.SYK173         G---GG-AC--.....T-G-CAT-ACAC--CTGGACC-C-CATACTG-CC-G...................---CCC--A--T--G--T--A---T--CA--TG-ACC--AAAG............................................... 5700
RCM.GA.x.GAB1           ----GG-A-A-........---CTAGTG...-----CC-TC----GG-------GA...............GATCAG-----GA----G----TT------A--CA---G-GGGCTAAA...AGTCAGTTAGATAGGT............................ 5012
RCM.NG.x.NG411          -----G-GCA-........--GCTG-T-...-----GG-T-G-GAA-------GA...............GATCAG-----GAGC--------TC-----A--AA---T-GGGCTAGGCAGCAGCTAAAGG................................. 5033
GRV.ET.x.GRI_677        G---GG-G-AG.....TA--CGTGG...--CAC--TCA-G-A---------AGGTCAG...............---CA---A--G---CTT---------TT----AT-C-A............................................... 5682
SAB.SN.x.SAB1C          GG---G-A-AG.....TA-CTTTGG...---A--C-C--G-TG------C-GCCAACA........GGGCAA---CC---AT----G--T--A-----TAGG-T--AT-C-A............................................... 5910
TAN.UG.x.TAN1           G---GG-A-GC.....T--TTGTGG...--CAC---CA-G-A------CGTAGCAGAAAACCATTGGGGTCAG---C-G--C--G-G-TT--A--C--GA---T--AT-C-GACTTCCTAAGAA...................................... 5738
VER.DE.x.AGM3           -G-----G-GG.....T-C-C-TTC...---C-G-TC-CC-AG-----C--G--AACA........GGACAG---CCC---T-G--------A--T--CCTT---CACCA-A.............................................. 5243
VER.KE.x.9063           GG---G-G-G-.....T-CGT-TTC...---C---TC-CC-AG---------A-CCACA........GGAAAG---CCC--A--G--G-----A--C--TCTG--CCATCAGA.............................................. 5745
VER.KE.x.AGM155         GG---A---GG.....T-C-T-TTC...---C---T--C--G---------ACTAACA........GGTCAG---CCC--CT-G-----T--A---T--CT---CCATCA-A.............................................. 5740
VER.KE.x.TYO1_patent    GG--C----G-.....T--GT-TTC...---C-G-T--C--A---G-----A-GCACA........GGACAG---CC----T-G--G---C-A--T--GCTC---CATCA-A.............................................. 5745
COL.CM.x.CGU1           --A-GGG-T--......CCCT-GAA-CCT-----T-TC--G-T------CCTA...............GGA...--GTTGAAGT-G---G-AC-ATGT--TCTG--CC---G-GAGAAA...GAGAGG--C-ACC-G---.............................. 5280
```

```
H1B.FR.83.HXB2              AAACTGACAGAGGATAG.................................................................................. 5558
Vif                         _K__L__T__E__D__R#__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__
H1A1.UG.85.U455             ---T-A-----------.................................................................................. 5004
H1B.US.90.WEAU160           -----------------.................................................................................. 5557
H1C.ET.86.ETH2220           ---T-AGT--------A.................................................................................. 4950
H1D.CD.84.84ZR085           --G--A-----A-----.................................................................................. 5080
H1F1.BE.93.VI850            -----AGT------C--.................................................................................. 4898
H1G.SE.93.SE6165            G--T-AG----A-----.................................................................................. 4955
H1H.CF.90.056               --G--AGT---------.................................................................................. 4905
H1J.SE.93.SE7887            -----AGT---------.................................................................................. 4872
H1K.CM.96.MP535             -----AGT---------.................................................................................. 4754
H101_AE.TH.90.CM240         G--A-A-----A-----.................................................................................. 5129
H102_AG.NG.x.IBNG           --GT-AG----A--C--.................................................................................. 5083
H103_AB.RU.97.KAL153_2      -----------------.................................................................................. 4784
H104_cpx.CY.94.CY032        -----AGTG--------.................................................................................. 4924
H1N.CM.95.YBF30             --G--A-----A-----.................................................................................. 5152
H1O.BE.87.ANT70             ---T-A-----A-----.................................................................................. 5613
H1O.CM.91.MVP5180           -G-T-A-----A-----.................................................................................. 5588
H1P.FR.06.RBF168            -T-T-----A-A----T.................................................................................. 5019
CPZ.CM.01.SIVcpzCAM13       ---T-A-----------.................................................................................. 5141
CPZ.CM.05.SIVcpzEK505       --GT-A-----------.................................................................................. 5130
CPZ.CM.05.SIVcpzLB7         ---T-AG----A-----.................................................................................. 5113
CPZ.CM.05.SIVcpzMB66        ---T-AG----------.................................................................................. 5104
CPZ.CM.05.SIVcpzMT145       --GT-AG----A-----.................................................................................. 5100
CPZ.CM.-.SIVcpzMB897        ---T-AG----A--C--.................................................................................. 5124
CPZ.CM.-.SIVcpzDP943        --GT-AG----A-----.................................................................................. 5144
CPZ.CM.98.CAM3              --GT-A-----A-----.................................................................................. 4972
CPZ.CM.98.CAM5              --GT-AG----A-----.................................................................................. 5260
CPZ.GA.88.GAB1              ---T-A-----------.................................................................................. 5622
CPZ.GA.88.GAB2              -TCT-AG----A-----.................................................................................. 4951
CPZ.US.85.CPZUS             ---T-------A-----.................................................................................. 5618
CPZ.CD.90.ANT               -TCT-A-----A-----.................................................................................. 5015
CPZ.TZ.00.TAN1              -----A-------C--................................................................................... 5220
CPZ.TZ.2001.TAN2            TTG-CAT-T-TTACA-AGCTGACAGAGGACAG................................................................... 5225
CPZ.TZ.2002.TAN3_1          TTG-CAT-T-TTACA-AGCTGACAGAGGACAG................................................................... 5225
GOR.CM.2007.SIVgor2139_287  ---T-AG---CA----C.................................................................................. 5095
GOR.CM.2007.SIVgorCP2135con ---T-AG---CA----C.................................................................................. 5094
GOR.CM.2004.SIVgorCP684con  ---T-AG---CA----C.................................................................................. 5014
H2A.DE.x.BEN                .........AT-ACAGAC.....CCCAGAGAAAGGGTACCG.....CCAGGAAACAGTGGAGAAGAGACCATTGGAGAGGCCTTCGAGTGGCTAGAGAGGACCATAGAAGCCTTAAACAGGGAGGCAGTGAACCATCTGCCCCGAGAGCTCATTTTCCAGGTGTGGCA 6046
H2A.GW.x.ALI                .........AT--CA-AC.....CCCAGAGAAACAGTACCA.....CCAGGAAACAGTGGCGAAGAGACTATCGGAGAGGCCTTCGAGTGGCTAGACAGGACAGTAGAGGCCTTAAACAGAGAGGCGGTGAATCACCTGCCCCGGGAGCTCATTTTCCAGGTGTGGCA 6039
H2A.SN.x.ST                 .........AT--CAG-C.....CCCAGAGAGACAATACCA.....CCAGGAAACAGTGGCGAAGAAACTATCGGAGAGGCCTTCGAGTGGCTAGACAGGACGGTAGAAGCCATAAACAGAGAGGCAGTGAACCACCTGCCCCGAGAGCTTATTTTCCAGGTGTGGCA 5491
H2B.CI.x.EHO                .........AT-...GAT.....CCCAGGGAGAGAGTACCA.....CCAGGAAACAGCGGCGAAGAAACAGTAGGAGAGGCATTCGAATGGCTAGAGACAACATTAGAACATCTCAACAGAGTAGCAGTCAATCACTTGCCCAGGGAACTTATTTTCCAGGTCTGGCA 6020
H2B.GH.86.D205              .........AT-...GAT.....CCCAGGGAGAGAGTGCCA.....CCAGGAAACAGCGACGAAGAAACAGTAGGAGAAGCATTCGCTTGGCTAGAAAGAACAATAACAGAGCTCAACAGGGTAGCGGTCAACCATTTGCCCCGAGAACTTATTTTCCAGGTCTGGCA 6022
H2G.CI.92.ABT96             .........AT-...GAT.....CCAAGGGAGAGAGTMCCA.....CCAGAAAACAGCGACGAGGAGACAATCGGAGAGGCATTCGAATGGCTAGAAAGAACAGTMACAGAGATCAACAGGATAGCAGTCAATCATTTACCCAGGGAGCTTATTTTCCAGGTTTGGCG 5373
H2U.FR.96.12034             .........AT--GAGAT.....CCAAGAGAGAGAATCCCA.....CCAGGAAATTCAGAAGAAGAAACTGTAGGAGAGGCTTTCAGTTGGCTAGAAAGAACTGTGAGAGATATCAACAGGGCAGCAGTGCAACATCTTCCCAGGGAACTTATTTTCCAGGTCTGGCA 5521
SMM.SL.92.SL92B             .........AT-ACAGAC.....CCAAGGGAAAGAATCCCA.....CCAGGAAACAGTGGAGAAGAAACAATAGGAGAGGCCTTCGAATGGTTACACAATACAGTGGAGGCCCTGAATCAAACAGCAGTACAACACCTACCAAGAGAGTTAATTTTCCAGGTCTGGAG 5397
SMM.US.x.H9                 .........AT-TCRSRT.....CCCAGGGAGAGRATCCCA.....CCTGGAAACAGTGGAGAAGAAACRRTAGGCGAGGCYTTCGACTGGCTACACAGAACASTAGAAGAAATAAACAGGGCAGCAGTGAATCATTTGCCRAGGGAGCTAATTTTCCAGGTCTGGCG 5433
SMM.US.x.PGM53              .........AT-TCAGAT.....CCCAGAGAGAGGATCCCA.....CCTGGAAACAGTGGAGAAGAAACAATAGGAGAGGCATTCGAATGGCTAAACAGAACAGTAGAAGAAATAAACAGGGCAGCAGTGAATCACTTGCCGAGGGAGCTAATTTTCCAGGTTTGGCG 5877
SMM.US.x.SIVsmH635F_L3      .........AT-TCAGAT.....CCCAGAGAGAGGATCCCA.....CCTGGAAACAGTGGAGAGGAAACAATAGAAGAGGCTTTCGAATGGCTAAACAGAACAGTAGAAGGAATAAACAGGGCAGCAGTAAATCACCTGCCGAGGGAGCTAATTTTCCAGGTTTGGCA 5963
                                     Vpx start
MAC.US.x.239                .........AT-TCAGAT.....CCCAGGGAGAGAATCCCA.....CCTGGAAACAGTGGAGAAGAGACAATAGGAGAGGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCGGTAAACCACCTACCAAGGGAGCTAATTTTCCAGGTTTGGCA 5960
Vpx                           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  N  R  T  V  E  E  I  N  R  E  A  V  N  H  L  P  R  E  L  I  F  Q  V  W  Q
Vif                         _.__.__##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  Q  N  S  R  G  D  K  Q  R  G  G  K  P  P  T  K  G  A  N  F  P  G  L  A
MAC.US.x.251_1A11           .........AT-TCAGAT.....CCCAGGGAGAGAATCCCA.....CCTGGAAACAGTGGAGAAGAGACAATAGGAGAGGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCGGTAAACCACCTACCAAGGGAGCTAATTTTCCAGGTTTGGCA 5959
MAC.US.x.251_BK28           .........AT-TCAGAT.....CCCAGGGAGAGAATCCCA.....CCTGGAAACAGTGGAGAAGAGACAATAGGAGAGGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCAGTAAACCACCTACCAAGGGAGCTGATTTTCCAGGTTTGGCA 5936
MAC.US.x.EMBL_3             .........AT-TCAGAT.....CCCAGGGAGAGAATCCCA.....CCTGGAAACAGTGGAGAAGAGACAATAGGAGAGGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCGGTAAACCACCTACCAAGGGAGCTGATTTTCCAGGTTTGGCA 5428
STM.US.x.STM                .........AT-TCAGAT.....CCCAGAGAGAGAATACCG.....CCCGGAAACAGTGGAGAAGAGGCAATAGAGGAAGCATTCGAGTGGCTACACAGAACGGTAGAGGACATAAACCGAGAGGCAGTAAACCATCTACCGAGGGAACTGATTTTCCAGGTCTGGCG 5607
MNE.US.x.MNE027             .........AT-TCAGAT.....CCCAGAGAGAGAATCCCA.....CCTGGAAACAGTGGAGAAGAGACAATAGGAGAAGCCTTCGAATGGCTAAACAGAACAGTAGAGGAGATAAACAGAGAGGCAGTAAACCACCTACCAAGGGAGCTGATTTTCCAGGTTTGGCA 5428
DEB.CM.99.CM40              ................................................................................................... 5327
DEB.CM.99.CM5               ................................................................................................... 5321
DEN.CD.x.CD1                -GTT-C-AGA......................................................................................... 5495
LST.CD.88.447               ................................................................................................... 4777
LST.CD.88.485               ................................................................................................... 4777
LST.CD.88.524               ................................................................................................... 4774
LST.KE.x.lho7               ................................................................................................... 5858
SUN.GA.98.L14               ................................................................................................... 5908
MND-1.GA.x.MNDGB1           ................................................................................................... 5226
OLC.CI.97.97CI12            .........AT--CAGAA..............AGAGTTCCAGAGGAT...CAGGGGCCACCCAGG.........GAGCCACTGGAGGATTGGGTGAGAGACACAATGGAAGAAATTATTGTAGAGGCAAAAAGGCACTTTCCCAGGGAGTTATTGGTCCGAGTTTTTCT 5493
WRC.CI.97.97CI14            ................................................................................................... 5826
WRC.CI.98.98CI04            ................................................................................................... 5841
WRC.GM.05.Pbt_05GM_X02      ................................................................................................... 4758
MND-2.CM.98.CM16            .........AT--CAGAA..............GGGGCCCCAGAGATCCCAGAAGGAGCAGGAGAG.........GTGGATCTAAATACCTGGTTGGAAAGATCTCTGGAGAAAATCAATCAGGAGGCAAGATTACACTTCCACCCAGAGTTCCTTTTCCGTCTCTGGAA 5656
MND-2.GA.x.M14              .........AT--CAGAA..............AGAGCACCAGAGGCACCACAGGGGGCAGGAGAA.........GTGGGACTAGAACAGTGGCTGGAAAGATCATTGGAACAAGTGAACAGAGAGGCTCAATTACACTTCCACCCAGAGTTCCTTTTCCGTCTTTGGAA 5583
MND-2.x.x.5440              .........AT--CAGAG..............AGGGCACCAGAGGCGCCAGAAGGAGCAGGAGAG.........GTAGGACTGGAGCAATGGCTGGAAACGTCACTGGAGAGAATCAACCGGGAGGCCCGGTTACACTTCCACCCAGAGTTCCTTTTCCGTCTCTGGAA 5219
DRL.x.x.FAO                 .........AT--CAGAAAGACAGTCAGTGGAGAGAGCTCCAGCGGAGCCAATGGGAGCAGGAGAG.........GTAGAGTTAGAAGAATGGCTACAGAGGAGTCTCTTAAGAATCAACCAGGAGGCTCGATTACACTTCCACCCAGAGTTCCTCTTCCGTCTTTGGAA 5237
MON.CM.99.L1                G.................................................................................................. 5389
MON.NG.x.NG1                ................................................................................................... 3999
GSN.CM.99.CN166             ................................................................................................... 5336
GSN.CM.99.CN71              ................................................................................................... 5318
TAL.CM.00.266               ................................................................................................... 5517
TAL.CM.01.8023              ................................................................................................... 5015
MUS-1.CM.01.1085            ................................................................................................... 5292
MUS-1.CM.01.CM1239          ................................................................................................... 5314
MUS-2.CM.01.CM1246          ................................................................................................... 5431
MUS-2.CM.01.CM2500          ................................................................................................... 5384
SYK.KE.x.KE51               ................................................................................................... 5357
SYK.KE.x.SYK173             ................................................................................................... 5700
RCM.GA.x.GAB1               .........AT--CAGAA..............AGAGCCCCC.....GAGGTCCCCACTGGGGCTGGAGAAGCAGAGTTCCAGCCTTGGCTACGGGACATGCTAGAAAAGGTCAACTTGGAAGCC...AGGTTACACTTTCATCCAGAGTTCATTTTCCGTCTGTGGCG 5149
RCM.NG.x.NG411              .........AT--CAGAG......GGCAGAGAAAGAGTGCCA.....GAGGCCCCCACTGGGGCTGGAGATGTAGAGTTTGCCCCCTGGCTTCACAGAATGCTAACAGAAGTCAACTTAGAAGCC...AGGTTGCACTTTCATCCAGAGTTCATTTTCCGTCTGTGGCG 5179
GRV.ET.x.GRI_677            ................................................................................................... 5682
SAB.SN.x.SAB1C              ................................................................................................... 5910
TAN.UG.x.TAN1               ................................................................................................... 5738
VER.DE.x.AGM3               ................................................................................................... 5243
VER.KE.x.9063               ................................................................................................... 5745
VER.KE.x.AGM155             ................................................................................................... 5740
VER.KE.x.TYO1_patent        ................................................................................................... 5745
COL.CM.x.CGU1               ................................................................................................... 5280
```


PLV Complete
Genomes

```
H1B.FR.83.HXB2               ...................................................................................................................................................... 5558
Vif                          .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._
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.x.IBNG            ...................................................................................................................................................... 5083
H103_AB.RU.97.KAL153_2       ...................................................................................................................................................... 4784
H104_cpx.CY.94.CY032         ...................................................................................................................................................... 4924
H1N.CM.95.YBF30              ...................................................................................................................................................... 5152
H1O.BE.87.ANT70              ...................................................................................................................................................... 5613
H1O.CM.91.MVP5180            ...................................................................................................................................................... 5588
H1P.FR.06.RBF168             ...................................................................................................................................................... 5019
CPZ.CM.01.SIVcpzCAM13        ...................................................................................................................................................... 5141
CPZ.CM.05.SIVcpzEK505       ...................................................................................................................................................... 5130
CPZ.CM.05.SIVcpzLB7          ...................................................................................................................................................... 5113
CPZ.CM.05.SIVcpzMB66         ...................................................................................................................................................... 5104
CPZ.CM.05.SIVcpzMT145        ...................................................................................................................................................... 5100
CPZ.CM.-.SIVcpzMB897         ...................................................................................................................................................... 5124
CPZ.CM.-.SIVcpzDP943         ...................................................................................................................................................... 5144
CPZ.CM.98.CAM3               ...................................................................................................................................................... 4972
CPZ.CM.98.CAM5               ...................................................................................................................................................... 5260
CPZ.GA.88.GAB1               ...................................................................................................................................................... 5622
CPZ.GA.88.GAB2               ...................................................................................................................................................... 4951
CPZ.US.85.CPZUS             ...................................................................................................................................................... 5618
CPZ.CD.90.ANT                ...................................................................................................................................................... 5015
CPZ.TZ.00.TAN1               ...................................................................................................................................................... 5220
CPZ.TZ.2001.TAN2             ...................................................................................................................................................... 5225
CPZ.TZ.2002.TAN3_1           ...................................................................................................................................................... 5225
GOR.CM.2007.SIVgor2139_287   ...................................................................................................................................................... 5095
GOR.CM.2007.SIVgorCP2135con ...................................................................................................................................................... 5094
GOR.CM.2004.SIVgorCP684con  ...................................................................................................................................................... 5014
H2A.DE.x.BEN                 AAGGTCCTGGAGATATTGGCATGATGAACAAGGGATGTCAGCAAGCTACACAAAGTATAGATATTTGTGCCTAATGCAAAAAGCTATATTTACACATTTCAAGAGAGGGTGCACTTGCTGGGGGGAGGACATGGGCCGGGAAGGATTGGAAGACCAAGGACCTCCCCCTC 6216
H2A.GW.x.ALI                 AAGGTCCTGGAGATACTGGCATGATGAACAAGGAATGTCACAAAGTTACACAAAATATAGATACTTGTGCTTAATGCAGAAAGCTATGTACACACATTTCATGAAAGGGTGCACCTGCCTGGGGGGAGGGCATGGGCCAGGAGGATGG...AGATCAGGACCTCCCCCTC 6206
H2A.SN.x.ST                  AAGGTCCTGGAGATACTGGCATGATGAACAAGGAATGTCAATAAGTTACACAAAGTATAGATATTTGTGCCTAATGCAGAAAGCTATGTTCATACATTCTAAGAGAGGGTGCACTTGCCTGGGGGGAGGACATGGGCCGGGAGGATGG...AGATCAGGACCTCCCCCTC 5658
H2B.CI.x.EHO                 GAAGTCTTGGGCATATTGGCGTGAGGAACAGGGCATGTCAATTAGTTATACCAAATATAGGTATTTGTGTCTAATGCAAAAAGCAATGTTTATACACTTTGCAAAGGGTTGTGGATGTCTGCGGGAAGGCCATGGGCCAGGGGGATGG...AGATCAGGACCTCCTCCTC 6187
H2B.GH.86.D205               GAGGTCTTGGGCATACTGGCGTGAGGAACAGGGCATGTCAATTAGCTATACCAAATATAGATACTTGTTGCTAATGCAGAAAGCAATGTTTGTGCACTATACAAAGGGCTGTAGGTGCCTGCAGGAGGGCCATGGGCCAGGGGGATGG...AGATCAGGACCTCCTCCTC 6189
H2G.CI.92.ABT96              AAGGTCCTGGGAATATTGGCATGATGAACAGGGCATGTCAGCTAGTTACACTAAGTATAGATATTTGTTGCTGATGCAAAAGGCTATGTTTGTACATTGTAAGAAAGGATGCACATGCTTGCAGAAGGGCCATGGGCCTGGAGGATGG...AGACAAGGACCTCCTCCTC 5540
H2U.FR.96.12034              AAGGTCCTGGCAATACTGGCGTGATGAACAGGGCATGTCAGAAAGTTATACTAAGTATAGATATCTCTGTCTGATACAGAAAGCTTTATTTATGCATTGCAAAAAGGGATGTCAATGCTTGGGGGAAGGCCATGGACCAGGAGGATGG...AGATCAGGGCCTCCTCCTC 5688
SMM.SL.92.SL92B              AAGGTGTTGGGAATACTGGGTAGATGAGCAAGGATACTCACCAAGCTATGCAAAATATAGATATGTACAGTTAATGCAAAAAGCAATGTTTCAGCACTTCAGAAAGGGATGCACATGCAGAGGAGAAGGGCATTCACAGGGGGGCTGG...AGAACTGGACCACCACCTC 5564
SMM.US.x.H9                  AAGGKCCTGGGAATACTGGCATGATGAAATGGGGATGTCAGTCAGCTACACTAAATRCAGATACTTGTGCTTGATACAGAAAGCTCTGTTTATGCATTGCAAGAAAGGGTGTAGGTGCTTAGGAGGAGAGCATGGGGCAGGGGGATGG...AGACCAGGGCCTCCTCCTC 5600
SMM.US.x.PGM53               AAGGTCTTGGGAATACTGGCGTGATGAAATGGGGATGTCAGAGAGCTACACAAAATACAGATACTTGTGCTTAATACAGAAAGCTCTGTTTGTGCATTGCAAGAGAGGGTGTAGGTGCTTAGGAGAAGAGCATGGGGCAGGGGGATGG...AGATCAGGGCCTCCTCCTC 6044
SMM.US.x.SIVsmH635F_L3       AAGGTCCTGGGAATACTGGCATGATGAAATGGGAATGTCAGAAAGCTACACAAAATATAGATACTTGTGCTTGATACAGAAAGCTCTGTTTATGCATTGCAAGAAAGGGTGTAGGTGCTTAGGAGAAGGGCATGGGGCAGGGGGATGG...AGAACAGGGCCTCCTCCTC 6130

                                                       Vif end
MAC.US.x.239                 AAGGTCTTGGGAATACTGGCATGATGAACAAGGGATGTCACCAAGCTATGTAAAATACAGATACTTGTGTTTAATACAAAAGGCTTTATTTATGCATTGCAAGAAAGGCTGTAGATGTCTAGGGGAAGGACATGGGGCAGGGGGATGG...AGACCAGGACCTCCTCCTC 6127
Vpx                           _R__S__W__E__Y__W__H__D__E__Q__G__M__S__P__S__Y__V__K__Y__R__Y__L__C__L__I__Q__K__A__L__F__M__H__C__K__K__G__C__R__C__L__G__E__G__H__G__A__G__G__W__.__R__P__G__P__P__P__
Vif                           _K__V__L__G__I__L__A__*
MAC.US.x.251_1A11            AAGGTCTTGGGAATACTGGCATGATGAACAAGGGATGTCACCAAGCTATGTAAAATACAGATACTTGTGTTTAATACAAAAGGCTTTATTTATGCATTGCAAGAAAGGCTGTAGATGTCTAGGGGAAGGACACGGGGCAGGGGGATGG...AGACCAGGACCTCCTCCTC 6126
MAC.US.x.251_BK28            AAGGTCTTGGGAATACTGGCATGATGAACAAGGGATGTCACAAAGCTATGTAAAATACAGATACTTGTGTTTAATGCAAAAGGCTTTATTTATGCATTGCAAGAAAGGCTGTAGATGTCTAGGGGAAGGACACGGGGCAGGAGGATGG...AGACCAGGACCTCCTCCTC 6103
MAC.US.x.EMBL_3              AAGGTCTTGGGAATACTGGCATGATGAACAAGGGATGTC............................................................................................................... 5467
STM.US.x.STM                 AAGGTCTTGGGAATACTGGCATGATGAGCAAGGGATGCCCGGGAGCTATGTTAAATATAGATACTTGTGCTTGATACAGAAAGCTTTGTTTATGCATAGCAAGAGGGGATGCAGATGCTTGGGGGAAGGACATGGGGCAGGAGGATGG...AGACCAGGCCCTCCTCCTC 5774
MNE.US.x.MNE027              AAGGTCTTGGGAATACTGGCATGATGAACAAGGGATGTCACCAAGCTATGTAAAGTACAGATACTTGTGTTTAATACAAAAGGCTTTATTTATGCATTGCAAAAAAGGCTGTAGATGTCTAGGGGAAGGACATGGGGCAGGGGGATGG...AGACCAGGACCTCCTCCTC 5595
DEB.CM.99.CM40               ...................................................................................................................................................... 5327
DEB.CM.99.CM5                ...................................................................................................................................................... 5321
DEN.CD.x.CD1                 ...................................................................................................................................................... 5495
LST.CD.88.447                ...................................................................................................................................................... 4777
LST.CD.88.485                ...................................................................................................................................................... 4777
LST.CD.88.524                ...................................................................................................................................................... 4774
LST.KE.x.lho7                ...................................................................................................................................................... 5858
SUN.GA.98.L14                ...................................................................................................................................................... 5908
MND-1.GA.x.MNDGB1            ...................................................................................................................................................... 5226
OLC.CI.97.97CI12             ATTCTGTGCTCAAAGACATATCA............................................................................................................................... 5516
WRC.CI.97.97CI14             ...................................................................................................................................................... 5826
WRC.CI.98.98CI04             ...................................................................................................................................................... 5841
WRC.GM.05.Pbt_05GM_X02       ...................................................................................................................................................... 4758
MND-2.CM.98.CM16             CGCATGCATAGAACACTGGCATGACAGACATCAGAGAAGCTTAAGCTATGCAAAGTATAGATACCTGCTGTTGATGAATAAAGCAATGTTTACTCACATGCAACAAGAATGTCCTTGTAGAAGTGGG...CATCCG...............AGA...GGA.........C 5796
MND-2.GA.x.M14               CACCTGTGTAGAACACTGGCATGATAGACTTAGAAGAACCTTAAATTATGCCAAATATAGGTATCTTCTGCTGATGCAGAAAGCCATGTTTGTTCATATGCAACAAGGATGCCCATGTAGAAGTGGA...CATTCA...............AGA...GGA.........C 5723
MND-2.x.x.5440               CACATGTGTAGAACACTGGCATGATAGACATCAGAGATCTCTTGATTATGCCAAGTATAGATACCTGCTGTTGATGCATAAGGCCATGTATACTCACATGCAACAGGGATGCCCATGTAGAAATGGG...CGCCCAAGG..................GGA.........C 5359
DRL.x.x.FAO                  CACCTGCATGGAGCACTACCATGATGCTCTTCAGTTATCTTTTACTTACAGCAAGTATAGATACCTACTTTTGTTACAGAAGGCCATGTTCATGCACTTTCAGCAAGGATGCTCATGTCTGCAGGGAAGGCATCCACCTCCCCTCAGACCAGCAGGAGATAGACTTCCTC 5407
MON.CM.99.L1                 ...................................................................................................................................................... 5389
MON.NG.x.NG1                 ...................................................................................................................................................... 3999
GSN.CM.99.CN166              ...................................................................................................................................................... 5336
GSN.CM.99.CN71               ...................................................................................................................................................... 5318
TAL.CM.00.266                ...................................................................................................................................................... 5517
TAL.CM.01.8023               ...................................................................................................................................................... 5015
MUS-1.CM.01.1085             ...................................................................................................................................................... 5292
MUS-1.CM.01.CM1239           ...................................................................................................................................................... 5314
MUS-2.CM.01.CM1246           ...................................................................................................................................................... 5431
MUS-2.CM.01.CM2500           ...................................................................................................................................................... 5384
SYK.KE.x.KE51                ...................................................................................................................................................... 5357
SYK.KE.x.SYK173              ...................................................................................................................................................... 5700
RCM.GA.x.GAB1                CACTTGTGTGGAACACTGGCATGACGTACATCAGAGAAGTTTAGAATATGCTGCTTATAGGTACTTACTGCTGATGCAAAAGGCTTTGTTTATCCACTGTCAGACAGGGTGTAGTCAGAGA........CATGGACCCAATCCTAGG...GCAGTAGGAGAAAGGATAA 5307
RCM.NG.x.NG411               TACTTGTGTGGAACACTGGCATGATAGGCTTGGAAGAAGCCTTGAGTATGCAGGCTATAGATATCTGCTTCTGATGCAAAAAGCTCTGTTTATTCATTGCCAATCAGGGTGTTCTCAGAGA........CATGGACAGGGACAAGCAAGGGAAGCAGGAGAAAGAATCC 5340
GRV.ET.x.GRI_677             ...................................................................................................................................................... 5682
SAB.SN.x.SAB1C               ...................................................................................................................................................... 5910
TAN.UG.x.TAN1                ...................................................................................................................................................... 5738
VER.DE.x.AGM3                ...................................................................................................................................................... 5243
VER.KE.x.9063                ...................................................................................................................................................... 5745
VER.KE.x.AGM155              ...................................................................................................................................................... 5740
VER.KE.x.TYO1_patent         ...................................................................................................................................................... 5745
COL.CM.x.CGU1                ...................................................................................................................................................... 5280
```

```
                                                    Vpr start                                                                                    Vif end
H1B.FR.83.HXB2             ..........................ATGGAACAAGCC...............................................................................CCAGAAGACCAAGGGCCACAGAGGGAGCCACACAATGAATGGACACTAGAGCTTTTA 5627
Vif                        _.__.__.__.__.__.__.__.__. W  N  K  P_.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.  Q  K  T  K  G  H  R  G  S  H  T  M  N  G  H  *
Vpr                                                  _M__E__Q__A__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._P__E__D__Q__G__P__Q__R__E__P__H__N__E__W__T__L__E__L__L
H1A1.UG.85.U455            ..........................------------...............................................................................-----------G---------------GT--GC-----A-G--T-----T-G--- 5073
H1B.US.90.WEAU160          ..........................------------...............................................................................-----------------------------T------------------------------ 5626
H1C.ET.86.ETH2220          ..........................------------...............................................................................-----------G----G--------A---T-----------G-------------- 5019
H1D.CD.84.84ZR085          ..........................------------...............................................................................--G-C------------G-------------T--------------T----------- 5149
H1F1.BE.93.VI850           ..........................------------...............................................................................---G------G-------------------T-----------G-------AA----- 4967
H1G.SE.93.SE6165           ..........................------------...............................................................................-----------G---------A--A--CT-----------G--T----A--G--- 5024
H1H.CF.90.056              ..........................------------...............................................................................-----------G------------------------------------------------ 4974
H1J.SE.93.SE7887           ..........................------------...............................................................................-----------C---------A------T--C-------------------------G 4941
H1K.CM.96.MP535            ..........................------------...............................................................................-----------G----------------A--------------T-----A----G 4823
H101_AE.TH.90.CM240        ..........................------------...............................................................................--------..G---------A--A--CT--------------T----A--G--- 5195
H102_AG.NG.x.IBNG          ..........................------AG----...............................................................................-----------G---------------GTC------------------A--G--- 5152
H103_AB.RU.97.KAL153_2     ..........................------------...............................................................................-----G--------A--------------T-------G------------------ 4853
H104_cpx.CY.94.CY032       ..........................------------...............................................................................-----------G----G----A--A---A-T-------------------------G 4993
H1N.CM.95.YBF30            ..........................-------G---A...............................................................................--------TGC-------------A--CT-T---------G-------AT-A--- 5221
H1O.BE.87.ANT70            ..........................-----------A...............................................................................--T--GA-T--G--A---GCT-AA-----CTT----------G----------C--- 5682
H1O.CM.91.MVP5180          ..........................------------...............................................................................-TG---A-T--G--A---GCT---------TT--------------------CC-G 5657
H1P.FR.06.RBF168           ..........................-----G------...............................................................................-----G----------G-CA--A------TTT-----G----TG---C-AACC--- 5088
CPZ.CM.01.SIVcpzCAM13      ..........................--------C---...............................................................................-----G------------CC--A--A---T----------G---G--T----AA-A--- 5210
CPZ.CM.05.SIVcpzEK505      ..........................--------G--A...............................................................................--------T----A---------------T----C------G-TT----AT-G--G 5199
CPZ.CM.05.SIVcpzLB7        ..........................----------T-...............................................................................-----------------------------T-T-----G---G--T-----G-A--- 5182
CPZ.CM.05.SIVcpzMB66       ..........................---------CT-...............................................................................-----------G-A-----A-----A--T-T---------------------A-A--- 5173
CPZ.CM.05.SIVcpzMT145      ..........................------------...............................................................................-----G-----G--A---TCA--A-----CT-TG-GC-G---CTGT----CACA--G 5169
CPZ.CM.-.SIVcpzMB897       ..........................------------...............................................................................--G-----------------A--A------T-----------------------AT-A--- 5193
CPZ.CM.-.SIVcpzDP943       ..........................-----------A...............................................................................--------T--G--A---GCA--A------T-----------G--T----A--A--G 5213
CPZ.CM.98.CAM3             ..........................--------G--A...............................................................................--------T--G--T----CA--A------T-----------G--T----A--GC-G 5041
CPZ.CM.98.CAM5             ..........................---------ATA...............................................................................-----------G--------CA--A--A---T-----------R--------AA-A--- 5329
CPZ.GA.88.GAB1             ..........................-----------A...............................................................................-----G--------T----CA--A--A--TT--C-G------G--T-----ACC--G 5691
CPZ.GA.88.GAB2             ..........................------------...............................................................................-----------G--A---GCAG-A--A--TT-------G---G-------AA-A--G 5020
CPZ.US.85.CPZUS            ..........................----------TG...............................................................................------------------GCC--A--A---T-------G---G--T----TG-C--- 5687
CPZ.CD.90.ANT              ..........................------------...............................................................................--GC-G--TG-G-----------A--A---ATG-----G---TTGT----AAC---- 5084
CPZ.TZ.00.TAN1             ..........................-----------A...............................................................................--GA-C---A-C-----G-----A--A---T---C---G---CTGT----CA-CC-- 5289
CPZ.TZ.2001.TAN2           ..........................-----------A...............................................................................------GA-T--C-----C--A-----A---T-T---------CTGT----CACC--- 5294
CPZ.TZ.2002.TAN3_1         ..........................-----------A...............................................................................------G------C---------A-----A---T-----------CTG-----CACCC-- 5294
GOR.CM.2007.SIVgor2139_287 ..........................--------GAT-...............................................................................---AC---TG-G-------A--A--A---TT--------------T----AA-G--- 5164
GOR.CM.2007.SIVgorCP2135con..........................-----G--GAT-...............................................................................---AC---TG-G-------A--A--A---TT---C-----------T----AG-G--- 5163
GOR.CM.2004.SIVgorCP684con ..........................-----G--GAT-...............................................................................---AC---TG-G-------A------TT------------------T----AT-G--- 5083
H2A.DE.x.BEN              CTCCCCCTCCAGGTCTAGTCTA....---.............................................................ACTGAAGCACCAACAGAGTTTCCC--------TGGGACC----G------CTT-GGG-G---C---GT-A----AAC-C-G 6322
H2A.GW.x.ALI              CTCCCCCTCCAGGTCTAGTCTA....---.............................................................ACTGAAGCACCAACAGAGTTTCCC--G-CC-GAATG--A--C--CCA--GAG-CAGGG----G---GT-A----AG-CC-G 6312
H2A.SN.x.ST               CTCCCCCTCCAGGTCTAGTCTA....---.............................................................ACTGAAGCACCAACAGAGTCTCCC--G--G--TAGGACC----C------------GGGG----G---GT-A----AACCC-G 5764
H2B.CI.x.EHO              CTCCTCCCCCAGGTCTAGCCTA....---.............................................................GCAGAAGCAGTCCCAGAGATTCCT-----G--TA--AAC-----A--A--A--GTGGG-AC-G---GT-G-G--CG-CC-G 6293
H2B.GH.86.D205            CTCCTCCCCCAGGCCTGGCCTA....---.............................................................GCAGAAGCAGCCCCAGAGATCCCT-----GA--G-GAAC-----A--A--A--GTGGG-A--G----T-GGG---A-CC-G 6295
H2G.CI.92.ABT96           CTCCTCCTCCAGGTCTAGCATAG...---.......................................................................GCAGAAGAAATCCCC--------TG-G-CC--C-----A------TGGG--------GT-G-----AG--C-G 5638
H2U.FR.96.12034           CTCCCCCACCAGGATTGATCTA....---............................................................................GCAGAAATACCT-----G---GGG-CC-----A--A-----TGGGG-C--G---GTGAG---AA-AC-G 5782
SMM.SL.92.SL92B           CTCCCCCACCAGGTTTGGCCTA....---...................................................................................GAACATGCT-----G--TG-GACCAAC-CA--A--A---TGGG----G----T-AG---TG-AC-G 5655
SMM.US.x.H9               CTCCCCCTCCAGGACTAGCATAA..A---............................................................................ACAGAAAGACCT--------TG---CC-----------A--GTGGG--------GTRG----AG--C-G 5696
SMM.US.x.PGM53            CTCCCCCTCCAGGACTAGCATAAA.A---............................................................................ACAGAAAGACCT--------TG---CC-----------A--GTGGG--------GT-G--A-AG--C-G 6141
SMM.US.x.SIVsmH635F_L3    CTCCCCCTCCAGGACTAGCATAAA..---............................................................................GCAGAAAGACCT--------TG---CC-----------A---TGGG--------GT-G-G--AG--C-G 6226
                                             Vpx end   Vpr start
MAC.US.x.239              CTCCCCCTCCAGGACTAGCATAA...---............................................................................GAAGAAAGACCT------A-TG----A-----A-----A---TGGG--------GT-G-G---G--C-G 6222
Vpx                       P__P__P__P__G__L__A__*_
Vpr                                                _M__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__._ E  E  R  P  P  E  N  E  G  P  Q  R  E  P  W  D  E  W  V  V  E  V  L
MAC.US.x.251_1A11         CTCCCCCTCCAGGACTAGCATAA...---............................................................................GAAGAAAGACCT------A-TG----C---TGA-----A---TGGGG-------GT-G-G---G--C-G 6221
MAC.US.x.251_BK28         CTCCCCCTCCAGGACTAGCATAA...---............................................................................GAAGAAAGACCT------A-TG----C-----A-----A---TGGG--------GT-G-G---G--C-G 6198
MAC.US.x.EMBL_3           ......................................................................................................................................................................... 5467
STM.US.x.STM              CTCCCCCTCCAGGCTTAGCATAG..A---............................................................................ACACACAGACCT-----G--TG----C--------A--A---TGGG----G----GT-G-G---AG--C-G 5870
MNE.US.x.MNE027           CTCCCCCTCCAGGACTAGCATAA...---............................................................................GAAGAAAGACCT--------TG----C-----A-----A---TGGG--------GT-G-G---G--C-G 5690
DEB.CM.99.CM40            ..........................------AGGTA-CCACCCTCACACCCA..........................................................---C-TTT-ACCTCAAG-ACA-TACCCATGACA-GGCT-G--TTG-AGC--GCCA-G 5411
DEB.CM.99.CM5             ..........................------AGGTATCCACCCTCACACCCA..........................................................---C-TTT-AC-TCAAG-ACAGTACCAATGACA-GGCT-GCACT--A-C-AGCCA-G 5405
DEN.CD.x.CD1              ..........................----C-............................................................................................--AA-CACTC-AAGGCA-CT--AG--AGCAG-- 5534
LST.CD.88.447             ..........................---TCCAG-A-AAGG...........................................CAGCAAAGGAGTCCAGAAAAACCT-----G---TTT-------CA---------TGGGG-C-----CT-GC---CACCA-G 4873
LST.CD.88.485             ..........................---TCCAG---AAGG...........................................CAGCAAAGGAGTCCAGAAAGACCT-----G---TTT--A----CA--A------TGGGG-C-----TT-GC---CACCA-G 4873
LST.CD.88.524             ..........................---TCCAG-CAAAGA...........................................CAGCAGCGCAGCCCAGAAAGGCCC-----G---TTT--T--T-CA--A------TGGGC-C-----CT-GC---CACAA-G 4870
LST.KE.x.lho7             ..........................---TCCAGGCAGAGG...........................................CAGCAAAGAAGTCCAGAGAGACCA-----G---TTT--A----CA---------TGGGG-C-G---TT-GC---CACCA-G 5954
SUN.GA.98.L14             ..........................----CCAGC.................................................AGGAGAGAGCCAGTAGAACAGCCA---------G-G------C---A------TTTGCAC-----CTGGC---TACCA-G 5998
MND-1.GA.x.MNDGB1         ..........................---.......................................................GGCCAGAAGAGAGATGAGCAAGTAT-------T-----A--T-CC--A------T-----C-G---CT-GC---TAC-A-G 5310
OLC.CI.97.97CI12          ......................................................................................................................................................................... 5516
WRC.CI.97.97CI14          ...........................................................................................................................GGCGCCTT--GGAG-ACCT----CA-T-CC-GA-GAC--TCTAGG-A-G 5875
WRC.CI.98.98CI04          ...........................................................................................................................AGC-CCTA--GGAG-ACCT----CA-C-CC-GA-GAC--TCTAGG-A-G 5890
WRC.GM.05.Pbt_05GM_X02    ........................................................................................................................-AGG-CCCCAA--GGAG-ACCC-AGGCA-TACC-GA-GAT--TCTAGG-A-G 4810
MND-2.CM.98.CM16          CTCCTCCTCCAGGGATGGTCTAA..............................................................CCTCTGTGTTACAGATGGAGCAACCA--C--G---G-G-CC--T--A--A--A---T-----------CT-A-T--CACC--- 5902
MND-2.GA.x.M14            CTCCTCCTCCAGGATTGGCCTAA.............................................................CTTCTGTCTTGCAGATGGAACAGCCG-----G--TG-G-CT-----A--A--A---T-T-----G---CTGA----TACA--- 5829
MND-2.x.x.5440            CTCCTCCTCCAGGGATGGCTTAA.............................................................TTTCTCTGTTTTACAGATGGAACAGCCA--T--G--TG-G-CT--C-----A-GA---T-T---A-----CTGG---GTAC---G 5467
DRL.x.x.FAO               CTCCTCCTCCTCCTCCA...TGA.......................................................TTGGTCTAATAATTGTGTCTTACAGATGGAAAGGCCA--G--G--TG----A--G-CA-----A--CT-TG----G---CT-G----TAC---- 5521
MON.CM.99.L1              ..........................----TT..........................................................................................CCCAT-G-A-GAC-GGCCCT--AG-CCGCCA-T 5428
MON.NG.x.NG1              ...................................................................................................................CGT-GTTCCCAT-G-GCGC-CCGCTCTG-AG-CCGCCA-- 4039
GSN.CM.99.CN166           ..........................------AG-AT-CCACCCAGTCATCCT.......................................................ATGCCTTGG-TCA-C-G-AG-GT-CCCA--AC--TG-CTATCG-T-A-A-CGC-C-G 5420
GSN.CM.99.CN71            ..........................------AG-A--CCACCAAGTCACCCT.......................................................-T-CCTTGGATCA-C-G-AG-GT-CCCA--AC--TG-C-GTTG-T-A-AGCGCCA-G 5402
TAL.CM.00.266             ..........................------AGGTTG...............................................CCACCCAGTCATCCT-T-CCGCTGACCTCCAG--TC-TTCCAA---CA-GGCCTCA-CT-GA-AG---CA-G 5601
TAL.CM.01.8023            ..........................------AGGTTG...............................................CCACCCAGTCATCCT-TGCC-ATAAC-TCAAGG-TC-TACCAA---CA-GA-CTCAATT-GA-AGA--CA-G 5099
MUS-1.CM.01.1085          ..........................------AGCTTGCCCCCCAGCCACCCA.......................................................G-GCCCTGG--CTCCAG-GT-GTCCCAA-CAC--TGCGCCA-G---A---AGCCA-G 5376
MUS-1.CM.01.CM1239        .......................................................................................................................-TGG--CTCCAG-GT-GTCCCAA-TAC--TGC-GCA-G-C-A-C--GCCA-G 5366
MUS-2.CM.01.CM1246        .......................................................................................................................-TGG-T-A-T-GT--A-C-CCAA--ACT-TG-TTCA-G-G-A--CCGCCC-G 5483
MUS-2.CM.01.CM2500        .......................................................................................................................-TGG-T-A-C-GC---GTACCAA--AC--TG-CGCA-G---A-A-CGCCA-G 5436
SYK.KE.x.KE51             .........................G----C---G--TTTC..........................TTCAGAGTTTACCAGCGGCAGCCAGGGGAACCACTATTTTACATTCC-AGAA-T-TCAATTG-GAC---G--GGAGTAAT-A-TG-CACT-CCAAAGC- 5472
SYK.KE.x.SYK173           ..........................----CGG-G--TTTC...........................TTCAATCCCTCCCAGCATGTGCAAGGAACACCATGGTTCTTCATTCCAAG-AATGT-G--TT-ACC---A--GTGATCAATGTCACT-TAAAGGC- 5811
RCM.GA.x.GAB1             CCATCCTACCGGGGATG...TAA................................................TGATGGCCCTCTCTTTACAGAGATGGAGCTGCCC--T--G--TG----T--G--AC-A--A--TT-TG--------TT-A-G--TACCC-- 5418
RCM.NG.x.NG411            AGATTCTTCCGGGAATG...TAA..............................................TCAAGCACTCCTCTCATTGCAGAGATGGAGATGCTC--G-----TG---AA--------A--A--CT-TG--------CT-A-G--TACCC-T 5453
GRV.ET.x.GRI_677          ..........................----C-TC--GAAGAGATCCAAGAGAA...............................CCTTTACCAGGATGG-TG---AT-TGG-AT-T-G-C---------TGGG-C------CT--A---CA-GC-C 5781
SAB.SN.x.SAB1C            ..........................----CCTC--GAGGGTGGCTCCCACCAGTAGGAGGGGATCCTCCCAAGGATCCCCCCAAGAATCCCAGAGAAGAGATACCAGGATGG-T----ACATGG-ATTTG-CC--A-----GTT-G-C------CTC-G---CA-G--- 6054
TAN.UG.x.TAN1             ..........................----C-G---GAAGA...................................GATTCCAGGGAAAGAAGACCCGGATGGTT----ATTTGG-AT-TCAGC--------GTGGG-C-----ATT-AG---CA-GG-- 5837
VER.DE.x.AGM3             ..........................----CCTC--G-AGA...................................GATCCCAGAGAAGCAAGACCGGGGGAA-T----ATATGG-ATTTGAGC--------GTGGG-C------CT-AG---CA-GC-- 5342
VER.KE.x.9063             ..........................----C-TC--GGAGA...................................GGTCCCAGAGAAAACAGGCCAGGAGAAGT---GAT-TGG-AT-T-A-C--A--A---TGGG-C------CT-AGG--CA-G--- 5844
VER.KE.x.AGM155           ..........................----CCTC--GAAGA...................................GATCCCAGAGAGGAGAGACCAGGAGGA-T----ATTTGG-AT-TCAGC--------GTGGG-C------CTCAG---TA-GG-- 5839
VER.KE.x.TYO1_patent      ..........................----CCTC--GGAGA...................................GATCCAAGAGAGGCAAGACCAGGAGAAGT----ATTTGG-AT-T-AGC--------GTGGG-C------CT-AG---TA-G--- 5844
COL.CM.x.CGU1             ..........................---AG-GTCCTG........................................................................A-AGGGACTTCAC-.---TTG-AG--C-AGG-- 5324
```

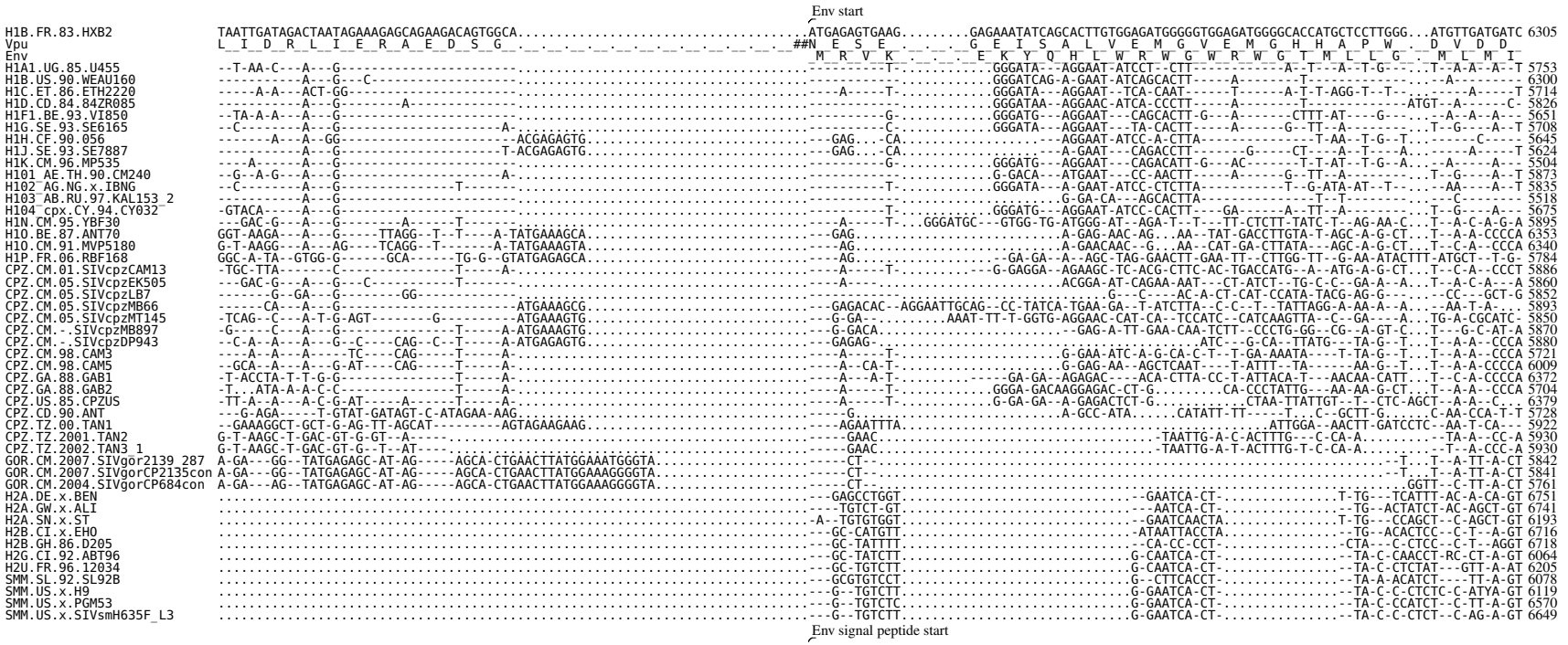
```
H1B.FR.83.HXB2            GAGGAGCTTAAGAATGAAGCTGTTAGACATTTTCCTAGGATTTGGCTCCATGGCTTAGGGCAACATATCTATGAAACTTATGGGGAT.......................ACTTGGGCAGGAGTGGAAGCCATAATAAGAATTCTGCAACAACTGCTGTTTATCCAT.. 5771
Vpr                        E  E  L  K  N  E  A  V  R  H  F  P  R  I  W  L  H  G  L  G  Q  H  I  Y  E  T  Y  G  D  .  .  .  .  .  .  .  .  T  W  A  G  V  E  A  I  I  R  I  L  Q  Q  L  L  F  I  H ##
H1A1.UG.85.U455           --A---------C------A--------------------CAG-----------A------A------------A-C--C---------.......................--------AG-----T-----T-------------T--------T-A--------T---.. 5217
H1B.US.90.WEAU160         -----A--------G-------------------------G-A-----------A------A---------------------------.......................--------A-----------------------------C---------------T---.. 5770
H1C.ET.86.ETH2220         --------C---C-G--------C-----C---------ACCA----------AA------A---T------------C----------.......................--------T-G-----A-----TT---------C-------------AA----------.. 5163
H1D.CD.84.84ZR085         --------------G--------C--------------------A-----T---A------A---T----------------------.......................---C------------T------T-------------------------A-------T---.. 5293
H1F1.BE.93.VI850          -------------------------G------------CCA-----T--C--------C------------A-C--C--------C.......................--------AG-----T-----T--------G--AT---------A-------------.. 5111
H1G.SE.93.SE6165          --A---------A-------------------------C-------T-----------A--G----------A-C--------------.......................--------A------T-----------------A--A--------A-------------.. 5168
H1H.CF.90.056             ------A----------G--------G-----------AG-A----------CAA------A--G----------A-C--C-----A---.......................--------T------T-----TT--------CG-----------A-------T---.. 5118
H1J.SE.93.SE7887          ------------A-----------------C-------CCA-----T--A-T------A--GT---------AGC---------A---.......................--A----A------C---------------------A-----------AT-------T---.. 5085
H1K.CM.96.MP535           -------------GA-----------------------CCA----------AATC-G--A--G----------AC--------------.......................--------A-----C-T-----T--------------AT----------A-------T---.. 4967
H101_AE.TH.90.CM240       ------------A-------------------------CCC----------A-------A--GT-C------A-C-A-----------.......................--------A---G--T-----T--------------GT----------A------G-T---.. 5339
H102_AG.NG.x.IBNG         --A---------AC----------------C-------CCG-----------A------A---------------------A------.......................--------A---G--TA----T-----------------T----------AT-----G-T---.. 5296
H103_AB.RU.97.KAL153_2    --------------G---------------C-------G-G-----T---A-------A---T-C----------------A------.......................----------------------------------------------------------T---.. 4997
H104_cpx.CY.94.CY032      ------------A-----------G-------------ACCC------------C----A--G---------A-C---------A---.......................--C----A---G--T-----T----------------T-----------A-------T---.. 5137
H1N.CM.95.YBF30           --A--AT-A--A-----------GC-C--------A-----------A-----G------A-----C-----A-C--A------A--C.......................--C----AG--G--A--G--A--T--C--G--A--A------T-A-------------.. 5365
H1O.BE.87.ANT70           --A-----A--AGCA------A--A---------C-------CC------A--C-C---G--A---T--C--T-----G----------C.......................--------T-------TATG--A--T----------CT-A------T----------C---.. 5826
H1O.CM.91.MVP5180         --A-----G--AG-A------A--A---------C-------CC-----T-A--A-C--GT------GT-C--T------G--------A--C.......................--------A-------TATG--A--T----------CT-A------------A-------C---.. 5801
H1P.FR.06.RBF168          -----A--C---GCG------A--A-----C--C----TACAA---T-A--CTCA------A------C----A-T--C--------C.......................--C-----AG-----AACT--A--T--T--G--C--A-------T-AA-C---------.. 5232
CPZ.CM.01.SIVcpzCAM13     --A--CT-A--------------A-A---C-A---C--A-G------T--CCAA------C--G-----T---ACC--C--------C.......................--G-----A---G-----G--A--T----------C--A--------------A-------G-T---.. 5354
CPZ.CM.05.SIVcpzEK505     --A--C--A---------G---C-GC-C--C-----C--CC------A------AC-------T-CT------A-T--A-----A--C.......................--C-----AG-----A--G-----C--T--G-CA--A-----------T-------A---.. 5343
CPZ.CM.05.SIVcpzLB7       --A--A--C--AC-G------A--A-----C-------ACCA---------CCAA--G--A---T----A----C--C------A--C.......................--C-----AG-----TC----T--C--------------C---------------C------T---.. 5326
CPZ.CM.05.SIVcpzMB66      --A--A-----A-GG------A--A-----C--C--C-AGA----------CAGC----A--G----------ACC----------A--C.......................----------------G--G--T------------------C-----------------T-------T---.. 5317
CPZ.CM.05.SIVcpzMT145     --A--AA----AC-A------A--A---------C-----C--ACCC---T-A----CAC-----AG--------T--------A-----A--C.......................T-C------G------AC----T-----------------T---------T-A--------G------.. 5313
CPZ.CM.-.SIVcpzMB897      --T--A--C--AC-G------A--A-----------------ACAG-----T----AT------A--G--C--T--A-C--A-----A--C.......................------G-G--T--G--T--C-----G--C--------T-------------T-------T---.. 5337
CPZ.CM.-.SIVcpzDP943      -----A--G--A-----G------AC---C------C-AGCA----T-A---------A------A------C--A------C--A-----A--C.......................--A--------T--A-----T--C--T----A---------G--CT--------T---.. 5357
CPZ.CM.98.CAM3            -----A--A------------------AC-C-------C-AGCA-------A-----A------A--GTT---T-----C-------A---.......................--A--------T--A-----A--C--T----A---------G--T-------------.. 5185
CPZ.CM.98.CAM5            --A-----A--A-C---G---C--A---------C--C-CC------A--A-AC------A--GT-C--------C---------A---.......................--A--AT----G--A-----A--T--T----AT--------T--T------G------.. 5473
CPZ.GA.88.GAB1            --A--AT-A--A-------A--A-------C-------ACCA---T-G---CAA------A--TTC--T-----C---------A---.......................--A------T-----A--G-----C-----------CT-A-------CT-------------.. 5835
CPZ.GA.88.GAB2            -----CT-A--A--C--G----C---------------ACCA-----A---CAG------A--TTA-----CA-T-A------A---.......................--A-----A-----A--G------C-----A--T------CT----------G-G------.. 5164
CPZ.US.85.CPZUS           --A--A--A--AG-A--G---T-GC-C-----------ACCC------A--A---------A--GT-C-----CAGT-----------A--C.......................--------AG------A------A--T------G--CT-A---------TT------------.. 5831
CPZ.CD.90.ANT             --A--AA-A--A--------G--A--A-----------C--ACA-CC-ACAT-A--ACA-C-T--AA-TTGGG-A-----C--A--C--A---.......................T-A----AG-------C-G-TTT-G----TTC-C--T---A--GCT--------CA-----.. 5228
CPZ.TZ.00.TAN1            --A--AA-A--AC-A------A--G-A---C-----A--ACCAATAT-A--G--GG----AA-TTGGG---TCACC-T-------A--C.......................T-C-----AG------AC-G-AAT-----C-AG--CT-----GAG-GCTT------C---C.. 5433
CPZ.TZ.2001.TAN2          --A-----A--A---C--A--------AA-G-----------A--ATCAGTAT-A--G--AG----AA-TTGGG-T----CT-T-C-----A--C.......................T-C-----A-------AC---AAT-----------C-----GAG-GCTA-C-----CT--C.. 5438
CPZ.TZ.2002.TAN3_1        --A-----A--A--C--A------A-------------A--ACCAATA--G--G--AG----AA-TTGGG-T-T-ACC-T-C-----A--C.......................T-C-----A-------AC---AAT-----------CT----GAG-GCTT-A--C-CT--C.. 5438
GOR.CM.2007.SIVgor2139_287  --A--A-----C-A-----A--A-----C---------CCC---T-G--CCAA--G--A--G---G-T-----T---------A--C.......................--------A--------AC---A------C-----CT-A---G---G-AA-C--G-T---.. 5308
GOR.CM.2007.SIVgorCP2135con --A--R-----C-A-----A--A-----C---------CCC---T-G--CCAA--G--A--G---G-T-----T---------A--C.......................--------A--------AC---A------C-----CT-A---G---G-AA-C--G-T---.. 5307
GOR.CM.2004.SIVgorCP684con  --A--A--A--C-A-----A--A-----C---------CCC---T-G--CCAA--G--A--G---CG-T-----T---------A--C.......................--------A--------AC---A------C-----CT-A---G---G-AA-C--G-T---.. 5227
H2A.DE.x.BEN              AG---AA-A---G-A-----CT-A---------GA-CCCCGC-T---AAT--CTC-T--CT-CT------C--A-T-GAC-----A--C.......................--CCTT-A---C-CCAG--AGC-C--T-A--CC--A----G-GCC--C--CG-G--C.. 6466
H2A.GW.x.ALI              AGA--AA-A--AG-A------T-A--G------GACCCTCGCAT---AAT--CTC-T--CGGCT--------ACT-GAC-----A--C.......................--CCTT-A-A---CCAGG-AGC-C--T-ATGCC--A----G-GCC--C--C--G--C.. 6456
H2A.SN.x.ST               AGA--A-A--AT-A-------T-A-AG--C---GACCCTCGC-T---AAT-ACTC-T--CA-CT----------CT-GAC-----A--C.......................--CCTT-A---C-CCAG--GGC-C--T-A--G--C--A----G-GCC--CC-CT-G--C.. 5908
H2B.CI.x.EHO              -----AA-A--AC-A------CT-A-AG-----GACCCCCGC-TAT-AAC--CAC-T--CA-TTT------CA-T-GGC-----AA---.......................--CCTT-A-----CA-G--AGC-C--T-A-C-C--C--G-GGGCA--C--CC----C.. 6437
H2B.GH.86.D205            -----AA-A---C-A------CT-A-AG-----GA-CCTCGC-T---AAC--CGC-T--TA-CTT------CAGT-GGC-----A--C.......................--CCTT---------CA-G--AGC-C--T-A---C--C----G-GCC--C--CC----C.. 6439
H2G.CI.92.ABT96           --A--AA-C--AG-A------CT-A-A------GACCCTCGC-T---AAC--CGC-T--TA-CT----------T-GGC-----A--C.......................--CCTT-A-----CA-G--AGC-C--T-A---C------GAG-GCT--C--CC-G--C.. 5782
H2U.FR.96.12034           -----A--A--AG-A---------A--G-----GA-CCTCGC-T---AAC--CGC-T--TA-CT--------A-T-GGC-----A---.......................--CCTT-A-----CA-G--AGC-C--T-A---C-------AG-GCAA-C--CC----C.. 5926
SMM.SL.92.SL92B           --A--AT-G--AG-A--G--AC-G-A-------GACCCTCGC-T---TAC--CT-----AA-TT-CG-A------T------C-----.......................--AATA-A---G-CA-GG-AA-----T-AG--AT-A---A-GGCAT-----CT-A--C.. 5799
SMM.US.x.H9               -----A-KA--AG-A------CC-G-AC--K--GA-CCTCGC-T---AATY-CGC-T--TA-CT----------T-GAC-----A--C.......................--CCTT-A-----CA-G--AGC-C--T-----C--C--GGG-GCA--C--C---------.. 5840
SMM.US.x.PGM53            -----AG-A--AG-A-------C-G-A---C--GA-CCTCGC-T---AAC--CGC-T--TA-CT----------T-GAC-----A--C.......................--CCTT-A-----CA-G--AAC-C--T--G--C--C----G-GC---C--C--A--C.. 6285
SMM.US.x.SIVsmH635F_L3    -----AA-A--AG-A------CC-G-A------GA-CCTCGC-T---AAC--CGC-T--TA-CT-----T------T-GAC-----A--C.......................--CCTT-A-----CA-G--AAC-C--T-----C--A--GAG-GCA--C--C---------.. 6370
                                                                                                      Tat exon 1 start
MAC.US.x.239              --A--A--G--AG-A-------T-A-A------GA-CCTCGC-T---AAC--CAC-T--TA-T---------A-T-GAC-----A--C.......................--CCTT-AG----CA-G--AAC-C--T-----C--C----G-GC---C--C--G----.. 6366
Vpr                        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  .  .  .  .  .  .  .  .  T  L  E  G  A  G  E  L  I  R  I  L  Q  R  A  L  F  M  H  .
Tat exon 1                                                                                          M  E  T  .  .  .  .  .  .  .  .  P  L  R  E  Q  E  N  S  L  E  S  S  N  E  R  S  S  C  I
MAC.US.x.251_1A11         --A--A--G---G-A-G----T-A-A-------GA-CCTCGC-T---AAC--CAC-T--TA-T---------A-T-GAC-----A--C.......................--CCTT-AG----CA-G--AAC-C--T-----C--C--G-GGGCA--C--C--G--C.. 6365
MAC.US.x.251_BK28         --A--AT-G--AG-A------T-A-A-------GA-CCTCGC-T---AAC--CAC-T--TA-T---------A-T-GAC-----A--C.......................--CCTT-AG----CA-G--AAC-C--T-----C--C----G-GC---C--C--G--C.. 6342
MAC.US.x.EMBL_3           ........................................................................................................................................--C-----G-GC---C--C--G----.. 5491
STM.US.x.STM              -----AA-A--AC-A-----CC-GC-G------GA-CCTCGC-T---AAGC-CGC-T--TA-CT----------T-GAC-----A--C.......................--CCTT-A-----CA-G--AGC-C--T-AG--C--C--G-G-GCC--C--C--G--C.. 6014
MNE.US.x.MNE027           -----A--G--AG-A------T-A-A-------GA-CCTCGC-T---AAC--CGC-T--TA-TT----------T-GAC-----A--C.......................--CCTT-AG----CA-G--AAC-C--T-A---C--C----GGGC---C--C--G--C.. 5834
DEB.CM.99.CM40            C-A--T--A--CG-G------CC-A-A------CA-C--AGAAGAA--ATT---AG-TT--A-T---TGTGTA--CTTACCC-CCC-ACCAGATTGGACGGGAGAACAAGCATGGGCT---TCG--CATT-A-TAC-----A---AG-T---A--A--T-A-GGT-G--C.. 5579
DEB.CM.99.CM5             C----T--A--CG-G------T-A-A---C---CA-A--AGAAGA-T-ATGG--AG-CT--A-CT---CTGTGTA--CCT-CC-CCC-ACCAGATTGGACAGGAGAACAAGCATGGGC---GTC---AATT-A-TAT-----A---AG-A---AG-A----T-GGC-T---.. 5573
DEN.CD.x.CD1              TG---TA-C-CAG-G--G--AAGA-A---C---CAGC-AAGAAGA-A--AC---AA--T--G-T---CTG--GGTC-TTACCA-C-CTCCCACATTGGACAGAAGGACAAGTC-TGGCA--C-CC--CAT--ATT-C---------CA-----GA--GAAA-A-GG-AG---.. 5702
LST.CD.88.447             -----AA-A--AG-G--G--CAGAC-C------C--AGAA-CAATTT-GGCCCAAG-CTCAG--T-CTGTGTC.......................AGCAGTACCGGGTCAGAGGA-AA--CCTGCATGAAAT-C--T-CC--AG-AA-TAG-GCA--C-----T---.. 5017
LST.CD.88.485             -----AA-A--G-G--G--CAGAC-C------C--AGAAGCCATTT-GGCCCAAG-CTC-G--T-CTGTGTC.......................AGCAGTACCGGGTCAGAGGA--A--CCTGCATGAGAT-C--T-CC--AG-AA-TAG-GCA--C-----T---.. 5017
LST.CD.88.524             -----AA-A--C--G-A--G--AAGAC-C------C--AGAAGCAATTT-GGCCCAAG-CTC-G--T-CTGTGTC.......................AGCAGTACCGGGTCAGAAGA--AG-CTTGCATGAGAT-C--C-CTT-AG-TA-TAG-GCT--C-A-T-A---.. 5014
LST.KE.x.lho7             -----AA-A--AG-G--G--AAGAC-C------C--AGAAG-AAT-T-GGCCCAAG-CTC----T--TGTGTA--C-G-ACA---.......................TCAGAAGA--AG-CCTGCATGAAAT-C--C-CCC-AG-CA-TAG-GCTT-----C-T--C.. 6098
SUN.GA.98.L14             --A--AA-C-G-G-C-----G-AA--G------C--AGCAGCAATAG-TGCCCAAG-CTT----T-CTGTGCG.......................GACTCTACAGGGTCAGAAAA--A---GTGTATGAGAGC---C-CTT-G--CA-TAG-GCCT----C--T---.. 6142
MND-1.GA.x.MNDGB1         -----AA-A---G-A-----AAGA-AG--C--C---CTC---ATC--AA---CAG--TCAG--T--TGTGTGC---ACACA---AG-GAG.......................GAAGA---CT-T-A-A--TTT--T-CC..T-AA--A-TAG-GCCA-T-GGG----C.. 5454
OLC.CI.97.97CI12          ..............................................................................................................................................................................AT 5518
WRC.CI.97.97CI14          C-A-----CTTTGCAATG--AAAACAG-----C--A---GAAGTC--TGTAA-AG--TTT-TTT--TG--C-C-.......................CAGAGAGAGTATGAGGAAGTCC-TACTA--C-TTATT-CCAC--GG-A-------AG-GG---C--C--G---.. 6022
WRC.CI.98.98CI04          C-A--A--CTTTGCAATG--AAAACTG-----C--A---GAAGTC--TGTAA-AG--TTT-TCT-CTG--CGC-.......................CAGAGAAATTATGAAGAAG-TC-CACTA--C-TTATT-TCAC--GG-A--A---AG-GG---C--C--G--C.. 6037
WRC.GM.05.Pbt_05GM_X02    C-A-----GTTTGCCATG--AAAAC-G-----C--CC--GAAGTC--TGTAA-AG--TTT-TCT--TG--CGC-.......................CAGAGAAATTATGAAGAACA---CACTA--C-CTATT-TCAC--GG-C--T---AG-GGAT-A-A-C-A--C.. 4957
MND-2.CM.98.CM16          -CA--AG--C-AG-A------T-A-AG------GA---ACGC-TAT-G----CAG----AAGCTGG-----CA------C-------.......................--A-TA-A---G---CG-CAGC-T--CTC-------T---A--GCT-----C---G--C.. 6046
MND-2.GA.x.M14            -CA--AA--C-AG-A------T-G-AG------GA---CGC-TAT-G----CAG----AAGCTGG--T---------A--------C.......................--C-T--AG--G---C-GAAGC-T--C-CCC-CT-A--GAG-GCT--A---C-G--C.. 5973
MND-2.x.x.5440            -CA--AA-AC-AG-G------T-A-AG------CGA---CGC-T-----A----CAG----CTC-TGGG-G-------GCAAC-A--A--C.......................--TA-A----T--CC--CAGC-------C----T-----AG-GCA-----CT-A--C.. 5611
DRL.x.x.FAO               -TA--AA-AC-AG-G------T-G-AG------GA---CGC-TAT-G----AATG----TC-TGGG--------GA------A---.......................T-A-TA-A----------C-CCAGC-T----T-C-CT-----AG-GCA--T--CT-G---.. 5665
MON.CM.99.L1              TG-----A--CG-G-----CC-A-A---C--CT-A---GAAGA-----G-----A-TT--G----GG-AACG--GCTGCC--CT---CCCGGCCTGGAATGCAGACCAGGCATGGGCA---T-T-CCATT-ATTAT-CT---TGGG-C----ACCA-C--T-A--GG--C.. 5596
MON.NG.x.NG1              TG---C--A--CG-G--G--CC---TYG--C--CT-A-A-GAAGA-----G-----A-CT--G-T--GG--AC----TTGCC---CC---CCCAGGGTGGAACGCAGACCAGGCATGGGCA---T-Y-CTATT-ATTAC-CC---GCGGG------AC-A-RA-A-----GG--C.. 4207
GSN.CM.99.CN166           TG---AA-A--CG-G--G--A-AG-A---C--CT-A---GACGA-----G----AA-TT----TG--G-GACG--GCTCCC---CC---CCAGATTGGACAGTGGATCAGGCAG-AATA----T-T-CAAT--ATTAC--C...-GGG-----AC---T--C-----GA--C.. 5585
GSN.CM.99.CN71            TG---AA-A--CG-A--G--A-AA-A---C--CT-A---GAAGA-----G----AA-TT----TG--G-AACG--GCTCCC---CC---CCGGATTGGACAGTGGATCAGGCAG-AATA----T-T-CAAT--ATTATG-C--G-GAG-----GAC-T-A--C-----GA--C.. 5570
TAL.CM.00.266             AGA--AG--CGAG-C---ATGC--T---C--CA-A---AGACGA-G--AT---AG-CT--A-T---TGTGTA---TTGCCA-CA---CCCTGATTGGACAGGAGAACAAGCATGGGCT--CTC---CAT--ATT-TGCT---G-AGGGTA-T---A--T-AC-CC-A---.. 5769
TAL.CM.01.8023            AGA--AG-A-G-G-C---GATGC-C-T---C--CA-A---AGAAGA-G--AT---AG-CT--A-T---CTGTGTA--GTTACCA-CA---GCCTGATTGGACAGGCGAACAAGCATGGGCT---TTC---CAT--ATT-TGCTC-G-AGGGTA-CG-GA--T-A---CC-A---.. 5267
MUS-1.CM.01.1085          TG---AG-A--CG-A--G--A-AA-A---C---T-C--AGAAGA-----G----AA-TT--A-TG-GG-AACG--GCTGCCA-CC---CCTGATTGGACAGTGGATCAGGCAG-AATA----T-T-CCAT--A-TATG-T--G-GAG-----AC---TG-C-----GG---.. 5544
MUS-1.CM.01.CM1239        TG---C--A--CG-A--G--A-AA-A---C---T-C--AGAAGA-----G----AA-CT--A-TG--G-AACG--GCTACCA-CC---CCTAATTGGACAGTGGATCAGGCTG-AATA----T-T-CCAT--ATTAC--T--G-GAACA---AC---T--C-G-GG---.. 5534
MUS-2.CM.01.CM1246        TG---A--A--CG-A--G--A-AA-A---C--CT-C--AGACAA-----G----AA-TT--G-TAGAG--ACA--GCTCCCA-CA---CCACAATGGTCTGTAGATCAGGCAG-CATAA---T-T-CCATT-ATTAC--T-A-TCAG-C---ACC--CT-----C-GA--C.. 5651
MUS-2.CM.01.CM2500        TG---T--A--CG-A--G--A-AA-AG--C--CT-C--GACGA-----G----GA-TT--G-C-GAG--ACG--GCTCCCA-CA---CCCTAATTGGACTGTTGATCAGGCAG--ATA----T-T-CCAT--ATTAC--T---TGGG-A--GGGT--C--C--C--GG---.. 5604
SYK.KE.x.KE51             --AC--G-AGCCC-A-----ATGC-AG------A-AGC-GAAGAAA-ATT---AG-TT--A-T---ATG--TA--GGTGG-A-CA-G-CCAGAAGAAACACCCACCATGGCTTGGATTAG-T-CA--CT--AT--GCAT----GCA--CA-CTTTA-----T---GAA---.. 5640
SYK.KE.x.SYK173           --ACT-G-AGT--CA-----ATCC-AG------A--CCTCAAGAAA-AT----AG-TT--A-T---ATC--TAA-CGAAG-G-CA-GGACTGAT...TCCCCGACAATGGCTTGGGAAAG-ACCA--TTG-A---GG-----GCA--CA-T-TTA-----C---GAA---.. 5976
RCM.GA.x.GAB1             ATA---T-GC-AG-A------AAG-A-------A-ATATGC--T---AACGCAAA----AG-TT--G-A------GCAAC-----------.......................--GCATA-AG------C-G--A--G--TC-GC-G--T---AG-GCCT------C-T---.. 5562
RCM.NG.x.NG411            ATA---A--AC-AG-A------AAG-A-------A-ACCTGAA-T---TACACAAGA---AA-TT----A-----GCAGC-----A---.......................--G--TA-AG------AA----T--G-----CCT-G---TA-TA--GCA--T--CT-G--C.. 5597
GRV.ET.x.GRI_677          AG---T--A--CG-A------CAGA--G--C--GGA-T--ACAT---AATCC-AG--T--A-TT-CTGTGTA---GGAGGGAA--AGA............CATAATACCCCATGGAAT-AGAT--GCT-CAAGTACTAT------G-T----A-GTCTA------G-A---.. 5937
SAB.SN.x.SAB1C            C----C--C--TTCA--G--CCAGT-C--C--C--A---A-CTC--TTTCC-TC-TT--TGGA----TGTG---GAACCA-CCAT-.........GATCATGGGCAAACC-GGCT--A---CTG-T-TAAATATTGT-------------A-GGCTT-A---G-G---.. 6213
TAN.UG.x.TAN1             -CA--A--C--TC-G-----ACAAC-C------GGG--AGAACTC---TTCCAAG--T--A-CTTCTGTC-G---GAAGGG-AAGG.........AATGGGGCACCA-TGATA-A-A-G-CTT-TAGATATTAT--GT-GG-A--GA-GGCTT-A---G-G---.. 5993
VER.DE.x.AGM3             --A--TA-C--TC-G-----CAAG-TG------GGGC-CGAGCTC--GTTCCAAG--T--A-CT--TGTC-G--GGAGGA-AAAGG.........AATCGCACTCCC-TGCTA-A-A-G-CTT-TAAATATTAT-A-T-GG------A---GCT--C---G-G---.. 5498
VER.KE.x.9063             --A--T--C--TC-G-----CAAATT-------GGC--AGAGCTC---TTCCAAG-TT--A-TT--TGTC-G--GGAGGA-A-AGA.........CGGGGAGCCCCA-TGAC--AGA-G-CCT-TAGATACTAT--GT-GG-T--GCA--C---G-T---.. 6000
VER.KE.x.AGM155           --A--AA-C--C--C-----CAAACTG------GGCC-AGAACTC--AT-CCAAG--T--A-TT--TGTC-G--GGAGGGG-A-AGA.........CAGGGAAGACCC-TAGC--A-A-G-CAT-TAAGTATTATC-CT-AG-T--GA--GCT--C---G-G---.. 5995
VER.KE.x.TYO1_patent      C-A--T--C--CC-G--G--GAGGCTG------GGG--AGAACTC--GTTCCAAG--T--A-CT--CTGTC-G--GGAGGA-AAAGA.........CATGGTACTCCC-TGAT--A-A-G-CCT-CAAATATTAT--GC-AG-A---A-GGCT--C---G-G---.. 6000
COL.CM.x.CGU1             ---A-CA-A-CCCGG--G--CAGA--G--C--C---TT--C-CACA--TCA-TGA--AT-TGGGG-TGTAC-AG--GGC-CT---GGAGA.......................C-G---A-G-CT-GTCTG--TTATG-G------CA--G-GTGTGCC--C-ACT-G--C.. 5471
```

```
                          frameshift insert in HXB2
                          Vpr premature end (HXB2 only)                                           Tat exon 1 start  Vpr end
H1B.FR.83.HXB2            TTTCAGAATTGGGTGTCGACATAGCAGAATAGGCGTTACTCGA..............CAGAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAG.................................... 5850
Vpr (frameshifted)         _F__R__I__G__C__R__H__S__R__I__G__V__T__R__.__.__.__.__.__Q__R__R__A__R__N__G__A__S__R__S__*
Tat exon 1                                                                                     _M__E__P__V__D__P__R##__.__.__.__.__.__.__.__.__.__.__.__.__.__
Vpr                       _F__Q__N__W__V__S__T__*_
H1A1.UG.85.U455           .--------------C-A----------------A---T--C-..............GG---A----GC--G-------------------A.................................... 5295
H1B.US.90.WEAU160         .---------------------------------A------C-..............-----A-------------------------A---.................................... 5848
H1C.ET.86.ETH2220         .-----------A--C-AG---------------A--TTA---..............-----A-----------------------------A.................................... 5241
H1D.CD.84.84ZR085         .----------------A----------------A------C-..............-----A--------------T-------------A.................................... 5371
H1F1.BE.93.VI850          .--------------C--T---------------A--GT--C-..............-----A----T-------------------------.................................... 5189
H1G.SE.93.SE6165          .--------------C-A----------------A----A-C-..............-G---A--G-T---GG-----C--G----------A.................................... 5246
H1H.CF.90.056             .--------------C-A---------------AA--------..............-----A----T----------C-------------A.................................... 5196
H1J.SE.93.SE7887          .--------C-----C-AT---------------A---TA-C-..............--A--A----GG--G--------------------A.................................... 5163
H1K.CM.96.MP535           .--------------C-AC---------------A---TA-C-..............-----A----GG--------T--------------A.................................... 5045
H101_AE.TH.90.CM240       .----------------A----------------A---TG-C-..............GG---A----GC--G------A-TG----------A.................................... 5417
H102_AG.NG.x.IBNG         .----------------A----------------A---T--A-..............GG---A----GC--G--------TG------------.................................... 5374
H103_AB.RU.97.KAL153_2    .-----G----------A----------------A---T--A-..............-G-----------------------------------.................................... 5075
H104_cpx.CY.94.CY032      .--------------C-A------T---------A------CT..............--A------AG-G-C-GG---TGG...--CCAG---.................................... 5212
H1N.CM.95.YBF30           .-AT--G-----C--C-AG--C-----------GA-C----CT..............--A...-GAG---G------A---------------.................................... 5440
H1O.BE.87.ANT70           .-AT--------A--C-A------T--------AA---AC-C-...........AGAGGA--AG--AG---------T------------GA.................................... 5907
H1O.CM.91.MVP5180         .-AT--------A--C-A------T--------AA--CTC-C-...TCTAACACAAGAGGA--AG--AG---------T------------GA.................................... 5891
H1P.FR.06.RBF168          .-AT--------A----AT-----T--------TA-ATTG-C-CCCTCT.........-GA--AG--AGG---C-----C-----------GA.................................... 5316
CPZ.CM.01.SIVcpzCAM13     .--T---G--------C-AC-------------TA-A-TA-C-..............--A------AG-GC-----------A----C---GA.................................... 5432
CPZ.CM.05.SIVcpzEK505     .-AT--G-----C----A---------G-----AA-C----CT..............--A...-GAG---G--------------------GA.................................... 5418
CPZ.CM.05.SIVcpzLB7       .-A---------C----A----------------A----A-C-..............--A--A---AGG---G---------A----------.................................... 5404
CPZ.CM.05.SIVcpzMB66      .-A--------------AT---------------A-C-A-ATT.........AATCAG-----A---AG---------------A--------A.................................... 5401
CPZ.CM.05.SIVcpzMT145     .-----------C--C-AT---------------A-A-AC-CT..............------A---AGG---G-----A---A----G---GA.................................... 5391
CPZ.CM.-.SIVcpzMB897      .-A---------C--C-A----------------A--CTG-C-..............--AG-A---AGG---------T---A---------A.................................... 5415
CPZ.CM.-.SIVcpzDP943      .-AT--G-----C----A---------------GA---AC-CT...............G----GAGG-----------------------GA.................................... 5432
CPZ.CM.98.CAM3            .--T--GC----C----A---------------GA-C-TC-CT..............GG-...--GAGG-----------------------GA.................................... 5260
CPZ.CM.98.CAM5            .--T--GC----C----A---------------GA--CT--YT..............G--...---AGG---------A----------Y--A.................................... 5548
CPZ.GA.88.GAB1            .--T---C--------C-AG---------------A-CCTC-C-..............--A--A---AG-TCC------T---A--------GA.................................... 5913
CPZ.GA.88.GAB2            .--T--G-----C----A---------G-----GA--GAA---.............----------AG---G-GA-.......................................ATGGATCCTA 5235
CPZ.US.85.CPZUS           .-AT--G-----C---AAT--------G-----GA--T-CTTGGCAAGGAGAACTCCT--A...G-GAG---G------T------------A.................................... 5921
CPZ.CD.90.ANT             .-AT---CAC--A---GCT--------G-----ACA-GGA-CC.........................................C-----CG--GA.................................... 5282
CPZ.TZ.00.TAN1            .-ATC-CCA---T---ATC--C-----------ATCA.......................................................................TGAATCCCA 5478
CPZ.TZ.2001.TAN2          .-ATC--CA---T---GTT--C-----G-----ATCA..................................................GA...........................ATCCTA 5483
CPZ.TZ.2002.TAN3_1        .--TC--CA---T---GTT--C-----G-----ATCA..................................................GA...........................ATCCTA 5483
GOR.CM.2007.SIVgor2139_287 .-AT-------------AT--------------GA--TTG-C-.........CCCTCTAGA--AG--AG---------C---AC------GA.................................... 5392
GOR.CM.2007.SIVgorCP2135con .-AT------------C-AT--------------GA--TTG-C-.........CCCTCTAGA--AG--AG---------C---AC------GA.................................... 5391
GOR.CM.2004.SIVgorCP684con .-AT-------------AT--------------GA--TTG-C-.........CCCTCTAGA--AG--AG---------C---AC------GA.................................... 5311
H2A.DE.x.BEN              .------GCG--A---AAC-GCTCA-----T---TAA--AA-GAGAAGA.........ACTCCTT-CC--GCTGCACCGA--CC----G--AT-CACTAA............................ 6556
H2A.GW.x.ALI              .------GCA--A---G-C-G-TCA---G-T---CAA--AA-GGGAAGA.........A-TCCTCTCT--GCT-TACCGA--CC----AA-AT-CAATAA............................ 6546
H2A.SN.x.ST               .------GCA--A---CG-C-GCTCA--G--T--TCAGC-CA-GGGACGA.........A-TCCTTT-T--GCT-TACC-A--CC----GG-AT-CGATAA............................ 5998
H2B.CI.x.EHO              .------GGG--T--C-A---CTCA--G--T--ACAAC-GG-GGGAGGA.........A-TCCTCTCT--GCT-TACCGC--TC-TA-GG-AT-CAGTAA............................ 6527
H2B.GH.86.D205            .------GCC--T----A---CTCA--G--T--ACAAT-AG-GGGAGGA.........A-TCCTCTCT---CT-TACCGC--CC-TA-GG-AT-CGATAA............................ 6529
H2G.CI.92.ABT96           .--TC--GGA-AC---A--TGCTC---G--T--A-GG-ACGC-GGCGGA.........A-TCCTCT-T---CT-TACCGC--TC----GG-GT-TTCTAA............................ 5872
H2U.FR.96.12034           .------GGG--A-----G--CTCA--G--C---AGC-GCA-GGGAGGA.........A-TCCTCTCA---CTGTACCGC--TC-T--GG-GT-CTCTAA............................ 6016
SMM.SL.92.SL92B           .--T--GCA---A---CACC---TCA--G--T--ACAGC-AA-GGGACCA.........A-TCCCCTG-G-TCT-TACCGT--GC----GA-GT-TTATAA............................ 5889
SMM.US.x.H9               .------GG---A---CA----CTC------C---CAGT--G--GGARGA.........A-TCCTCTCT---CT-TACCGC--K-----GG-GT-TTATAA............................ 5930
SMM.US.x.PGM53            .------GGA--A---GCT--CTC------C---CAA-G-A--GGAGGA.........A-TCCTCTGT---CT-TACCGC--TC----GG-GT-CTCTAA............................ 6375
SMM.US.x.SIVsmH635F_L3    .-------G---A---GCC--CTC------C---CAAT--A--GGAGGA.........A-TCCTCTCT---CT-TACCGC--TC----G--AT-CTATAA............................ 6460
                                                                                                                Vpr end
MAC.US.x.239              .------GGC--A--CATC--CTC------C---CAAC--G-GGGAGGA.........A-TCCTCTCT--GCT-TACCGC--TC----AG-AT-CTATAA............................ 6456
Vpr                        _F__R__G__G__C__I__H__S__R__I__G__Q__P__G__G__G__.__.__.__N__P__L__S__A__I__P__P__S__R__S__M__L__*
Tat exon 1                .__S__E__A__D__A__S__T__P__E__S__A__N__L__G__E__E__.__.__.__I__L__S__Q__L__Y__R__P__L__E__A__C__Y__N##__.__.__.__.__.__.__.__.__.__.__
MAC.US.x.251_1A11         .------GGC--C--CATC--CTC------T---CAAC--G-GGGAAGA.........A-TCCTCTCT--GCT-TACCGC--TC----AG-AT-CTATAA............................ 6455
MAC.US.x.251_BK28         .--T---GGC--A--CAAC--CTC------C---CAAC--G-GGGAGGA.........A-TCCTCTCT---CT-TACCGC--TC-T--GG-GT-CTATAA............................ 6432
MAC.US.x.EMBL_3           .--T---GGC--A--CAAC--CTC------C---CAAC--G-GGGAGGA.........A-TCCTCTCT---CT-TACCGC--TC-T--GG-GT-CTATAA............................ 5581
STM.US.x.STM              .------GGC--T--C--C-CTC---G--T---CAAC-AG--GGAGGA.........A-TCCTTTG----CT-TACCGC---C----GGAGT-TTGTAA............................ 6104
MNE.US.x.MNE027           .------GGC--C--CAAC--CTCT-----C---CAAT--G--GGAGGA.........A-TCCTCTCT---CT-TACCGC--TC----AG-AT-CTATAA............................ 5924
DEB.CM.99.CM40            .C-G---GAA-CT--CTTT--C--GGAG-G--AG-C---CA--GGG...TATCCTAACATT---CC-TTG-C-GGAA--AAT--AGA-GTTAG-GACGGAGAATAA.............CCTTTCACCT 5691
DEB.CM.99.CM5             .C-G---GAA-CA--CTTC-----AGAG-GG-AA-C----A-GAGA...TACCCCAATATT--ACCCCT--C-GGAA--AA---GGAGGT-AG-GATGGAGAATAA.............CCTCTCTTTT 5685
DEN.CD.x.CD1              .-AT---G-A--A---TTC-----GGA-GC--AAAGAGT-A--CACTAT.........-CA-AC-T-AG-CCCTTAC-CC-TC-C---GAAG--CCATGA.......................... 5792
LST.CD.88.447             ....TTGCC-ACC----CCAGCTATGC-GC--CAAGA.........GCAGGAACAGCT--CCC--C--A-GC-GCACC-AGACCC-CTAAAAGAGACATACAAAGAGGCCGTTACTAA........... 5122
LST.CD.88.485             ....TTGCC-ACC----CCAGCTATGC-GC-ACAAGA.........GCAGGAACAGCT--CCC--C--A-GC-GCACC-AAACCC-CTGAAAGAGATATACAAAGAGGCCGTTACTAA........... 5122
LST.CD.88.524             ....CTGCC-CAA----CCAG-TATGC-GC--C-AGA.........GCAGGAACAGCA--CCCA-C--A-GC-GCACCTAGACCA-CTGGAAGAGATATACAAAGAGGCCATTACTGA........... 5119
LST.KE.x.lho7             .C--CC-...ACT----CCAGCTATGC-...-CA-CC-GA......GCAGGAACAGCC--TCCA-C--A-GC-GCACC-AGACC--ATGAAGGAGACATACAACGAGGCAGTGACTAA........... 6203
SUN.GA.98.L14             .C-GCC-...AT-----CT...............................TCTT-C-TC---ATC----G----T-AG-A-ATGT-A-CCCAAGGCCACCAACAAGACCAAGACCAGGGCAAGGGGACATTAGAAGAGGCCTACAAGACTAA.... 6259
MND-1.GA.x.MNDGB1         .C-AGCTCAA------GATGGA-C-TTC............A-GGAAAGAAGACCACAA-T-CCCCCCT--G--TTCA-GC-A--AG--GATAGATTATAA........................... 5541
OLC.CI.97.97CI12          GA-........................G-G-T-A--AC---TTTGCTGTTCATTAGGCT-CTAC-A--AGCCTTG-TTCAA--T...TATAG-AATGGATGC................CCAGGAGTGGAGCTTAGAAGA..............AACCCTCACAGAAG 5628
WRC.CI.97.97CI14          .C---CCCAA-------CTA-ACAACA-GAGAAAAGGGAA-A-CAATTTCAAAGAGCAG-ACAATTGTAT-----CTTG-AG--A-C-G-AGGAGATCAGGGAGAAGTTGAGCAAAGA............. 6139
WRC.CI.98.98CI04          .C-T-CCCAA-------C-A-ACAACA-GA-AAGAGAGAA-A-CAATTTCAAAGAGCAG-ACAATTCTTT-----CTTG-AA--AGC-G-AGGAGATCAGGGAGAGGTTGAGCAAAGA............. 6154
WRC.GM.05.Pbt_05GM_X02    .C---T-CAA--A----CCA-AGAA-A-GAGAA-AGAGAAA--GATTTTCAAAGAGCAG-ACAATTTTAT------TTG-AA--GGCT--AGGAGATCAGGGAGGAGTTGAGCAGAGA............. 5074
MND-2.CM.98.CM16          .-AT---TA---A--CA-GG-A---C-T--C--ACAAGAAG--GGGGGATGT......T-CCCCTT-CG-TCCTT-CCGAGATCGGACAA-CCCCTGTGA........................... 6139
MND-2.GA.x.M14            .--T---CA---A--CA-GG-A---C----T--ACAAGAAG--GGCAAA......TATA-CCCCCTTAG-TCCTT-CC-AGGCCG-ACAA-CCCTTGTAA........................... 6066
MND-2.x.x.5440            .--T---TA---A---A-GGGA--TC-C--C--ACAAGAAG--GGAAGATAT......A-CCCACTTCG-TCCTT-CCGAG-CCG-ACAA-CCCTTGTAA........................... 5704
DRL.x.x.FAO               .--T---TA---A--C-AGG-A--T---C-T--A...C-AGC-GGACGCTGG......A-CCCATTTCG--CCTT-CC-AGGGCG-ACCAAC-ATTGTAA........................... 5755
MON.CM.99.L1              .-A---GGAA--C---TAC-----ATATGC--AGCAA-TAA--AGATACCCTGTTCTTAG-CCA-TGAG-G-G-C--CTC--G-CCC-A--AGTTCTGTTCCACAGGCCGATCCAGACAATCCCAGAAGACCTTCCAGATATAGAATGGATGAATGA....TGGAACCTG 5761
MON.NG.x.NG1              .A-G--GGA-------TAT--C--GCATGCCC-A--GRT---CAGGTACCCTYGRCTG-GACCACT-AG-G-C-C--CCC-AG-ACCCA--AGTTCTGTCCCTCATGCAGACCCAGAGAGGCCTTTGCACCCGTCCAGATACCACCACGATGAATAG..TGCAASGATAG 4374
GSN.CM.99.CN166           .--T-A-GA---A--CTTT--C--GGTC-ATTTATCA...G-GAGATACCCAGCAATTAG-CCCTC-AGGG-C-CC-CCC-ACCAGACAGTA-CTCTGTTTCTCATGCCGATCCTGAGCAGCCTAGGAGACCCTCAAGATATCGCATGGATGAATAG.....AGACCCGG 5746
GSN.CM.99.CN71            .--T---GA-------TTC--C--ATAC-ATA-AA--GTCA--AGGTACCCAGTAATCAG-CCCTT-AGGG-T-CC-CCC-ACCAGATAGTA-CTCTGTTCCTCATGCTGATCCTGAGCAGCCCAGAAGACCTTCAAGATACCGCATGGATGAATAG.....AGACCCGG 5734
TAL.CM.00.266             .---CAGCAA-----CTAT---GAGCAGCA-A-AAGA----CGCACTACCCCAATATC-GACCCCTTT--G---GA--G-AG--G-CCATGC-ATAA............................ 5865
TAL.CM.01.8023            .---CAGCAA-----CTTT---GAGCAGCA-A-AAGA---GCCCACTACCCCAACATA-GACCCCTTT-CG-C-GGA-G....AC-CCATGC-ATAA............................ 5360
MUS-1.CM.01.1085          .------GAC--A--CTAT-----GTAT-ACA-A--A-TGA--AGATATCCAACCATAAG-CCTTTGAG-G-C-CACA-C--CC-CCT--TA-TTCAACCCCAAATGCTGACCCTGTACCTCCTTTGAGTCCTTCTAGGTACAGGATGGATGAATAA....TGGATCCAT 5709
MUS-1.CM.01.CM1239        .-ATC--GAG--A---TAT-----GTACTCTAA-ACA-TCA-GAGATACCCCAACATAAGACCCTTGAG-G-G-CACA---TCC-CCT--TA-TTCAATGCCTAATGCTGATCCTACTCCTCCTTTGAGACCCTCTAGATATAGGATGGATGAGTAA....TGGATCCTT 5699
MUS-2.CM.01.CM1246        .-AT--GGAC--A---TAC------TATGCTCAAACA-TCA-GAGATACCCACCGTTG-GCCCATTGAG-G-G-CGCAGC--CCAGAC--TA-TTCCATGCCAAATGCTGACCCCACACCATCTTTGAGACCCTCCAGATATAGAATGGATGAATAA....TGGACCCTT 5816
MUS-2.CM.01.CM2500        .---C--GAG--A--CTAC-----GTAT-GTCAG----T-A--AGATACCCACAGGTG-GCCCCTT-AG-G-G-CGCA-C-TCCAGAC--AA-TTCAACACCAAATGTTGACCCCTTTCCTCCACGAAGACCCTCCAGGTACCGCATGGATGAGTAA....TGGATCCAG 5769
SYK.KE.x.KE51             .--TGCTGCA--A----C---G--G-C--AGTA--CG-GA-ACAGGGGATACCCACAT-CTTCTTAG............................................................... 5706
SYK.KE.x.SYK173           .--TGCTGCA--A----CT--AC-G-CC-G-TAT-CC-GA-ACAGAGGGTACCCCCAT-CTTCTTAG............................................................... 6042
RCM.GA.x.GAB1             .--T----A---A---GC-GGG--T--G--T--AACAT-CA--GGAAGT.........A-TCCTCT-CG-TCC-T-CCGCAA-CG---AA-ATCATGTAA............................ 5652
RCM.NG.x.NG411            .------GCA---A---GA-GGG-----G--T--A-GAG-GA--GGAGGA.........A-TCCTCTTAG-TCC-T-CTTCA-TCG---AA-ATCTTGTAA............................ 5687
GRV.ET.x.GRI_677          .------TG---T---A--AGG--AG--CCTTTTTCCC--TACGAAGAGAGGAGAAATGGACAAG---G-G--...-CCC-ACCCCCT..................CCTCCAGGACTTGCATAG.......... 6039
SAB.SN.x.SAB1C            .A-G-A-GG-A-A---T-TA-GCCA-AG-CTCA-CC-G-GTATGGACCAGGAGCAGGAGGCCC-CCCC--G-TCTG----GAGC-TC-GGAGGAGCTGCATCGGCCGCTCCAGGCCTG...TGA.......... 6333
TAN.UG.x.TAN1             .--TC--TG---C--CA-GAGA--A-CTCCTTTT-AGC-CTACGAAGAAAGGAGAAATGGAGTCG---GGG--CGG-AT-G---AGAG...............CCTCCTCCAGGACTTGCATAG.......... 6101
VER.DE.x.AGM3             .---C-GTG---A--C--CAGA--ACA-CCCTTT-AAC-ATACGAAGAAAGGAGGGATGGACAAG-G-G-G--CGA-C--GGC-CGT-................CCACCAGGACTTGATTGA.......... 5603
VER.KE.x.9063             .---C-TTG---T-----TAGA--ACAGCCCTT--AAC--TACGAGGAAAGAAGAGATGGACAAG-G-G-G---GA-CGAA-C--GCT................CCACCAGGACTTGATTAG.......... 6105
VER.KE.x.AGM155           .---C-GTG---A-----CAGG--ACA-CCCTTT-AGC-ATACGAGGAGAGGAGAAATGGACAAG-G-G-G---GACC--GAC--GTC................CCACCAGGACTTGATTAA.......... 6100
VER.KE.x.TYO1_patent      .--TC--TG------CA-GAGA--GCAGCCCTTT-AAC-ATACGAGGAGAGGAGAGATGGACAAG-G-G-G-C-GA-C-AATC--GTC................CCACCAGGACTTGAGTGA.......... 6105
COL.CM.x.CGU1             .G-TCATGCG---CAAT-T-GCTTTCCCTCTACGCAA-AC-CTTAG.................................................................................... 5516
```

```
                                                                                                                              Rev exon 1 start
H1B.FR.83.HXB2             ..........ACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAAACTGCTTGT......ACCAATTGCTATTGTAAAAAGTGTTGCTTTCATTGCCAAGTTTGTTTCATAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCT 6004
Tat exon 1                 .__.__.__.__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__L__G__I__S__Y__G__R__K__K__R__R__Q__R__R__R__A
Rev exon 1                                                                                                                               _M__A__G__R__S__G__D__S__D__E__E__L

H1A1.UG.85.U455            ............CCTAGAGCCCTG-A-A-ACCCGG-AAGT-AGCCT--AA--GC-TGT...-G---C--T---------GT--------GG---------C----C--TC-G-A-----G---------T----------A-----------A--CC------G--C-- 5449
H1B.US.90.WEAU160          ............-CTAGAGCCCTG-A-G-ATCC-G-AAGT-AGCCT-AGA--GC-TGT...-----------------GA-----------------------G------G-----------------------------------------------T-- 6002
H1C.ET.86.ETH2220          .........C-----A--------C-------------------G----------.....-ATC-A--T-----------A---A---A------T-T------C--TC-G--.----G---------T---------------------C----------------- 5394
H1D.CD.84.84ZR085          .........CA---T--------C------------------GG----------.....-A-----T------------------A---------------C-------G---G----------------------------------------------C-- 5525
H1F1.BE.93.VI850           .........CT----T--------T-----------C-------C----C----.....-----A--T---------CGA-----------------TGG--C--T-C---G--G-G---------------------------------------AC---A-- 5343
H1G.SE.93.SE6165           .........C-------------T----G--G--------------C-C---.....-A---A--T-T-------GT--------GG--------------C--TC-G-AC---G-----------------------------AG--C------G--A-- 5400
H1H.CF.90.056              ..........--------------C-----------------C-----------.....-A-----T--------------C---A-----------A-G--C--TT---AG---G------A--T---------------------C-------AC---A-- 5350
H1J.SE.93.SE7887           .........CAG-----------T---------------------------------.....--T----T-------------------A------T-----G--C--TC--CAG--G-G---------T-----------------------------TC---- 5317
H1K.CM.96.MP535            .........CA------------C--A---------------------------.....-A-C-G---------------CG--------A-----------A-A--C--TT---AG--G-G---------T---A----------------A------CC-------C-A-- 5199
H101_AE.TH.90.CM240        .........C-------------T----G-----------C-----------.....-G---G--T--C---------A-------GG---------C-A--C--TC-G-A-----G-----------------------------AG--C------G--A-- 5571
H102_AG.NG.x.IBNG          .........C-------------C--C--G-----------C-----------.....-G---G--T--C-------T---C---G.---------T-G--C--TC-G-AC--G-G----------------------------GC------G--A-- 5527
H103_AB.RU.97.KAL153_2     ........-----------------------------C----------------------.....-----------------------------TC----------------------G-A-----G------------------------------------------------ 5229
H104_cpx.CY.94.CY032       ...ATCCTGAC-------------C-----G-----------C-----A---.....-A---G--T-TC---------------GG--------------C--TC-G-A-----G-----------------------------A---T------G--T-- 5373
H1N.CM.95.YBF30            .........-T-----------T-----------C--A--------A-----C.....-AT-----------------GA-------A---C---TT-TA---C----C--AG----G-----------A-----------------------T--A----------A-- 5594
H1O.BE.87.ANT70            .........GG-GCCC--T---C-C--C--T-----------CC-G-TCC-----.....-AT------------C---GA--C----A--------T-T-----------G---G---G-GT--G--A-----------------------AG---GA-C-GCTGCT--A 6061
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-GT--G--A------C-----------------------A---AGA-C-GC-GCT--- 6045
H1P.FR.06.RBF168           .........C--GCCT--A---C----G-----G---------CTC--GCC-A---.....-A---------C--C---GCC--C----A--------T-T-----------C--AG--G-GT--G--G------------------------AC--AGA-CTGCG------ 5470
CPZ.CM.01.SIVcpzCAM13      .........GT-G-----A---------------C--C-----A-----A------.....-AT----------------GCT-----T--C-----TATGTA---C----C--AG----G------A--T-----------------------AGC--GA-----C-CG--- 5586
CPZ.CM.05.SIVcpzEK505      .........-T-------------T--C------------A-----G--A-----C.....-AT------C--------GTT--------A---C--TGTGTA---C----C--A-----G-----------A-----------------------T--A----------A-- 5572
CPZ.CM.05.SIVcpzLB7        .........C-----T--------T-------------------------C-C---.....-A--------T-------------------A---C--T-C-T---C----C--AG----G------A--T---------------------T-----TC-----------A-- 5558
CPZ.CM.05.SIVcpzMB66       .........TT-------------T--------------------------------.....-AT----------------C-A--C----AC--C--T---C----C----CG-AG----G------T--T--------------G---------A--------------A-- 5555
CPZ.CM.05.SIVcpzMT145      .........CT----A--A-----C--------G--CA-A----G---A-AA--C.....-AT--G-----T---C-------------T-A-----TAT-C----C--T-C--AG----G-----------------------------AG-GAGA---AA-CCT--- 5545
CPZ.CM.-.SIVcpzMB897       .........CT----A--------T-----------C-----------CC-C---.....-A-------------------T---------A---C--T---T----C----C--G-----G------A--T-A-----------------------AG-GA---G--GCTA-- 5569
CPZ.CM.-.SIVcpzDP943       .........T------------C-A-----------C--A----G---------C.....-AT-----------------TGT--C-------------ATG--C-----TGAG-A-----G------A------------------------A--GGGA--CA--CACC-G 5586
CPZ.CM.98.CAM3             .........T------------CTA-----------C--------G---------C.....-AT--C--------------T---C-------------AT----------TGAG-A-----G------A------------------------A--GGGA------C-CTTG 5414
CPZ.CM.98.CAM5             .........T------------CTA--------------------------------C.....-AT--C--T----------TGT--C-------------ATG-----C--TGAG-A-----G------A-----A------------------A--GGGA---A--CACT-G 5702
CPZ.GA.88.GAB1             .........C--G-----------A--C-----C--------A-G--A-T-----.....-AT------------C---GCT-----T-A---C--TAT-TA---C----C--A-----G------A------------------------A-CCACAA-----C-CA-G 6067
CPZ.GA.88.GAB2             TAGATCCTAGT--T-----------T-------------------A-----A----C.....-ATTCA------C--C----G---C----A-----TATGC----C----CC-GC----G------T--T------------------------G----GAA------CG--- 5399
CPZ.US.85.CPZUS            .........CA---------------G--------C--------G--------C.....-AT---------C--------------------C--TGTG----------T-CG-AG--G-G------A--T------------------------G-G-GA--CA--C--A-A 6075
CPZ.CD.90.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---G--G--T------------------GAA-AGCT-GAA-GAAC---CGA 5436
CPZ.TZ.00.TAN1             TAGATCCTCAGG---CA--A---G-A---------GC-GCA--G----AC-----.....--A--C--T--C---------A--C------------------C-------C--T-CG-A------A-----A-----------------------GAA-AG---GCAA-TCTGCT-TA 5642
CPZ.TZ.2001.TAN2           TAGACCCTCAGG---CA--A-----A---------GC-GCA--G-G--AC-A---.....--A--C--T--C------C-A--C-----------------CGC----C--T-CG-A------A-----A-----------------------GAA-AG---GCAA-TCTGCT--A 5647
CPZ.TZ.2002.TAN3_1         TAGATCCTCAGG---CA--A-----A---------GC-GCA--G-G---C-A---.....--A--C--T--C---------A--C-----------------CGC----C--T-CG-A------A-----A-----------------------GAA-AG---GCAA-TCTGCTA-A 5647
GOR.CM.2007.SIVgor2139_287 .........G--GCCA--A---C----A-----G--C-----CCC---CC-A--C.....-A--CA--------------------C--T-A---C---T-T-----------C--AG--G-GT--G--A------------------------AC--AGA-CTGC----A-- 5546
GOR.CM.2007.SIVgorCP2135con .........G--GCCA--A---C----G-----G--C-----CCC---CC-A--C.....-A--CA--------------------C--T-A---C---T-T-----------C--AG--G-GT--G--A------------------------AC--AGA-CTGC----A-- 5545
GOR.CM.2004.SIVgorCP684con .........G--GCCA--A---C-A--G-----G--C-----CCC---CC-A--C.....-A--CA--------------------C--T-A---C---T-T-----------C--AG--G-GT--G--A------------------------AC--AGA-CTGC----A-- 5465
H2A.DE.x.BEN               ..............................................................---A---------GCGA--CA-T--C---------GC-G------TCG-A---G-GGC-C--A--A--A----AGC-A---GGCA-AC--AGAA-GACTCC-AGG 6662
H2A.GW.x.ALI               ..............................................................--CA-----------GG-A--C---AC--------GC-G-----TT---AT--G-GAC-C--G--A-GG----A-C-A---GGCA--C--AGAA--TCTCC-AAG 6652
H2A.SN.x.ST                ..............................................................---A--T--C-----------C---AC--------GA-G-----TT---AC--G-GGC-C--G--A-GG----AAC-A---GGCA-A---AGAA--ACTCCGAAG 6104
H2B.CI.x.EHO               .............................................................T-CA-----C---------A---C--AC--------GC----C--TC-T-A---G-G-C-G--G--A-GT----AAC-CTC--GAAA----TCTTC-AA---G--- 6633
H2B.GH.86.D205             .............................................................T-CA----C-----G--A--C----AC--------GC------TC-T-A---G-GTC-T--G--A-GT----A-C-CTC--GAA-----TCTGC-AA------- 6635
H2G.CI.92.ABT96            .............................................................T--A--T------C-----A--C----A------T--GT-A--C--TC-T-AC--G-GAC-C--AG-A-GT----AGCAACCT-GAA-A---AGAACT-CG-A-AAG 5978
H2U.FR.96.12034            .............................................................T--A-----------------A-------A-------GC-A-----TC-T-A---G-G-C-T--G--A-GT----AGCAACCC-GAA-A--GAGT-C-AAG-AG-T- 6122
SMM.SL.92.SL92B            .............................................................--GC--T-T-------------------C-------GC---------TC-T-GG--G-G-C-T------TA-T----CA-AACCC-GAA-A---GTTAAGAAG-AG--- 5995
SMM.US.x.H9                .............................................................---G--Y--C-----G--A-------A------T--GCA------TC-Y-A---G-G---G--A--A-G----AGCA-C-ACGAA-A---ACT-CGAAG-AGA-- 6036
SMM.US.x.PGM53             .............................................................T-CA--T--C--------A--C--T-AC--------GC-------TC-T-A---G-G---G--G--A-GT----AGCA-C--CGCA-----AGAA--ACTCCGAAG 6481
SMM.US.x.SIVsmH635F_L3     .............................................................---G-----C-----G-G--------A-------T--GCA------TC-T-A---G-GT-----A--A-G----AGCA-C-CCGAA-A---ACT-CGAA--AGA-- 6566
                                                                                                                              Rev exon 1 start
MAC.US.x.239               ..............................................................--CA---------------------AC--------GT-------TC-T-A-----G---G--G--A-GT----AGCAATCACGAAA----AGAA--ACTCCGAAA 6562
Tat exon 1                 .__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__T__C__Y__C__K__K__C__C__Y__H__C__Q__F__C__F__L__K__K__G__L__G__I__C__Y__E__Q__S__R__K__R__R__R__T__P__K
Rev exon 1                                                                                                                               _M__S__N__H__E__R__E__E__E__L__R__K

MAC.US.x.251_1A11          ..............................................................--CA---------------------AC--------GT-------TC-T-A-----G---G--G--A-GT----AGCA-TCACGAAA----AGAA--ACTCCGAAA 6561
MAC.US.x.251_BK28          ..............................................................--CA---------------------AC--------GT-------TC-T-A---G-GA--G--G--A-GT----AGCA-TCACGAAA----AGAA--ACTCCGAAA 6538
MAC.US.x.EMBL_3            ..............................................................--CA---------------------AC--------GT-------TC-T-A---G-GA--G--G--T-GT----AGCA-TCACGAAA----AGAA--ACTCCGAAA 5687
STM.US.x.STM               ..............................................................---G---T---C-----A-------A---C-----GC-G-----TG-T-------GTC-T--A--AA------AGC-ATCA-GAA-A---GTTAAGAA--AG--- 6210
MNE.US.x.MNE027            ..............................................................--CG-------C---------------AC--------GT-------TC-T-A-----G---G--G--A-GT----AGCA-TCACGCA-A---AGAA--ACTCCGAAG 6030
DEB.CM.99.CM40             TTGCTTCAG..A-G---AG-ATAG-T-CATTT-A-GAAAGAG----T--AC-CCAC.....C--GC---TATG--C---GCC--CATTC-A---------GC-G-----T---CA---G-GTC-----G--T---------CCACGC-G-A---AGAG-AAGT-C---AGA 5853
DEB.CM.99.CM5              ACATCATAG..A-G--AGAAATAG-T-CATTCAAGGAAACAG----T--AC-CCA-.....C-T-----TATG--C---GCC---ACTA-G---------GC-G--C--T--GCA-----GTC-----G----A--------TCACGC-G-A---AGAG-A----CG--T-A 5847
DEN.CD.x.CD1               .........ACGCG--CAG-ATAG-C-CCTTC-CTG-CA-TAAA...--CC----C.....-A--CA-----C-------C-C---GCT-AC----------T-G--C--TC--CA-----GAC-----T--TGTT-----------GCTCCAA-A-AGAGA-C-A--GCTCG- 5944
LST.CD.88.447              ..........................................---ACTG-G---C---..AAC-A---G----GC------------------------TATTC----C--TCAGCAG--G-GA-----G--ACGA---TATGTCC-C-G-A------A-AA--GTCGCCCAG 5249
LST.CD.88.485              ..........................................---ACTG-G--TC---..GAC-A---G----GC--C----G--------------TATTC----C--TCAGCAG--G-GA-----G--ACGA---TATGTCC-C-G-A-----AGAA--GTT-CCCAG 5249
LST.CD.88.524              ..........................................T--ACTCCAAT-C--C..CAG-AT--G---C-C--C---GTT------------C--TAT-C----C---CAGCAG--G--G------G--ACGG---TATGTCC-C-G-ACC---AG-A--GTT-CCTAC 5246
LST.KE.x.lho7              ..........................................AG-CCTGCAA--G---..CAG-AT-CC--T-GC---------A--------------TATTC----C--TCAGCAG----GG------GC-GCA----TATGTCC-C-G-AAAC-GAGA---GTT-CCT-G 6330
SUN.GA.98.L14              ...............................................CTTGGAG--C---..GAT-A---A----GG----GG-GA--C---------C--T----C----C--TC-GCAG----GG------G---CA----TATGTCT-C-G-G-AC------AT-A-TC-ATA 6383
MND-1.GA.x.MNDGB1          .................--AACT-CTAG-AGAGTATTA-------TTGCAA-------..GAG-AT--A----GG--C--G--A--C-----------TATGC----C--TCA--AG--G-GT-----A--AAGG--CCATGTCT-C-G-AAAC-TGTA-C-G---CTAA- 5692
OLC.CI.97.97CI12           ATCTGAAGGA-TGC..............................................................................................................A-GAATA-ATGTTATTGT--A-AA- 5667
WRC.CI.97.97CI14           ..........................................-A-CTT-T--C-C---..GATCATCGA--T-C----GTTGCT---GTG-A-AGC--TACGC-----C--TC-GCA---G-GG-----G--T--T---C-TTCTT-C-GCAA--A-ACTAAG-C-GATA-A 6266
WRC.CI.98.98CI04           ..........................................-A-CTT-T--C-C---..AATCAT--A--TGGA---T-T--A---GTT---AGC--T--GC----C---C-GCA---G-GG-----G--T--T---C-TTCCT-C-GCAA--AGATTG-TGCG--TA-A 6281
WRC.GM.05.Pbt_05GM_X02     ..........................................-A-CTT-T--C-C---..AATCATC-A---AGA---------A---AT---CTCA--T---T-A-----TC-GCA---G-GG-----G-----T---CATTCAT-C-GCAA--AGACTAATAAGGATAAG 5201
MND-2.CM.98.CM16           ..............................................................--CA--T--C-------GA--------A---C-----GC-A--C---T-GCAG--G---------AG-ACAT---CATGCCT-C-GAAAC--GAGATCTA--CAGAGA 6245
MND-2.GA.x.M14             ..............................................................---A--T-----C--G--A-------A---C-----GTG---C---T-GCAG-----T-----G--ACAT---CATGTCT-C-GAACC--GAGAACT--G-AGAGA 6172
MND-2.x.x.5440             .............................................................T--G----TC-------GTT--------A------T--GC-G--C---T-GCAG-----T-----GG-ACAT---CATGTCT-C-GA-A---G-GA-CT---C---GG 5810
DRL.x.x.FAO                .............................................................T-CA---GG--C---T-C--CGCT-AC--C--T--GC----C--TC-TCA-----G------G--ACGT---AATGTC-GC-G--CC---A--A--ACCGCCA-C 5861
MON.CM.99.L1               TTGACCCAGACT--CCCAAAGAAC-A--C--CCC-GCAACC--ACGC--AC------..ACT..---C----T-----G-GTT------------C--T--GA----C---C-C--------G-C-C--G--T--A...TATG-C-G---A----AGAGA---A--GCGA-G 5922
MON.NG.x.NG1               TGAACCCAGC-T--CCC---GA-C-------CTC-GCGACA--GCTC--CC-G---..AAC..---C----T---C--G-TT---GC-C-G--CY-T--G-----C---C-CCA-----GAC-T--A--A------------GGCCGAAA---G-GC--CG-TCC-AGA 4538
GSN.CM.99.CN166            AAGAGTGTCTCT-GCCCAA----C----G------GCGGC---CGCC--CC-G--C.....T-TGC------C--C--G------GC-C----C--T--GT-G--C--TC---GG--G-GAC----TC-T---CTCCATG-C-GATCA--C-AGAG-AA----CC-AAA 5910
GSN.CM.99.CN71             CAGAGTGTCTTT-GCCTAAG---C-A--G------GCCGC---A--C--CC-G--C.....T-TGC---------C----G---GCT---------T--GT-G-----TC---GG--G-GAC----TC-T---CT-CATG-C-GATCC--C-A--G-AAGA--TC--AA 5898
TAL.CM.00.266              ..............................................................A--------C--G--A-----T-AC--C-----GC----C---T-GCAG--G-GT-----T---CGG...TATGTCCCTG-CA----GTTC--GAAG--GAGA 5965
TAL.CM.01.8023             ..............................................................G--T--C--C-----A-----T-A---C-----GC----C---T-GCAG--G-GT-----T---CGG...TATGTCCCTG-CA-----ATC--GAA---GAGA 5460
MUS-1.CM.01.1085           CAGTAGAGGAGT-GCCAAAGGAAC-A-GA--G--TGCCGCA--GGCC--GC-----.....--ATC-----TC-------------GCT---------GT-C-C-----TCA--AG-----A-----T--T-------------CGTTTCA-A-A--GATCGA--GCT-A- 5873
MUS-1.CM.01.CM1239         CAGTAGAAAAC--TCCC--GGAAC---GC--C--CGCAGC---G-CC--GC----C.....-G-TCC--------------CG---GCT---------ATGT-G--C---CA--AG-----T-----T------------------CGCTTCA-A-A--GT-C-A--TCT-A- 5863
MUS-2.CM.01.CM1246         CAGTTGAAGGC--GCCT--GGACC---G---C---GCGGCA--CCCC--AC----C.....-----C--T--C--C---TGT--CGCG-----C--TAT-T-G--C---CAG-A------A-----T-----A----------CGC-TCA---A--GA-C-A---CT-A- 5980
MUS-2.CM.01.CM2500         AAGTAGAGAATT-GCCC--T---C--AGA---TCCTCAGCT--A-C---CC-A--C.....--CC------C------GCT---GC--A------TATGCAG-------C-T-----G--A-----G--T--T------------C-G-AA-------T---A---CTC-- 5933
SYK.KE.x.KE51              .................AAG-A-CATT--TCTC--A-GG----C-G---C------.....-G--GA-------C-------------A------TT-TAGA--C--TC-TCA------G-------G--AA-A----T-C--G---GAAA----AGA-ACGC-----A-A 5852
SYK.KE.x.SYK173            .................AAG-A-CATTC-T--AG-AAGGA--ACC---AC-A----.....-AT--A----T---C-----C--------A----------GC--------C-TCA---G-GA--G--A---A-T-----C---ACCA-G-AA-C--GCAGCT--C--TAT- 6188
RCM.GA.x.GAB1              ..............................................................--C---T-----------A-----T-A---C------C-C---C---C-GCAG---GAC--------AAAT----CTTCC-G-GCAA-AC--AGAA--TCT-AG-AA 5758
RCM.NG.x.NG411             .............................................................T-GC--T-----C-----GA---GTG-A---C--T--GC----C--TC-TCAG--G--T---------AAA-----CTTCT-G--GAA-A---AGAG-CA--CC--GA 5793
GRV.ET.x.GRI_677           ..........................................---ATT-CAAC-A---..ACT-AT--A------------C-----A-C--T-A---C--TG-GC----C---C-GCAG--G-G------AG-GCGT---CATGTCTCT-G-AAA---AGAAA-AC---CA-- 6166
SAB.SN.x.SAB1C             ............................................................T-AT-C-----TC--C---GTT--C-----------TATTC-C--C---CAT-A---G---C---------TCG---TATGTCCCT-G--CA---AGAGCTTCT-AGAAG 6442
TAN.UG.x.TAN1              ..........................................----TT---AC-A---..ACT-A---A----T---C----GA--C----A---C--T---------C--TC-GCAG--G-G-----------A-----CATGTCTCT-G-ATC---AGA-CTAAG-AGAA- 6228
VER.DE.x.AGM3              ..........................A-A-CTC-AAGCA--CCTG-AGCGG---..ACA-A---G------------TGT--C--T-A---C--T--GC----C--TT--CA---G-GA-----TG-TA-----CATGCCCCT-G-ATC---AGAAAGAAG-TT--- 5742
VER.KE.x.9063              ..........................A-A-TAT-AGA----GCT---AA------..AGA-AT--G----T---C---------------A---C--T--GC-------TC-GCA-----G-C-----TG-TA-----CATGCCCCT-G-ACC---AGAAAGAA---T-T- 6244
VER.KE.x.AGM155            ..........................A-A-TAC--GA-A--CCTTGAGA-A---..ACA-AT--A----T---C-----A--C--T-A-----------T-C--C---T--CGG----GAC-----T--TA-----CATGCCTTT-G-ACC---AGAAAGAAG-TT--- 6239
VER.KE.x.TYO1_patent       ..........................AGACTATCAGA-----CTGCAGA------..AAA-AT--A--T-T---C-----A-----T-AC--C--T--GC----C--TC-GCA-----G-C----TG-TA-----CATGCCCCT-G-ACC---AGAAAGAAG-TT--- 6244
COL.CM.x.CGU1              ..................................................AG--G---..AATGA---A------GG-----GGCC------------C--T--GC-A--C--TC-TT-------GGC-G--A---CG-----AAG-A---T---A-TGCTACA-A-......... 5626
```

```
                                                   Tat Rev exon 1 end   intron start
H1B.FR.83.HXB2             CATCAGAAC..............AGTCAGACTCATCAAGCTTCTCTATCAAAGCA...GTAA......................................................................GTA....................GTACATGTA.. 6061
Tat exon 1                  H  Q  N  .  .  .  .  .  .  S  Q  T  H  Q  A  S  L  S  K  Q
Rev exon 1                   I  R  T  .  .  .  .  .  V  R  L  I  K  L  L  Y  Q  S
H1A1.UG.85.U455            -C---A..............GGC---A--GA-------A--CT-A--C-------..----......................................................................---A.................TTA-CTT------. 5510
H1B.US.90.WEAU160          -C---A..............AAC-----------------A-------------..----......................................................................---................................. 6059
H1C.ET.86.ETH2220          -C---A..............AGC---A--GA-------AA-CT-A----------..----......................................................................---CCAAATAATAG..................------. 5458
H1D.CD.84.84ZR085          -C---C..............AGC----------------A-C--A--C-------..----......................................................................---.............................------. 5582
H1F1.BE.93.VI850           -C---A..............AGC-----AGTG------AA-------C-------..----......................................................................---GTGTTA...........AATA---T------. 5410
H1G.SE.93.SE6165           -C----..............AGC---A-AGGC-------A-C--G--C-------..--G-......................................................................---GAAAAA...........ATTA---T------. 5467
H1H.CF.90.056              -C-GCA..............AGTTTG--AGA-------AA----A----------..----......................................................................---.........................T--TCATA-. 5407
H1J.SE.93.SE7887           -C--CT..............GGC---A------------A-CT-A--C-------..----......................................................................---AATCAA............GTA-C-T------. 5383
H1K.CM.96.MP535            -C-T-C..............AAC---G---AC-------A-C-----CG------..--G-......................................................................---GTGCTA...........AGTA----------. 5266
H101_AE.TH.90.CM240        -C----..............AGC---A--GA-------AA-C--A--C-------..----......................................................................---..................ATAA---T------. 5632
H102_AG.NG.x.IBNG          -C----..............AGCC-----GA-------AA-C--G--C-------..--G-......................................................................---GTAATA...........ATTA-C-T----G. 5594
H103_AB.RU.97.KAL153_2     -C---A..............GAC-A-------G------T-------C-------..----......................................................................---...........................GTAC. 5281
H104_cpx.CY.94.CY032       -T---A..............GGC--CA--GGC------AA-CT-A--C-------..----......................................................................---..................TTAA---T------. 5434
H1N.CM.95.YBF30            -C----..............AGC---A---G--------A-CT-A--C--G----..----......................................................................---AAACCTGTATAT...................... 5654
H1O.BE.87.ANT70            ...AGC..................CA--CAGA----A---A-C--G--C-------..----......................................................................---ACGCTA............................. 6109
H1O.CM.91.MVP5180          GCAAGC..................TA--CAGA-A--A---A-C--G--C--G----..----......................................................................---ACGCTG............................. 6096
H1P.FR.06.RBF168           G--A-C..............AAC-A-----A--------A-C--G--------T-..----......................................................................---GAAATA...........ATAA--TGCAT--. 5537
CPZ.CM.01.SIVcpzCAM13      TT-GCA-G...................-A-A-A-A------A-C--G--CG------..----......................................................................---TATTATCTGACCCT.................... 5648
CPZ.CM.05.SIVcpzEK505      -C-----G...............-A-A---G--------A-C-----C--------..---C......................................................................---AAGCTATATAT......................... 5631
CPZ.CM.05.SIVcpzLB7        TC-A---G...................-----C-CA-T------A-C--A----------..----......................................................................---CTCCTATAAGTTGTAGTCAATA-............ 5629
CPZ.CM.05.SIVcpzMB66       TC-G---G...................---------A--------A-C--G--C-------..--G-......................................................................---ATTCATGCAGATATATGCTTA-GCTT-A... 5631
CPZ.CM.05.SIVcpzMT145      TCGAGC--T...............CAGG-T...A--A-G-A-C-----CG------..----......................................................................---AAAGGCTCAC.......................... 5600
CPZ.CM.-.SIVcpzMB897       G-GAGC-GT...............-AGG-TCA---AG-TA--ATA-C-AAGC-..----......................................................................---ATAAGTGTAATATAATAT.................. 5632
CPZ.CM.-.SIVcpzDP943       -C-GCT-CA...........AGT---AC-CA-A------A-C--T--CG-C----..----......................................................................---CTTTAC.............................. 5643
CPZ.CM.98.CAM3             TC-GCTGCA...........AGC---AC-CA-A------G-------CAGC----..----......................................................................---ACTTCACATGCT........................ 5477
CPZ.CM.98.CAM5             -T-GCTTCA...........AGC---AC-CA-A------A----A--CG-C----..----......................................................................---ATTTTGCATGCT........................ 5765
CPZ.GA.88.GAB1             GC--CTGCAGGC...........---A-A-A-A------A----A--C-------..----......................................................................---T................................. 6119
CPZ.GA.88.GAB2             -T-A--GG...............---A-A-GCA------A-C-----CG-C-----..--G-......................................................................---TCCCCTCTTT.......................... 5457
CPZ.US.85.CPZUS            AC-GCT--GAGC...........---GC-GA-A------A-C-----CGC-----..----......................................................................--GCCCTTTTGTGCT........................ 6138
CPZ.CD.90.ANT              ACAACTGCTGAA.........AGC---G-A-A-A------A-C--G-------T-..----......................................................................---CTCTTA.............................. 5496
CPZ.TZ.00.TAN1             ....-C-GT...............AC--A-A------A-C--G--CG-C-----..----......................................................................---CCC................................. 5687
CPZ.TZ.2001.TAN2           A-CA-T--G...............-A-A-TCAAG---CT-TACGA-AGCA.......----......................................................................---CTC................................. 5692
CPZ.TZ.2002.TAN3_1         G-CA-T--G...............-A-A-TCAAG---CT-TACGA-AGCA.......----......................................................................---CTC................................. 5692
GOR.CM.2007.SIVgor2139_287 GC-G-C--A..............GA-----A-A----G---C--G----------..----......................................................................---GAAGCTTTTACTGTAA.................... 5610
GOR.CM.2007.SIVgorCP2135con GC-G-C--A..............GA-----A-A----G---C--G----------..----......................................................................---GAAGCTTTTACTGTAA.................... 5609
GOR.CM.2004.SIVgorCP684con GC-G-C--A..............GA----GA-A----G-A-C--G----------..----......................................................................---GAAGCTTTTACTGTAA.................... 5529
H2A.DE.x.BEN               A-A.....................-C-A------C-TCGC-----GC-C--G-CA-..--G-......................................................................--.................................. 6703
H2A.GW.x.ALI               A-A.....................-T-A--G-----TCGT-----GC----G-CA-..--G-......................................................................--.................................. 6693
H2A.SN.x.ST                A-A.....................-C-A--G-----TCGT-----GC----G-CA-..--G-......................................................................--.................................. 6145
H2B.CI.x.EHO               A-G..........................---AC-AC-T-----GC-C----CG-..--G-......................................................................--.................................. 6668
H2B.GH.86.D205             A-G..........................---AC-GC-C-----GC-C--G-CA-..--G-......................................................................--.................................. 6670
H2G.CI.92.ABT96            AC-A--.......................G--A--TC-TT----GC------CG-..----......................................................................--.................................. 6016
H2U.FR.96.12034            A-G..........................G--T--TC-T--G---C------CA-..----......................................................................--.................................. 6157
SMM.SL.92.SL92B            A-G..........................G--A---C-TT-GA-AC------C--..----......................................................................--.................................. 6030
SMM.US.x.H9                A-G..........................G--A--AC-TT----GC----G-CA-..----......................................................................--.................................. 6071
SMM.US.x.PGM53             A-G...................GC-A--G--A--AC-TT----GC-C----CG-..----......................................................................--.................................. 6522
SMM.US.x.SIVsmH635F_L3     A-G..........................---A---C-CT-C--GC------CA-..----......................................................................--.................................. 6601
                                                   Tat Rev exon 1 end   Tat Rev intron start
MAC.US.x.239               A-G...................GC-A--G--A--AC-T-----GC------CA-..----......................................................................--.................................. 6603
Tat exon 1                  K  .  .  .  .  .  .  .  .  A  K  A  N  T  S  S  A  S  N  K
Rev exon 1                   R  .  .  .  .  .  .  .  .  L  R  L  I  H  L  L  H  Q  T
MAC.US.x.251_1A11          A-G...................GC-A--G--A--AC-T-----GC------CA-..----......................................................................--.................................. 6602
MAC.US.x.251_BK28          A-G...................GC-A--G--A--AC-T-----GC------CA-..----......................................................................--.................................. 6579
MAC.US.x.EMBL_3            A-G...................GC-A--G--A--AC-T-----GC------CA-..----......................................................................--.................................. 5728
STM.US.x.STM               A-G..........................---T---C-AT----GC------CA-..----......................................................................--.................................. 6245
MNE.US.x.MNE027            A-G...................GT-A--G--A--AC-T-----GC------CA-..----......................................................................--.................................. 6071
DEB.CM.99.CM40             G-AT-C-CGCCAGCTGCTGAAAGT-A----C--G------A--T-G--C--G----..----......................................................................-C-AATCTGCATGTAAATGCTAAA-GGA-CA-T.. 5946
DEB.CM.99.CM5              AGCTTTGCT.........GCGAATGA-A--GA--------AG-T-A--C--G----..----......................................................................---AAGCTGCATATGTATATAGTATC---...... 5927
DEN.CD.x.CD1               ..................CAAAGCG--AC-C--G------AAG--G----------..-----.GTTTGATGGATCTCCTCTCATCACTGATTTTAAGCTTTTCAATAATCATTATTTGCTTAGTTTGTTATATT--TATTATTACAGTGATAGTTCTT-AC-CATAGG. 6091
LST.CD.88.447              GTATTT........................GAAG--A-GCAAGA--CTGTGG-CGG..----......................................................................-A................................... 5287
LST.CD.88.485              GTATTT........................GAAG--A-GCAAGA--CTGTGG-CGG..----......................................................................-G................................... 5287
LST.CD.88.524              ATA-TT........................GAAAC-G-GCAGGA--CTGTGG-CAG..----......................................................................AA................................... 5284
LST.KE.x.lho7              ATA-TT........................GAGAC---GCAGGA--CTGTGG-CAG..----......................................................................-G................................... 6368
SUN.GA.98.L14              TC-...........................-AGG--A-GCAAGAGGCTGT-CGAAG..----......................................................................--.................................. 6416
MND-1.GA.x.MNDGB1          A-GA--..................-TA-CTGG-AG-GGT-AAGAAGCTAT-CGAAG..----......................................................................--.................................. 5736
OLC.CI.97.97CI12           GC-GTT-T-ATTGTCAGCTTTGCTTT-TGC-GAAGGGGCTTGGTATCTCTT-TT-TACA-AGTCACCTAGGAAGAGGAGAAAGAAGGTGGTA.......................TCGCATCAACAAGATTCTA-GGACAC...GTAA.................... 5789
WRC.CI.97.97CI14           ACAACA.....................................GCTG-TTCAAG..----......................................................................--.................................. 6289
WRC.CI.98.98CI04           GCAA-A.....................................GCTA---CAAG..----......................................................................-A................................... 6304
WRC.GM.05.Pbt_05GM_X02     TCAG-C.....................................GCT----CAAG..----......................................................................-C................................... 5224
MND-2.CM.98.CM16           -T--T-............GAAAAA-TATCAGAAG--AGT--AG-G-CTGTGGGAAG..----......................................................................................................... 6293
MND-2.GA.x.M14             -T--T-............GGAAAA-TATCACAGG--AGT-AAG-AGCTGTGGGAAG..----......................................................................................................... 6220
MND-2.x.x.5440             -T--T-............GAGAAAGTA-CA-GCA--AGTCAAG-AGCT-TGGGAAG..----......................................................................................................... 5858
DRL.x.x.FAO                GTGGTT............TCAAGAGTATCTGAGAC-GGT-ACA-G-CTGTGGCAAG..----......................................................................--GAAAAAG............................ 5918
MON.CM.99.L1               TC--CTGT-CCT........GGCCT-AGC-GCAG-A-GAA-C--GC-AG-------..--G-......................................................................---................................. 5976
MON.NG.x.NG1               ----CT-CG.........ATTGGCGA--G--GCT----G-A-AA--CGGG-GTCA-..----......................................................................--G.................................. 4592
GSN.CM.99.CN166            GC-ACAG-ACCT...GATCCTAGCGTGTC-T--A------AC---GC--AGGTC--..----......................................................................---AGGC............................... 5974
GSN.CM.99.CN71             GC-TCAG-ACCT...GCTCCTAGCGTG-C----A------ACC--A--CA--TC--..----......................................................................---AGGC............................... 5962
TAL.CM.00.266              AGATCC-CTCCT......AGCAGC-A----GAGG--A-G-A-C--G----------..----......................................................................................................... 6019
TAL.CM.01.8023             AGATCC-CTTCT......AGCAGC-A-----AGG--A-G-A-C--G----------..----......................................................................................................... 5514
MUS-1.CM.01.1085           TCGAGC.................CTG-C--GC---TC--A-C--G--CA--TCA-..----......................................................................---GC................................. 5920
MUS-1.CM.01.CM1239         --GA-C.................GTG----A-----CGAA-C--G--C---TCA-..----......................................................................--GAT................................. 5910
MUS-2.CM.01.CM1246         TCG-TT.................CTG-C-T-AG----GAA-C--G--CG-TTC--..----......................................................................---AT................................. 6027
MUS-2.CM.01.CM2500         G-AGGC.................CTG-CA-ACA------A-C--A--CA-CTCA-..----......................................................................--GAT................................. 5980
SYK.KE.x.KE51              AC-GCT.................GAGG-CTT-GC-GCGCA-AGCTC-A---CT-G..----......................................................................................................... 5894
SYK.KE.x.SYK173            .............................T--G-GG-T-A-AGCGC-C-T-CAGG..----......................................................................................................... 6218
RCM.GA.x.GAB1              G-AA-T--G....................G--G--A--TT-C--G--C-T--T--..----......................................................................................................... 5797
RCM.NG.x.NG411             A--A----G............AAG-A-----T-A---C-TT-GA-AC---TG----..----......................................................................................................... 5841
GRV.ET.x.GRI_677           --A..........................GA-A------ACC--A--CGGC----..----......................................................................--.................................. 6201
SAB.SN.x.SAB1C             ..........................ATTT-G---A-TCAAGT-TCTCT-C-CA-...C--G.......................................................................TA.................................. 6477
TAN.UG.x.TAN1              ---.......................TCA-A------GAA-CT-G-----C----..----......................................................................--.................................. 6266
VER.DE.x.AGM3              -CG..........................CT-G---GCTT-C--GA-CA---A--..--G-......................................................................--.................................. 5777
VER.KE.x.9063              --G..........................C--A---GCCTA----A-CAGG-C--..----......................................................................--.................................. 6279
VER.KE.x.AGM155            TCG..........................G--G---GCAT-C--G--C-GC----..----......................................................................--.................................. 6274
VER.KE.x.TYO1_patent       TCG..........................CT-G--TTG---C--T--CAGC-C--..----......................................................................--.................................. 6277
COL.CM.x.CGU1              .................................................-CA--..----......................................................................--.................................. 5637
```

```
                                    Vpu start (ACG in HXB2)
H1B.FR.83.HXB2               .......ACGCAACCT.........ATACCAATAGTAGCAATAGTAGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATAT.............................AGGAAAATATTAAGACAAAGAAAAATAGACAGGT 6185
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__L__R__Q__R__K__I__D__R__
H1A1.UG.85.U455              .......-T-AC----.........T-GGA---CTGG------AC--GGC-GA--------C----CT----------A------A-T-----GGT--------AAA............................-AA---T-GC---AG----A------------- 5636
H1B.US.90.WEAU160            .......-T----T-..........T-G-A----T------------C-----------GG---------------------------C----------..............................-------------------------------- 6183
H1C.ET.86.ETH2220            .ATGGTTGATTT--TAGCAAAAGTAGATTAT-G-A---T------------CA--------C----C-----------------A------C-TAT---------..............................------T-G------------G-----T---- 5597
H1D.CD.84.84ZR085            .......-T----T--.........T---A----T--------------------T----GC------------------------A-------T-----------.............................C-T-G----AA---G---------------T--- 5706
H1F1.BE.93.VI850             .......-T-TC-TA-.........T-GTT-GC-A---GG----C-------A--------C----------C------------A-T-----TAT---------..............................-A----C-GG----G-----------A-T-AA- 5534
H1G.SE.93.SE6165             .......-T---GT-A.........T--GT----T-----------G----A--------T-C---GC---C------------A--------T-T---------..............................---G-----AG--A---G-A--G-----GA-A-- 5591
H1H.CF.90.056                .......-T-T-TATA.........T--GG-T---G-AT-GG--CGCT-G-----AC-TTT--C---GCC...G--A--------A--------TAT---------..............................-AA---T-GG----G----AG-----------C 5528
H1J.SE.93.SE7887             .......-T-AT----.........T-G-A-----C--------------TA----------TT-CC-T-------G-A-----A-T-----GTAT---------..............................-A----C-G-----------G--------T-A-- 5507
H1K.CM.96.MP535              .......-T-GTGT-C.........T--G----TTC-AT-G---C-TT-G------C-CT----T--GC-...---A----A---A--------TAT---------..............................------T-GG---A-----A--G----A--T--- 5387
H101_AE.TH.90.CM240          .......-T-AC----.........T-GGA---TAGT----------G-C-GA-------GC----CT----------A-----AA-T------GCT------GT-..............................-A-------C----G----------------- 5756
H102_AG.NG.x.IBNG            .......-T---------.......T--A-----ACT----------G-C-----------T-C---GC---C------------AG--------TAT---------..............................---------AGG-A---G-AG------------ 5718
H103_AB.RU.97.KAL153_2       .......-T-T--TGCAA....TCCT--G------C----------------------T-GG-----------------G----------T-----------..............................-------------------------------- 5410
H104_cpx.CY.94.CY032         .......-T-TT-TTC.........TGGGA---CTGG----------G-C-G--------GC----T----T------A------A-TT-----T-T---------..............................-A----T-GAG---G-----G-G---------C- 5558
H1N.CM.95.YBF30              .......-T--TG..................TC-T-G-G-T-CA----G----G--C-----G-T-GC------G--A-A--C---G-AT--C--TAT-G-------..............................-A-------AA-TTG--GGA-------A-ACACA 5769
H1O.BE.87.ANT70              .......-T---T-A-.........-GGGACC-GC--------A--ATTA-TAGT-CTTTGC-GT-T---AATG--A--T-A---GGGT-TA-T......CTTAGAAAATAT...........TTAGAACAAAAGGAAC--GACAG--AGG-G---G-----CTTGA-A 6245
H1O.CM.91.MVP5180            .......-T---T-AA.........GAGAACC-GC----CT--A----T---AGT-CTTTGTGTC-T---AATG--C--A-A----TGT-TAAC......CTTAGAATTTAT...........TTAGTGCAAAGA-AAC--GATAG---GG-GCAGG-----CTTGAAA 6232
H1P.FR.06.RBF168             .......-T---T---.........-G-GATGAG-C--T-T--A--AT-GCT-GG---TT-C----TTGCATT-----A--T---GGG-AG--CC-GC--TTTGTGCTA.............................---GA-AG-GA-AGGGAT--GT-T-TACAAA 5661
CPZ.CM.01.SIVcpzCAM13        .......-T-AT-TTA.................----CTTT-GGTTGCTT-GCCA-T-CCTT---TT--C-TAAC--TT--A-A---AGG-ATT--......TGG............AGATTATGTAAA............C---ATAAGCTTG--GT-G----T--G-ACC 5766
CPZ.CM.05.SIVcpzEK505        .......-T-TTGTTG.........C-TAT--AGT---G-T-CA---G-----C-A---AG-CTC------T------A--T---G-A----------......T-TAGAATATAC............AGA............G--G--AA-GT-G--GA-------TCTCA-C 5755
CPZ.CM.05.SIVcpzLB7          .......-T-G-C-TA............AT-GA-T---G-T--A---GC-----TA---A-C----C----T------A--A----TA-A--C-......T-TC-GCTTTAT..........................-AAG-G-ACA---------AGGC----A-T-AA- 5750
CPZ.CM.05.SIVcpzMB66         .......-T-G-TATA......GTGCAG-A-G---GCTT-T------TTC--A-TA---A-T--G-T----TC-----AA-A---GTA-A---T......T-T............AAACTCTGTAAA............G-GGATAGG--G----AG--------T--A- 5758
CPZ.CM.05.SIVcpzMT145        .......-T---G-TA............GA---T--GCT----A--TT---CA---CTTT---GC-TG----T---T---CT---ATAGC--CT......T-T..............................AAAGAGTAT-AA--GT--CA-CA-GT---GCG------A--AC 5721
CPZ.CM.-.SIVcpzMB897         .ATGGAA-TATTCATA..........--CTT-GG-C--AT-GG-A---TTA---A-T---T----GC-----T-G--TGGT-AAAAG-TTAT-A-TGT...T-T..............................AAGGCACTT-AA-G-......CA-G----G-G-...--TCAA- 5756
CPZ.CM.-.SIVcpzDP943         ..ATGCA-T--T-A--.........TGGGA-CA-A---GCT-GA-T--T----G-A---A-GGC--T--T--T-CT--A--C---GGG----CTT-T---ACG-GG...........AGAAGAAGAAAA-TAG--GA-GAC---GC-CA------CATTGAT- 5784
CPZ.CM.98.CAM3               .......-T--T-A--......TGG...GA-CA-A---GCT--A-T-GTA---G-A---A---T--T-----T--T--A-CC---GGG----CT......TTTA-AGAGTGG...........AGAAAAGGAAAAGAAG--GA-AAC---GC-CA------TAGG-CT- 5613
CPZ.CM.98.CAM5               .......-T--T-AT-......TGG...GAGCA-A---GTT--A----T----G-A---A-C-T--T--T-TT--T--A--T---GGA----CT......T-TA-AAAGTGG...........AGAGAAATAAAAGAAG---GGAAC---ACCCA--G---TT---AT- 5901
CPZ.GA.88.GAB1               .......-T-ACT-TG........TTAGT-.GGCT---TTC-CA--TT-G---GCT--....--TGC-TGGAAC--TTG-A-A---GGATACA-T......ATTA-ATGGGGATAT......AGAAGGTATAAAAGACAT-GGC-TGAG-C-G--........--T--G--AC 6252
CPZ.GA.88.GAB2               .......-T--TTT-C......ATGTGGGT-GCTA---G-C--A-T-G-A---GTACTCTTT--G-T---AAT-----G------GGT--T--.........-G-ATCTCTGTCTAT......AAAAGGTGGAAAAGACAC---GA-GAGCA-AG--TC....--T--TTT-A 5599
CPZ.US.85.CPZUS              .......-T--TTAA-......TGGT-TGA-----GTTT-----C-TT-GG-A--...-A-GGT--TT---T-G-GA-AA-C---GGAC----T......-CTAGATTGTGG...........AGACAAATAAAA-TA---GA-AAT-CT---CA-G-G--TC-A-ATC 6274
CPZ.CD.90.ANT                .......-T-ACTAA-ATATTTGAGTATG-TT-TC-T--CT-TAGTAT-G--C--TGGATT--CTGT---C-T......A-ACTC-AT-AGC-T......T-TA-AATATAT..........AAGCAGCAGCAAATAGAT--T-AGAG--AT-------TC-----AGT-- 5638
CPZ.TZ.00.TAN1               .......-T-AT-AAA......ATAG--GTGGG-AGT-TGTC-ACTAATG-CA---GCATTC-TTGT---TT-C-GA--T-AATAGGGGG-..........-GC-TGCTAATAGGT......ATAGGTATAAGAAGAGA-TT-GA-AGGGA-AGGCA-C-TCA-AGAGTA- 5829
CPZ.TZ.2001.TAN2             .ATGGTA-AA-T-GTA.........G-GGG--GT--GCTT-CTAAT-TTA---GTA-CTTTTGT---T--TT---TC-AA-AG-AGGAGGCC-GC-A---ATTACAATTATAAAAAGAGAGATAGAAAGAGAAAGACAAC-TCA-AG--T--T-GA--G-C--ATA--AA 5852
CPZ.TZ.2002.TAN3_1           .ATGGTA-AAAT-GTA.........G-GGG--GT--GCTT-C-AAT-TCA---GT-CCTTTTGT---T--TT---TC-AA-AG-AGGAGGCT-GC-A---ATTGCATTTGTAAGAAGAGAGTTAGAGAGAGAAAGACAAC-TCA-AG-GTGAT-GA--G-C---TA--AA 5852
GOR.CM.2007.SIVgor2139_287   .......-T---T--A..........-G-GAT...A--AT-G-TA--ATTA---G-A--A--T-GC--GCT-T--CT--GA-AATT-GGC--AA----T-T-C--TATACTTAAGAGATAGAAGGGAGAGACGCTTTTTTG-TCGC---GA-AGGCT-CT--GCA-T-AAG 5761
GOR.CM.2007.SIVgorCP2135con  .......-T---T--A..........-G-GAT...A--AT-G-TA--ATTA---G-A--A--T-GC--GCT----CT--AA-AGTT-GGC--AA-GC-C-T-C--TATACTTAAGAGATAGAAGGGAGAGACGCTTTTTT--TCGC---GAGAGGCT-CT--GCACT-AAG 5760
GOR.CM.2004.SIVgorCP684con   .......-T---T--A..........-G-GAT...A--CT-G-TA--ATCA---G-A--AT-T-GC--GCT-T--CT--GA-AA-T-GGC--AA-GC-C-T-C--TATACTTAAGAGATAGAAGGGAGAGACGCTTTTTTG-TCGC---GA-AG-CT--T--GCA-T-AAG 5680
H2A.DE.x.BEN                 ....................................................................................................................................................................... 6703
H2A.GW.x.ALI                 ....................................................................................................................................................................... 6693
H2A.SN.x.ST                  ....................................................................................................................................................................... 6145
H2B.CI.x.EHO                 ....................................................................................................................................................................... 6668
H2B.GH.86.D205               ....................................................................................................................................................................... 6670
H2G.CI.92.ABT96              ....................................................................................................................................................................... 6016
H2U.FR.96.12034              ....................................................................................................................................................................... 6157
SMM.SL.92.SL92B              ....................................................................................................................................................................... 6030
SMM.US.x.H9                  ....................................................................................................................................................................... 6071
SMM.US.x.PGM53               ....................................................................................................................................................................... 6522
SMM.US.x.SIVsmH635F_L3       ....................................................................................................................................................................... 6601

MAC.US.x.239                 ....................................................................................................................................................................... 6603
MAC.US.x.251_1A11            ....................................................................................................................................................................... 6602
MAC.US.x.251_BK28            ....................................................................................................................................................................... 6579
MAC.US.x.EMBL_3              ....................................................................................................................................................................... 5728
STM.US.x.STM                 ....................................................................................................................................................................... 6245
MNE.US.x.MNE027              ....................................................................................................................................................................... 6071
DEB.CM.99.CM40               ....................................................................................................................................................................... 5946
DEB.CM.99.CM5                ....................................................................................................................................................................... 5927
DEN.CD.x.CD1                 .......-T--TTAAG.....................ATC-AGT---G-CC-A---A-TACG-TTGCT-G-T-T-T-CA--TGTTATT-CCTGG--AGC--CCATAGGT...........GTAGGGTATCTTGCTTAT-G-GCC-AT-A-AGTTAT-G-GA---GCTTA 6221
LST.CD.88.447                ....................................................................................................................................................................... 5287
LST.CD.88.485                ....................................................................................................................................................................... 5287
LST.CD.88.524                ....................................................................................................................................................................... 5284
LST.KE.x.lho7                ....................................................................................................................................................................... 6368
SUN.GA.98.L14                ....................................................................................................................................................................... 6416
MND-1.GA.x.MNDGB1            ....................................................................................................................................................................... 5736
OLC.CI.97.97CI12             ..................................................GT..................................................................................................................... 5791
WRC.CI.97.97CI14             ....................................................................................................................................................................... 6289
WRC.CI.98.98CI04             ....................................................................................................................................................................... 6304
WRC.GM.05.Pbt_05GM_X02       ....................................................................................................................................................................... 5224
MND-2.CM.98.CM16             ....................................................................................................................................................................... 6293
MND-2.GA.x.M14               ....................................................................................................................................................................... 6220
MND-2.x.x.5440               ....................................................................................................................................................................... 5858
DRL.x.x.FAO                  ....................................................................................................................................................................... 5918
MON.CM.99.L1                 .......-T-A-TTA-...TGGTGGTC--TGG---C-AT--CTTACAGCC-GA-CC-CAT-GCCC-CCCC-T-GC--C-TG-GCT-GG...TGG.....................CGATATTATAAAATCACTAAG---TTT-A-AG--T-G-CCA-G-----C-G---C 6112
MON.NG.x.NG1                 .......-T-A-TA-C.............TATTGGAT-GC--CCAT--CT--TTGGTGTG-T--TG-GCTTT---CGCCT--C..................ACCCTATACCTTCTCTATCAGTCTTATAAGGAGCAC--GCGGG-GTCC-GCTTCG--CAG---TTTC 4722
GSN.CM.99.CN166              .......-T---T---............G--GC---TTGGTGGTGG-GTGCT-CTA-TAT--CCT-C---TATC-TTG-...GTAG-GC-TC-T......-CTCTCTAC.....................TTAGCTT--G-T-AG-GGGT-A--G-----CC-A-ACCAA 6095
GSN.CM.99.CN71               .......-T-AGTG--............G--GCTC-CTGGTGGTGG-GGGCT-CT---AT--CCT-T---TATT-CTGC...-TAG-A--CT-T......-CTCTCTAC.....................TTGGCGT--G-C-AG-GG-T-A-GG----GCC-A-G...A 6080
TAL.CM.00.266                ....................................................................................................................................................................... 6019
TAL.CM.01.8023               ....................................................................................................................................................................... 5514
MUS-1.CM.01.1085             .......-T-A-CTA-...TGGTATT--G--GCT--CATTG--AC--GCA-CTACT-T-TG--TGCT---TTTGCCT------TAG-TTACCA-...................AGGTGGTGTAAGCCCAAGAAG....................GT-GAG-T---TG 6038
MUS-1.CM.01.CM1239           .......-T-A-TTA-...TGGTATC--G--GC--CTTT-G--AC--GCA-CTACT-TCTG---GCCT--TTTGC-T-C---CTAG-ATATCAG...................AGGTGGTGCAGACCAAAG.......................GT-GA--TG--CG 6025
MUS-2.CM.01.CM1246           .......-T---TTAC...TGGTATC--GG-GCT-CCAT-G--AC---TA-TTATC-T-TT--TGCTT---T-GCCT--A-CCTAG--TATCA-...................AGGTGGTGTCAACCTAAG..................-C-CC--T-GA--TG-ACG 6148
MUS-2.CM.01.CM2500           .......-T-A-TTGG...TGGTGGT-TG--GC--C--T-G--ACT--CA--TATT-T-TG---GCTT---TTGC-T-----CTTG-ATATCAG...................CGCTGGTGCCAGCCTCAG................-A-GGGCAGGT-GAG-T--ACG 6104
SYK.KE.x.KE51                ....................................................................................................................................................................... 5894
SYK.KE.x.SYK173              ....................................................................................................................................................................... 6218
RCM.GA.x.GAB1                ..................................GTAT-G-G-G...CCCTGG-CT-GT-CT--TT-TAGA-CC-CAG-T-A-AAG............GTTTTTATAGAT..................CTTTT-G-TAC--T-ATT-A-GG--A-T-G-ATA 5892
RCM.NG.x.NG411               ..................................GTAT-G-A-G...TCCTGGTCTG--CCT--TT-TAGATGC-TAG-TTATAAG............GTGATAGGAGATATAATTAGAGAAGTTTTT-TTTGT-G-G-TTATAG--A-T-G-AG-TA-ATC 5954
GRV.ET.x.GRI_677             ....................................................................................................................................................................... 6201
SAB.SN.x.SABIC               ....................................................................................................................................................................... 6477
TAN.UG.x.TAN1                ....................................................................................................................................................................... 6266
VER.DE.x.AGM3                ....................................................................................................................................................................... 5777
VER.KE.x.9063                ....................................................................................................................................................................... 6279
VER.KE.x.AGM155              ....................................................................................................................................................................... 6274
VER.KE.x.TYO1_patent         ....................................................................................................................................................................... 6277
COL.CM.x.CGU1                ....................................................................................................................................................................... 5637
```


Env start
H1B.FR.83.HXB2
Vpu
Env
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.x.IBNG
H103_AB.RU.97.KAL153_2
H104_cpx.CY.94.CY032
H1N.CM.95.YBF30
H1O.BE.87.ANT70
H1O.CM.91.MVP5180
H1P.FR.06.RBF168
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMT145
CPZ.CM..SIVcpzMB897
CPZ.CM..SIVcpzDP943
CPZ.CM.98.CAM3
CPZ.CM.98.CAM5
CPZ.GA.88.GAB1
CPZ.GA.88.GAB2
CPZ.US.85.CPZUS
CPZ.CD.90.ANT
CPZ.TZ.00.TAN1
CPZ.TZ.2001.TAN2
CPZ.TZ.2002.TAN3_1
GOR.CM.2007.SIVgor2139_287
GOR.CM.2007.SIVgorCP2135con
GOR.CM.2004.SIVgorCP684con
H2A.DE.x.BEN
H2A.GW.x.ALI
H2A.SN.x.ST
H2B.CI.x.EHO
H2B.GH.86.D205
H2G.CI.92.ABT96
H2U.FR.96.12034
SMM.SL.92.SL92B
SMM.US.x.H9
SMM.US.x.PGM53
SMM.US.x.SIVsmH635F_L3
Env signal peptide start
MAC.US.x.239
Env
MAC.US.x.251_1A11
MAC.US.x.251_BK28
MAC.US.x.EMBL_3
STM.US.x.STM
MNE.US.x.MNE027
DEB.CM.99.CM40
DEB.CM.99.CM5
DEN.CD.x.CD1
LST.CD.88.447
LST.CD.88.485
LST.CD.88.524
LST.KE.x.lho7
SUN.GA.98.L14
MND-1.GA.x.MNDGB1
OLC.CI.97.97CI12
WRC.CI.97.97CI14
WRC.CI.98.98CI04
WRC.GM.05.Pbt 05GM_X02
MND-2.CM.98.CM16
MND-2.GA.x.M14
MND-2.x.x.5440
DRL.x.x.FAO
MON.CM.99.L1
MON.NG.x.NG1
GSN.CM.99.CN166
GSN.CM.99.CN71
TAL.CM.00.266
TAL.CM.01.8023
MUS-1.CM.01.1085
MUS-1.CM.01.CM1239
MUS-2.CM.01.CM1246
MUS-2.CM.01.CM2500
SYK.KE.x.KE51
SYK.KE.x.SYK173
RCM.GA.x.GAB1
RCM.NG.x.NG411
GRV.ET.x.GRI 677
SAB.SN.x.SAB1C
TAN.UG.x.TAN1
VER.DE.x.AGM3
VER.KE.x.9063
VER.KE.x.AGM155
VER.KE.x.TYO1_patent
COL.CM.x.CGU1

```
                              Vpu end
                             Env signal peptide end Env gp120 start
H1B.FR.83.HXB2             TGTAGT......GCTACAGAAAAATTG..............TGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCAACC........ACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTG 6445
Vpu                        L  *
Env                        _C_ S_ . _. _A_ T_ E_ K_ L_ . _. _. _. _. _. _W_ V_ T_ V_ Y_ Y_ G_ V_ P_ V_ W_ K_ E_ A_ T_ . _. _. _T_ T_ L_ F_ C_ A_ S_ D_ A_ K_ A_ Y_ D_ T_ E_ V_ H_ N_ V_ W_ A_ T_ H_ A_ C
H1A1.UG.85.U455            ---.........AA-G--C--C---...............--------G-----C-------------------A--T---GTT.........-----C-------------------------G---A--G--------C-----T----------- 5890
H1B.US.90.WEAU160          ------.......---G-------C---...............---------------------------------A--------..........-------------------------------A---------------------------- 6440
H1C.ET.86.ETH2220          ----A-......-GA-TG-GG--C---...............-------------------------A-----A--T--T-G-.........C-T---------------------------A-----------C----GT---TT------ 5854
H1D.CD.84.84ZR085          -A----.......---G----T--C---...............-----------T-----------------A-----A--------..........-----G-----------------T-----A-A-------C-------A-C-----T---------- 5966
H1F1.BE.93.VI850           ----A-.......---G----C--C---...............---------------------------------A--------.........--T--------C-------------------A-G-----C------C----------------- 5791
H1G.SE.93.SE6165           ------.......--CT---A-T--T---...............-----T-------------------------G-A---T---GAT.........--------------------C---------T-----AG-G-T--AAGC---------C-----T---------- 5848
H1H.CF.90.056              ------.......---G--C----C---...............-----T-----A--------------------A--G----AA.........-------------C------------------G--------G-----AAAG---------C-----T--------A-- 5785
H1J.SE.93.SE7887           ----AG.......--A-A----G-C---...............---------------------------------A--T----AG.........-------T-------C-----T------------AG---T--AAA---------C-----T--------T-- 5764
H1K.CM.96.MP535            -----C......AA-G----C------...............----------------------------------A--------A.........C-----------C-------------------G-AG------------C-----T--C---------- 5644
H101_AE.TH.90.CM240        ------.......--CT----C--C---...............-----T-----T----------G----------GA--T---GAT.........-----C-----------------C--G---C----G-----A--G--C------C----------------- 6013
H102_AG.NG.x.IBNG          ---.........AA-G-T---C------...............--------G-----C------------------ACG---GAG.........-----C----------------------------------A-----------C-----T---------- 5972
H103_AB.RU.97.KAL153_2     ------.......-----------T--A...............-----------T-----------------------A--------..........--------------------------------AG--AG------------------------T------- 5658
H104_cpx.CY.94.CY032       ------.......--CT---A-C--C---...............-----------T----------------------GA--C---GAG.........-----C-----------------A-------------G-A---A---------A-C-----T---------- 5815
H1N.CM.95.YBF30            AT-G-G.........T-T---C--CAT...............-----A-----G--C----------A--A-----GA------GAG.........--A-----T--C-----T--------------CC--AG--------CT--C--CA-C-------------A--A-- 6032
H1O.BE.87.ANT70            ---TTGAGC......CTTAG-C-GC-A..............-AT-CA---------GC------G-----A---G-A---T-----A.........C-AGTA-----C-----T------------CCT-ACAAGC--T--AAAG-------A-------AT----A----- 6493
H1O.CM.91.MVP5180          ---TTGAGT...TA--GTA--C-----...............-AT-C------T-----C------------A---G-A--G---G-A.........C-AGTA-----C-----T------------CCT-ACAAGC--T--ACAG-------A-------AT----A----- 6483
H1P.FR.06.RBF168           ---GC-AGT...-GA-T-C-C................-AT--T--G--A-----------T--A--T-----GAA-T--TGAA.........GTT--CT-G--C-------G--------ATCGCT-AC-AGC-A----CAG-------A------------C--G----- 5921
CPZ.CM.01.SIVcpzCAM13      ---TTGACA.........TCTG-TGAA...............-----A-----G--------A-----------------GA----TC-A-..........--AGTA--C--C-----T-----------------C--AGC----------CT-----CA-A------------C--A-- 6023
CPZ.CM.05.SIVcpzEK505      AT-G-G............TCTG-----...............-----A-----G--C---------------C----GA--T---GAG.........--AGT------C-----T-------------G--CC--AG---------CT--C----A-----------G--A--A-- 5994
CPZ.CM.05.SIVcpzLB7        --CTCATTG...ATA-TTAGT-GCACTAAAAATGAAAATTTA-----G-----G--C-----A--G-----C------A-----GAA.........G-A------C--------G----------A------C-AA-GG------C---C---A-A-------T-T--A--T-- 6010
CPZ.CM.05.SIVcpzMB66       ---GAA...........--T--CA--...............-----G-----T--C-----A---------C----GA--G---GA-..........--A-----G-------------------A------C-GA-CC----A--G--C------A------TT-T--A--T-- 6027
CPZ.CM.05.SIVcpzMT145      ---TT-GGT......GA-------T-G-...............---------T--T----------C--A--A-----G---G----AA.........-------C--T--C-------C--C-------GAGT---AG------A-CT------CA-C----------C--A--T-- 5990
CPZ.CM.-.SIVcpzMB897       G-GGT-ATTAGAT--GA-A-T--T--A...............-----A-----G--C-----A--------------A-----A--G----AA.........--A------T--------------------C---------C-GATAG-T----C---C--CA-A------TT-T--A--T-- 6016
CPZ.CM.-.SIVcpzDP943       ---TTGACTGGT...-GT--TC-G---...............----CA-----G--C-----A-------------------A--T--GGAG.........--A--A--C--------------G--------TC-----CT-A-C-A---A-CC--C---A-----------T-G--A----- 6023
CPZ.CM.98.CAM3             ---TTGACT...-G--GT--TC-G---...............----CA-----G--C-------------------------A--G---GA-.........--A--A--C------------C------G--GTC---TCT-A-C-A----CC------A------------T----A----- 5864
CPZ.CM.98.CAM5             ---TTGACT...-GCCAT--G...-G-...............----CA-----G--C-----A---------A-----A--G---GAA.........--A--A--T--C---------C------G--CTC----CT-A-C-A---A-CC--C---A------------T----A----- 6149
CPZ.GA.88.GAB1             ---TTGACC.......T-T--G...--A...............-----A-----A--------A---------T---C-T--T--TGA-.........C-GGTA--C-----------C------C-----G---C--AG------------CT-------A-------------G--A-- 6509
CPZ.GA.88.GAB2             ---TTGAAGACA-AAT-T--TC---G-...............-----A-----A--C-----------T--A--C---------G--T--T.........--AC-AT-G--C-------G--A-----A-C---T-TAA-AGAG--ACCT-------A---------------A----- 5850
CPZ.US.85.CPZUS            ......CCA...-GATTGTCT-GC--A...............----CT-----A------------------A-----GA--T-T-GAG.........--A--CT----C-------C--T-----A--G------CA-GCAG-----CC--C--CA-----------------G--A-- 6516
CPZ.CD.90.ANT              ATAGAGAAG...-GG--GA-TG--GAC...............-AT--A-----A-TC-----A--C------C----GAA-T--G--A.........C-T----------------CA--A----CTCCATGACAAG----------G--C------A-----A--TACCAGT-- 5871
CPZ.TZ.00.TAN1             CTAG-GATT...-GATTTAGC-CT-AT...............-ACAC------G-T------A--------T------A--G--CCAA.........C-A--CT-G----------C--T------G-TATTAC-AG--G---TAA---C--CA-A-----A--------AA-- 6065
CPZ.TZ.2001.TAN2           G-G-T-GGA......T-TA--G--AT...............-ATAC-------A-T----------------CA-T------A--G--T--A.........C--T-CT----------T--TA-----G-TATCGC-AGC-G---TAA---C--A-A------A-------AAT-- 6070
CPZ.TZ.2002.TAN3_1         G-G-T-GGA......T-TAG-G-C-AT...............-ATACT------A-TC-----T-----CA-T------A--T--T--A.........C-AT--T----------T--T------G-T-TCAC-AG--GG---AA---C--A-A------A--------AA-- 6070
GOR.CM.2007.SIVgor2139_287 --C-T-GAG.........-CT-CGCAT...............-AT-CT--T--G--------A--G--A--C---CGA--G----AT.........GTA---T-----C-------G----------TTATGTGAGC-A---ACA---C---A---------A--G--G--A-- 5979
GOR.CM.2007.SIVgorCP2135con --C-T-GAG.........-CT-CGCAT...............-AT-CT--T--G--------A--G--A--C---C-A--G----AT.........GTA---T-----C-------G----------TTATGTGAGCCAG--ACA---C---A-A-------A--G--G--A-- 5978
GOR.CM.2004.SIVgorCP684con --C-T-GAG......-AGACTG-GCAT...............-AT-CT--T--A--------A--G--A--C---CGA--G----AA.........GTA---T-----C-------G--------C-GTTTTGTGAGC-A---ACA---C---A---------A--G--G--A-- 5904
H2A.DE.x.BEN               GC-T-CTTA...-TATATTGT-GCCA-...............-AT--G--T--T-TC-------CA----C-C------AA-T---T-T.........-TTC-CT------------------------------TA-AAATAGAG-CAC-----GG--CATACAG-- 6873
H2A.GW.x.ALI               GC-T-CTTA...-TATATTGT---CAA...............-AT--G--T--T-TC------C--G--A-CA-----AA-T---T--.........-TTC-C--C-----------------------------CA-AAATAGAG--AC-----GA--CATACAG-- 6863
H2A.SN.x.ST                GC-T-CTTA...ATATATTGCGTCCAA...............-AT--G--T--T-TC------C--G--C--------GAA-T---T--.........-TTC-C--C-----------------------------TA-AAATAGAG--AC-----GA--CATACAG-- 6315
H2B.CI.x.EHO               ATCTA-GGG...TA--TG-GC--GAAC...............-TT-----T----TC------TA----C-CA-----AA-T---T-A.........-TTC-C--C-----------T..................CAGAAACAGAG--ACC----GA--TGTACAA-- 6838
H2B.GH.86.D205             ATC---GGG...TT-GT-TGT---CAA...............-AT--T--T----TC------CA----C-CA----G-A-C-----A.........GTTC-C--CA----------------------------CACAAACAGAG-CACC----GA--TGTACAG-- 6840
H2G.CI.92.ABT96            AC-TA-GGG...TTCTATTGT-CCAAA...............-AT-----T--T-T------CA----C-CA----G-A-C--G-G-.........GTAC----C--C----------------------------CA-AAATAGGG-CACA----G---TACACAG-- 6186
H2U.FR.96.12034            GCAT-CGTG...ATATACTGT-CTCA-...............-AT-----T----T-------CA------CA----G-A-C--G--A.........-TTC-C--C-----------------------------CA-CAATAGAG--ACA----GA--CACACAG-- 6327
SMM.SL.92.SL92B            GTGTTAGGG...A-CTGGTGTGC-CA-...............-AT--A---A---T-------TA-T----CA----G-A-C--T--G.........-T-C-C--C--C------G.....................CC--AATAGAG-CACA----GG--CGT-CAG-- 6200
SMM.US.x.H9                GC-TCAGGG...AT-TATTGTGTTCAA...............-AT--A-----A-TC-----T-----A-CA-----G-A-T--G--A.........GTTC-C--C--C------------------------CAGAAATAGGG--AC-----GA--CACACAA-- 6241
SMM.US.x.PGM53             GC-TA-GGG...ATCTATTGC-CTCAA...............-AT--------T-T------T-----A-CA------A-T--G--A.........-TTC-C--C--C-----------------------------CA--AATAGGG-CAC-----GA---AC-CAA-- 6692
SMM.US.x.SIVsmH635F_L3     GT-TTAGAG...AT-TGTTGTGTTCAA...............-AT--A-----A-TC-----T-----A-CA------A-T--G--A.........-TTC-C--C--C-----------------------------CA--AATAGGG-CAC-----GA---ACACAA-- 6771
                               Env signal peptide end Env gp120 start                                                      V1 loop start

MAC.US.x.239               GTCTA-GGG...ATCTATTGT-CTC-A...............-AT----------T------T-----A-CT----G-A-T--G--A.........-TTC-C--C-----------------------------------CA--AATAGGG--AC-----GA---AC-CAG-- 6773
Env                        _V_ Y_ G_ . _I_ Y_ C_ T_ L_ . _. _. _. _. _. _Y_ V_ T_ V_ F_ Y_ G_ V_ P_ A_ W_ R_ N_ A_ T_ . _. _. _I_ P_ L_ F_ C_ A_ . _. _. _. _. _. _. _. _. _T_ K_ N_ R_ D_ T_ W_ G_ T_ T_ Q_ C
MAC.US.x.251_1A11          GTCTA-GGG...ATCTATTGT-CTC-A...............-AT----------T------T-----A-CT----G-A-T--G--G.........-TTC-C--C--C-----------------------------CA--AATAGGG--AC-----GA---AC-CAG-- 6772
MAC.US.x.251_BK28          GTCTA-GGG...ATCTATTGT-CTCAA...............-AT----------T------T-----A-CT----G-A-T--G--A.........-TTC-C--C--C-----------------------------CA--AATAGGG--AC-----GA---AC-CAG-- 6749
MAC.US.x.EMBL_3            GTCTA-GGG...ATCTATTGT-CTCAA...............-AT----------T------T-----A-CT----G-A-T--G--A.........-TTC-C--C--C-----------------------------CA--AATAGGG--AC-----GA---AC-CAG-- 5898
STM.US.x.STM               GC-T-CTTG...A-CTATTGC-CTCA-...............-AT-----G--A-T------T-----A-CA----G-A-T--G--G.........-TTC-C--C--C-----------------------------CA--AATAGGG--AC-----GA---AC-CAA-- 6415
MNE.US.x.MNE027            GCCTA-GGG...ATCTATTGC-TTCAA...............-AT----------T------T-----A-CT----G-A-T--G--A.........-TTC-C--C--C-----T---------------------CAG-AATAGGG--AC-----GA---AC-CAG-- 6241
DEB.CM.99.CM40             CAGG-GAAA...CAG-A-A-GC--CAA...............-ACT-G-----A--------A-----A--C---GTA--T--C-AA.........GTTGA-T-G--C---A--G-CA-CT-ATC--AG...............TCA-GA-----AGTGACA--A-- 6125
DEB.CM.99.CM5              G-GGCAGCA...C-A---A-GC--CAA...............------G-----A--------A--G--C--A---C-CA-T--C--T.........GTTGGCT----C---A--G-TA--AG-TC--A-...............TCA-GA-----TGTGACA--A-- 6094
DEN.CD.x.CD1               ATA-TATGT...A-A--------GCA-...............-AC--------A-TC-----CA------CA----G-A-CAGTG-A.........G-TC-G--G--C--CA--.....................G-GAT--ACAGAGGA--A-----AG--GCAT-A-- 6439
LST.CD.88.447              CAAG-GGAG...AAA---A-G--GCA-...............-AT--A-----G-TC--------G---AAA---G-A------GTA.........GTGC-AT--A-------T....................T---CTAATA----G-T-A----T----ACAT-A-- 5454
LST.CD.88.485              CAAG-GGAG...--A...A-G--GCA-...............-AT--A-----A-T---------G---AA----G-A--T--TGTA.........GTGC-AT--A-------T....................T---CTAATA---GCT-A----T----ACAT-A-- 5451
LST.CD.88.524              ACAGCAGGA...-AA-AG--G---CAA...............-AT--A-----A-TC-----A-----AAAA---G-A--T--TGTA.........GTAC--T--A-C-----T....................T---CTAATA---GCT-G----T---TACAT-T-- 5454
LST.KE.x.lho7              G-C-TAATA......TGG-G----CAA...............-AT--A-----A-T----------G--AAAT---G-T---TAATGTTTCA.........GTGC-CT-GA-------T....................T---CCAATAC--GCC-C----T---GACAT-T-- 6532
SUN.GA.98.L14              CTGTTAGGT...AG--T--C--T-CA-...............-AT--A-----A-TC------ACT---AA----G-ACC----GTA.........GTGC-AT--A-A-----T....................T---CTAATA---GCT-G----TG--TAC-T-A-- 6579
MND-1.GA.x.MNDGB1          ......GGT...AT--T--G----CAA...............-AT--G-------TC-----A-----A--A-----------T-AA.........--ACA-T-GA-------T.........................TAATTCA-G-C-C-----TA--CAC-AAT-- 5899
OLC.CI.97.97CI12           ATCTTAGGAATCAG--T--GG...CAA...............-AT--G-----G-TC-----AAC---AAA-----GA--TAGT-AT.........GAT--GT--A-G----TG........................CTAACA-A-G-C-G-----T----ACA--T-- 5952
WRC.CI.97.97CI14           ATAGC-ATC...AG-TT--CTGT-CAA...............-AT--A-----G--C-----A-----ACA----G----CAGT--A.........GA-TTC---C-G------------------------------AAATAGATCAT-A----TA-GCACACG--- 6444
WRC.CI.98.98CI04           ATAG-ATTA...-GG-T--GG---CAA...............-ACA-A-----G--C-------------CAA---G----GAGT-AA.........GATTTCT-GC-C------------------------------AAACAGATC-C-G----T--GCAC-CA--- 6459
WRC.GM.05.Pbt_05GM_X02     ATAG-ATTA...-GATT--G---GCAA...............-ACA-A-----G-------------G--ACA----G---------AG.........GA-TT-A-G--G------------------------------AATAGATC-C-----TA---ACACAT-- 5382
MND-2.CM.98.CM16           GTAG-ACTA...-GGTT--G---CCAA...............------G-----A--------CAC----AAA---C-TA-G---GAG.........--ACAC--G--C-------------------------------AAATG--TCCT-A----TA---AC-AGT-- 6460
MND-2.GA.x.M14             ATAG--ATA...-G--T--G----CAA...............------A-----A--------CAC----AAA---C-CCC-----GA.........--GCA---T-----------------------------T--CAACA-CTCCC-------TA--CACAAGT-- 6389
MND-2.x.x.5440             GTA---GTA...ATAGT--G-G--CAA...............------G-----G--C-----CAC---CAAA---C-TA-G----GG.........--ACAC--G--------------------------------CAATA--TCAT-C----TG--CACAAG--- 6033
DRL.x.x.FAO                G-A-TAAGT...ATAGG-...---CAT...............------A--G--G-TC-----AACC--CAAA----GACC----T--.........--ACA---GA-C-----T.....................--G--TAAT---TCAT-C----TG--TACAAGT-- 6082
MON.CM.99.L1               ATG-C-TTG...-TG-AT-GGTCTAAC...............---AC---C-----C--C--T--T--A--A-----ACCG-----T.........C-TC----C---------C--T---C-A--TTACGGGTC--A------C-GGG--CAA----CTGG-GTC-T---- 6360
MON.NG.x.NG1               G-GTC-RAT...T-AGG-A-C----CA...............-----------A-TC-----A--G--G--C-AT--ACCT--G--G.........C-TC-AT-G--C-----T-----CC--G-YTATGGCTCA-A------C-GGG---AA----CTAG--TC-T---- 4973
GSN.CM.99.CN166            ATCCCGGGT...TTG-AGTGTG--AAT...............---ACG-----G--C-------------G--A----G---C-----T.........C-AC-AT-G-----C--C------C-AG-TAT-GC-AG--AT--ACC-GGA--CA-----ATTT-TACA--A-- 6334
GSN.CM.99.CN71             CCGG--TCG...AGGT-TA--GGGAAC...............---ACG-----G-TC-------------G--A----GA--T----AA.........C-AC-AT-G--C-----C---------AG-CAT-AC-AG--AT--ACC-GGA--CA-----AT-T-TACG--A-- 6319
TAL.CM.00.266              G-A--GTTT...ATA-T--GG--CCA-...............-AT-----CA---------A-----A---------AA-T-----A.........C--C-AT-G-------C..........................CAT--C-AGGG-CCAA---T-A---TCCCAA-- 6197
TAL.CM.01.8023             A--CT-ACC...T-CG-----C-GCA-...............-TCC-------G--C--C--A--G--A--A------AA-T-----G.........C-TC-AT-G--C--C--C..........................TAT--C-AGGG-CCAG---T-T---TCCCAA-- 5717
MUS-1.CM.01.1085           A-----ATA...AT-TGG-G--G-CAA...............------------T---------------G---------A--T--T---.........C-AC-A--C--C---------------C--CAT-GCAA---G---ACC-GGG------A---ATA--CACA--A-- 6289
MUS-1.CM.01.CM1239         A-----ATA...ATATGG-GC--TCA-...............-----------G-TC-------------G--A------A--T-----T.........C-AC-A--C--C------------------CAT-GCAA-C-GG--ACC-GGG------G---AT---CACA--A-- 6276
MUS-2.CM.01.CM1246         ATATTGAGA...AAACAT-CC--GCAA...............--------G--G-T----A-------G--A-CA---A-T--C--A.........C-AC----C--C---------A-------TAT-GCAA-C-G----CC-GGG--C--G-T-ATA--GACA--A-- 6408
MUS-2.CM.01.CM2500         ------CTG...ATATGG-G----CA-...............-----------G-T------A--C--G--C--------T-----A.........C-AC-AT----C-----C-----------CAT-GCGA-C-GG---CC-GGG------G---ATA---ACA--A-- 6352
SYK.KE.x.KE51              G-C-TAGTA...----GGTT-G--CAA...............-AT--------A-T------A--T--CAAT---G-A--T--TCAA.........GTAC--A-G------------------------------CCCCCATACGGGA-G------A---A-GAA--- 6072
SYK.KE.x.SYK173            -TAG-ATTT...ATGGA-A--C--CAA...............-AT--A-----A-T-------A-----ACAT---G-A--T---TAT.........G-AC----C------A-C.......................CTCCCATA-AGGA-GA-----T--TA-GAAT-- 6394
RCM.GA.x.GAB1              GTACAAGGA...T-ACA-A--CCGCAA...............-ATA-A-------TC-----T--C--A--A----GAA-CAGC--A.........GTGC--A-G--------TG.......................T--TAATAC-C-ATCA----GA--T-TAAAT-- 6072
RCM.NG.x.NG411             ATA-T-ATA...-TAGG---G---CAA...............-AT--G-----A-T------T--C--A--A------A-CAGT--A.........GTTC--A-G--------T.......................T--TAACGCAC-ATCA----GA--GTTAAAT-- 6128
GRV.ET.x.GRI_677           ATA...GGA......-T--GT--CC--...............-AT--G-----G-T------AA-C--A--A-----AA-TT----A.........GTTCAGGC------CATG......................GCCCAATACC---A-G-----A--CACCAA--- 6365
SAB.SN.x.SAB1C             -TAGTCCAG.....................................-AT--T-----G-T------TA-C--A--A-----AA-TAGTT-A.........GT-CAAGCC------AAG......................CCCCAATACG--CC-A------T-TAC-AAT-- 6641
TAN.UG.x.TAN1              ATAG-ACTG...TTATAT-GG-C-CA-...............-ATA-T-----G-T------TA-C--A--A-----AA-CAGCT-A.........GTGCAAGCT--C---ATG......................CCTAATACC---CC--------A---AC-AAT-- 6442
VER.DE.x.AGM3              GTAGTGCTT...AA----AGGC--CAA...............------------A-T------A-----A--A-----AA-CAGCT-A.........GTACAGGCT--C--CATG......................CCCACCACC-GAC-A-----A--TAC-AA--C 5944
VER.KE.x.9063              -CCTTA......TTA-GT----G-CAA...............-----------G-TC------A-------A------A-CAGCT-A.........GTACAAGCC--C---ATG......................TCCTACTACC-GGC-A-----T---AC-AAT-- 6446
VER.KE.x.AGM155            ATAG-GATA...AG----A--C-GCA-...............---A-A-----G-TC------A-----A--A-----AA-CAGCT-A.........GT-CAAGCT-----CATG......................CCTACTAC--GGT-G-----A--TAC-AAT-- 6444
VER.KE.x.TYO1_patent       ATA.........-TGTT-AGT--GCAA...............---A-A-------T------AA----A--A-----AA-CAGCT-G.........GTGCAGGCT--C--CATG......................TCCCACCACA-GCT-A-----T--TAC-AAT-- 6443
COL.CM.x.CGU1              G---TAGTG...TTA...AGT--TAAA...............------A-----A---C-A--A-------CA---G-A------GATGTAAATGATCAGCAGT-C--C--CTTT--TTCCT-CCC--AG....................A--CAACAGGTGTTG-GG-- 5801
```

```
H1B.FR.83.HXB2           TGTACCC...ACAGACCCCAACCCACAAGAA........GTAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGG......AAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCAC 6597
Env                      _V__P__._T__D__P__N__P__Q__E__.__.__._V__V__L__V__N__V__T__E__N__F__N__M__W__.__._K__N__D__M__V__E__Q__M__H__E__D__I__I__S__L__W__D__Q__S__L__K__P__C__V__K__L__T__P_
H1A1.UG.85.U455          -------..--------------------........A---AT----C-----------G-A-----------......------A----------T-----------A--------C------C-------------------------G-------T- 6042
H1B.US.90.WEAU160        -------..--------------------........------------A---------------------.....------A------------------------A--------T----------------------A-------------------- 6592
H1C.ET.86.ETH2220        ---C---..-----------G---------........T-G-GT----A----------------------.....--------------G--G---------C-----------------------------G-------------------G--G------- 6006
H1D.CD.84.84ZR085        -------..--------------------........A---A-C--AA------C----------------.....-----CA-------G--C-----------------------T-----------T-----A---------------------------- 6118
H1F1.BE.93.VI850         -------..--------------------........---T-TC--AA------A------------G-T------.....------A-------------A------ACA-------------------------------------------G---------- 5943
H1G.SE.93.SE6165         -------..--------------------........A--A-TA---A------A------T---------.....------A-------G-----A--------------------------G------------------------------GC-------T- 6000
H1H.CF.90.056            -------..------------------G........A-G--CA---AG-----A-----G-GC-----T-------.....G-----A-------G--G---------ACA--C-----------------------T-G--A--------------------- 5937
H1J.SE.93.SE7887         -------..------T--A-G---------G........A-GAATC--CC-------A---------C---------.....-----------------C--------G--A------------G--------G-----T-----------------GA-------T- 5916
H1K.CM.96.MP535          -------..--G---------------------........----AGA---A------A---------C---------.....------A-------------G--------ACA-------------------G-----------A--T-------G-G-------T- 5796
H101_AE.TH.90.CM240      -------..--------------------........A--CACC---A------A----------------.....------A-------------G--------G------G--------------------T---------------------G------T--T- 6165
H102_AG.NG.x.IBNG        -------..--------T-------------........A--CAT----A------A----------G----------.....------A-----------------------A---------T---C------C-------------------------G-------T- 6124
H103_AB.RU.97.KAL153_2   -------..--G--------G-----------........A--CC----AA-------------------------G--.....------A------------------------------------------------------------------------G---------- 5810
H104_cpx.CY.94.CY032     -------..--------------------........----CTC--A-------A-----G--C--------.....--------------------------------------------------A--G--G-------A-------C---GC------T-T- 5967
H1N.CM.95.YBF30          ---T--T..--T--T-----T----------........--GC-----ACCC-----A--T-----A-----T--------.....G-----A-A----C---C--A-----A---------T------C-G-----A--G---T-------C------T--------------T 6184
H1O.BE.87.ANT70          ---T--T..-----------CT---T-T---........TATCC-----CAC-----------TG-C-----T--A----.....-------T----------------A-----G--A--C--T--T-----------C--G--T--T--A--T-----TC--A-G--TTTC- 6645
H1O.CM.91.MVP5180        C--T--T..------T------T-----T---........T-TCC-C-A-GC-----------T--C----G-T--A----.....-------T----------G--C--A-----------A--C--C--T-----G-----A--G--TT--------T----AG---A-G--TTTCT 6635
H1P.FR.06.RBF168         ---G---..---------T-CA---AT---G........--CAA-A-------A-----GTC-------T--T-----.....-------T-T------GAC---A-----G--A--C--C--T--C-------------G------T--A--T--Y---------G--AATY- 6070
CPZ.CM.01.SIVcpzCAM13    ------T..---------T--T--T----------.......----TC-TCAT-----A-----GG-C-----T-------.....---T---C-G-----------------------A--A--C--C-GT---C-C------C--G--T--T--A---------------GC----T--TT 6175
CPZ.CM.05.SIVcpzEK505    C--C---..--T-----------------------........---C--A-ACC-------------CG----G-T-----.....------CA-T--------T--A------A--A--C--T--T--C--G------A--G--TT----A--C-----T--G---------T 6146
CPZ.CM.05.SIVcpzLB7      -------..------T--TG-T--TA---C--........A---A-C-AAGC-----A-------------------T------.....C------CA-------A-T--------AC-----G----T--C--------------------A-------A-----GC--------T- 6162
CPZ.CM.05.SIVcpzMB66     ----T--..----A----T------G-G--G........A---A-C-AAC-------A--T---T-C--C--TGCA----.....G-----CA-T---------G--A--------------------G----A-C-T-----C--G--TG----A-----------------TT--TT 6179
CPZ.CM.05.SIVcpzMT145    ------T..--T--T----CT--T-----G........----C-C--ACCT------C--T----G-G--------------.....G------T------------------T--A------G--A--C------A----C-C-----A------T--T--A-----------------A--CT 6142
CPZ.CM.-.SIVcpzMB897     ---G--T..------T--T-----TG----------........A--ACG--A-A------A-----G-G---G-TGCA---.....---C---A---------------C--A------A--A---C-C--T--CC---------------------------------C-G------C- 6168
CPZ.CM.-.SIVcpzDP943     ---T---..-----------T-------------------........--GCATG-ACC----A-A--T----CC----G--------.....G------A---------------A------GACT------T--T--C------------------T------A---------T--G--G--A---- 6175
CPZ.CM.98.CAM3           ---T--T..-----------T-------------------........--GCAGA-CCC------A--T------C-----T-------.....---C---AC-----C--------A---A--ACT--C--T--T---C-G-----------TT-G--A---------T--GC-G--G---A 6016
CPZ.CM.98.CAM5           ---C--T..GT---T--T-----C-----G........--GTATA--CC-------AT-T----GC-----------------.....---------A-T-----C---------A------G--A------T-GT---C-G-------------T-------A---------T---C-G--A---A 6301
CPZ.GA.88.GAB1           ------T..-------T----GT--T--G----........--T-TC-TCC-------A-T-----TCA----------------.....---------A-T------G--C--A---------A--C--T--------C-C-------------G--T--C------T----------G--G------TT 6661
CPZ.GA.88.GAB2           ---G--T..----------TCT---G----G........--CA-CC-AA-T--------------G-A------TG-A-----.....G-C-----CT----------G--A--------G---------T-CT---CC--------C--G------T-GA--T---------------G-A----TT 6002
CPZ.US.85.CPZUS          ---C--T..--T-------T-----------------........--CAT---CC-----------T-----G---G-------------.....G------A-T------C------A------G--A------T--T--CC--------C--G----T----A--T---A----G----------T 6668
CPZ.CD.90.ANT            ---G--A...-T---T--AG-T--TATT..............---TAG-C-C---...--CTC-GTC-GGTTT-ATGCTTAT.......--------T-T----------AGT---ACA--A-----GN-ACAA----TCC-A--------AT-------------------C----A--TA 6020
CPZ.TZ.00.TAN1           ---G--T...TT---T-----T--TT-T-----------........--ACCC-A-CC------T--AT--GG------T---GAA.....G-------T---------GC--G------A-A--A------T-ATCAC-T-TTC-A--G--TT-T-------T------------------A---T 6217
CPZ.TZ.2001.TAN2         --------...TT---T--T--T--TT-T--G........--CCYC-A-C-------T--AT-SMY---G-T----GAA.....--G---T----------R-G-A---A-A--A---T--C-TTCAC-T-TCC-A------TT-T--A--C---------------------A---T 6222
CPZ.TZ.2002.TAN3_1       ---T---...TTG--T------T--TT-T--G........--GCCTA-A-C-------AT--AT---A-----T---GAA.....G-G---T-T---------A--G-A---A-A--A---T--C--TCAC-T-TCC-A------TT-T--A--C-----------------------A---T 6222
GOR.CM.2007.SIVgor2139_287 ---G--T..--C-----A-GG---AC---G........T-TCC-----A-------A--------CC------T--C------------.....---------T-----------------T--A------G--T-------G-T--C-------------G------T--------T-----------C---AGTTA 6131
GOR.CM.2007.SIVgorCP2135con ---G--T..--C-----A-G----AT---G........T-TCC-----A--R---A-----G-CC-----G-T--C-----------.....---C---A-----------T--A------G--T-----------T--C-------------G------T--------T-----------C---AGTTA 6130
GOR.CM.2004.SIVgorCP684con ---G--T..--C-----A-G----AC---G........T-TC------A-------A-----G-CC-----G-T--C-----------.....---------T-----------------G-----------------T--C-------------G------T--------T-----------C---AGTCA 6056
H2A.DE.x.BEN             CT-G--A...GACA-TGATG-TTAT--G----------........A--A-TT-------------GGC-----G-TGCA------.......--T---ACAG--AC------AGCAGTA--A---G-CTGGCA-C---TT---GAC-TCAA----A-----------C--GC----A--T- 7022
H2A.GW.x.ALI             CT----A...GACA-TGATG-TTAT--G----------........A--GCTT-G-----------GGC----CG-TGCA------.......G-T---ACAG-AAC------AGCAGTA--A---G-CTGG--AC---TT---GAC-TCAA----A-----------C--G------A--TT 7012
H2A.SN.x.ST              CT-G--A...GACA-TGATG--TAT--G----------........A--GCTT--------------GGCC---CG--GCA------.......--T---ACAG-AAC------AGCAGTA--A---G-CTGG---C---TT---GAC-TCAA----A-------C--C---C-----A--CT 6464
H2B.CI.x.EHO             CC-C--A...GATA-TGATG--TATACC-----------........A-CCAAT-----A-A------GGC-----G-TGCA------.......G-T---ACAG--AC---T---GCAACAA-----G-GTGG---C-C-TT--GACCTCAA----A-----------T------------------ 6987
H2B.GH.86.D205           -C-C--A...GACA-TGGTG--TACACT--G........A-CAG-C-----CA-A-----GGC-----G-TGCA------.......G-T---ACAG--AC-C------GCAGTA--T---G-GTGG--AC-C-TT--AACCTC-A-----A-----------C---C------------ 6989
H2G.CI.92.ABT96          CC-T--T...GATA-TGATG-TTACAGT-----------........C--GCAT-G-----A-------GC-----G-TGCC-------.......G-T--CACAG-AAC--------GCAGTA--A---G-TTGG-A-C-G-TT--GAC-TCTAC---A--T------T---------A--T- 6335
H2U.FR.96.12034          CC-T--T...GATA-TGGTG-TTATTC---G........T--GCAT-------C-------GGCC-----G-TGCC-------.......--T--CACAG-AAC---G--AGCAGTG--A---G-CTGG---C-T-TT--GAC-TCTA----A------------G--G--A--C- 6476
SMM.SL.92.SL92B          CT-G--A...GATA-TGGAG--TATTC-----------........T-GGCCC-T------T------GGCC----CG-TGCT---GAT......---ACAG-AACT------AGCAATA--A---G-TTGG-A-C-C-TT--AAC-TCTA-T--A--T------------G--------T---T 6349
SMM.US.x.H9              CT-G--A...GATA-TGGTG-TTACTC-----------........T-GGCAA-T------C------GGC-----G-TGCT---GAT......---ACAG-CAC-------AGCAATA------G--TGG-ACC-C-TT--AAC-TC-A-T------C-----------C-C-------- 6390
SMM.US.x.PGM53           CT-G--A...GATA-TGATG-TTATTC-----------........T-GGCCA-------T---------GA---G-TGCC---AAC......---ACAG-CAC-------GCAATA------G--TGG-A-C-C-TT--GACCTCTA----A--C------------------------CT 6841
SMM.US.x.SIVsmH635F_L3   CT-G--T...GATA-TGATG-TTACTC-----------........T-GGCA--C---A-C------GGC-----G-TGCT---AAT......---ACAG-CAC-------AGCAATA------G-GTGG-ACC-C-TT--AAC-TC-A-T-----C---------C-T-------- 6920
                 V1 loop end                                                                                                                                                    V2 loop start
MAC.US.x.239             CC----A...GATA-TGGTG-TTATTC-----------........--GGCCC-T-----T--------GC---G-TGCC----------.......--T---ACAG-CAC--------GCAATA------G---TGGCAAC-C-TT--GACCTCAA--------T------------T-----T 6922
Env                      __L__P__._D__N__G__D__Y__S__E__.__.__.__._V__A__L__N__V__T__E__S__F__D__A__W__.__._N__N__T__V__T__E__Q__A__I__E__D__V__W__Q__L__F__E__T__S__I__K__P__C__V__K__L__S__P__
MAC.US.x.251_1A11        CC----A...GATA-TGGTG-TTATTC-----------........T-GGCCC-T-----T--------GC---G-TGCT---GAG......---ACAG-CAC--------GCAATA--A---G--TGGCAAC-C-TT--GACCTCAA--------T-------------------T 6921
MAC.US.x.251_BK28        CC----A...GATA-TGGTG-TTATTC-----------........T-GGCCC-T-----T--------GC---G-TGCT---GAG......---ACAG-CAC--------GCAATA-----CG--TGGCAAC-C-TT--GACCTCAA--------T------------T-----T 6898
MAC.US.x.EMBL_3          CC----A...GATA-TGGTG-TTATTC-----------........T-GGCCC-T-----T--------GC---G-TGCT---GAG......---ACAG-CAC--------GCAATA-----CG--TGGCAAC-C-TT--GACCTCAA--------T------------T-----T 6047
STM.US.x.STM             -C----A...GACA-TGGTG-TTACTC-----------........T-GGCAA-------C-----GGC----G-TGCC---GAC......---ACAG-CAC-------AGCAATA-----CG--CTGG-ACC-T-TT--AAC-TC-A-T--A--T-------------G--T---- 6564
MNE.US.x.MNE027          CC----A...GATA-TGATG-TTATTC-----------........T-GGCCC-T---A-T--------GC---G-TGCT---GAG......---ACAG-CAC--------GCAATA------G--TGGCA-C-C-TT--GACCTCAA--------T------------------T 6390
DEB.CM.99.CM40           CC-G--ACATG-CTTAGTG-GAGA-G---TCCCAATGCCAAAT--GACCCA---C...TTTA-TGC----G-T-ACCCC.................-A-AG--G---C-ATGGC-------G-CT-----G-ACA-A--GTCAT-C--A------------GC----A---T 6277
DEB.CM.99.CM5            CT-G--ACATG--ATGGTT-GAGA-----TCCCTATGCCTAAT--G-CCCAG...TATTTT--CTC----TCA-ATGTT.................---A-C----AT-ATGG--C--C--G-CT--C----ACAGA---TCTT-T--A--C------------G---------T 6246
DEN.CD.x.CD1             ---T--T...GA-------T-T-A-AGG--GGTCTGGATG-A-AATGCCAC-G--TACTTT--C-TC-A-C---ATCCA.........TA-TAGACCAG-T-GT-CA-A--AT-A-AGAT......C-G-ATA-A--G-C-T-C--A--T-----C--------A---TG 6591
LST.CD.88.447            CA----A...GACTTG-AGTCTTATGC-C-G.........A-TCCTA-TCAC--CA-A-GTTT------C-CT-AAGA-ATAAAAG-T---C-A--TA--C-G--AGCATGGAGT-C---G-C----A-GGT----ACT-T-A-G-----T--------G---AT---T 5612
LST.CD.88.485            CC----A...GATTTA-AGTCTTATGC-C-G.........A-TCCTA-TCAC--CA-A-GTTT------C-CTCAAGA-ATAAAAG-T---C-A--CA--C-G--AGCATGGAGT-C---G-C----A-GGT----ACT-T-A-G-----T--------GA---AT---T 5609
LST.CD.88.524            -C-C--A...GATCTG-AATCTTATGCTC-G.........A-CCCTA-TCAT---A-T--TCT--------CTCAAGAAATTAAGG-C---C-A--AA--C-G---GCATGG-GA-C---G-CG--CA-GGT----GCT-TTA-G---------AC-----A---AT---TT 5612
LST.KE.x.lho7            CT-G--A...GATTTG-AATCTTATGC---G.........---CCTA-CTAT---A-TT-------C-----CA--ACCAGTAAAAG-C---C-GG-TA--C-G--AGCCTGGTCT-C---G-ATGC-A-GGTA--CAGT-TTA-G--A--------------A-C-AT---T 6690
SUN.GA.98.L14            CT-G--A...GATTTA-AA-CTTATGC------------..--GCCCA-AAC-GGGT-AGA---G-------CAGA-G-AATAAGC--T---C-A--A--GC-G--AGCATGGC-A-C---G-C-TC-A-GGTA--GCC-TTA-G-----T---------GA-C-AT---CT 6737
MND-1.GA.x.MNDGB1        CA----T.....TCATTGCCAGATT-T--TGAG.........-A-A-TCCTG--A-A-AG-----------CAGGACTTATAAGGG-----C-G--A--TT-T--AGCATGGC-T-C----GGGA---A-G-TA---ACC-TA--T-------------GA-T-A----T 6057
OLC.CI.97.97CI12         CC------...T-TTTTGATG-AGATG-GAT..........-A-CA-CCCA-CCC--GAGT----G-GG-ATTTTATAATGTGTCAG-C-----AG-A---AG---AGCATGGTTT-GC--TGCTTCC--G-TA---GC-GTT--G--A---------G--G-----T--TT 6110
WRC.CI.97.97CI14         -A-T--A...CATTTGGAGC-TTTGG----G.........CAGTACA--CCTG-GA-A-A----------C-CA-GAAAAGTAGCAG-----C-A--A--GC-G----GCATGGACA-GAG---CTTCAA-GATT--CACCTTAT------------------A-C--A--CT 6602
WRC.CI.98.98CI04         CA----A...CATATAGAGC-TTGGG---------........CA-TATC-TCCT------A-A--------------CT-G-AA-GTAGCAG-C---C-G--AA-CC----AGCATGGACA-GGG-T-CATC-A-GATT---ACT-TAT-----A--T------T--GA----A---T 6617
WRC.GM.05.Pbt_05GM_X02   -C----T...CATATAGAAG-AAT-AC--------........CA-TATA---A-GG-A-A-A----------CA-AAAA-GTGGAAG-------GG-A----C-G--AGCATGGACA-GAG-G-CT---A-GATA---ACTCTGT-G--A--C--------------A---AT---T 5540
MND-2.CM.98.CM16         ---T---AGCTTGTTG-A-T-TGA-G-GC---------........CC-A-TCC-AAC-TAACATGGA-TTT...C-CAGGTCCCATGGAAG-------GG-A---ATG---AGCATGG-GA-C---CTC-TCCA-GATA--CGC-GTAT----A--T-------G---C-G------T 6618
MND-2.GA.x.M14           ------A...GCTTATTGC-TTATG-G--G.........CA-C-CA-CCCC---A-A-------------CAG-CCCTATACAAG-G-----AG-----AGG--AGCATGG-GA-C---TTC-TC-A-GATA---GC-GTT-----A--G-----------G---A---T 6547
MND-2.x.x.5440           ---C----...GTCTGTTGC--TATG----------.........CAGCATA-CCCC---A-A------G---------CAGGCCCTATAGAAG-G-----A--A-GAC---AGCTTGG-GA-CA--TTCTTCCA-GATA---GC-GTAT-----A-----------GG------A---TT 6191
DRL.x.x.FAO              CA----A...GTTTATTGC-TTATG----------.........AC--A-A-TAAC---A-AGAG----------CTG-CCCTATGACAG--------GG--A--A-----GCATGG-GT-CGC-GTCATCCA-GATA---GC-GTG--T--A--------T----A-C-AT---TT 6240
MON.CM.99.L1             CC-C--G...--------T-C---G---TCT..........C---ATT-G--CA-C---T--GG-G--C--TGCT-ATCAA.......--CT-------G---G-A----GTA----------G-AGTC-C-G-TCTCG--GGCTT--------T-----G---------A--CA 6509
MON.NG.x.NG1             CC-C--T...T-----T--TGTT----GC-CT..........C-CCCTC-C---A-A-------GC------TGCA-ATAAC.......--CT-------G---G------GTG--A--C-----AATCAC-C-TCTCA--GGC-T------------------G---------TA 5122
GSN.CM.99.CN166          -T-G---...T-G-------TCA--GGC---G..........--GCC-T-----CA-A------G--A--C------C-AT..........---------T-----------TG--G----GA---T--C---GG-ATCAC---TTA-----GGC----G--A--------C------G--T---A 6483
GSN.CM.99.CN71           CT-G---...T---------TCA--GTCT--G..........--GCCAC-----CA-A------GG-A-----T--C-AT..........-----------T-T----------TG--G---AGA---T--C---GG-ATC--CC-C---TTA-----GGC------GA---------C---C-G--T--GA 6468
TAL.CM.00.266            C--C--A...T-----T--TC-GAAT-TC--G..........---CCA--G---C-C--T--GT-C---CG-TG---AT..............--T---------G--C--C--AGCA--A--C---GGAT---CC-C-TTTTGAC-GC-AC--GA-------ACT--G------A---A 6346
TAL.CM.01.8023           ------A...T---------TC-GAAT-T---G..........---GCCA--C----T-----T--GT-C-----G-TG---AT..............--CT---------G--C--C--AGCA--T------GGA-GCA--G-TCTTAAC-GCAACC-----T--CACC----C------T--TA 5866
MUS-1.CM.01.1085         CC-C-----...--G--T--TGCA--C--G-----------..........--GCA-T-G----A-CT-T-GGG-AAAGTTTGACATCTATAAG....---T-T---------C-------AC---T----G-G-GGTC-A-C-TTA----GGC-T----A---------C------------TA 6441
MUS-1.CM.01.CM1239       -C-T---...------T---GCA--C--G----------..........--CAAT------A-CT-T-GGG-A-A-TTT-ATGTCTATAAA....---T--------GACC-------GT----T----G----GGTCCA---TCA-C--GGC--------A--C------C------G------TA 6428
MUS-2.CM.01.CM1246       CC------...--G-------GCG------G--G..........--CAAC-----CA-CT--A-TG-GCACTTT-ATATATATAAG...--CT-T-----------T---------AC---T----G-G-AGTCCA-C-TTAC---GGC---T------G------G---C-----A--TA 6560
MUS-2.CM.01.CM2500       CT------...-----------AGC----T------G..........--CAAT-----CA-C-----GGG-A-A-TTTGATATCTATAAA...--CT-T------G--T---A---AC---T----G-GCGATCCA---TCA----GGCT---T--A--C----------------C-----T--CA 6504
SYK.KE.x.KE51            ------A...T-TTCTGAAC-GAT-G---TC..........AG-G-TAAT-T-AGTGGG---T-----C-CTGC-----AATTCTTCCC---G---AAG-C-----C-TTTGC-----C--GTCTGA-C-T-TCCTG--GGCTGAT-G---T-----------------T-T--TA 6227
SYK.KE.x.SYK173          ---G--G...T---CAGATC-AAT-G---T-...........AGGG-CAAC-T-ACAGG-----T-C----CCAGCA----AATAGC-GTC-CATG--AAGGC-G--A--ATTA------C--GTCAGC------TCTTG---GC-AAT-GA--C------G--GC--G-A--CA 6549
RCM.GA.x.GAB1            -A----A...GAG-GAGGA-TAT-T-C---G..........---T-CAA----------T------CGA----G-TGCT---AAT......---AG-T-ATAT------AGCAA-A--TA--G-GTGG-A-C-T-AT----TCC-CT------A------------C-G------G----T 6221
RCM.NG.x.NG411           -A----A...GA--GAGG--TAT---C-------------..........---CCA----------T-----G--A---G-TGCT---GAA......---AG-T--TAT------AGCAA-GTCTA--G-GTGGCA-C-T-AT---TCC-CTT-G--A-------------GG-------A---A 6277
GRV.ET.x.GRI_677         CA----A...GAT--T-ATG-TAATAC---G..........--GCCTC-A...--CA-T--------GC----CG-GGCT---GAT......----CCGC------A----AGCAG-GAGTA------CATCTAC-C-TT---A-----CGA-G-G---T------T--GC-CT-----CA 6514
SAB.SN.x.SAB1C           CA-C---...GAT--TGAACCAGA-GG-AC-ATAGCAGAA---CCCA-ACCT---A-C----------A---G-TGCT---AAAAATCGG--CCCTC-A----GC---AGCAG-GAGTA-C----CA-CT---G-TT---AAGT-CTT-------------T--GC-CT-A---A 6808
TAN.UG.x.TAN1            -A----A...GAT--T-ATG-TTATACT--------------..........---CAG--A...------CT-T--G--A---G-AGCA---AAGGAC-GG---CCAT-A--G-C-----GC-G-GAGTA-C----CATTTGC-C-TT--AAGC-CT--G----------C--------C-G--A--TA 6597
VER.DE.x.AGM3            GA----A...GAT--T-ATG--TACAC---G..........---CCAT----CA-C--T---CCA---G-AGCA---GCTGAC-G---CCC-T-A----C---AGCAGGAAGTA------CA-CTGC----TT--G--G-CT--G-------------------------T-A--TT 6099
VER.KE.x.9063            CA----A...GAT--T-ATG--TTATAC---G..........--GCC---A...--CA-T--T--GCCC----G-AGCC---GCAGAT-GG---CCAT-A--G-C-----GCAGGAAGTA-----TCATTTG--G-TT---A-----CAT-G--A-------C-----------T-T--T- 6601
VER.KE.x.AGM155          CA----A...GAT--T-ATG--TATAC---------------..........---CC-C----...---A-A-----GCCA---G-AGCA---GCAGAC-G----CC-T-A-----C---AGCAGGAAGTA-C--TCA-CTGC-G-TT--A--G-CAT--------C---------GC--T-A--T- 6599
VER.KE.x.TYO1_patent     CA----A...GAT--T-ATG--TATAC---------------..........---CCTC-A...---A-C--T---CCA---G-GGCA---GGTGAT-G----CCCAT-AA---C---AGCAGCCAGTA-C--CCATTTAC-C-TT--G----CTA-G--A--T-----G--G---T-A---- 6598
COL.CM.x.CGU1            -T----ACCTC-GCCTGGA--G---GT-----------........CA-AACA--CC-------A-------GC-----G-TC-T-TT........-G---AG-T-CAGT-GGG--G-ATGGATAATA-C-CAG-C-ACC-TA--A------G-----GA--------CT--G--G--AG--T 5956
```

```
                              V1 loop start
H1B.FR.83.HXB2             TCTGTGTTAGTTTAAAGTGCACTGATTTGAAGAATGATACTAATACCAATAGTAGTAGCGGGAGA............................................................ 6662
Env                        L  C  V  S  L  K  C  T  D  L  K  N  D  T  N  T  N  S  S  S  G  R  _._._._._._._._._._._._._._._._._._._._._._._._._._._._._._.
H1A1.UG.85.U455            ----C----CG---G-T---CA-A-CA-C-CC-T-A-C-AC-CC-AT--C-AC-CC-AT.................................................................. 6101
H1B.US.90.WEAU160          ---------C------T------A--G----TGTGAC--A-TTG-AG---GAG-C--ATACC-ATAGTAGTAGTGGAGGG............................................. 6672
H1C.ET.86.ETH2220          ---------C------C--T-A--C-A-C--A--CA-------GGTT-CA-A--A---T.................................................................. 6065
H1D.CD.84.84ZR085          G------C-C------T---------GCATC--GAA-C-GC-C-GAT----AC--C-CA.................................................................. 6177
H1F1.BE.93.VI850           -------G-C------C--T---A--GCC-CC---A-C-G-C-GGAA--GCCAG-GGCA.................................................................. 6002
H1G.SE.93.SE6165           ---------CC-----C--T------G-A-CC--CA-AGGC--C-AA-GA-A--A----ACTGATAATAGCACTGAAACCAACAAT....................................... 6086
H1H.CF.90.056              ---------C-C----C--T---A--G-C-GA--CA----CTC--A--GC-C---C--T.................................................................. 5996
H1J.SE.93.SE7887           -A-----C-C------T--T-G-A-CA-C-CT-G-A-C-G------A-C-----AC--TA-T............................................................... 5978
H1K.CM.96.MP535            -------C-C------T--T-------AC--AGGAACC-A--GC-------A-GCA-CTA-C-CT............................................................ 5861
H101_AE.TH.90.CM240        ----C----C------T--T--CA--GCT--TTTGACC-A-GGC-GT-GC-AA-CC-AT-TCTCTAAC......................................................... 6233
H102_AG.NG.x.IBNG          ----C----C----G-T--TCA-A----C--C--CAGCTAC-GC-A--G---C-ACCTAACT--T............................................................ 6189
H103_AB.RU.97.KAL153_2     ---------C------T-----------------G--GGT--C--GT-CC........................................................................... 5860
H104_cpx.CY.94.CY032       ---------C---T-CA--T-T-A--GCA-CT-C-AC--A--G-------G-C-C-..................................................................... 6023
H1N.CM.95.YBF30            -A-----A-C-A-GCTT--T-AC---AGCT-TGGG--GGAA-GG-A-----CA-A-...........................................................ATGACA 6246
H1O.BE.87.ANT70            -G-----ACAAA-GG----T--AA-CA-AGCTGGAACA--A---GAA.............................................................................. 6692
H1O.CM.91.MVP5180          -A-----ACAAA-G--C--TGTA---C--C-A-CAA---AA-CAGG-.............................................................................. 6682
H1P.FR.06.RBF168           -A-----C-C-A-G--T--T-G---A......TGCC-C-AGC-GG-T--AT--.................................................AATAACCCCTCT 6129
CPZ.CM.01.SIVcpzCAM13      -A-----A-CAA-G--A--T-G-A-CG-T-CT-GAC-A-A--GC--T-GC-ACTC--A-.....................................AATAACCCTACCCTCAAGAACAATACT 6261
CPZ.CM.05.SIVcpzEK505      -A-----A-CCC--TCA--T-GCTCA-G--G--G--TG-AC----GT.............................................................................. 6193
CPZ.CM.05.SIVcpzLB7        -T-------C----C-T----G-ATCCCT--ATT-----ACTC--G-.............................................................................. 6209
CPZ.CM.05.SIVcpzMB66       -G-------C------C--TT-ACTG--T--ATGCAT--AAG-A-ATGGC-AC-CC-C-AACT-C..........................ACTGTTCAAATTAGCACTGGTAATGATAGCACT 6277
CPZ.CM.05.SIVcpzMT145      -A-----G-CCC-T-CT--T-A-A--CCC-CT---AC-T-CTGC--T.............................................................................. 6189
CPZ.CM.-.SIVcpzMB897       -------G-CC-----C--TT-CACAAAT-CA-C-ACAGAGGCAT-T.............................................................................. 6215
CPZ.CM.-.SIVcpzDP943       -A-----C-C-A-GC-T--T-T-CCA-AC--C-TAAGC-GC-C-GGG---GAC-C---GAAT--CTCT......................................................... 6243
CPZ.CM.98.CAM3             -A-----C-CCA-GG-A--T-GAA-GG-T-CTTTCA---GC-C-T-T-----G-A--AAACTTCTACTATGACTACT................................................ 6093
CPZ.CM.98.CAM5             -A-----C-CCA-GC-A--T--GACCC-T-----CAT--GC-C---T----C--AC--T.................................................................. 6360
CPZ.GA.88.GAB1             -A-----A-CC--GC------G-A-GGCT--CTT-AGCCAGGCA-AA--CCTA-CA..................................................................... 6717
CPZ.GA.88.GAB2             -A-----GCAA----C---T---TC-A-A--TGC-AC-GGA--CGAA-C--A-GC--CT--A..............................................................AAC 6067
CPZ.US.85.CPZUS            -G-------C-A-G-CT--TCT-A-CCCCG-T-G-A--....................................................................................... 6706
CPZ.CD.90.ANT              -G---A-A-AAA-G--T--T--A-GA-AC--TGGAACAC---CA--ACCA----CA-CAACA--TACA.........................................GTAACACCAAAG 6100
CPZ.TZ.00.TAN1             -T--CA-A-AGA-G-CA--T...ACAA---CT---ACC--A----AA-CC.........................................CTGAATTCGGCAACAACAACCTTAACACCAACA 6297
CPZ.TZ.2001.TAN2           -T-----A-CCA-G-----T----TAG-A--T-CAACAGTAGTA-ATGCA-CAGCACCAACA.......................CCGAAAGTAAGTGCAACAGTAGCAACAACACCA...ACA 6320
CPZ.TZ.2002.TAN3_1         -T-----AGTCA-G--T--T--CA-A-CCC-A----TC--A-CAC-AGCA--C-CA--TACAGCGGCA.................................CCAGTAAACGCGACAACAATACCA...... 6314
GOR.CM.2007.SIVgor2139_287 -G-----C-C-C-G--T--T--AC-AC-A--T-CAACA-A--CA--AC-G-C-GAG-TGACA-ATACA............................ACTATACCACCACATACAACTATACCA...... 6226
GOR.CM.2007.SIVgorCP2135con -G-----C-C---G--T--T--AA-GACA-CAG-GACG--A-CRCAG-CA-A-RYA-CAAAT.............................................ACAACTATACCA...... 6204
GOR.CM.2004.SIVgorCP684con -G-----C-C-A-G-CT--TT-AA-CGAC-CA--G-GG-GC----AATCA-C-GCA-CTACA-CTAGCGCAAACACA....................CCTACAAGCACAGAGACAACTATACCA...... 6160
H2A.DE.x.BEN               -A-----GGCAA-G--T--T-GCAGGG-TC-AGGGA----C-CG---CCG-A-CCC--GACCTCGAGTTCCACAACCTCGAGACCACCCACATCC..........GCAGCCTCCATAATAAATGAAACTTCTAACTGCATAGAAAAC 7159
H2A.GW.x.ALI               -A---A-AGCAA-G-----T-GCA-CA-A-GC-CA--G-GC-CA----CATCCCCG---CCAG-GAGCACACTC..........................AAACCCCTGATAAATGAGAGCGATCCATGCATAAAGGCA 7125
H2A.SN.x.ST                -A-----AGCAA-GCGT--T-ACAGCACA-CTGCAA-A-AC-CA---TCC-CACCA-CAACC-CCACAACAGCA..........................AACACAACAATAGGAGAGAATTCTTCATGCATACGCACA 6577
H2B.CI.x.EHO               -G-----G-CAA-G-----T-A-A-GAC-TG--GCTCAG---GC-AAG-G-CC-C--CGTCCTCCGCGTCCCTCAGATCTTCTACTCAG...........ACCCTACTCAACGAAGATAGCAAATGTATTCAAAAT 7112
H2B.GH.86.D205             -G-----GGCAA-G--C--T-G-A-AACCG-A-CAA-CC-AGGG-ATGCC-------CTACC-CCACTAAGCCTACTACCACCTCT..........CGTGGGCTGAAAACGATTAACGAAACAGACCCATGCATAAAAAAT 7120
H2G.CI.92.ABT96            -G--CA--GCCA-G-GA----A-A-AACTG---CAA-C-AATGGGGTCTA-CAG-G-C-ACC-CTGTAACCCCTGCAACAGTTCCCACAACTAAAATGGTGACAGCT..........GAACTAGTGAATAGCACTAGCCAGTGCTTAATGTAC 6478
H2U.FR.96.12034            -T---A-AGCAA-G-GA--T-G-A-AACTG-A-CAA---GATGGGGATTA-CAG-A-CA-CAGC-ACAACAACTAGCCCAGCAATGCAAGAATCAAGT..........AAAGTAATAAATGAAACAGAGCCTTGTATACGAAAT 6610
SMM.SL.92.SL92B            -A---A--GC-A-G--A--T-A-A-AAATG---CA--C-GATGGGGGTTG-CA--AGCA-CT-CTACTACTAGCTCACCAACTACTACTAGCCCCTTAACTGCTGCTAGCCCATCAGGAGAG..........GAAATCGTTAACGACACTATGTCTTGTACAAAGAAC 6507
SMM.US.x.H9                -A---A-A-CAA-G-G---T-A-A-AAGTG--MCA-MC-GATGGGGTTTA-CAG-AGSACCAGC-CCAACAACAACACAAACATCAACAACACCACCATCACCAATAATAGCA..........AAGGTTGTAAATGACAGTGATCCTTGTATAAGAAGT 6539
SMM.US.x.PGM53             -A--CA---CCA-G-GA----A-A-GACAG-A-CA--C-GATGGGGGTTA-CA-AG-AT-AA-T-CCAACAACAACAACAACAACAACAACAAAAAGCCCAAAAGCA..........GAAGCAATAACAGCAAAAGTTATAAATGAAAGTGATCCTTGTATAAGCAAC 6999
SMM.US.x.SIVsmH635F_L3     -A---A-AGCAA-G-GA--T-A-A-AACTG---CA----GGTGGGGTTTG-CAG-A--A-CAGAGACAACAACAACAGCGAAATCAACAACATCAACAACAACAACAACA..........GTAACACCAAAGGTTATAAATGAAGGTGATCCTTGCATAAAAAAT 7075

MAC.US.x.239               -A--CA---C-A-G-GA----A-A-AAGTG---CA----GATGGGGATTG-CA-AATCAATA-C-ACAACAGCATCAACAACATCAACGACAGCA..........TCAGCAAAAGTAGACATGGTCAATGAGACTAGTTCTTGTATAGCCCAG 7065
Env                        L  C  I  T  M  R  C  N  K  S  E  T  D  R  W  G  L  T  K  S  I  T  T  T  A  S  T  T  S  T  T  A  _._._._._._._._._._._.  S  A  K  V  D  M  V  N  E  T  S  S  C  I  A  Q
MAC.US.x.251_1A11          -A--CA---C-A-G-GA----A-A-AAGTG---CA----GATGGGGATTG-CA-AATCATCA-C-ACA...ACCACAGCAGCACCAAAGGCAATG..........TCAGAAAAAATAAACATGATCAATGAGACTAGTTCTTGTATAGTTCAT 7061
MAC.US.x.251_BK28          -A--CA---C-A-G-GA----A-A-AAGTG---CA----GATGGGGATTG-CA-AATCATCA-C-ACAATAACAACAGCAGCACCAACATCAGCACCAGTA..........TCAGAAAAAATAGACATGGTCAATGAGACTAGTTCTTGTATAGCTCAG 7047
MAC.US.x.EMBL_3            -A--CA---C-A-G-GA----A-A-AAGTG---CA----GATGGGGATTG-CA-AATCATCA-C-ACAATAACAACAGCAGCACCAACATCAGCA..........CCAGTATCAGAAAAAATAGACATGGTCAATGAGACTAGTTCTTGTATAGCTCAG 6196
STM.US.x.STM               -A---A--GCAA-G-G---T-A-A-GAATG-A-CA--C-AATGGGGATTG-CAG-G-AAACAGT-ACAACAGTGACACCAACAGCAGCAGCAGCAGCAACA..........AAGCCAGAGTTAGTAAATGAAACAAGCTCCTGTGTAAGTAAC 6707
MNE.US.x.MNE027            -A--CA---C-A-G--A----ACA-AAGTG-A-CA----AATGGGGATTG-CA-AATCATCA-C-ACAACAGCAGCACCAACAACAAAAACAACAACAACA..........AAGGAAATAGAAGTGGTCAATGAAAATAGTACTTGTGTAAATCGT 6536
DEB.CM.99.CM40             AT-------CA-GC------TAA-AACATCA-CAA--C-A-CAC-A-CA-AC-CATCAACA-C-ACAACAACAATAGCAACAACAACA..........AAAACAACTACAGAT 6381
DEB.CM.99.CM5              AT-----GC-GA-G--C--T-G-AG-ACA-CA-CAACA--A-CAG-AGCA-CA-CA-CAATA..........AAAACAACAACTGAC 6323
DEN.CD.x.CD1               -A-----C--AA-G--T--T--AAGAAAAG--G--A-C-GC-CA-A-GCCGTC-CCGTGTCCTCTACCACAACCACAACCACAACCATAAAGACCCCAACCACA.......................... 6695
LST.CD.88.447              A------G--GA-G--A--TGAA-GGGAAGGA-TA-GA--A-CA-AAGCA-CA-AA-CAAAA-CG...TCAACAACACCC...ATGCCTTGCTTTATAACAGAAAAAACCACAACCACCACCAAAAGATCTACAACTGGGGGACCAGAATTAATAGAGGACATAAATTGT 5776
LST.CD.88.485              A------G--AA-G--A--T--A-GGGGAGGA-TA-GAC-A--A--A-CACCATCACCAATAG--AAAACAACAACACCC...ATGCCTTGCTTTATAGCAGAAACA...ACCACAACCACCACCAAAGGGAATACAACTATGGAAAGTTTAATAGAGGACACAAATTGT 5773
LST.CD.88.524              AT------C-CA-G--A--T--A-GGGAAGGACTA-G--AG--G-GA-CA-CA-CA-CAACA-C-CCAACAACAACACCC...ATGCCTTGCTTTATAACAGAATCTACTACTACTACTACAAAGAGTTCAGGGACTGAACAACCATCTTTAATAGAAGATACAAATTGT 5779
LST.KE.x.lho7              AT-----G--AA-GC----TGGG--AG-T...................-A--A-CACCA-CAACA-C-CCAAAAACTACTACACAGATGCCTTGTTTTATCAACGAACAGGTA.........ACAGTTAAGAACCCGGGGAATGAAACAAGATTAGAGGAAGATCTTAATTGC 6836
SUN.GA.98.L14              AT-----C-AAA-G-----TC-ACCGAAACC--CAAC-C-CTCCT----CTCA-CAGT-AAA--TAGCTGTGATTATTGGACGACCACCACGGCTAAGACTACAACTCAGACTACAAGTAGTACCAGCTCCACTGCTAGTACAACTACTCCAATGCCTCTAGATTGGAAT 6907
MND-1.GA.x.MNDGB1          AT-------AGA-GC-A--TCAG--AACAG-A----TAT-AGCA--A-CAGC--AGCCTATA-CTACACCTACTACTACATCTACAGTTGCAAGTAGTACA..........GAGATTTACTTAGATGTAGAT 6179
OLC.CI.97.97CI12           -T-------AAA-GC-A--TGAA---CCATCAGCAA--TT--CAGAACC--T--CACCTATC--TACAGCTGCTCCA..........ACAACTGTCACTAGCAGTACAGGAGGTCAGACCACTACAATAGGGTTAGGAGTTGAT 6244
WRC.CI.97.97CI14           AT-----A-AAA-GC-A--T---CCAGGAG-AGGGAG-GAG-CATTT............-ATCC-TGTGAATCTAATACTACAAGAGCT.................ACTACAAGTGCTACTACAACAGTTGGTACAGTTAGTACTACTCCAGTTATGCAGGTAGAAAAC 6742
WRC.CI.98.98CI04           A------G-AAA-G-CA--T--ACCAGGTGCACCA---T----C-AT-CG-AGGACTTTAACCCCTGTGAAATTAGCACAGGTGCACCTACTATTAGTACCATAAGTGCTGCACCTACTACTAGTACCATAAGTGCTGGTACCAGGACACCAGCTATGCCAGTAGATAAT 6787
WRC.GM.05.Pbt_05GM_X02     AT-----G-AAA-GTCT--T--ACTAGG-G-TG-GA-A-GC---TAT-CAGAAGAC-TA-AACTGACAACAGTAAAA..........ACAACAACACCAGTAACTCCGCCACAGAATACTACCCCTAGTATGGATGTAGAGAAT 5674
MND-2.CM.98.CM16           A------C-AGA-GC----T--AA-AGGA--AG-CCT-G---CCC-T-TACCC-C--CATCC-CCACAACTACCACAACTACCACC..........AAGACGGTGGCAAATACTACTGAGTTGGACATAGAC 6740
MND-2.GA.x.M14             A------C-AGA-GC-A--T--AA-GGGAG-A--CAG--G--CC-G--C--CC-C-CC-AT-CCCACTACCACCACTCCCACCACTCCA..........GCAGTGACGGGGCTGGAGGTGACG 6660
MND-2.x.x.5440             AT-----G-AGA-G--A--T--A--AGGTC-A-----A--AG-GCAAGCA-CAGC--AGACC-C-ACACCTGTACCCACAACAACAACCCCCTCCACTACCACT..........AGCTCCAGCACAAATAAGACAACAACTCCTGTGTTGGTTGTTGAG 6340
DRL.x.x.FAO                AT-----G-AAA-GG-A--T-G----C-C--A-TA-TA-A-GCAG---GC-CA-CACCAACACC-ACTACA..........GTAGCTACTACTGTTAATACTAGTGTTGCCTCAACAACAACAGCAAATGACACTTTGGAAATAGCTGTAAAT 6383
MON.CM.99.L1               -G--------GA-GCTC--TGT---AG-T--C-CC-T-T-G---G---G--CC-C-CCT-CCCCTAGTACCCCCACACCC..........TGGGGA 6595
MON.NG.x.NG1               -G-----A--GA-GCG---TCAAARGA-C--TG-AAGC--A-CAGTG-CRCCA-C--C-ACC-CYWCC...GCCACCACACCA..........TGGGGA 5208
GSN.CM.99.CN166            -G-----GG-GA-GT-T--T-ACCTCACA--C-CATCC--A-GCGAACCA-CA-CA-CTCCA--GCCACCAAATGTCTCCACAACCACG..........CAGTGGGGAAGT 6584
GSN.CM.99.CN71             -G-----AC-AA-G--A--T-AGCTCCCA--T-CGAGC--A-CA--A-CACCGTCC-CAACA-C-GCCACAACATCACCCTCATCCCCAACGCCAACACCA..........TGGGGA 6575
TAL.CM.00.266              -A-----A--GA-G-----TGAG--CA-A-CA--GA-CCT--CA-T--G-GCC-A--C-ACA-C-CCAGCCCCGACG..........GCCATGAGCAATAGCAATATG 6444
TAL.CM.01.8023             -G-----G--GA-GGC---T--A--CA-C-CA--AA--CTC-C-GG--CAGTAGCACC-CCA-C-ACAAATTCAACAACAAACTCAGTAACATCC..........AAGAGAACCAATGACACCAAT 5982
MUS-1.CM.01.1085           -G-----A-AAA-G--A--TG-GA-AG-A-C----T-C--C-GAT-T..........ACGGGAGGGAATAGCACCGGAACCACAGCACCCCCAACAACACCATGGGGATCATGG 6545
MUS-1.CM.01.CM1239         -G-----C-AAA-G--A----G-AGAA-A--TGTAACA-GGCCAT-A..........ACAACAACAACAGCATCAACACCAACAAAAACAACACCAGTGACCCCTTGGGGG 6529
MUS-2.CM.01.CM1246         -G-----A-AAA-G-----TGACCT-C-C--T-T-AG--TGTC--GT..........ACCACCGCAGCAAGCGCCACAACAAGCGGGCATACAACCGTCACCCCTTGGGGCAGA 6664
MUS-2.CM.01.CM2500         -G-----G--AA-G--A--T--CA-GC-A--TTCAACA-GGCCG---..........ACAAAGCCGACCACGCCCAGTACCACTCCATGGGGAAGG 6590
SYK.KE.x.KE51              -G---A-C--AACGTTA--TGTA--GGAC---G-CAG--G--G-GGT---CT--CC-CAACA-TC..........AGTCCAACAACAACAACAAACAAAGCAGCAACACCACCT 6331
SYK.KE.x.SYK173            -G---A----GA-GTT---T--ATTGGAC--C-GCCC-G---CA-GT..........ACCCCTACTACAAGTCCCCCTACTACACCTCCAAATGAA 6635
RCM.GA.x.GAB1              -A---A---CCA-G--T--TT-A-CAA-A--TGG-AGCTGGG--GGA-TCCC--CCTCA-CACC-CCAACAACAACAAAAACAACAACACAAAGAACT...ATAGGTGTAGAAAAGGAATGTACTGCTGGC..........AACGAAACATGTGAG...... 6364
RCM.NG.x.NG411             -G---A-A-CAA-GT-T--T----CA-----TGG-AG-TGGG--GGT-C-GCC-CCCAATCACCTACAACAACAACAAGAGCAACAACAAGAACAACA...AAAGAAATAGGGAAAGATTGTTATGCGAAT..........CAAACAGAGTGCAGT...... 6420
GRV.ET.x.GRI_677           -A---A---AAA-GTCC--TGTA--GC----TGG-ACAG-C-CG--A--GGCC-CC-CTACTGC-ACTACAACAATGACT..........ACCCCCTGTCAGAAT..........TGCAGTACAGAG 6621
SAB.SN.x.SAB1C             -G--CA-A-AAA-G--C---TA-CGC--AG-AGGA-GAG-AGCA--A-CA-CATCACCATCA-C-TCAACAGCACGCCCCGAGGTAGTGAGT..........GTGGGCTTTAATGACTCA 6918
TAN.UG.x.TAN1              -G--CA-A-AGA-G--T--T---A-A--A-CA-G-ACCG-CCC---ATCA----CCCCTACATC-AGCAGCACTACG..........GATCCCTGTCCAAATACC..........GACGAAAGTAGCTGTAACGCCACC 6716
VER.DE.x.AGM3              -G--CA---AAA-GTCC--TGTA--A-----CTCCTC-GAGCC-----CC-C-CC--AAA-T-CCACGGCCTCAACAACCAATATCACAGCCTCAACAACC.........ACTTTGCCGTGTGTCCAGAACAAG...ACAAGTACTGTGTTAGAATCATGTAATGAAACA 6257
VER.KE.x.9063              -A---A-A-AAA-G-GC--TGTA--G-----CTCAAG--GA-GA-A-CCA-C-CCA-CTTCAGCTTCAACAACAAAAGCTCCAAAAACAGGT..........GATCCCTGCATTAAATCTACCAATAATAATGTTAATTTACAGCCCTGTAATGCCTCC 6750
VER.KE.x.AGM155            -A---A-C-AAA-G--T--TGTA--G--A--AGGCTCCG-A-CCT-T-CCCCAGCA-C-TCT-CTACGGCAGGAACCAAA..........CTACCCTGTGTTAGAAATAAAACAGAC...TCCAACCTACAGTCATGCAACGACACC 6733
VER.KE.x.TYO1_patent       -A--CA-C-AGA-G--T--TGTA--G--A--TTCCACA-GAG-A-GGGCG-CA-CACCTACA-CGACGCCGAAATCTACCGGC..........CTACCCTGTGTAGGCCCGACGTCAGGT...GAAAATCTACAGTCCTGTAATGCAAGC 6735
COL.CM.x.CGU1              A-----C-CCAA-G-TT-----CA-AG-A--T-GGAC-GAA---GG--C-----CAGTA-CCCC-ACAACA..........ACAAATAATAGTGCTTCAGAT 6048
```

```
                           V1 loop end┐  ┌V2 loop start
H1B.FR.83.HXB2             .........ATG...ATAATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGC...ACAAGC...ATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAATACCAATAGATAATGATACT......................... 6788
Env                        _._._._._M_._I__M__E__K__G__E__I__K__N__C__S__F__N__I__S_._T__S_._I__R__G__K__V__Q__K__E__Y__A__F__F__Y__K__L__D__I__I__P__I__D__N__D__T_._._._._._._._._._._._._.
H1A1.UG.85.U455            ......ATC-CA...GATGGA-T--GG-A---A--G-----T--------------G-C-..---GAA..C------A---AAA-------T----T--C---------G---------G---A---CA---AAC-GA-AAC......................... 6233
H1B.US.90.WEAU160          .........GAA...-AG------G-G-----A--G-----------------G---C-..---CT-..-------AA---AGAA---CT---------C-----------------G---G--------C-----A-........................... 6798
H1C.ET.86.ETH2220          ...........-C...-AT-GT-CC--T-AT--A--G-----T--T-----------A-C-..---GAA..C------A---AAAAG---G-C-------C-------------------G-----C-TA----C-G-T-......................... 6191
H1D.CD.84.84ZR085          ......CTGCCA...-CGG--A--CC------A--G--------------------A-C-..---GTA..G---CG-A----AG-A--C---T-C-----C---------G-----G-G---A-----------GGG-AAAAATGAAATC............... 6318
H1F1.BE.93.VI850           ...............................-GC-------T-----------G-C-..---GAA..G-----A---AA--T----TT--C---GC---------GG--------G--------GC---A-C-A-AGC......................... 6110
H1G.SE.93.SE6165           ......AGC-CT...G-GGATA-CCC------A-------------------G-A-C-..---GAA..-------A---AAAA---------C--G-----C----G------G--G------TA-----CA.............................. 6212
H1H.CF.90.056              ...............-GGA--CAGG---G--AC---C---T-----------G-A-CT..---GTA..C------A----CA--------T-C-----C-C-------G------G--G------T-----CA---G-......................... 6119
H1J.SE.93.SE7887           ...............G---GTAGTCCT-ATATC--G-C-----------------A-CT..---GAA..------AA---AAG-A--C---G--C--GC----C----G--AA---G--G------T----G-A-C-A-......................... 6101
H1K.CM.96.MP535            ...............G-GG-AAGCCC--C---A--------T-------------A-C-..---GAA..----A--A----AA-A-A--G----C----C----C----G-------G--C-G---C-TA--GGC--GGG-AATAACAGTAGT.............. 5996
H101_AE.TH.90.CM240        ......ATA--A...GG--ATATA-C--AT--AG---G------TA----T-----G-C-..---GAA..C---C--A----AA------G-TCC-----C-C--------G-------G---A---T--AG--A-G-AGACTAGT.................... 6368
H102_AG.NG.x.IBNG          ..................GACAT---T--G--A---------------------A-C-..---GAA..G-----A----A----AAAA---GATGC-----C---------G------G-G---A---TA--G-AA---A-GGT...................... 6312
H103_AB.RU.97.KAL153_2     ...AATACT-GT...-GC--AA-A-TGATG--A--G-----------------C----C-..---GA-..C-G----A---A---A-A-----------C-C------------------G-G---A-......................................... 5989
H104_cpx.CY.94.CY032       ..................G--ATT----A--GA------------------G----A-C-..---GAA..-----G-A----AA-A------------GC----C----G-A--------G-G-----TA--GC-AGAGTGCCAATTAATGGT.............. 6155
H1N.CM.95.YBF30            ACAAGAGAACCA...GAC--A-GAT-CAA-C-A--G-----T-----A------GCA-C-..---TGAG..C---C--A---AAA-A--C---TT--CT-TC-G------GT-GAA---G--G-------CA--GCC........................ 6375
H1O.BE.87.ANT70            ..........................AACCTT--G--G--G--TGAG--T---G-A-C-..---TGTT..--C-A--AC---AAA-G-----A--C-G--TC-A--C---GT-TCA---T-G--GGA-C-GA--G-GACA-GCAGCACAAAT.............. 6812
H1O.CM.91.MVP5180          ......CTAT-A...-ATGA-ACA-T-AAT-----G-G---T--TAG---T---G-A-CT..---GT-..C-C-C--AC---AAA-G--C--A--C-G--TC-A--C---GT-TCA---C-G-GTAAGG-TA--G-CTCA-A-GCAGTAAAT.............. 6820
H1P.FR.06.RBF168           GAGAATAGA-CC...GCCGCCA-TCC-TT----C--TTT--G---AG---T----CA-C-ACAGT-TTG..-A-GATAAA---CA-G-AC-GC--GCACTC--C-A-AGA-C-GACCTA---GC--TG-AT--G-CA-CA-A-AATACA................. 6273
CPZ.CM.01.SIVcpzCAM13      TCAGGTAAG-AT...GA--CT-GAGTTTT-C-A--G--------TA-A--T--C-CA-CA..---GAA..T-G----A---AAA-A-AC-G-TC---T----A------GT-GA----C-GCA-AGCC---GA-G-.......................... 6390
CPZ.CM.05.SIVcpzEK505      ......GTT-AC...CA--CAA-CC-C-T-C-A--GC-----T------A--T---G-G-C-..--TGAG..T------A---AAA-A--C----TC--CT-TC-G---------TGGGA-----------CT----C-C-A---A-AGCAGTGGT.......... 6331
CPZ.CM.05.SIVcpzLB7        .........--T...-AT-GCTCT--T-AGACT--G-----T--TAGC-------CA-CA..---GAA..C------A---AA-A-AC-GACT--CT----G-------GT-GA----T---CT-AG---A-C--G........................ 6329
CPZ.CM.05.SIVcpzMB66       GCTAACAAC--T...-CGG-T-GC-C-AT---C--GT-C-----T----------GCA-C-..---GAA..T------AC-GAAA-A-AC-G-TG---T----G-------GG-AA---C-TGAG---C----AGGAA---AGCCACCAGAA.............. 6421
CPZ.CM.05.SIVcpzMT145      ...AACTCT-CT...GACGATCGCTTG-----C--G-G---T--TAGC--T---G---CA..--TGAA..C------A---AAAAGAC---TG---T-C--A--C---GTGGAG-------C-G------GA--............................ 6315
CPZ.CM.-.SIVcpzMB897       .........................CAG-T-CAGT-T--T--T----------G-A-CT..---GAA..T------A---AA-A--C---TG--CT-C--A------G-GAG--C----C-AGTC-T----GCA---AGACA.................... 6329
CPZ.CM.-.SIVcpzDP943       ......................ACA-GCCAGAT-G--------A----TC--T-----G-CT..--TGAA..T------A---ACA-A--C----T----T-T--A------GT-GA-------G---SMK--T............................. 6348
CPZ.CM.98.CAM3             AACAGTCCT-AT...GA--AAAYAG-CA-CACAG-----G--T--TA-C--T------G-CT..--TGAG..G------A---A-A-A-A----T----T-CC--------GT-GA----C--G--TTG--T-----C-----CGAT............... 6231
CPZ.CM.98.CAM5             ...ACAGAGT-A...-A--ATATA----A-ATA--G-----T--TA-A--T------G-CT..--TGAA..G------A---AAA-A--G---T----T-TC-A-------GT-GA-------G-----T--T-GAGG-A---AGAGTAATGAT............. 6501
CPZ.GA.88.GAB1             ...AACCAG-CA...TCTTCTCC-CCTCTC--A--G--------T--A--T---G-G-CA..---GAA..T------A----AA-A-AC-G-TC---T--C-A-------GTGGAG---G-GG--AACC---GG-----G-ACAAC.................... 6852
CPZ.GA.88.GAB2             GGAATTGAA-A-...GATGG-TT-GC-CA---C--G-G---T--TA----------CA-C-..---GAA..T------A---AAAA---C--ATT---T-C--A--C-GG---AA----C-TG--GGT-CTA-C-............................ 6196
CPZ.US.85.CPZUS            ...AGTAGTGCT...G---ATACT-CT-ATATA--G-G-------T-----T------A-CT..--TGAA..T------AC---AAA-A-AC---TG---T-C--A-------GT-GA---C---GCT-AT--CA-----A--...................... 6835
CPZ.CD.90.ANT              ACAACAACACCA...----G-A--TGGCATGA--C--C--G-A--TAAC--T---CAG----..---G-A..T-T-A--A---AAA--A---ATGA-A--CA-A---------GGA---C-T--GAAGTGTC-GG-CA-C-A-GAGACTAAC............. 6244
CPZ.TZ.00.TAN1             GTAAATTTG-GT...TCT--ACCT--CTAT---G-GT-T--T--T--A--T---CAG-CA..--TGAG..T-T---A---AAAA-AC--AT----T-C--G-------G-GAA-----TG--AA-GAG---GG-A-C......................... 6429
CPZ.TZ.2001.TAN2           ATAAATTTG-CA...-AT--AA--G--TAT--A---TTT------T--A--T---CAG-CA..--TGAA..T-T---A---AAAA--C--AT----T-T--GK------GGGAA-----T--GAA-G-T-GC-G---A-G-ACT.................... 6458
CPZ.TZ.2002.TAN3_1         GTAAACTTG-CA...GAT--ACCT--TTAT--AT--TTT------T--A--TC-GCAG-CA..--TGAA..T-T---A---AAAA--C-GAT----T-T--G-------GGGAA-----T--GAAGGATA-C-GGAG--TGCAGGATAACAGGAGCAACAGCAGTAA 6475
GOR.CM.2007.SIVgor2139_287 .........CCA...CAC--CCCAC-GTT---C--TT-C-----TAGC--T---G-T-C-..---GGTT..C---A--A----ACAACA--GC-GC-A--T--A-------G--AA---------GA--C--G-G-AACA....................... 6349
GOR.CM.2007.SIVgorCP2135con .........CCA...TAC--CCCAC-GTT---C--TT-C-----TAGC--T---G-T-C-..---GGTT..C---A--A----AAAACA--GC-GC-A--C--A-----------G-AA--------AA--C----G-GACG..................... 6327
GOR.CM.2004.SIVgorCP684con ..........CA...-AC--CCCAG-GTT---T--TTTC--T--TA----T----G---C-..---GGTT..C---A--A----AAAACA--GC-GC-A--C--A--C--C--G-AA-------C-GA--C-------AA--GAG.................. 6286
H2A.DE.x.BEN              AACACATGCGCA...GG-T-A-G-T-T-AG-----G-TGC-A--TGAG---------G-AG...GGGTTA...GA-CA--A----AAAAG--GGT-TA-G-ACACA-GG---TT-GAA---G-GG-TTGTGACA-C........................... 7285
H2A.GW.x.ALI               GACAACTGCCCCAGGGG-C-A-G-G-T-A------GGTC--T--TCGG------C--G-CA...GG-TTA...CAG-----A----CCAA-AC-GT-TA---A-ACA-GG---CTC-AAA---G-GG-TTGTGA-CCA.......................... 7254
H2A.SN.x.ST                GACAACTGC-CA...GGGT---GAG---A------GGTCG-----TCAG---------G-CA...GG-TTA...GAG--G-A----AAAA-ACT-T-TA---A-ACA-GG---CTC-AAA---G--G-CTGTGA-TCA.......................... 6703
H2B.CI.x.EHO               GACAGCTGTGCA...GGG--A-GACT--A---A--G-T-G-----TCAA------A--G-CA...GG-TTA...-A------A-G--TCAA-AC--T-TA-G-ACACC-GG--------AA---T--G-GTGTGA-A-GGGGACA-GGAGTAAT............ 7253
H2B.GH.86.D205             GACAGCTGC-CA...GG-C-A-GAG---AG--A----TGC-A--TAA--T-G---G-CG...GG-CTA...-G------A-G--C-AA-AC--T-TA-A-ACACC-GG---CTC-GAA---T--GAGTGT-ATA--............................ 7246
H2G.CI.92.ABT96            GATAATTGT-CA...GGT--AC--TC--A-TCC--GGTTGG----AAA--T--C--G-CA...GGGCTA...-A---G-A-C--AGAA--G--T-CA---AGAC--GG---TC--AA---T-GCA-TGYGA-C-G............................. 6604
H2U.FR.96.12034            AATAGCTGCTCA...GG-T-A--AC---A-CCAT-GGTG-GT--TAGA--T------G-CA...GGCTTA...-A---G-A----AAA--AG-GT-TA-A-A-ACA-GG---TCT-A---C---G-TTGTGA-C-G............................ 6736
SMM.SL.92.SL92B            AACAATTGCTCT...GGC--A---C-G-A-CCA--G-T-GGT--TCAA--T--C--G-CA...GG-CTA...-A------ACC---AAAAG-C-GT-CA---A-ACC-GG---CTCCAGA---C--G-CTGTGAGC-A........................... 6633
SMM.US.x.H9                AATAATTGT-CA...GGCT----AC-G-AGCCC---GGT--GT--TAAA--T--C--G-CA...GGGTTA...-A------AC---AAA-AGAG--T-TA---A-ACA-GG---TC-AGA---T--G-GTGTGA-C-A..................AAT............. 6668
SMM.US.x.PGM53             AATAATTGT-CA...GGCT-A---C-G-A-CCA--GGT--G---TAGA--T--C--G-CA...GGGTTA...-A--AG-AC---AAA-A-AG--T-TA---A-ACA-GG---TC-AGA------G--TGTGA-C-A........................... 7125
SMM.US.x.SIVsmH635F_L3     AATAGTTGTGCA...GGCT----AC-G-AGCCC--G-T-GGT--TAAA--T--C--G-CA...GGGTTA...-AG--G-AC---AAA-ATAG--T-TA---A-ACA-GG---TC-AGA---T----CTGTGAGC-G........................... 7201
                        V2 loop end┐                                                                                          ┌V3 loop start

MAC.US.x.239               GATAATTGC-CA...GGCT----AC---AGC-A--G-T--G---TAAA-----C--G-CA...GGGTTA...-A-----AC---AAAA-AG-GT-CA---A-AC--GG--CTCTGCA---T-GG--TGTGA-C-A.................................. 7191
Env                         _D__N__C__T_._G__L__E__Q__E__Q__M__I__S__C__K__F__N__M__T_._G__L_._K__R__D__K__K__K__E__Y__N__E__T__W__Y__S__A__D__L__V__C__E__Q_._._._._._._._._._._._._.
MAC.US.x.251_1A11          GATAATTGC-CA...GGCT----AC---AGC-A--G-T-GG---TAAA-----C--G-CA...GGGTTA...-A-----AC---AAAA-AG-GT-CA---A-AC--GG--CTCTGCA---T-GG--TGTGA-C-A.................................. 7187
MAC.US.x.251_BK28          AATAATTGC-CA...GGCT----AC---AGC-A--G-T--G---TAAA----CC--G-CA...GGGTTA...-A-----AC---ACAA--G-GT-CA---A-AC--GG--CTCTACA---T-GG-TTGTGA-C-A.................................. 7173
MAC.US.x.EMBL_3            AATAATTGC-CA...GGCT----AC---AGC-A--G-T--G---TAAA----CC--G-CA...GGGTTA...-A-----AC---ACAA--G-GT-CA---A-AC--GG--CTCTACA---T-GG-TTGTGA-C-A.................................. 6322
STM.US.x.STM               AACAACTGC-CA...GGCT----AG---A-TCAC--GT-GGT--TAAA---------G-CA...GG-CTA...-A-----A---AAA-AGAG-GT-CA---A-ACG-GG--CTCTAGC---C----CTGTGA-C-A.................................. 6833
MNE.US.x.MNE027            GATAATTGC-CA...GGCT----AC---AGCCA--G-T--G---TAAA-----C--G-CA...GGGTTA...-A-----AC---AAAAGAG-GT-CA---A-AC--GG--CTCTGCA---T-GG-TTGTGA-C-A.................................. 6662
DEB.CM.99.CM40             TGGTCAGGGGAA...-AC--AACC-TGAC-C-ATATTGG--T--T--A--T---G--TCA...GGTCCT...TAT----A----AAAG-A--GAGC-CA---G-A-GGCTAG-GGA------CCAGTGGGCG------A-AGA-GGTAGTGGGAAC.............. 6528
DEB.CM.99.CM5              TGGTCAGGAGA-...-AT-A-ACAGTGAC-C--TATTGG--T--TA-C-------G-T-CA...GGGCCA...TAT--G-AC---AA-G-A-TGC--AG---G-G-GG-TAG-GGA---C---CAGTGGGAGACAG-AA-CCAAACAAATACTAAC............. 6470
DEN.CD.x.CD1               GAGTGGTATGGAGGA-AT-A-ACT-CC...ATCT-CCTT--G--TAA---------G-C-ACAGG--AAGGCT-T--G-A----AGAGTA-TGT-TG-A--TCA--G-AGA-T-GAA---G----GGA-................................. 6821
LST.CD.88.447              ATTAAATGGG-A...-ATGA-ACT-CT-AGATT-AT-G-G-G--TAGA-AT---G-T-CA...GG-CTA...TGT--G-ACTGTAG-ACAG--AT-A-GCAGAG----AGGT-TGA---G-C--C-TGTGATA-G----GA......................... 5908
LST.CD.88.485              ACTAGATGGT-A...-ATGA-ACT-CT-AGATT-AT-G-G-A--TAAA-AT---G-A-CA...GG-CTA...TGT--G-ACTGTAAAG-AG--AT-A-GCAGAG----AGGT-TGA---G-C--CCTGTGATA-G----GA......................... 5905
LST.CD.88.524              ACACGATGGT-A...-ATGA-ACT-C--A-ATT-AT-GGG-A--T--A-AT---G-G-CA...GGGCTT...TGT----A-TGCAA-G-AG-GAT-A-GCA-AA----AGGT-TGA---G-C--C-TGT-AT--G----G-......................... 5911
LST.KE.x.lho7              ACAAGAGGGC-T...-ATGA-ACC-C--AGAG--ATGC-G-A---CAG-AT---G-A-CA...GGGTTA...TGC----ACTGCAG-AC-G--AT-A-ACA-AGC----AGAT-TGA---G----CGTGCTC--GAG-GAGG........................ 6968
SUN.GA.98.L14              TGTACTGAC-CA...GAG-ACATAGCG-AGAGC-AT---GTA--TAAA-AT---G-G-CA...GGGCTT...TGT----A-TGTAAAC-G-G-TGG-ACA-AAC----AGGG-TACA--AG-T-CTTGC-AT-G------C........................ 7039
MND-1.GA.x.MNDGB1          AAAAATAAT-CA...GA-GAAA--GT--AGAG--ATC-TGTA--TAGG-AT--C--A-CA...GG-CTA...TGC--G-A-TC-AA-G-AG--AT-GTAA--AA----AGAGGGGA----G-G-A-TGTGA-A--............................. 6305
OLC.CI.97.97CI12           AATAGTACAG-A...-AGCCA--A-C--A-A--C-T----C-A--TAGG-AT-----T-CA...GGGTTA...TGT----ACTGTAAAATAG--ATGA--CA-ACA---G--TTGGAA---G--GC-TGT-C-A----C......................... 6373
WRC.CI.97.97CI14           AATTGCTCACA-...G-T-ATACTTC--AGATT-ATC-TGTG--TAGA--T--C-C--CA...GG--TG...TGT-AG-ACTGTAGAATAG--AT-A-A-A--C----AGAGG-GA----G-GG-TTGT-ATA-----AG-........................ 6874
WRC.CI.98.98CI04           AATTGTACA-CA...G-G-ATACCTCT-AGATA-ACC-TGTA--TAGA--T----C--CA...GG--TG...TGC-A--A-TGCAG-ATAG--ATTA-A-A--C----AGAGG-GA----G-GG-TTGC-ACA-A----GG---..................... 6922
WRC.GM.05.Pbt_05GM_X02     AACTCTACTTCC...T-G-ATCTC-C--ACACA-ATC-TGT---TAGA--T---G-A-C-...GGGTTA...TGC----A-TGTAAA--AG--AT-GTA-AG-CC----AGAGG-GA----G---G-TGC-C-C-AC--........................... 5803
MND-2.CM.98.CM16           ACCAATAAC-CT...GA--CTACA-C-CA-C---AC-GGGT---TAAA-------CT-CA...GG-TTA...TGT--G-A-TGCAAATT-G-GAT-G-A-A-AA----AGAT-TGAA---G---CTTGT..................................... 6860
MND-2.GA.x.M14             GTACAGAAT-AT...GAG-GTACC-C-CA---A-AT-GGGTA--TAAA--T--C-CA-CA...GG-TTA...TGT--G-A-TGTAAAATAG-GATTA-A-AGAGC---AGAT-TGAG---G-G-CCTGT...................ACTAAA.............. 6786
MND-2.x.x.5440             AAACAAAAT-AT...GAG-CAACA-C-CAGC-A-AT-G-GTA--TAAA-----C-CG-CA...GG-TTA...TGT--G-ACTGCAAATT-G--AT-G-G-A-AAC----AGAT-TGAA---G---CTTGC-C-A-GCT-A---AAACTGGTAGTGCT.............. 6487
DRL.x.x.FAO                AATACAGAACCA.........-TA-TG-A-ACA-AC-G-GTA--TAAA-A----G-G-CA...GG-TTA...TGC--G-ACTGTAA-G-AG-GATCACA-A-AAC----AGGT-TGA----G--G-TTGT-A-TG-GGCA--GAAACTGAATCACCAGAAAGC........ 6530
MON.CM.99.L1               AATTGGGGAGG-...-ATGGAACAGG-CAGCCAG--T-T--T-----A--T--CCAG-CT..----GAA..T-C--G-A----AAA-A--C--ATG---AGC--A----GG--GGAA--C-----GAAGGAG--AGGAAGC-A-GGC.................... 6733
MON.NG.x.NG1               AGTTGGGGAGGC...-ATGG-AC-GG-CAGCCTC-TT-C--T--T--A--T--CCAA-C-..----GAA..T-T----AC--AAR-A-AC--ATG---AGC--G---KGG--GGAG--C--T--G-A-GACA-G................................ 5334
GSN.CM.99.CN166            TGGGGGGGGGAA...-ATGGAACAGG-CA-CCCT--T-T-----T--C--T---CAG-C-..----GAA..T-T----ACC-AAA-A-AC-GATG---AGCC-C--C-GG-G-GAA--C--T--GGA-GAGACCC---G-........................... 6716
GSN.CM.99.CN71             AATTGGGGAGGT...-ATGGAACAGG-CAGCCA---T-T-----T--C--T--CCAG-CT..----GAA..T-T--G-AC--AAA-AGAC-GATG---AGCC-C----GG---GA------T--GAGGGC-C-GG---G-.......................... 6707
TAL.CM.00.266              CCGTGGTGGGGA...GAC-ATACA-C--A-CCAG--T-T--T--TA-A-GG--C--G-CA...GG-GCA...T-T--G-A----AG-T-T-CCT-C-T-T-TCAC-GG---CCT-GCA--CC----GAAGGA-CCAGGCA--GGATCC.................. 6582
TAL.CM.01.8023             GTGTGGTGGGGA...-AC-AT-CC-C--A-CCAG-GT-C------TA-A-GG-----G-CA...GG-G-A...T-T--G-A----AA-T-C-CCT-T-T-T-TCA--GG---CT-GCA--CT----GAAGGA-CCAGGCA--GGCAAT.................. 6120
MUS-1.CM.01.1085           GGCAATGGC-AT...GAT-GT-CTGG-CA-CCT--GTTC--T-----C-------CAG-C-..----GAG..T-C--G-A---AAA-AGAC-GATG---T----G--C-T-GTTGA----C-G--GAG-GACAGCTCA--G-G-GAGAATGAT.............. 6689
MUS-1.CM.01.CM1239         AAATGGGACGCC...-ATGGTACAG-GCA-AGC--GTTC--T--T--A-------CAG-CT..----GAA..T-C----AC--AAA-A--C-GATG--CT-C--A-----C-CTGA----C-G--G-GTGAG---------....................... 6661
MUS-2.CM.01.CM1246         TGGGCAGAC-AT...-ATGA-ACAC--TTGACC--GT-C--T--T--C-------CAG-C-..---CGAG..T-C----A-GTAAAAA-AC-GATG--CT-TC-C--C-T-GT-AAA---CC-C--GAA-GGG--G-----G.......................... 6796
MUS-2.CM.01.CM2500         TGGTCTGATGGC...-ATGATACA-TGCA--TC--GT-C--T--T--C-------CAG-C-..----GAG..T-C--G-AC--AAA-A--C--ATG---T-TC-G----T-GTGGAC---T----GAGGGAT-GG-GC---.......................... 6722
SYK.KE.x.KE51              GCATGGTGGGG-...GAC-ACTCT-C--A-ACT-GGT-C--T--T--A--T--C--G-CA...GGGG-T...T-T-AG-A----AAA--TC-GT-TA-------C-----C--GTCA--------GAG-GAG--A................................ 6457
SYK.KE.x.SYK173            ACGTGGTGGGGA...GAC-ACAGT-C--AGCCA-G-TTC--T--------T--CT-G-CA...GGGG-G...T-C-AG-A---AAAA---C--T-TAGA--T--C--C-----GGA---CT-G--GAA-GAG--A............................... 6761
RCM.GA.x.GAB1              ...............GA-G-TC--G-T-C---TG-G-TGTCT--TGAA--TGC-G-AGCA...GG-CTA...-AG----A-G--AAA--C--GT-TA---ATACC-GG---TCTAGA--CC-TTGGTGTGA-A-G...............GAAACA.............. 6484
RCM.NG.x.NG411             ...............-GTG-CCCTG-C-C---TG-C-TG--T--TGAA--T-CGG-GGCA...GGGTTA...-A-----A-G-AAAAAT-----T-A---ACACA-GG--CTC-AGA---C-CTGGTGT-AG-CA............................... 6534
GRV.ET.x.GRI_677           CAGATAGAAGGA...GA-----CAG-G-A-CCAGC-TCC------A----TGCA--TGCA...GG-TAT...CA-----A-GTAAAAA----TT-TAGCATGACC-GG---G-T-AG--GT--G-CTGC-ATA----AACAGGAAGTGAAAAGGGA.............. 6768
SAB.SN.x.SAB1C             GTGATAGAACAA...GAG-----A----A-C--GCC-TG-----T--C--TGC---GGCA...GG-TAT...-G---G-A-GTAAA-A-A--TT-C-CAA--G-A-GGG--G-T-AA--GG-GG-TTGTGA---GGGAAGGGAAAAATCAAACGCA.............. 7065
TAN.UG.x.TAN1              TTAGTTACA-AT...-GC-----TT-T-AGA-T-GTTCT-TA-----C--TGC---GGCA...GG-TAT...-GG----A-GTAAAAA-----T-TA--AGTAC--GG---G-T-AG--GT-GG--TGTGAGA-GG-AA-C-AC...ACCACAGGC............. 6860
VER.DE.x.AGM3              ATCATAGAA-A-...GA-T-AA-TG---AGCCTGCTTCT--T--TA-A--TGCA--GGCA...GGGTAT...G------A-C---AAAA----GT-T-CA-TGG-G-GGA--G-TGCA--A--C--GTGT.................................... 6377
VER.KE.x.9063              CTTATAGAAGA-...GA-C-----G---A--CAGCTTC---T--TA-A--TGC---GGCA...GGGTA-...--T-----A-C--AA-A-A---TT-T-CA-T-G-A-GGA--G-TGCA--G--CTATTGT-AGA-C-GCTC--GCTCTAACTCTAGC.............. 6897
VER.KE.x.AGM155            ATCATAGAA-A-...GAG---A-TG-C-AG-CAGCGTC-------A-C--TGC---GGCT...GGGTA-...--T--G-ACC-AAA-A----TT-C-CA-T-G-A-GGA--G-TGCA--A--CT-TTGT-AGCG--G-ACAT-GCATAATGGGACA.............. 6880
VER.KE.x.TYO1_patent       ATTATAGAA-G-...GAG------G-T-AGCCCGCCTCT--T--TA-A---GCA--GGCT...GGCTAT...G------A-C--AA-A-A---TT-T-C--TGG-G-GGA--G-TGCA--A--CTATTGC-A-A----GAC--A-AGCACTAGCAAA.............. 6882
COL.CM.x.CGU1              TGGGATGAGTCA...-ATTG-A-AG--TATCC-TGGT-T--T---AGAA-G---TCA-CAGCTTTCTTA...C-T-A--A--GAAA-G--TT-----TA-G-----CAGTAG--GA-TTA-CTG-G-TTGG-A----AA--GACAGTAATAGT................. 6195
```

V2 loop end

```
H1B.FR.83.HXB2              .....................ACCAGCTATAAGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGAC 6936
Env                         __.__.__.__.__.__.__.__T__S__Y__K__L__T__S__C__N__T__S__V__I__T__Q__A__C__P__K__V__S__F__E__P__I__P__I__H__Y__C__A__P__A__G__F__A__I__L__K__C__N__N__K__T__F__N__G__T__G__
H1A1.UG.85.U455             .....................-AT--T----GAC-A-T--A------T------AC------------T-----C-----G-----------------------------A-------------------G-----GG--CCTGA-----------A---G- 6381
H1B.US.90.WEAU160           .....................-A-------C-----T--A------AT-----AC--------------------------------A-------------------------------------G-----G------A-----------AG---- 6946
H1C.ET.86.ETH2220           .....................--TGAT----GA--A-T--A------T------AC---A-----A---------------C--T--G--C-------T---------------T--A-------A-------------G----GAG-------A----C------------ 6339
H1D.CD.84.84ZR085           .AAT.........GATACTTATGGT-CT----G---A-T--A------T-------C---------A--T-------A-----------------------------------A-------------A-------G----CG----A-G-------------C- 6478
H1F1.BE.93.VI850            .....................-GTGAA----G-C-A-T--A------T------AC----------T--------------T-GG--T-----C--T---------------T--A-------A---A-------------G----A-GA--------G-----G- 6258
H1G.SE.93.SE6165            .....................--T-AT----G-C-A---CA------TGT----AC---C-A------T-------------A-T-----C--------------------T--AA-------------T----G----GGG-----GA------------------ 6360
H1H.CF.90.056               .....................--TCAG----G-C-A-T--A------T--------------------T--C--------G--------A--T------------------T--A-----C-------------G-----C-----A--A--------------T 6267
H1J.SE.93.SE7887            .....................-AA-AT----TTC-A-T--AC-----T-------------A---A--T-----------G--T---C-A----------------------A-------------------G--------AT----------------T 6249
H1K.CM.96.MP535             .....................--TGAA--C-G-C-A-T--A------T------AC---------A-T--------A---A-----------------------------A-----C-------------G-----GG----A-G---------G------- 6144
H101_AE.TH.90.CM240         .....................-GTGAG----G---A-T--A------T--T----------AG-----T---------A---------T-------T-----------A-T--A-------A-----T----G------G--------AT--------G-----G- 6516
H102_AG.NG.x.IBNG           .....................-GTCAG----G---A-T--A------T-------C-----------T--------------------------------------------A--------------------G-----GG----AGG----------------G- 6460
H103_AB.RU.97.KAL153_2      ..........................G------T---------T--------G-------A-------A------------------------------------------A--------------------G-----CG----A-A-------C---------G- 6134
H104_cpx.CY.94.CY032        .AGTAATAGGAATAATAGTACAGAAGAG----T---A-T--A-------G-----AC------A------T--C--------G--T-------------------------A-------------A---T----G--------G-G--A-AT-----C-----TT---G- 6324
H1N.CM.95.YBF30             ...............TATAAT-AA-CA----G-C-A-T--A------T---A---CTG-G-----A--T-----T---ACT-----------------A--------C-----A--AC-A--C------C---A-G---------G-AGGA-AC--T-G------AT---A 6529
H1O.BE.87.ANT70             .........AAGACAAACAGC-AA-TG----CA--A--T-A-------T--A--AC---C--G--A----------------T-------A--------A-------C--------T--A--A--A-A---T--CT-T--G------C-GC-CAGAA--T-----------CA 6972
H1O.CM.91.MVP5180           .................GGA--A-CA----T---A--T-A-------T--A--A-T--C-AG-----------G------AGT-------C-----------C--------T--AA-A--A-A---C--CT-T--G-------G-C-CAGAC--T-----------C- 6971
H1P.FR.06.RBF168            .......................CTA----GA--A-TT-A---C------A-TAC---------A--A--C-----A--GA-T---------C-A-------G---------A--A--A--G-A---AC-AA-G------C--KC-G-MAGGA--T------T-------- 6418
CPZ.CM.01.SIVcpzCAM13       ............GGAACTGGTGAT-CA----CAA---TT-A---------A---C---A---------A---------A--G-----------T------A---------A------------A--A--A--A-----A-----C----------G--GTA-AC--TTCA--G-A----A 6547
CPZ.CM.05.SIVcpzEK505       ...............AATAAC-GTCAA----G---A-T--A---------A---CTG-G---------T--AA-T-----------------A---T-----------A--AC-A--A------C---A-------------G--C-AGAT--T----------T--GG 6485
CPZ.CM.05.SIVcpzLB7         ...............ACAACCGAG---------C-T-T--A------T--TA---C---A--C------T--C------A-T--T-----------A--T-----------T--A-------------TC-------GG------C------A--A-AT------T-GT--T- 6483
CPZ.CM.05.SIVcpzMB66        .........GGAGATAAGAACG--TTA----GA--ATAT-A---------A---C---G--T--A--T--CT----A---C--TC----A-------T------------T--A------A-----CC-CT-G--G------G-----A-AT--T-C---G-T----- 6581
CPZ.CM.05.SIVcpzMT145       ..................AAT-GT-CT--C-G-----TT-A------T---A---C---C--T--A--T---------ACT--------A-------A---------C-------A--A--A--A-----AC-C-------------G---TAGAT-A---A--G-AT-A-A 6466
CPZ.CM.-.SIVcpzMB897        ............GTTAAAAATGGA-CT----G---A-TT-AC-----T---A---C---A--C------T-------AAC---A-----A--------T--TT-C---------A--A-----A------CC--T-------------G-CC---AT------G----A--GA 6486
CPZ.CM.-.SIVcpzDP943        ............AATGGTAMT-AKTC-----G-C-A-T--AC-----T---A---CT---------A--A--C--T--AACC-----------------------------A--CC-A------------A--CA-G------C---G-GCCAGAT--T-------------A 6505
CPZ.CM.98.CAM3              ........................-CA----GAC-A-T--A---C--T---A---CT-----------G-----T--ACC-----------------A---------C-----AA-CC-A----A---A--CA-G------C----TGCCA-AC--T-------------A 6376
CPZ.CM.98.CAM5              ........................-C-----GA--A-T--A------T--TA---ACTG-A---------A------T--AACT-----------------A---------C-----A--CC-A----------A--AA-G----------G--C-A-A---T------------- 6646
CPZ.GA.88.GAB1              ........................-CA----G-A-A-TT-A---C--T--TA---C---A------A--A--------AACG------------T-----A-------C---------A--A--A--G-----A-------T------C---G----AGAC--TTCA-----A----TA 6997
CPZ.GA.88.GAB2              ........................-CG-T--G---A-TT-A---C--T--TA---CA--C------A--T--------AACC-----------A------------A---------C-------A--A------C------CC-GT-G-------------G----AGAT-ATCCA-----A-----A 6341
CPZ.US.85.CPZUS             ........................-C-----GAC---T--A---C------A-C-CT--C--------T------T--ACC---------------------A----------C------A--AC-A--C------C---C------------G-A--AGAT-------A------AG---G 6980
CPZ.CD.90.ANT               ......................T-----T-C--ATGGCAC--C-------A--ACT--C-----AT-------GA----TCTA-T------A--------T-----------------T--A--A--A--A---T--AT-G-G-----G-AG--G--GAT--T-CA---GT---GA 6389
CPZ.TZ.00.TAN1              ..................AAT-AT--T---T-T--ACAT-A---C--T--------------T--A-AA---GAT--ATCTA-T------A------------CAGA--C------T--A--A--C------CC-GT--------GAG--C----AT------CA--G-A------ 6580
CPZ.TZ.2001.TAN2            ................AAGAAT--TGAT---T-T--ACAT-A---C--T-------C-------T--A-AA---GAT--ATCTA-T------A------------CAGA--C------T--A--A---------C--GT--------------G--C-AGGT-AT-CA--G-AT--GA 6612
CPZ.TZ.2002.TAN3_1          GAAAAGCAATAATAGCAATAAT-GTG-A---T-C--ACAT-A---C--T-------C-------T--A-AA---GAT--ATCCA-T-------------------CAGA--C------T--A--A------A----T--GT------------GG--C-A-AT-AT-CA--G-AT---A 6645
GOR.CM.2007.SIVgor2139_287  .......GGTACAAATGGGATC-AAG-G----GA--A-TC-A------T-----TACT--AT-G--A--T-------C--A--C-------------C-A--T------G-------A--A--A--C-A---CC-GA-G-----------G---TACAA--T------T------G 6512
GOR.CM.2007.SIVgorCP2135con ....AATGAGACCAAAGAGTAT-GTGAG----G---A-TC-A------T-----TACT--AT-G--A--T-------C--A--C-------------C-A--T------G-------A--A--R--C-A---CC-GA-G-----------G-----ACAA--T------T------G 6493
GOR.CM.2004.SIVgorCP684con  ........................-----G---A-T--A-------------TACT--AT-----A--T-------C--A--C-------------C-A--T------G-------A--A--A--C-A-C-CC-G--G-----------G--G-ACAA--T------T------G 6428
H2A.DE.x.BEN                .............ACAACAGCTGG--CA-G-T-CA---G-CA---C-----A---A----C-A-G--T-A---GAT---CAC-AT-GG--TG-T--GAGGT-TAGA--C-------A--AC-G--C-------CC-AT---G-------G---CC-AC-ATTCA--CTTT-A-- 7442
H2A.GW.x.ALI                .TTTAACACCACCACAAACCAG-----G-G-T-CA---ACCA---C-----A--------C---G--T-A---GAT---CAC-AT-GG--TG-T--AAGGT-TAGA--C-------A--AC-------AC--CC-A----G---CG--G---TC-AT-ATTCA--CTTT-C-- 7423
H2A.SN.x.ST                 ....AATGACACCAAGAAAGAG-AA-CA-G-T-CA---ACCAC--C-----A--------C---G--T-A---GAC---CAC-AT-GG--TA-T--GAGGT-TAGA---------A--AC-G--------CC-G-----G-------G---CC-AT-ATTCA--CTTT-AG- 6869
H2B.CI.x.EHO                .................GAA-G--AA-G-T-TA-A-A--CC-----T--A---A-T--CCA-G-AT-A---GAC---ACAT-AT-GG--TAGTT-AAGAT-TAGG---------T--CC-A--A------TT-G-----G-----C--G-----CT-AA-ATTCA--CTTCATG- 7404
H2B.GH.86.D205              .......ACCAGGAAGTATACC-G---A-GCT-TA-A-G--CC--C-----AA--A-T--CCA-G--T-A---GAC---ACAT-AT-GG--CAGCT-AAGGT-TAGG---------T--CC-G--G---TTTC-A----G---------G---CC-AC-ATTCA--CTTCATG- 7409
H2G.CI.92.ABT96             .TCCAATAAGAGTGAGAATGAG-GT--A-GCT-TA---GGCAC-----T--A-----------CAGG-AT-A---GAT---ACAT-AT-GG--TG-T--AAGAT-TAGA---------T---C-A--G-------TT-G-----G---------G---CT-AC-ATTCA---TTT-C-- 6773
H2U.FR.96.12034             ....AATACTAGTGAAAGTGAA-GT-AA-G-T-CA---ACCAC--C----------T--T--CCA-G--T-T---GAC---ACAT-AT-GG--TG-C--AAGAT-TAGG----------A---C-A--C-A----TC-GT---G-------G---CT-AT-A-TCA---TTTATG- 6902
SMM.SL.92.SL92B             .......GGAGGAAATGAAAGC-GT--A-G-T-CA---ATCA---C--T--AAGT--T--ACA-G--T-A---GAC---CAC-AT-GG--TG-C--CAGATATAGG--C-------A--AC-A--C-A---TT-G--T-G-------G---CA-AT-ATTCA--CTTT-C-- 6796
SMM.US.x.H9                 .AGCAATGGAAATGAAACTGAC-GT-AA-GCT-TA---ACCA-----------AGT--T---CA-G-AT----GAC---CAT-AT-GG--TG-T---AGAT-YAGA----------A---C-A----A---TT-G--T-GG-------G---CC-AT-ATTCA--CTTT-CT- 6837
SMM.US.x.PGM53              .AGTACTACTAGTAGAAATGAA-GT-AA-G-T-TA---ATCA-----------AGT--T---CA-G-AT----GAC---CAT-AT-GG--TG-T---AGAT-TAGA--C-------A---C-A----A---TT-G--T-GG-------G---CA-AT-ATTCA--CTTT-CT- 7294
SMM.US.x.SIVsmH635F_L3      ....CCAGCAAATGGAAGTGAG-GT-AA-G-T-CA--CAGCA-----------AGT--T---CAGG-AT----GAC---CAT-AT-GG--TG-T---AGAT-TAGA--C-------A---C-A----A---TT-G--T-GG-------G--TCA-AT-ATTCA--CTTT-CT- 7367

MAC.US.x.239                .GGGAATAACACTGGTAATGAA-GT--A-G-T-CA---ACCAC--------T--T--T--CCA-G--T-T---GAC--ACAT-AT-GG--TG-T---AGAT-TAGG--------A--TC-A----A---TT-G--T-G-------G-C-CA-AT-ATTCA--CTTTATG- 7360
Env                         __G__N__N__T__G__N__E__S__R__C__Y__M__N__H__C__N__T__S__V__I__Q__E__S__C__D__K__H__Y__W__D__A__I__R__F__R__Y__C__A__P__P__G__Y__A__L__L__R__C__N__D__T__N__Y__S__G__F__M__
MAC.US.x.251_1A11           .GGGAATAGCACTGGTAATGAA-GT--A-G-T-CA---ATCAC--------T--T--T--CCA-G--T-T---GAC--ACAT-AT-GG--TG-T---AGAT-TAGG--------A--TC-A----A---TT-G--T-G-------G-C-CA-AT-ATTCA--CTTTATG- 7356
MAC.US.x.251_BK28           .GGGAATAGCACTGATAATGAA-G---A-GCT-CA---ATCAC--------T--T--T--CCA-G-AT-T---GAC--ACAT-AT-GG--TA-T---AGAT-TAGG--------A--TC-A----A---TT-G--T-G-------G-C-CA-AT-ATTCA--CTTTATG- 7342
MAC.US.x.EMBL_3             .GGGAATAGCACTGATAATGAA-G---A-GCT-CA---ATCAC--------T--T--T--CCA-G--T-T---GAC--ACAT-AT-GG--TA-T---AGAT-TAGG--------A--TC-A----A---TT-G--T-G-------G-C-CA-AT-ATTCA--CTTTATG- 6491
STM.US.x.STM                ....AATGTCACTGGCGAAGAG-G---A-GCT-CA---G-CA---C--T---AGT--A---CA-G-AT----GAC--ACAT-AT-GG--TG-CT-AAGGT-TAGG--------A--AC-------A---TT-G--C-G---C---G----CA-AT-AT-CA--CTTT-C-- 6999
MNE.US.x.MNE027             .GGTAATAGCACTGAAGATGAA-GT--A-G-T-CA---ATCAC--------T--T--T--CA-G-AT-T---GAC--ACAT-AT-GG--TG-T---AGAT-TAGG--------A--TC-A----A---TT-G--T-G-------G-C-CA-A--ATTCA--CTTTATG- 6831
DEB.CM.99.CM40              .AGA.....................CAGGGT-CA---A-CAC-----TGA---T-----C--C---T-T---GA--CTAGCCG----A-A--TT--AAG--CAGG--C-------A--A-----A-A--GTC-A----G-----G--G-----A-AT--T------T-----GT 6676
DEB.CM.99.CM5               .AAA............AGTATA-GACATGGGT-TA---ACCAC--C----AT-----T--A------T----GA--CCAGTCG----A-----TT--AAG---AGA--C-------A--A-----A-A--GTC-GT---G-----G--G-C-GA-AC--T------T--C--CT 6627
DEN.CD.x.CD1                ....GATAGGGAAACTAACCAA--ATAT---T-TA---GGTA---C--TCAA--CAC-------C--A--T---GA-------TGT-------A---T-CAAA---AGGATG-------A--A--------C------TC-C--T--G-------G--C-TC-T-GG-G---GTAT--C- 6987
LST.CD.88.447               .........................CT-G-T-TA----CCAC-----TGATAGTA-T---CAG--A-AT--CAAT----G-GTAA-ATCTAATGC-TATT-TAGGCT-------A--A--A--C-A-AT-C--T-----G--C--CG-A--ATT-AATTTCTC-G-TAAT. 6052
LST.CD.88.485               .........................C--G-T-TA----CCAC-----TGATAGTA-T---CAG--A-AT---AAT----G-GTAA-ATCTAATGC-TATT-TAGGCT-------A--A--A--C-A-AT-C--T-------C---G-A--ATT-AATTTCTC-G-TAAT. 6049
LST.CD.88.524               .........................CT-GCT-CA----CCA---C--TGATAGTA-T---CTG--A-A---CAAT----G-GTTA-GTCAAATGC-TATT-TAGGCT-------A--A------C-ACAT-C--T-------C---G-A---CTAAA-TT-TC-G-TAAT. 6055
LST.KE.x.lho7               ...............GAGAAC-GA-C--G-T-CA----TCA------TGAT---A-A--A-----A-AT---AAT----GGGTAA-GC-AAATGC-TATT-TAGGCT-------T--A--A--C-A-AT-C-AT-G-G---C---G-AC-ATTAAA-TT-A-T-A-AA-. 7121
SUN.GA.98.L14               .........................CA-G-T-TA-----CA---C--TGAT---A-T----------A-A---C-AT----GTATTA-GC-AAATGC-TA-T-CAGGCT-------A--T--G--G-A-AT-T-AT---G-----G--G-A--ATTAAATGC-ACT-AGAA-. 7183
MND-1.GA.x.MNDGB1           ....................AT-CT-GCT-TA---ATCA------TGAG------T-A-----G-A-A-----AG----G-CTT--GATAAG-...TGT---TTTAGG-----TG--TC-A--A-A--TC--GT---G--A-----G-G---TTAAAT----AAT-A-TTG. 6449
OLC.CI.97.97CI12            .........................CA-G-T-CA--T--CA-------------T--G--A--T----AT----A-T-T-C-GT--CA--T-AGCAAATGG-GA-G-TG------A--AC-A---AGG-T-C--A----G------------G---AGAA---CC-....... 6511
WRC.CI.97.97CI14            .......................GAA-GCT-TA-A-ATCAC-----TGAG----------A-AG--A-A-----A---A-G--GGCAG--TGA-CAATAT---AGAATG------T--AC-A--A-CA-TTT-GT-G-----------G-G--ATT-AA-TCCTCT-AGCT-. 7018
WRC.CI.98.98CI04            .......................GAA-GCT-TG-A-ATCA------TGAA---T--G--A-A---A-AT----AG--A-GT-GGCAA--TGA-CAATATG--AGAATG------T--AC-A--A-C--TTT-G--T-----------G-GCCCTTAAA-TCCTCT-A-TT-. 7066
WRC.GM.05.Pbt_05GM_X02      ..........GGTAATGGTACACAA-CT-G-T-CA-A-ATCA------TGAG--T--A--CCAT----AT----A-----C--GGCA---TGAGCAATA---AGAATG------A--AC-A----A--TAT-AT---G--------G-A--ATTAAAT-G-A-T-AGCT-. 5962
MND-2.CM.98.CM16            .ACAAATGGCACAAATAACACATATTCT-G-T-TA-----CAA--C--T--A-----A--A---C----A---AAT--A-C----ACA--TGA---AAAAT-TAGACTG------A--TC-A--A-A---TTC-CT-G-GG--CCTAG-A-GATT-AATGTCTCT-A-AA-. 7028
MND-2.GA.x.M14              .AAGGAAAATGACACTGAAACTGAAGAA-G-T-TA-----CAC-----T--A-----A--A--------AT---AAT----CC---AACA--CAA---GA-AT-TAGA-TA--------TC-A--G-ACATTC-G----G-----CG-G-G--ATT-AAT--AACT-A-TT-. 6954
MND-2.x.x.5440              .ACTAATAGTACTGAACCTGAATATGAG-G-T-CA----TTCA------G-AA----A--A--------A-C--AAC--A-CT---ACA--TAG---GA-AT-TAGG-TG------A--CC-A--G----TCT-A-------------G-A--ATT-AA---GAC--A-TT-. 6655
DRL.x.x.FAO                 .AATAACTGTACTAATAAAAAA-GA-CA-G-T-TG----C-----C---T-AA----G-----C----AT---AAT----CC---ACA--TATG--GAAAT-CAGA-TG------T--AC-A--A-A-ATAT-G-----GG-------G-A--ACTAAAT-TAA-T-GTAAG. 6698
MON.CM.99.L1                ...................GTCA----T-CA-CCTG-A---C-----A---TA---A--C--A--T---GAG---TCCAAT-A-------G-----TT-G-----------A--CC-A--C-A----TT-AT-G-G-----G--G-CCCCG-C--T-CG---CA---GA 6881
MON.NG.x.NG1                .......GATAACACAACTGATGA-TAT---T-TC--CT--A------T----Y-TA---Y-C--A-------GA---TCCAAT-A------T-----T--T-----C-------T---C-C--C-A---T-AT-G-G-----G--G-CCAK-C-----CA--GCAGAACA 5497
GSN.CM.99.CN166             ...............AACCAG-GTG-A---T-TA-T-G--A------T------TAT--C--T--AAAA--GTC---TCC------CC-A--GG-----------C--C-------T--AC-C--C-A---T--G--------C---G--GTT-AT-----CA---GT---CA 6870
GSN.CM.99.CN71              ...............AATGAC--TGA----T-TA-T-T--A----C-----A---TATG-C--G--AAAG--GTC---ACC------CC-A--GG-----------C--C-------T---C-C--C------T--G--------C---G--GCC-AT-----CA---GTG--CA 6861
TAL.CM.00.266               .......................TT----T-TGCC---G-A--C--TGTG-----T--AG----A--T----GA---AACTCA--ACC---G--T-----T------G--C-------T--A--A--C------CC-GA----G-------G---GCC-A-GG-CA--T-G----A 6727
TAL.CM.01.8023              ...................GACT-ATA----T-TGCA---G-A--C--TTTG-----A--AG-T------T---GA----AC-CAT-A-C-AG-GT-----A--T--G--C------T--A--A--C------CC-GA----G------G-C--CC-A-GG-CA--T-G---CA 6271
MUS-1.CM.01.1085            .........................--T---T-TA-C----AC--C------A---TA---C--C-------A---GAG---AGCAG---CC-A--GG-----G------A--C-------A--C--G--G-ACAGCT-G---------C----G--GCC-AT---T-C---GCA----G 6834
MUS-1.CM.01.CM1239          ........................-CT---T-CA-ATT--AC-------------TA---A--T---CAA---GA---AGCAG---CC-A--GG-----G------A--C-------A--T--A--G-A-T-TC-AT-G------------G--GCT-AT---T-CA---CA-----A 6806
MUS-2.CM.01.CM1246          ........................-CA---T-CA-T----AC--C------T--TTAT--C-------CAG---GAG---AGCAG---CC-A--GG-C--A------A--C----------T--A--A-A-AGTC-CT---------C----G--GCT-AT-----CA---CA---CA 6941
MUS-2.CM.01.CM2500          ....................GA-G----CT-TA-C-----A---C------A---TAT------C--ACAG---GAG---AAGCAG---CC-A--GG--------------C----------A--A--A-A-AGCT-GT---------------G--GCT-AT-----CA---CA---CA 6870
SYK.KE.x.KE51               .............GGAAATGAA-GTTAT--CT-T--ACAGCA------T--TAGT---A--GCGCA-T--T---GA---ACAGA-T---CGC---T--------T--A--C------T--AC-G--A-ACT-AC-GT-G------------G----CAGAA---TG-A---GAT-CTG 6614
SYK.KE.x.SYK173             .............GGCAATAGT-G-TAT---T-CC--TT-CA------T--AAGT---T--AT-CGCA--A---GAG---ACAGA-----C-A---CT----G--T--A--C------T--AC-G----ACT-AT-GT-G------C----G----CA-AT---TG-A---GAC-ATG 6918
RCM.GA.x.GAB1               .AATTCTACAAATTCTACAAAA-AG-AA-G-TTTG-A-GGCAC--C--T--AA-TTC---ACA---ATTT---GA-CCAAAG-A--GG--A---T--AGGT--AGA--------T--AC-A--G------CT-A--GGTC--C--AG----A-AT-AT-CA--CTTT-ATA 6653
RCM.NG.x.NG411              .......GGTAATAATAATACT-GGCAA-G-TTTA-A-GGCA-----------AA--AGT---ACA----TT----GA-CCAAA--A--GG--A---T--AGGT--AGG--------A---C-A-----A----C--A--G----------CT-G--GCA-AT-AT-CA--CTTT-ATA 6697
GRV.ET.x.GRI_677            .AGTAAG...................GAT-G-T-CA---T-CA-----TGAT-----G--A-A-G-A--T---GAT--AAC--AT-GG--TA-TT-AAGAG--AGA--C-------A--A--A--G-A---TT-G------------------G------GAT-AT-GA--CTTT-CT- 6919
SAB.SN.x.SABIC              .ACA............CATACTGTGG-A-G-T-CA---TCCA------T----------A--C-A-G----T---GAT--AACT-AT-GG--TA-CT--AGGC-TAGG--------T--A--A-----A---TT-G----G----GCAG---CAGAT-ATTCA---CATAAGG 7222
TAN.UG.x.TAN1               .ACGAGA...................G-T-G-T-CA---TTCAC--C---GA---T--A--A-A-G-A--T---GAG--AACT-AT-GG--TA-CT-AAGAT--AGA--C-------A--A--A--C------T--CT-------------GG----CT-AT-AT-CA---TTT--TG 7011
VER.DE.x.AGM3               .AAGAAGGGTAACAATTCTAAC-GAGAA-G-T-TA---TTCA------TGAT-----T--A-A-G-A-------GAT--AAC--AT-GG--TGAGT-AAGAT--AGG--C------------A--G-----TT-AT--------C--CG--T-TGAT-ATGCA--GTTTAAGA 6546
VER.KE.x.9063               .ACAAAA..................GAA-GCT-TA---TTCAC-----TGAT-----T--A-A-G-A-------GA---AAC--AT-GG--TGAGC-AAGGC--AGA--C------------A--A--A------TC-GT--------CC--G--C-TGAT-AT-----TTATAAG- 7048
VER.KE.x.AGM155             .AAA.....................GAG-GCT-TA---TCCAC-----TGAT-----T--A-AGG-A--T---GAT---AC---AT-GG--TGA-T-AAGAC--AGA--------T--A--A--A-AC--TT-G--T-----------G--TG-GAT-ATGCA---TTTAAG- 7028
VER.KE.x.TYO1_patent        .........................GAG-G-T-CA---TTCA------TGA------T--A-A-G-A--A---GAC--AAC--AT-GG--T-AGT-GAGGT--AGG--------T--A--A----A---TT-G-------------G--G-AGAT-AT-----GTATAA-- 7027
COL.CM.x.CGU1               ......................-T---GGCA-CAA---A-GA----GCA-AT-ATACAG----C--A-TA---GAT-T-ACTATAG-G--T---G--AGG-C-GGG-TC------AG-TC-A--A-ACAT-C--T---G----G--G----A-A--GGG-----G------G 6343
```

```
H1B.FR.83.HXB2              CA......TGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGAAGAGGTAGTAATTAGA.......................TCTGTCAATTTCACGGACAATGCTAAAACCATA 7076
Env                         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  A  E  E  E  V  V  I  R  .  .  .  .  .  .  .  .  S  V  N  F  T  D  N  A  K  T  I
H1A1.UG.85.U455             --.......--C-GG-----------------------------C-A-----G-----------------------------------------AG--AA--AGG------.......................----AA--------AA------A--------- 6521
H1B.US.90.WEAU160           -C.......----A------------------------C--------A--------G-----C---T-----------------------------CA-----------.......................----AA--------------------A---- 7086
H1C.ET.86.ETH2220           --.......--CCAT-------------------C-----------AA-----G-------------AT---------T---A---------GT---AC-A---------.......................-T--AA--C-G--AA------C---TA--- 6479
H1D.CD.84.84ZR085           --.......--C-A----------T--------------------------G--G---------------------------------------A-C-----A---.......................----AA--C---AA----------T---- 6618
H1F1.BE.93.VI850            --.......--C-AG-----------------------------AA----G-------------T--T-----------C-----------GTA-----C---.......................---CAA--A--T-AA-T----A--------- 6398
H1G.SE.93.SE6165            --.......----A---------T---------------------A-----G-------------A--AC-G---------T---------G-A-AA--AA-G-----.......................----AA--------A-----A-C---GT---- 6500
H1H.CF.90.056               T-.......----------------T---------------------A----G-------------------A--------A--C-----------C--A-CA---------.......................A--AAA--A--T-A-----A-C---A---- 6407
H1J.SE.93.SE7887            --.......--C-A---------T--------------C--------A-----G-------------A---C----------A---------G---CA--A---------.......................----AA--CA--T-A-------C---A--- 6389
H1K.CM.96.MP535             --.......----A----------------------------------AA----G--------------------------C-----------------A--A-------G.......................----AA--A-T--A--T--A-A---A---- 6284
H101_AE.TH.90.CM240         --.......----A----------T---------C--------------A-----G-----------T---------------------------A--A---C---.......................----AAG--C----AA-------C-------- 6656
H102_AG.NG.x.IBNG           --.......--C-A---------------------C----------C-A----------------------------------------G---A-----G--C---.......................----AA--A----AA-------C-------- 6600
H103_AB.RU.97.KAL153_2      --.......----------------T----------------------A-------------------------------T---------------------------.......................-------------------------A---------- 6274
H104_cpx.CY.94.CY032        --.......--C-----------T-G----G---C--T-----------A-----G-----------T-------------A--CT-----ACG-------------.......................---AAA---A----A------A-C---AT--- 6464
H1N.CM.95.YBF30             GC.......----------G--T--T--------C-----------A-A-----GA----C-----GT-AA-CC-------A--CT--AATAC---T-GAA-T--T------......................-AT--T-G-CAC-GT-ATC-G 6657
H1O.BE.87.ANT70             --.......--C-G---CA-A-CGGT---TACT-----------C---------AC----AGT-----G--AA-A---------G-CA--CT-TA---G-A-AA--AG---G-TG.......................G-AAAAG--A-TTT---AGG--GA----AT--C 7112
H1O.CM.91.MVP5180           T-.......--CCAC---A-TTCAGTG--TACT-----------C--C-A----AC----AGT---------AA-AC-G-----G-CA--CT-TAG----A--A--AG-----TG.......................GGAAAA---A-T--A--ATCA--A--G-AT--C 7111
H1P.FR.06.RBF168            -T.......----ATC-GACAGT--T-AC---C-----------------A----ACR--R-----A---T-RA-AC-------A-C----------G-R----CCCT--G-A-TT.......................A--CAA---G-GT-A----CAAGAT-TGTT--C 6558
CPZ.CM.01.SIVcpzCAM13       A-.......----G------A---------C--T--------------C-AA----------A----C---T--A-CA-T-----G---------TAC----A-CA--AC-G-A---.......................GTAAA----GC-T-TA-A---A-CC-TGATTGG 6687
CPZ.CM.05.SIVcpzEK505       A-.......----AC-----A--T--T------------C---------A-A------GA----C--A--GT--A-CC-------------CT-----AC-TC-A-TA-C-----A---.......................AAT----CA-AA---...........--AC-C 6613
CPZ.CM.05.SIVcpzLB7         --.......----AC-----T--T--------C-----T-----------AA-----TA-T--T-------T--A-CC-G-----T-----------------A-AA-CA-G-----G.......................T-CAAGG-T--T.........-AGG-T 6608
CPZ.CM.05.SIVcpzMB66        AG.......----A------T--T--T--G--T------------C--C--A--T-----C--------GT-T--T--------A-CCT---A----A-G-T-AC---CC-AGAT.......................AGG---G--T-AA-TG--ATGG-C--T--- 6718
CPZ.CM.05.SIVcpzMT145       -T.......----A------A--T--------T------------C----AA-----G-C-A-C--C------A-T--G-----G---AC-------CA-TC--ACT--GGCC---.......................A-AGACCCTTCA-AG--TTTAGC- 6597
CPZ.CM.-.SIVcpzMB897        --.......----G------T--------G--T-----------T----A---T-----C--T--C---T-C--C--G-----A--C--G-------G-A-CAAGACTG-AGTT.......................AGG--A-GAAGT-AATCT---A-AG-G--T--C 6626
CPZ.CM.-.SIVcpzDP943        A-.......----G---A-A--T--------------T-----C--C-A------G---A-T--A--RT-AA-TC-T-----A--------T--GA-CAGCACGAYTGY---G.......................A---AAS-CA-G-MAR-A-TG-A-R-G-----C 6645
CPZ.CM.98.CAM3              --GGAAGA----AC---A-A---------C--------T-----C--C--A-----G---A-T--A---T--A-CC-T-----A---G----C---A-CA--ACTA-TGC---.......................AGG--C-GGT-C-R-T-C 6507
CPZ.CM.98.CAM5              --.......----------A-A---------------C--T-----C--A--A---------A-T--A---G--A-CC-T-----A--C-----T-----G-G-CCC-CTG--GCTAGA.......................-G-GAAGATCTA-GACAC----T---T 6783
CPZ.GA.88.GAB1              A-.......-----T-----A--T-----G--T-----------------AA-----G--TA-C-----GT-A--TA-T-----A---T----T----G-A-CA--ACTG-A---.......................GTG-AG---AAGT-AA-A---A--G-TGT-TGG 7137
CPZ.GA.88.GAB2              AG.......----A------A--T-----G--T--C-------------A----ACT---A-T-----GT--A-AC-C--------CT-------------A-AG--GG-GTTC...................AGAACTAAGAA-A-G-CAGCTC-A-GTCTCAG-G-C--AG-T 6490
CPZ.US.85.CPZUS             AG.......----A------T--T-----G--G-----T-----C--A-AA-----G--GA-T--A--G--CA-AA------------T-----ACTA--A-T--TACTG-A---.......................-G-AAA...A--TT---AG-C-TT--T 7111
CPZ.CD.90.ANT               TG.......----A--------TCAGT------T--C--T----------A--C---A-G--GG----ATGGT-A--A--------A-C-TACCA-AC-A-CACTTC----G-A-TG.......................AA--GTCGCAAA-AT--ATC--TGCTTGTA-G- 6529
CPZ.TZ.00.TAN1              A-.......--CT-C-------TCAGT---T--C-------------G---TAT--TA-GA--G-C--AGC-T-A-AC--------GTCC--G-A---------AAC-AA-GC-TAC.......................T-TGT-AAT--CTCAGT-AA--C-C--T-- 6717
CPZ.TZ.2001.TAN2            -C.......---T-C-----G-CAGTG--G--C--C-------------TTC--TA-GA--G-C--AGCTT---AC--------G-CAT---A----A-T--AAC-AA-GC-TAC.......................T-TGC-GATGTCA-TGTCAA-CCGC-TT-- 6749
CPZ.TZ.2002.TAN3_1          TC.......---TAC--C--G-CAGTG--G--T--C-------------TTT--TA-GA--G-C--AGCCT---AC--------A-CA--C-A----A-G--AAC-AA-GC-TAC.......................T-TGT-AATG-CACT--CAA-CCGC-TT-- 6782
GOR.CM.2007.SIVgor2139_287  A-.......--C-G---C--GTCA-T---T--T-----C-----G--C------AC---C--C--A--GT--A-AC-C-----A-C-T----------GGA-T-CCAC---A-TC.......................AGCAAA---AGT-GT--T-G--GA---GAT--T 6652
GOR.CM.2007.SIVgorCP2135con A-.......--C-A------GTCA-T---T--T-----C-----G--C------AC---C--C--A--GT--A-AC-C-----A-C-T-----A---GC-T-CCAC---A-TT.......................AGCAAA---AGT--T--T-G--GA---GAT--T 6633
GOR.CM.2004.SIVgorCP684con  A-.......--C-AC-----GTCA-T---T--T-----C-----G--C------AC---C--C--A--GT--A-AC-C-----A-C-T-----A---GGA-TAAT-CC-CA-T-.......................ATCAGCAG---AGT-GT--T-G---A-C-GAT--T 6571
H2A.DE.x.BEN                -TAAG...--C--T--A--AGTAG-T-CTTC---C---AGGATG--GGAAA-GCA-ACT--T---TGGT-TGGC--T-------C-AG-------A-TAGAAC-TAT--CTAT.......................TGGC--GG--GA--T---AGG-CT-T---T 7582
H2A.GW.x.ALI                -CAAT...--CT-T--A--AGTAG-T-CTAC---C---AGGATG--GGA-A-GCA-ACT--T---TGGT-TGGC--T-------C-AGG------A-TAGAAC-TAT--CTAT.......................TGGC--GGT-GA--T---AGA-CT-T---C 7563
H2A.SN.x.ST                 -CAAT...---T-T--G--AGTAG-T-CTAC-------AGGATG--GGAAA-GCA-ACC--C---TGGT-TGGC--T-------C-AGG------A-TAGAAC-TAT--CTAT.......................TGGC--GGT-G---T---AGA-CT-T---T 7009
H2B.CI.x.EHO                -CAAC...----GT--G--AGTAGT-TCT-TG...TACAGAATG--GGAAA--CAGACC--T--ATGGT-TGGC--C-----T-CAAGG------A-CAG-AC-TAT--CTAT.......................TGGC--GGT-AA------AGG-CT-T---- 7541
H2B.GH.86.D205              -CAAC...--C-GT--G--AGTAG-GTCCTCC--C---AGAATG--GGAAA--CAGTCC--T--ATGGT-TGGC--C-----T-CAAGG-----GA-CAG-AC-TAT--ATAT.......................TGGC--GAA-AA------AGG-CC-T---- 7549
H2G.CI.92.ABT96             -TAAT...--CT-T--A--AGTGGTCTCTAC-------AGAATG--GGAAA--CA-ACC--T--ATGGT-CGG---T-----A-CAAGG------A-TAGAAC-TAT--CTAT.......................TGGC--GGG-GA------AGG-CT-TA--- 6913
H2U.FR.96.12034             -TAAT...--CT-T--A--AGTAGTCTCTAC---C---AGAATG-GGG-AA--CA-ACC--T---TGGT-TGGA--T-------C-AG-T-----A-TAGAACTTAT--ATAC.......................TGGC--GGT-GA------AGA-C--TA--T 7042
SMM.SL.92.SL92B             -TAAT...----GT--G--AGTAGT-TC-TC-------AGGATG--GGAAA--CAGACT--T--ATGGT-TGGC--T-----T-C-AG------GA-TAGAAC-TAT--ATAT.......................TGGC--GG-GGAAGT---AGA-C--TA--T 6936
SMM.US.x.H9                 -TAAC...-----T--G--AGTAGTTTCTTC---C---AGAATG--GGAAA--CA-ACC--T---TGGT-TGGT--C-----T-C-AG-------A-TAGAAC-TAC---TAT.......................TGGC--RG--GAAG---AGA-CC-TA--T 6977
SMM.US.x.PGM53              -TAAC...---T-T--G--AGTGGTTTCTTC---C---AGAATG--GGA-A-GCA-ACC--T---TGGT-TGGT--C-----T-C-AGG------A-TAGAAC-TAC---TAT.......................TGGC--AGT-GAAGT---AGA-CC-TA--T 7434
SMM.US.x.SIVsmH635F_L3      -TAAA...---T-T--G--AGTGGTTTCTTC---C---AGAATG--GGA-A-GCA-ACC--T---TGGT-TGGC--C-----T-C-AGG------A-TAGAAC-TAC---TAT.......................TGGC--GG--ATAGT---AGA-CC-TA--T 7507
                                                           V3 loop end

MAC.US.x.239                -TAAA...---T-T--G--GGTGGTCTCTTC---C---AGGATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC.......................TGGC--GGT-G---T---AGG-CT-TA--T 7500
Env                         P  K  .  C  S  K  V  V  V  S  S  C  T  R  M  M  E  T  Q  T  S  T  W  F  G  F  N  G  T  R  A  E  N  R  T  Y  I  Y  .  .  .  .  .  .  .  .  W  H  G  R  D  N  R  T  I  I
MAC.US.x.251_1A11           -TAAA...---T-T--G--GGTGGTCTCTTC---C---AGGATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC.......................TGGC--GGT-G---T---AGA-CT-TA--T 7496
MAC.US.x.251_BK28           -TAAA...---T-T--G--GGTGGTCTCTTC---C---AGGATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC.......................TGGC--GGT-G---T---AGG-CT-TA--T 7482
MAC.US.x.EMBL_3             -TAAA...---T-T--G--GGTGGTCTCTTC---C---AGGATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC.......................TGGC--GGT-G---T---AGG-CT-TA--T 6631
STM.US.x.STM                -TAAT...--CT-T--G--AGTAGTCTCCTCG-----TAGAATG--GGA-A--CA-ACT------TGGT-TGGC--T-----A-C-AG-------A-TAGAAC-TAT--CTAT.......................TGGC--GG--G---T---AGG-CT-TT--T 7139
MNE.US.x.MNE027             -TAAC...---T-T--G--GGTGGTCTCTTC---C---AGAATG--GGA-A--CAGACT--T---TGGT-TGGC--T-----A-C-AG-------A-TAGAACTTAT---TAC.......................TGGC--AG--AA--T---AGG-CT-TA--T 6971
DEB.CM.99.CM40              T-.......--C-AC--C--G-CTG----GGC-------A--CTC--ACACA-CA-G-C---C--ATGGG-A-AA--T-----GTC-GAT-AG--GAGG-CA-A--AGC--CATATA.....ATAAGAAAAGAGGTAAAA-GTG-GG-GCA-A-TGGAAGC-TT--T--- 6834
DEB.CM.99.CM5               TG.......-----C-----G-CTG----GGC-------A--CTC--ACATA-CA-G-C------ATGGA---AA--C-----TTCAGAT-A----AGG-CA-A--AC--ACATATC..............ATAAGGAAG...GGCC-GGAAA-T--GA-C-TT--TG-C 6773
DEN.CD.x.CD1                TT.......----A----A-GTCAG--AG-GTG------G---AG--ACACA-CA-G-CCAGC--CTGGT--A-CA-G--CA-T-CAG---A-CTTCT---TAATCAT-CACA-ATT..............ATAAGAAACACA-GCAAAGATA-TC-GA-A-T-G----T 7136
LST.CD.88.447               ........-----C---A-A-CAG--ACC-CT--C---G-CTAT--G-TTAGTTC---GAGT-G-TTCT-TGGC--T-----A-CCAATCATAC---C--T-A-T-----CC-CTA.........ACTCCAAAAAAGATG-GA---C--CTA--AGCAAAGTTTGTATAT 6205
LST.CD.88.485               ........-----C---A-A-CAG--ACT-CT--C---G-CTAT--G-TTAGTAC---GAGT-G-TTCT-TGGC--T-----A-CCAATCATAC---G--TAAGT-----CC-CTA.........ACTCCAAAAAAGATG-GA---C--CT----GCGAAGTTTGTATAT 6202
LST.CD.88.524               ........-----T---A-A-CAG-TAC--C---C---G--TAC--G-TAAGTTC---TAGT-GCTTCT-TGGC--T-----A-CCAATCATAC------C-AGT-----CC-CTA.........ACTCCAAAAAAGATG-GAG--C-T-AT-GAGC-AAATTTGTTTAC 6208
LST.KE.x.lho7               ........---GA----A-A-CAG-TAC--C------TGGATAT--GCTTAGTAGT--AGT-G-TTCT-TGGC--T-----A-C-AATCATACCAGG--T-AGC-T---CC-CTA.........ACTCCAAATAAGATG-AAG--C-A-AT-GTGCAAAGTTTGTGTAT 7274
SUN.GA.98.L14               ........----A----A-A-CTG-TACC-CT-----TA--TAT--G-CAAGTAC---TAGT-GCTTTT-TGGC--T-----A-C-AG-CATA-G-----T-AGT---ACC-ATC............AATAACAAGGTA-GAC-AGGGG-T--GGGA-AAT-TGTATGG 7333
MND-1.GA.x.MNDGB1           ........----GC---A-ATCAG---G--G-----T--GCACT-AGTAG-CAC---AGT-GCTTTT-TGGC--T-----A-C-A-GCATA-G...-A-G--A-T-G-T-CCCATAGATGATAAATATAGGGGCC-A-AGG-A--TCATC-A-GGAAGTTTGT-TAT 6608
OLC.CI.97.97CI12            ........--C-G---CA-TTCAG---TACT--C--C-CGCC-C-GTCAG--AC---GAGT-G-TTCT-TGGC--T-----T-C-AG-CATA-G-C-AGCAG-C-G---CCTATA...ACAGGGGATTTGAATCATAAACCA-GAGAGGGAACAGGA-TAT-TGTT--G 6670
WRC.CI.97.97CI14            ........---GAG------TCAG----ATG------G-CATG--G-CAG--ACCA-TAGT-GCTTCT-TGGA--T-----A-C-AG-CATA-G--G-GA-A-T-----GA-GTA...ATAGGAAATATAAATCATAAACCA--GGAAGGATCAGGA--ATTTGTGTAT 7177
WRC.CI.98.98CI04            ........---GAG-----ATCAG----TT---C---G-CATG--G-CAG-TACTA-TAGT-G-TTCT-TGGA--T-----T-CCAG-TATA-GA---G--A-T--G-AGAGATA...ATAGGAAATATAAATCATAGGCCTC--GAAGGTACAGGA--ATTTGTATAT 7225
WRC.GM.05.Pbt_05GM_X02      ........---GAG-----TTCTG---TACT-----TG--TT---G-CAG--ACTA-TAGT-GCTTTT-TGG---T-----A-C-AAGTAT--TA---G--A-T-G--AGA-GTA...ATAGGAAATATAAACCATAAACCA--AGATGGAACAGGA--ATTTGTTTAT 6121
MND-2.CM.98.CM16            ........-----C---A---CAG-------G--------GCC-T-GCCAG-CACCACC------ATGTGTGGA-CT-G---G-CAAA-CAT--TT-TA-T-A-T----ACAGACA.....AATACAAAGAAAGGCAAA-AGG-G--TCAT---C--AAGT-TGTGTAT 7184
MND-2.GA.x.M14              ........---G-C------TCAG-------G-----TG-CCC---GCCTG-TACTA--------ATGT-TGGC--T-----A-CAAA-CATA-TT-T--T-AGT-----TTGGTG.....AACCCTCAA.........AAGG-G--TCAT---C--AAAT-TGTTTAT 7101
MND-2.x.x.5440              ........---GGG-----GTCTG---G--G--C---GCACCGC-GCCAG-CAC-A-C--T--AATGT-TGGC--C-----A-C-AAGCAT--TT-T--T-AGC-T---CAGACA.....AACCCACGAAAAGGAAAA-ATG-G--TCAT---C--AAAT-TGTATAT 6811
DRL.x.x.FAO                 ........----A----A-A-CAG---G--G-----TAGACAT--GCCTG-CAC-A-C--T-GC--TT-TGGC--C-------CAAAGCAT-GTC-G--T-AGT----AGA-ACA..............AGGGGAAAACCAG-C--TCT-A-TC--AAAT-TGTGT-T 6845
MON.CM.99.L1                GT.......---T-T-----TTCTG----AC-------T----CT--CCAA--TA-----G-T---TGGT-C-A---G--CA-T-CGGGTAAT-CTCCCA-TAC-AC-G-A-TG.......................ATGAAT--GCAG-AAA-TG-GT---T-GTAG-T 7021
MON.NG.x.NG1                AT.......--CT-R------TCAG----AC---C--G-----T--A-T----A-------T---TGGT-T-A-C-T----CT-CCTCT--CAC-A-TR--ACCAC-G---TGATG.......................AATGGGAAA-GAA-TG-AT-C-T-GTA--- 5637
GSN.CM.99.CN166             --.......--CTAC-----GTCAG----ACC--C---------C---CAA--CT-----G----ATGG--C-AT--------G-CCTATCAGCC--G-A-TAATACC-GA-T-ATG.......................ATGAATGG-A-G-A-A-TG-AT---T-GTT--- 7013
GSN.CM.99.CN71              AG.......---T-C-----ATCAG----AC---C--G-----C---CAA--TT-----G----ATGG--T-ATC-------G-C-TATCAGCC--GGA-CAATACC-GAGT-ATG.......................ATGAATGGCAAA-A-A-TG-AT---T-GT---- 7004
TAL.CM.00.266               AG.......---CAC--CA--TCAG-------C------G-CCCT--C-CCA--A---C-G-T--GTGGT-AA-TC-C-----G-CAAAG.........--A-TCACAC-CAGGTCATA.......................CA-CAGGC-CCA-A-A-TG-AAG-GCT-TA--C 6864
TAL.CM.01.8023              AG.......---CAT--CA--TCAG-------C--C---G--CCC--C-CCA-CA---C-G---ATGGT--G-TC-------T-CAAAG.........--A-TCAGAC-CAA-T-ATA.......................CA-CAGGCACC--AAA-TG-AT-AGCC-TA--C 6408
MUS-1.CM.01.1085            A-.......----AT-----G-CTG-TAGG--T--C--G-----C--ATTA--CT-G--TG-T---TGG--A-A-C-T-----A-C-TATTT---G-G-A-TAATAC-GCAGT-ATG.......................ATGAATGGGGAT-AAA-TG-AT-A-T-GGA--- 6977
MUS-1.CM.01.CM1239          A-.......----AC-----G-CAG-TACG--T--C--G-----C--ATTA--CT-G--TG-T--CTGGT-A-A-C-T-------CATATTT----G-A-TAATAC-GCAGT-ATG.......................ATGAATGGGGA--AAA-TG-AT---T-GGA--- 6949
MUS-2.CM.01.CM1246          TT.......----AT-----A-C-G-TAGG--C--C--T---AC--ATTA--GC--A-TG-T---TGG--A-AA--G-----T-CCTACCT-CC--GCA-T-ACAC-GC-GT-ATG.......................ATGAATGG-GAT-AAA-TG-AT---TTG----- 7084
MUS-2.CM.01.CM2500          A-.......----GG-----A-C-G-TACG--T--C--G-----C--ATT----C---TG-T---TGG--A-A-C-------T-CCTACTT-CCT-GCA-T-ATAC-GCAGT-ATG.......................ATGAATGGCAAG-AAA-TG-GT---T-G-T--- 7013
SYK.KE.x.KE51               AG.......---CA------A-CAG---TTC---C------CCTT-C-ACA-TT-G-CTAGT--ATGGT-T-A-----------CCTATAA--C-A----CAA---T-GGTTTATT.......................AGACAA--GGAT-AAA-TG-AT-AGTG-TT--- 6757
SYK.KE.x.SYK173             TC.......-----------A-CTG----TC---C-----G-A-T-C-ATA--T-G-CTAGC--ATGGT-C-AAC-G-----G-CCTACAA--C-A---TAA---T-GATTTATC.......................AAACAA--GGAT-AAA-TG-GT-AGT--TT--T 7061
RCM.GA.x.GAB1               -C.......---GTT-----T-CTG-CACTTC---C------ATG---ATA-TACT--GG-CT-AGGGT-TGGA--------ATCAA-TAAT-T-A-T---ACTTGG--ATAT.......................CAG-GAAGGCAAAGT---AGG-C-GTT--- 6790
RCM.NG.x.NG411              -C.......---CA------T-CAG--ACGTC---C-----CATG--A-ATA--AC------C--ATC-T-TGGA--------ATC-A-TAGT--GA-TAGAAC-TGG--CTAT.......................CAA-GAAAACAAAGT---AGG-C-GTT--- 6834
GRV.ET.x.GRI_677            --AAG...--C-AG-----TTCAGT---G--T-----TAGATT---C-ATA-TACTA--A-T--AGGGA-AGGA--------T---AG-T-----A-TAGAAC--AG--AT-G.......................CAGAAA-GAGGAAATGATA-TG--A-AGTT-TA--- 7065
SAB.SN.x.SAB1C              --.......----GG-----A-CAGT-TC-GCT--C--TAGGTT---C-ATA-TACT---AGC--AGG-A-AGGAA-T-----T---TAT-T--C-A-TAGAAC--A---CT-G.......................CAGAAAAATGGCAATT-CA-TG--T--GTT-TT--T 7365
TAN.UG.x.TAN1               TT.......----G------TTCAGT---TAGT-----TGGATTG--G-ATA-TAC---GAGCT-AGC-T-TGGCA-------------AG------A-CAGAAC--A---AT-G.......................CAAAAGCATGGAG-G-GCA-T---T--GTG-TA--- 7154
VER.DE.x.AGM3               --AAC...---T-T-----TTCAGTG--G--T-----TA-CTTG--A-ATA--AC---GA-T---GG----T----G-----G--CTACT---GA-TCGAACCCAG--AT-G.......................CAGAAACAT-GAG-A-GCA-TG-CT-AGTGTTAG-G 6692
VER.KE.x.9063               A-AAT...----GT-----TTCAGT------C-----TGGACTT--G-ATA--AC---T--T--AGG---A-----G-----A--CTATT-T--GA-TCGAAC-CA---AT-G.......................CAGAAGCAT-GAG-G-ATA-TG--T--GT--TAG-T 7194
VER.KE.x.AGM155             --GAA...---T-T-----TTCAGT---G--T--C---AC-TT---G-ATA--AC---A-C---GGT---T-A--G-----A--CTATT-----A-TCGAACCCAG--CT-G.......................CAAAAACATGGAG-G-GCA-TG-CT-AGTGTTA--C 7174
VER.KE.x.TYO1_patent        A-AAT...--CT-------ATCAGT---G--T------GGCTT---G-ATA--AC---GA----AGGGT--T--C-G-----A--CTATCAT--GA-TCGAACCCAG--AT-G.......................CAGAAACAT-GGG-A-ATA-T--CA-AGT-TTG--C 7173
COL.CM.x.CGU1               --.......--C-AC-----G-CAG---GTCC-----C----AGT---ATAT-AC---CATGT----TG-AT-AG-C----C-TCAAA--AGTTGTCT--TTGG-CG-AAGAT...................AGAGAGGGA---TGGAAG-AT---TCA-GG-C--TTGAG 6489
```

```
                                           V3 loop start                                                                                                                          V3 loop end
H1B.FR.83.HXB2             ATAGTACAGCTGAACACATCTGTAGAAATTAATTGTACAAGACCCAACAACAAT...ACAAGAAAAAGAATCCGTATCCAGAGAGGACCAGGGAGAGCATTTGTT...........ACAATAGGAAAA...ATAGGAAATATGAGACAAGCACATTGTAACATTAGTAG 7228
Env                        _I__V__Q__L__N__T__S__V__E__I__N__C__T__R__P__N__N__N__.__T__R__K__R__I__R__I__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
H1A1.UG.85.U455            -----------TGT--ATC-----A----C------T-C-----TT----T-CAAGA-A--AT-T-C-T-GGTA--GT...-T----T----ACA------CTA-...........GT--C---T---ATA-----GG----A----------------TG-C----- 6676
H1B.US.90.WEAU160          --------------TGT---CA-------------------------------------.....---------A---AAC-T-A.....................T-C--TA-...........-----C----G--ATA------G----A----G-------------C------- 7235
H1C.ET.86.ETH2220          -----------T--TGA-------------C------G-------G---T----......-------G----T--AA-G--A....................ACA-A----CTA-...........G---C----G-CATA------G----A--------------------------GA 6628
H1D.CD.84.84ZR085          --------T--T--TCA--------------------------G---T----A--G...GA-----C----G-CA-C---A..........A-----CA---GC-CTA-...........-----C-A-GT-TACA-C-A-G-TA--AG--------T-------T------G- 6767
H1F1.BE.93.VI850           --------C--T--TGA-------C-G-----------------------.......-----G---G-T--A-A-T-A......--------ACA-A----CTA-...........G---C----GC-ATA------G-C--A---A-G--------------------G- 6547
H1G.SE.93.SE6165           -----------T--T-A-A------------------C---------------ACA-TG-A--GG-T--GAATGGGA...-TT--------ACA-A-G--CTA-...........G---C---TGCCATA------G-C--A----------------TG---C--A 6655
H1H.CF.90.056              -----------T--G---C-A---A-C----CA-----C--G--T-----T---.......--G----C---T--A-A-T-A.......--G-----AC-------CTA-...........G---C---TG-CATC------G----A----------------T-------- 6556
H1J.SE.93.SE7887           --------A--T--T-A-A-----------GTG---TAC-----T--T---------......-----G---G-T--A-AC--G......------ACA--TGC-CTAC...........G---C----G--ATA-------------A---G--A-----------------GA 6538
H1K.CM.96.MP535            -----------T--TGA-A------C-----------------G--A--------......----------T--A-A---G....................A-A------CTA-...........-----C---TG-TATA------G----A-----G---------C---------G- 6433
H101_AE.TH.90.CM240        -----G--C--T--T-A-------------C--------C------TC-----------......------C---T--AAC---A.................AC---T---CTA-...........-G--C----G-TATA------------A---A-----T-----G-G----A-G- 6805
H102_AG.NG.x.IBNG          ----------T--GCT-ATC-----A------------C------T-------------......-----G-TG-A-A---A.................CA-------CTA-...........G---C---TG-CATA-----GG----A----------------TG-C----A 6749
H103_AB.RU.97.KAL153_2     --------------AGA-C----G--------------------------------......---------G-T--T-A---A.........................TA-...........G---C----G-CATA-C----G----A------------------------------T 6423
H104_cpx.CY.94.CY032       -----------TGCA-AGG-----A-------------C-----TGG--------......---------TG-A-A---A......--G-----ACT-A---GGTA-...........G---C---TG--ATA------G----A-------------------------------G- 6613
H1N.CM.95.YBF30            T-G--G--ATG---TGAGA-A--GCC---A-----------G--AGGA--T---.......---G--GG-CAGG-G-AG--A..........T-CT-TGA------TA-...........-AC----A----ATAG------G-C--T---------T-C------TG-CTC--A 6806
H1O.BE.87.ANT70            -----GACC--A---T-TA-CC--A-C--G-CC----GA------AC-A-TA......GAC-T-C--GAG--GA-A--A.......--T---AT-GCCTGG-ACAGCATGGGA.......-T-GGG----C-GCAGG-AAC-GCTCA--GGC---TT-----C--GTA--A-GC 7264
H1O.CM.91.MVP5180          ------ACC--A-----TC--A--A-C--G-CC--C-T----GAAGGA-TTGCA...GAGGT-C--GAT--ATA--CA.......--T---AT----TGGCGCAG-ATGACACTTAAA-G--GTAAC--TACATC-CC--GATCA--GGT---TT------CATA--A--A 7272
H1P.FR.06.RBF168           ------A-AT-A--T-ARAA---GTC-C-C-CC---GT-------AGGT--T--C.......-------RG-CA-G-G-AA--A..........-----CAT--CCTGG-A-...............-AT--GAA-TTCTAC-C---YG----C--RA----TT---------R--TCC-T 6704
CPZ.CM.01.SIVcpzCAM13      --------A--A-G----G-A--GA-TC-G-CG----AG--GGTAGGG--T--C.......-------GGC-A-G-A-AG--A.........---------A-TGA-C--CTA-...........-AT--G-ACC-TATAT-T--TG---CA--GA----CTT---CG-AC-C-A-G- 6836
CPZ.CM.05.SIVcpzEK505      T-G--G--AT-A--TGA-AG-A-CCC---A-----C---------AGGA--T--G..........---GG-CA-G-A-AA--A...........T--T-TGA----CTA-...........-AC----A----TATA------G-C-CT--G------T-----G-GG-A-AC-- 6762
CPZ.CM.05.SIVcpzLB7        C----------C--T---AG-A--C-C--C---------T--GGTAGGA--T--GACA-T--A-GG-........CCC-A..........---------ACAGCTG--CTA-...........GG--C--A--CTGTAG------G-C-CA---G-G--T------G-A----ACCA 6757
CPZ.CM.05.SIVcpzMB66       -----GA-AT-A---GA-A-----CGTC-C------C--T---A-AGGA--T--CACA-T--A-GGT...--A-C---A.................-A-TCAGAT----TA-...........GG-----A--CTGTA------G-C-CA---G--G---CTT---CC-GT-G-AC-A 6867
CPZ.CM.05.SIVcpzMT145      ---A-------C--AGATC-A---A-G----CA-----GG-----AGGA--T--C.......------GGTCA---A-AA--A..........----CA-TGA-T---TA-...........-AC----A---TGTAG----GG-C-CT---A------T-----G-G---A-A-G- 6746
CPZ.CM.-.SIVcpzMB897       --T--G-----AGCT----G-AA--T-T--C----------C-----TAGGT--T--AACC-T-GAGGGC...--A-C---A.......--G-----ACA-ATT---TA-...........-GG-CTAA--CGGTGG----CG-C-CC----GG----G-G--C-GA----A-G- 6775
CPZ.CM.-.SIVcpzDP943       ----------T-TC-T-A-AGG--CC-G--G-CA----T-------AGGA--T--C...-GC---GG-CAG--A-AA--A.............CA-TGT-----TA-...........-ACC---A---TATT-------G---CC---------T--------TR-C--CGC 6794
CPZ.CM.98.CAM3             C-GA-N--RT-TC-G-ANN-G--CAGC--A---------------AGGA--T-NC...-GC---GG-CA---A-AG--A..............C-TGA-C---TA-...........-AC----A----TATTG-----G----CC-----G----T------C-G--C-A--- 6656
CPZ.CM.98.CAM5             ---A----AT-T--AGA-GGGA-TA-C--A-T-----TG------AGGA--T--C...-GC---GG-CA-C-A-AG--A...............GT-T-C--CTA-...........-AT----A---CATAG-----GG--CC-----------------TG-C--C-A 6932
CPZ.GA.88.GAB1             ------------AGTAGA-G-A--TAGTC-A------CAT------AGGA--------......----GGG-GAGG-G-AG--A......--G-----A-TGA-C---TA-...........-AT----A---TGTAG-----G---CC---TCT--CT-C------G--A-A-G- 7286
CPZ.GA.88.GAB2             ----------T-A--A-GGG-AA-TCCC---------C-GT-----AGGA-----......---G---G-GC---AAACC-T.....A-C--------C-A----CT-C.............-AC-CG-A-GCCCTA---------CCCT--GA----GTCC--CC--C---A-G- 6639
CPZ.US.85.CPZUS            C--------AT-CTCAGAGGGA--CA-T--G-C-----TT-----AGGA-----......----GTGGG--ATG-A-AAC-A..........-------------A-TGA-T---TA-...........-AC---CC---GATAG-----G--G-A---G-----------------CTCA-A 7260
CPZ.CD.90.ANT              T-T-G-A-AGAATT-GA-AACT--AC-----CA-----T------AGGA--T-GGACAGT-----...--ATC-A-AA--A...........-------A-TGA-T---CTA-...........-ACG---A--T-GCA-C----G-C-CT--GA----GTTC-----CAG-C-A--A 6678
CPZ.TZ.00.TAN1             T-----A-AT-T--TGT---AA--A-TT-A-CG----GA----A-AGGA---------......-----GGTCA-G-A-AG--A.......--T------T-TGA-C---TA-...........-AT----A---TGTAG----GG-C-CC---GA----TT------TCAG-C-A-GC 6866
CPZ.TZ.2001.TAN2           C-----A-AT-T--TGA---A----TC-T-A-CC---GAG---A-AGGG---------......-----GGCCA-G-A-AG--A.......--T------C-TGA-C---TA-...........-AT----A---TGTTG-T---G---CC----A-G--TT-C--CC-TG---A-GC 6898
CPZ.TZ.2002.TAN3_1         T-----A-AT-C--TGA---A----CC-T-A-CC---GGG---A-AGGA-------...GT----GGCCAGG-A-AA--A.......--T------C-TGA-T---TA-...........-AT-C---A---TGTTG-T--GG-C-CC---A----TT-C----TGGG---A-GC 6931
GOR.CM.2007.SIVgor2139_287 -----CA-A--AGC--G-C-A-----------CC----GA----GA-AGGT--T---......-------GG-CAG----AA--A.......--G--CAT-...A--A--TA-...........-ACTC--A---CATAG-------C-CA---A-G---TTT---C--GTA--A-GA 6798
GOR.CM.2007.SIVgorCP2135con -----CA-A--AGC--G---A-----------CC----GA----GA-A-GT----......--------GG-CAG----AAG-A.......--G--CAT-...A--A--TA-...........-ACTC--A---CATAG-------C-CA---A-G---TTTC--C--GTA--A--A 6779
GOR.CM.2004.SIVgorCP684con -----CA-A--AGC--ACC-A---------C-CC----GA----GA-AGGT--T---......-------GG-CAG----AAG-A.......--G--CCT--C-ATT-A-AAC...........T--GA-AAC-T-ATAGG-AAC-CA-GA-AGGCTTTCTGCAAGT-T-A-RAG-C 6720
H2A.DE.x.BEN               -GCT--A-CAA-T-TTATAA-C-CAC---GCG-----AG-----AGGA--T--G..---GTTTT-CC---AACAC-T.....ATGT----ATT--TG---...............CACTCTCAGCC-ATC-ACAC--GGCCT--G--G---TGG--CCGGT--...G- 7725
H2A.GW.x.ALI               -GCT--A-CAAAC-TTATAA-C-TACT--GC------AG--G--AGGA--T--G..---GTTGT-CC---AACAC-T.....ATGT-----TT-AT----..............CACTCCCAGCC-ATC-ATAA--GACCC---------TGG--CTGGT-C...A 7706
H2A.SN.x.ST                -GCT--A-CAA-TTTTATAA-C-CACCG-AC------AG--G--AGGA-----G..---GTTGT-CC---AACAC--.....ATGT-----TT--TG---..............CACTCCCAGCC-ATC-ATA---GACCC--G------TGG--CTGGT-C...A 7152
H2B.CI.x.EHO               -GCT--A-TTCAT-TTATAA-C-GAC---GC-C-----A---G--AGGA-----A...-TGGTTGTGCC---AA-A-C-.......-TGT----T-TTCTC--C..............CATTC-CAGCCTATC-ATAA--GACCT-A-------TTGG--CTGGT--...A 7684
H2B.GH.86.D205             -GCT--A-TACAT--TATAA-T-GTC---AC-C----AG--G--AGGA-----G..---GGTTGT-CC---AA-A-C-.......-TGT----ACT-CTT--C..............CATTC-CAGCCTATC-ATAAG-GACCC---------TTGG--CTGGT--...A 7692
H2G.CI.92.ABT96            -GTT--A-TAA-T-TTATAA-T--AC---A-RA----GG-----AGGG--T--A..---GT-TT-CC---TAC---T.....ATGT-----CT--TC--C..............CACTCTCA-CC-ATC-AT-A--GACCC-----G--TTGG---TGGT--...G- 7056
H2U.FR.96.12034            -GCT--A-TAA-T-TTATAA-T--AC---GT-C----G---G--AGGG--T--G..---GTCTT-CC-G-AACC--T.....ATGT----ACT--TG--C..............CATTCTCA-GCTATC-ATAAC---CCA-A-------TTGG---TGGT--...G- 7185
SMM.SL.92.SL92B            -GCT--A-TAA-T-TTATAA-T-GAC---A--G--C-G------AGGA--T--G..---GTCCT-CC-G--ACC--T.....ATGT----TTTG-TC---..............CATTCGCA-CC-ATA-AT-AG-GGCCA---------TGG--CTGGT--...G- 7079
SMM.US.x.H9                -GCT--A-TAA-T-TTRTAA-C--AC---G-GA----G------AGGA--T--G..---GTCTT-CC-G--ACC--T.....ATGT-----TTG-TC--C..............CATTCGCA-CCCATA-AT-AG-GACCA-A---G--CTGG--CTGGT--...G- 7120
SMM.US.x.PGM53             -GCT--A-TAA-T-TTATAA-C--AC---A-GA----G---G--AGGA--T--G..---GTCTT-CC-G--ACC--T.....ATGT-----TTGATC--C..............CATTCGCA-CCCATA-AT-AG-GACCA-A-------CTGG--CTGGT--...G- 7577
SMM.US.x.SIVsmH635F_L3     -GCT--A-TAA-T-TTATAA-C--AC---G--A----G------AGGA--T--G..---GTTTT-CC-G--ACC--T.....ATGT-----TTG-TC--C..............CATTCGCA-CCCATA-AT-AG-GACCA-A---G--CTGG--CTGGT--...G- 7650
                                           V4 loop start                                                                                                                          V4 loop end

MAC.US.x.239               -GTT--A-TAA-T-TTATAA-C--AC---G--A---G------AGGA--T--G..---GTTTT-CC-G--ACC--T.....ATGT-T--ATTG-TT--C..............CACTC-CA-CC-ATC-AT-AT-GGCCA-AG--G---TGG---TGGT--...G- 7643
Env                        _S__L__N__K__Y__Y__N__L__T__M__K__C__R__R__P__G__N__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
MAC.US.x.251_1A11          -GTT--A-TAA-T-TTATAA-C--AC---A--A---G------AGGA--T--G..---GTTTT-CC-G--ACC--T.....ATGT-TCAATTG-TT--C..............CACTC-CA-CC-ATC-AT-AT-GGCCA-AG--G---TGG---TGGT--...-- 7639
MAC.US.x.251_BK28          -GTT--A-TAA-T-TTATAA-C--AC---G--A---G------AGGA--T--G..---GTTTT-CC-G--ACC--T.....ATGT-T--ATTG-TT--C..............CACTC-CA-CC-ATC-AT-AT-GGCCA-AG--G---TGG---TGGT--...G- 7625
MAC.US.x.EMBL_3            -GTT--A-TAA-T-TTATAA-C--AC---G--A---G------AGGA--T--G..---GTTTT-CC-G--ACC--T.....ATGT-T--ATTG-TT--C..............CACTC-CA-CC-CTC-CT-AT-GGCCA-AG--G---TGG---TGGT--...G- 6774
STM.US.x.STM               -GCT-GA-CAA-T-TTACAA-T--AC---G-GC---G------AGGG-----A..---GTCTT-CC-G--ACC--T.....ATGT----CTTG-TC--C..............CATTCGCA-CCCATT-AT-AG-GACCA-A------TTGG--CTGGT--...G- 7282
MNE.US.x.MNE027            -GTT-GA-TAA-T-T-ATAA-C--AC---G--A---G------AGGA--T--G..---GTTTT-CC-G--ACC--T.....ATGT-T--ATTG-TT--C..............CACTC-CA-CC-ATC-AT-A--GGCCA-A---G---TGG---GGT--...GA 7114
DEB.CM.99.CM40             -G---G-CCGCT--ATACAACT-GAC-T-A-CA---GT---G--AGGA--T--G..--CTACCG-GCC--A-AC--G......-CTA----TCTGT-C---TA-...........--T-CCTTT-T-CAA-GGCTG-GG--A-AGAG----------GAC-A...-A 6980
DEB.CM.99.CM5              -G-A-C-CA-CA--GTATAA-C--TCTT-GG-----GTT--G--AGGG-----G..--GTAT-G-GCTG-T-A---G......-CCA-----CT-T-C---TA-...........-----CCTTC-T-CCA-GGCTG-GG--A-AGAG-----------GG--G...-- 6919
DEN.CD.x.CD1               T--T-CAGAAA--GTT-TGTAA-CA--T-G--G------C--G--AGGG--TGTA...-CCT---GG--T-AAAC--G......-C-G----A-T--TCA-GAGA...........G-C-CTATT-T-CCA-G-A-GGC---T--GAG----GCA--C--G--AGTAG- 7285
LST.CD.88.447              -A----GCAGGA--ATGGGGAC-CATT----GA----TT--GAAAGGA--T-GG..T--GA-GT-TCT-C-AT--GT......TC-A----ACTTCTG---TAC...........TATGGCTTGG--CATGGTA-C-GGC-----TTG--T--G-----AT--...GA 6351
LST.CD.88.485              -A----TCAGGA--ATGGGGAC-CATT--C-GA----TT--GAAAGGA----GG..T--GA-GTCTCT-C-ATC-GT......TC-A----ACTTCTG---TAC...........TATGGCTTGG--CATGGTA-C-GGC-----TTG--T--G-----AT--...GA 6348
LST.CD.88.524              -AG---GCAGGA--ATGGGGAC--ATC----GA----T----AAAGGA--T-GG..T--GA-GTCTCC-CAATC-GT......TC-A----ACTTCTG---TAC...........TATGGGTT-G--CATGGGTCT-GAC-A---TTG-----A--C--AT--...GA 6354
LST.KE.x.lho7              -A----GCAGGA--ATGGGG-T--ATT----GA--C-T----AAAGGA--T-GA..T--GA-GT-TCC-CAAT--G-......TCTA----ATATTT----TA-...........TATGGCCT-G--CATGGCA-C-GAT-A---TT------G-----AT--...GA 7420
SUN.GA.98.L14              -A----GCAGCT--GTGGGGGT---T----C-G--C-T---GAAAGGG--T-GG..T--CA-GT---T-CTATA-GT......A-TA----ACTGCTG---TA-...........TATGGGCT-G-GCCAGGTTCT--GT-A----TG-----G-----AT--...GT 7479
MND-1.GA.x.MNDGB1          -AG--G-CAGGA--ATATGGCT--A-G--AG-A---CAC---AAAGGA--T-GG..T--GT-GTG--T-CT-CATCA......-CTA----ATT-TT----TA-...........CATGGGTT-G--CCTGG-AAG---T-A-AGA---GCATG--C-C-T-C...A 6754
OLC.CI.97.97CI12           GC-A--GCTGAT--GTATAAGT-------AG-G----TT--GAAGGG---T-GA..-GT-T-GTTTC-G-TAA-TCA......-CTT----TCTG-T--A-TA-...........G--GG-TT-G--CCTTAT--G-----AGATA---GCTTG----GGT--...A 6816
WRC.CI.97.97CI14           -AG--G-CAGAC--GTATAA---TA-G--AG-A---CAT--GAAAGGT--T-GA..T-TTT-GT-TC-G-AAA-TCA......-CTT----TTTGAT--A-TA-...........G--GG-TT-G--CCTCATA-G-----T---A---GCTTA--C--AT--...C 7323
WRC.CI.98.98CI04           -A---T-CTAGT-GGTATAAG-----G--AG-A--C-AT--GAAGGGG--T-GA..T-TGT-GTGTC-G-AAACTCA......-CCT----ACTTAT--A-TA-...........G--GGCTT-G--CCACATA----C--A--GA---GCTTA--C--GT--...A 7371
WRC.GM.05.Pbt_05GM_X02     -A-----CAGCT--ATATAA---G-CT--AG-A---CAT---AAAGGA--T-GA..-GTCT-GTCTC-G-AA--TCA......-CCT-----TTGATT-A-TA-...........G--GG-TTGG--CCTTATAAG-----T--GA-----TTA--C--AT--...C 6267
MND-2.CM.98.CM16           -G----G-TGA---GTACAACT--C--G-AGTA----G----AAAGGA----GA..T-TGTG-TCTC--CA-C--GT......-C-A-G--ATTGATC---TA-...........-GTGGGTT-G--CCAGGGAA----C-T-AGA---GTATG---C-GT-G...A 7330
MND-2.GA.x.M14             -GG--GA-CAAA--ATGGGGGT-GA-GG-GGTA---GAT--GAAGGGA----GG..T-T-T-GTTTC-G----CA-GT......-C-T----ATT-ATT--CTA-...........CATGGGTT-G--CCAGGCA-G---C---A-A---G-ATG--C-GGT-A...T 7247
MND-2.x.x.5440             -G---GG-TAAA--ATGGGGAT--C-GG-A-GG---G---GAAGGGA--T-GA..T-T-TC-TTTCC-CA-CA-GT......TCTA----ATTGCTG--CTA-...........CATGGGTTGG-GCCAGG-AAG---C-T---A-G-GCA-G---C-GT-A...GA 6957
DRL.x.x.FAO                -GG--TA-TAA---ATGGAACC-CA------TC----G----AAAGGA--T-GG..-GC-T--T-TCT-CA-CCTCA......-CCT-----TTGTTG---TA-...........CATGGGTT-G--CCAGG-AA----T-A-A-AG--G-ATG-----AT--...GA 6991
MON.CM.99.L1               -G-T-GGCTAA-C--CT-CA---CA-C----C---C-TT--G--AGGA-----A...--T-TT-GG-ATT-A-AG--A......---G----C-TGA----CTA-...........T-GCA-CTT-T-GTGGG---T----CTC-TA----CT-C--C--AG---A--A 7170
MON.NG.x.NG1               -GGC-TGM-AA--C-CTC--A--GA-C--A-C---C-TT-----AGGA--T--G...-CTT--G--ATT-A-AG--T......--TG----A-TGA----CTA-...........T-GCA-ATC-T-GTGGG------C-CY---A-G---T-C----C-GG-A-A--A 5786
GSN.CM.99.CN166            GGCT-TGGAGAAG-TTATCA-C-TA-TC---CA----TC-----TGGA--T--A...-C-T--G--ATT-G-AA--A......---G-G--A-TGA----TA-...........T-CCA--TG-T-GTAGG--------CT--GA----CT----CCGAC--GA-CC 7162
GSN.CM.99.CN71             GGCT-TGG-GAAG-TTATCAGC-TACTC---CA----TC-----TGGA--T--A...-CCT-----ATT-A-AG--A......---G----A-TGA----CTAC...........T-CCA-ATT-T-GTAGGG---G---CT--GA-G--CT----C--AC-C-ACCA 7153
TAL.CM.00.266              C-CC--GGAAA-G-ACACAACC-CAC---C-CA--C-T------AGGA--T-GG.....T---GGACT-A-AG--A......-C-G----ACT-ATG---CAC...........-GCCAGAT--T-GCAGGGAAG--CC-C-ACAGG--CT--------GC-A...GA 7010
TAL.CM.01.8023             C-CT--GG-AA-G-ACACAACA-CACC--AC-G--C-TT-----AGGA--T-GA.....TC----GATC-T-AG--A......-C-G----TCT-ATG---CAC...........-GTCAGATC-T-GCAGGTAA-G--T-A-AGAG---CT--------G--A...GA 6554
MUS-1.CM.01.1085           -A-T-TGG-GAA--TTTTAGA---A-TC---C-----T------AGG------G.....T--GG-ACC-T-AA--A......---G----C-TGA-T--CTA-...........T-GCAGCTT-T-GTAGAT--CG---CAC-CAG---CT-----C--A--A-A--A 7126
MUS-1.CM.01.CM1239         -AGT-TGG-GAA---TTTAGA--TA-CT-A-CC---CTT---G--AGG---T--G...--T-T--GG-ACC---AG--A......--GG----C-TGA-T--CTA-...........T--CAGCTT-T-GTAGG---GG-C-CAC--AG----T-C--C--G--A-AC-A 7098
MUS-2.CM.01.CM1246         -A-T-TGGAGAA---TTTAGG---A-TC-A-CC----TT--G--AGGA--T--G.....TT-G--ATC-A-AG--A......---G----C-TGA-C--CTA-...........T-TCAGCTG-T-GTGGAT---G-C-CCC-CAGG---T----C--G---C-A--A 7233
MUS-2.CM.01.CM2500         CA-T-TGGAGAA--TTTTAGG--CA-TT-A-CA--CCTG--G--AGGA--T--G...-T-TT-G--ACC-T-AA--A......---G----C-TGT-T--CTA-...........T-TCAGCT--T-GTAGAT---G---CCC-CAG---CT----C--A--C-AC-A 7162
SYK.KE.x.KE51              C----G-CAGAA--ATT-GGAT--C--C-GGTG---GA------AGGA--TG-G...T-T-T-----AC----AAT-G......-C-G----ATTTTTC--GCCAGTGATT...CAAGGG-G-TTG---ACAGGGAA-GC-GC--A-AG----TT---C--AG-C...A 6912
SYK.KE.x.SYK173            T----T-CAGAAGCACTTAGAT-GC-G----TA---GA------AGGA--TG-G...T-T-TC----AT--T-AAT-G......-CTG----CTATTTCC-CCCAGTAATA...CAAGGG-A-TTG---ACCGGGA-GG-CGCA-A--GT--CTTC---CGAG-A...C 7216
RCM.GA.x.GAB1              GGTC-CA-TAGTTTTTATAA-T-GTC-G-A-CA--C-GG-----TTCA--T-GA..---GTG---G-G--ATCGC-A......-C-A----AGTCTTT.................-TCTC-CT--G-GTAGAGAAG-GACCA-A-GG---TTGG----GAT--...GA 6930
RCM.NG.x.NG411             GGGT--A-TAGAC-TTTTAA-T--AC-G-A-C-----GG-----A-GT----GA..---GTG---G-----TCAT-A......-C-A----AGT-TTC................-TCTCTCT--GGGTGCA-AA--GGCCA-A-GG---TTGG----GAT--...GA 6974
GRV.ET.x.GRI_677           -AGT-GA-TAA-TTTTACAACT-GAC-G-G-GA--CCG------TGGT--T--A..---GTGTTGCC-G-AACA---.....ATGG-----TT--T---................CACTCTCAG---TAT-ATACC-GGT-A-A------GTGG--CC--T-C...CA 7208
SAB.SN.x.SAB1C             -GGC-CA-TAGAT-TTTTAA-T--ACT--C-GA----G------AGGA--T--G..---GT-TTGCCTG-AACC---.....ATGG-T--CTTG-TC--C..............CAT-GCCAG---TAT-ACACC-GGT---A------TTGG--CTGGT--...G- 7508
TAN.UG.x.TAN1              -A-C--A-TAAAC-TTATAAGT-GA-G--AGTG----G------AGGA--T--G..---GTCTT-CC-G-AACC---.....ATGG----TCTC-TG--C..............CACTC-CA-C--TAT-ATAC---AT-A---------TGG--CC-TT-C...CA 7297
VER.DE.x.AGM3              T--T-TA-TAAAC-TTACAA-C--AC-G---C---C-A------AGGA-----A..---GTCTT-CC-G-AACA---.....ATGG-----CT--TG---..............CATTCTCAG-GGTAC-ATAC--GGC---------TTGG---C--T-C...CA 6835
VER.KE.x.9063              C-TT-TA-CAA-T-TTACAA-T-GAC-G-C-C---C-A---G--AGGA--T--G..---GTCTT-CC-G-AACC--A.....ATGG----TTT--TT---..............CATTCGCAG---TAC-ATACG--GC-T-----G--CTGG---C-TT-C...CA 7337
VER.KE.x.AGM155            T-GC-CA-TAA-C-TTATAACC-GAC-G---CA--C-A---G--AGGG--T--G..---GTCTTGCC-G-AACG--A.....ATGG----ATT--TC--C..............CACTC-CAG--GTAT-ATAC--GAC-A--G--G--CTGG--CC--T-C...CA 7317
VER.KE.x.TYO1_patent       T-GT-CA-CAA-C--TATAA-C--TC-G-C-CC----GG-----AGGA-----G..---GTCCT-CCGG-AACG--A.....ATGG-G--ACTG-TT--C..............CACTCTCA----TAC-ACATG--GC-T-----G--TTGG---C--T-C...GA 7316
COL.CM.x.CGU1              TATTATTG-T-TCCA-A-GA-A---CTT-AGGG----T----AGAGGA--T-G-...T-TCAC-G--ATC-AAAC-CT......-CCAAT--AGC-AA---CTAC...........TATGAGCTT-TTCCTTATTC---AGGC-TTT-T-G-AGG---C-GT--GTACC 6638
```

```
H1B.FR.83.HXB2                AGCAAAATGGAATAACACTTTAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATC.................... 7304
Env                             A  K  W  N  N  T  L  K  Q  I  A  S  K  L  R  E  Q  F  G  N  N  K  T  I  I
H1A1.UG.85.U455               -AGGG-C-------GA--A--C-----G-----GAAC------AGA--A-G---AAT--C--A-C-.........................ATAA 6747
H1B.US.90.WEAU160             -A---GT-------------------------T-GAA-----------AT-AAACA-TT---G--T-A--C---A.................G 7312
H1C.ET.86.ETH2220             --A-----------A---C--C--A--G--AAGGAA------CA-A-G--C--CCCT---...------......................ATAG 6699
H1D.CD.84.84ZR085             --T---------------C---CG----G-------A------G--A-C-TT--GAACC-A-CA---..........................ATAA 6838
H1F1.BE.93.VI850              -A--C--------------C-GG-GT-TG--AAGGCAG------AGTCG--T--CCCT------...---GC-...................ATAA 6621
H1G.SE.93.SE6165              -AG---------AG-GG-C---C-GA-TG-C---GCAG--C--G-GA--ATC---AAT--G-GC-GCGAG-AC...................ATAA 6732
H1H.CF.90.056                 -A--G-C--------G------C-C---G---T--CAC-----G--AT---C--GAAC----GA-C-.........................ATAA 6627
H1J.SE.93.SE7887              -AG-G-T----G----------CGTAGAG------CA---C----------C---AAT--A-CA............................ATAA 6606
H1K.CM.96.MP535               --A---------C-TG-------GCAGAG--AAGGAA--GC---A------T---AAG---GGA-C--T-......................ACAT 6507
H101_AE.TH.90.CM240           -A-------------AGT----------G--A--GAA-------A---G--C---AAT--G-CA............................ATAA 6873
H102_AG.NG.x.IBNG             -A--G----------A------C-T---G---T--CAC------AGACGT-C---AAG--C-CC-C-.........................ATAA 6820
H103_AB.RU.97.KAL153_2        -A----------------A-------------T--T----------A----------G-----A-C-.........................ATAG 6494
H104_cpx.CY.94.CY032          -AATG-T------G----C------GT----AG-GAAG----G-A-AG--TC--CCCT-----A-C--T--A-....................... 6686
H1N.CM.95.YBF30               --A-CT----G-ACCA-TG-GG--TAGA-C-AGAGAGG--A---AGA--ATCC-G--G--A--C--C.................ATAACCTTCA 6883
H1O.BE.87.ANT70               CA-TG-T---GGA--A-TA--------A-C----GAA-GG-ATTT-----T-G-AAAC----CAGGT...............AGTATTAACATGA 7344
H1O.CM.91.MVP5180             GA-TGT----G-A--TG-CC--C----A-C-----TA-GG-ATTT-A-T-TTG-AAACC-A-CAG-G-AT.............GTTACCATAA 7352
H1P.FR.06.RBF168              GCA-R-K----CA--A--C---G--TCTG-CAG--AGGCCA-CT-GA--GCT-A-CCTC-ACCCCCC-ACCA--A-CAT.........ACATTTG 6790
CPZ.CM.01.SIVcpzCAM13         GA---C-------G-A--AC-GC--A-AG-TAGAGAA-GC----T-A-GG--A-AAA-GCA---GC--ATGG--C-TAC.........AATATAA 6922
CPZ.CM.05.SIVcpzEK505         GA--...---G-AC-A-TA-GG--TACA-C-AAACAA-T-A---T-A-CA-C......................-GAAAG.........AATATTA 6824
CPZ.CM.05.SIVcpzLB7           GA--C-T---T----A-TAC----C---G-GAAA--AG-GC---CCACGGTC--CAATG-ATC---C-A-..................ACTGTAC 6834
CPZ.CM.05.SIVcpzMB66          -A--GT-----C---T--A--T---A-AG--AGGCA-GCC----AT---AC--A-AAGGGC...T-TCTCGG--ATGAA.........ACTATCA 6950
CPZ.CM.05.SIVcpzMT145         GA--C-G---GCA--AG-CC-C--TG-A-C-AAAGAAGTTC----GA-CAT-C-GA-G--G...........................AACATAA 6814
CPZ.CM.-.SIVcpzMB897          GA--GCC------G-G---C-G-G----G--AAGGAGGC-----ACA-CAC--A-C-----CTC--TCTC-GTC--.........ACAGAAATAA 6861
CPZ.CM.-.SIVcpzDP943          C---G-----G-AC-A-GG--M--WG-C-CTRTG--GGCTA---A-AGT-TTAAGCC-GGC-G---C-TC-CCT--.................. 6870
CPZ.CM.98.CAM3                GA--GTC---G-CG-A-GAC----TG-A-CT-G-CAAGCC--GC----G-TT---AC---CCT--C-CA-G-C.......................A 6730
CPZ.CM.98.CAM5                G--CC-G----TG--A-AG-----CG---CC---TTAGCCA--GC-A-GT-TGA-AA--GC-TAGC--A-G-C.......................A 7006
CPZ.GA.88.GAB1                GA-C-C-------CGA--GG-GG--G-AG--AAA-AGGCT---GC-ACCTC--CAAATCGA-CAGC-G-C-AC--A....................A 7363
CPZ.GA.88.GAB2                TA--CT----------T-C----CAGA---AAACAA--GA-T-A-A-CAGCACAACGTGGC-C-GGGGGGAC---.....................A 6716
CPZ.US.85.CPZUS               -CTG-C----G-G--ACAGAG----T-C-CTTTAGAG-T-A---A-A-GG--GCAAACCTG-CA--GGT-..........................G 7331
CPZ.CD.90.ANT                 GA-GCT----G-AC-AG-ACGT--CA-A-C--AGCA-GTTC-TGCG--G--T-GGAA---AGTAG-C-AC-A--C-AAT.........GCGAAAA 6764
CPZ.TZ.00.TAN1                -A---C----T-C-GG-AC---G-TTG-GCTATGGCTGCCA---ACAC-ACCA-GA-GGCC-GA--TGA--CGG-A.............CAACAAA 6949
CPZ.TZ.2001.TAN2              -A--CT-----G--C-A-G-C---GCAG-GCTATGGAAGCTA---ACA--ACCC-A-CCCTGT----T-A-GCTG-A.................AATG 6978
CPZ.TZ.2002.TAN3_1            -A--CT----T----A-G----G-CA-TGCTATGGAGGCTA---CCA--AC-C-AAAG--A-GCGG-CT--ATG-A.............ACAAGAA 7014
GOR.CM.2007.SIVgor2139_287    GA----C---C-GGGTG-G--G---G-C-C--TGCAGGCTC-C-AC--GAGCC-GATCC-CTTCCGG............................ 6865
GOR.CM.2007.SIVgorCP2135con   GA----C---C-G--TG-A--G---G-T-C--TAC-GGCTC-C-A--CGA-TCAGATCCGCTTCC-G............................ 6846
GOR.CM.2004.SIVgorCP684con    ----TGGCA---GGCGTTACA-GCTACAG-GCGGGCTCTCAA--CGA-C-TGACCCACT-CTTC--C............................ 6787
H2A.DE.x.BEN                  --GC-G-----GGG-AG-CA-GC-GG--G-GAAGCAA-CCC-TGT-C----TCCCA--T-C--AGG--TC-ATGATACA.......GGGAAAATTAACT 7817
H2A.GW.x.ALI                  --GCG------GG--AG-CA-GC-GG--G-GAAGGAA-CCC-TGT-A----TCCCA-GT----AGG---C-ATGA-ACA.......AACCAAATTAACT 7798
H2A.SN.x.ST                   --GCG-G-----GG-AG-CA-G--GG--G-GAAGCTA-CCC-TGC-A----TCCCA-GT----AGG---C-ACGA-ACA.......GAAAAAATTCGTT 7244
H2B.CI.x.EHO                  --G---C----CAG-AG-CA--C-GG--G-GAAAGAG-CCA-T-A-A-T--TCCCA--T--TCAGG-----C--ATATC.......TCACAGATAAGGT 7776
H2B.GH.86.D205                G-G---C----CAG-AG-CA----GG--G-GAAA--G-CCA-C-T-A----TCCCA-GT----AGG-GGTGC--AAAATATCACA...AGCGTAAAGTTAG 7790
H2G.CI.92.ABT96               --GC--------GG-AG-CA-GC-GG--G--AAGGAA-CCG-GGTCA-G--CCCCA-CT----AGG-----ATGA-ACA.......AAGAAMATAACCT 7148
H2U.FR.96.12034               T-G---T-----AGGGG-AA-CC-GG--G-GAAGGAG-CCA-TGCCA-C--TCCCA-GT--C--GGT--C-AG-ATATT.......AGTCAGATAAATC 7277
SMM.SL.92.SL92B               --G--------GAG-AG--A-GC-GG--G-TAAG-AA-CCA--GTCA----CCCCA-GT---C-GG---C-ATGATACT.......AGGAAAATTAATC 7171
SMM.US.x.H9                   --G-G-------AG-GG-CA-CCGGG-AG-GAAGGMAMCC--GGTCA----TCCCA-GT----GGG---T-ATRAKAYT.......RAACAAATTAAGC 7212
SMM.US.x.PGM53                --G--GT----GAG-GG-CA-CC-GG--G-GAAGGAA-CCA-TGTCA----TCCCA--T---CGGG---T-ATGAGACT.......AAGAAAATTAATC 7669
SMM.US.x.SIVsmH635F_L3        --G---C----GCG-GG-CA-CC-GG-AG-GAAGGAA-CC--GGTCA----TCCCA-GT---CGGG---T-ATGATACT.......AGGAAAATTAATC 7742

MAC.US.x.239                  --G---------GG-TG-AA-----G--G-GAAGCAG-CCA-TGTCA----TCCCA-GT---C-GG---T-AC-ATACT.......GATAAAATCAATT 7735
Env                             G  K  W  K  D  A  I  K  E  V  K  Q  T  I  V  K  H  P  R  Y  T  G  T  N  N  T        D  K  I  N
MAC.US.x.251_1A11             --G---------GG-TG-AA-----G--G-GAAGCAG-CCA-TGTCA----TCCCA-GT---C-GG---T-AC-ATACT.......GATAAAATCAATT 7731
MAC.US.x.251_BK28             --G---------GG-TG-AA-----G--G-GAAACAG-CCA-TGTCA----TCCCA-GT---C-GG---T-AC-ATACT.......GATAAAATCAATT 7717
MAC.US.x.EMBL_3               --G---------GG-TG-AA-----G--G-GAAACAG-CCA-TGTCA----TCCCA-GT---C-GG---T-AC-ATACT.......GATAAAATCAATT 6866
STM.US.x.STM_                 --G-G------GAGGAG-AA-----G--G-TAAGGAA-CC--GGTCA----TCCCA-GT---CAGG---C-ATGA-ACG.......GCAAAAATAAGGA 7374
MNE.US.x.MNE027               --G---T-----GG-GG-AA-----G--G-GAAGCAG-CCA-TGTCA----TCCCA-GT---C-GG---T-AC-ATACT.......GATAAAATCAATT 7206
DEB.CM.99.CM40                C-G--GC---GCA---G-AAC---GG-A--GCGACAA--GA--TTG---AT---------GGCA--T-GG-C--ATAAT...................C 7060
DEB.CM.99.CM5                 --GT--T----CAGGAG-AG-T--GG-----AGG----G-----AG--GAT--A-..........--TGTG-C--AA......................A 6987
DEN.CD.x.CD1                  --GC--C----C--GA--AAGGC-GG-A----AA-AAGC-G-T-A---GG-G---AAC-TC---GGC---G-G-AATGG..................... 7364
LST.CD.88.447                 --G-C-G---GGA-GA-TG--TC-TA-TT---GA-A--TGC-T-A-AGG-TTAA--C--CAGCG-TG-ATTAT-C-CATGGAAACTGTATAGATAACACA..........AAGCCTTGTGGTAGACAACTAAAGGGATTACCTATAG 6488
LST.CD.88.485                 --G-C-----GGA-GG-TG--T--TA-TT---GG-AA-TGC-TCA-A-G-TTAA--C--CAGCA-TG-ATTAT-CAAATGGACAATGTCCTGAGGGAAAA..........CAGCCTTGTGGTAGACATCTAAAGGGATTACCTATAG 6485
LST.CD.88.524                 G-GTC-----GGA-GA-TG--T--CA-TT---GA--G-TG----AGA-G-TTAA--AG-C-GC--TG-ATTAT-AAGAAGGAAATTGTTCCACACCAGGA..........AAACCATGTGGTAGACAATTAAAGGGTCTACCCATAG 6491
LST.KE.x.lho7                 --GCC-----GGA-GA-TG--T--TA-TC---GC-AA-TGC---AG--GTTGAA--C-G-AGCC-TG-ATTAT-CAGAAGGGACAGGGACATGTGATTCTAAGAAA..........ACAACTTGTGGACGGAAATTAAAAGGCTTACCAATAG 7563
SUN.GA.98.L14                 --G-C-----GGA-GAG-C--TGCCACCT---GG-AACT---G--GC-GGTGGAACCTTTAGC---C-TG-GT-ATGGC......TGTACATTTGATAATATAACAAAGACA............TGCAGATTTACCAATGGTACAGATTTTAAAAAGATGATTAAGTTTA 7631
MND-1.GA.x.MNDGB1             --G-CGT---GGGTTAG-AC-TTGGAGTC------AAG--C---ATA--TT-AA--ACTCC-TC---GTG-ACCAGACCTGTAAAAATTTTACTAGCACTGGAGAGGAGAACAAACAAAACACGGACAAGCAAAAGGAGTTTGCCAAATGCATAAAGACTCTTAAGATAG 6924
OLC.CI.97.97CI12              G-G-C-----GGATGG-TGC--T-TGCAC---GA-AAGG-C--GAGA--G--AA-AAC-C-TGGGC--ACTAT-CAGGGGTTAACATAACAGGCATAGGGAGG...............AACTCTACAACCTTTGCCAAAGCAATCAAAAGA.........TTCAAGTTTG 6962
WRC.CI.97.97CI14              T-GGCTT---GGA--A-TGC-TT-TG-TT---GG-AG--T---GAGA--T-TAA-A-TTC-TGGGT--ACTAT-CAGAGGGGGACTGCAATAGCAAGAAT................AATTCAGCAGACTTTGCTAAATGTATTAAACGG.........TTCCAGATAA 7466
WRC.CI.98.98CI04              --GGTTG---GGAC---TG---T-TG-TT---GA-AG--TC-GCAG---T-TAA-AAT-CATGGGT--ATTAC-CTGAA...GATTGCAATCAGAAAAAC................GATTCTAGAGAGTTTGCTAAATGTATTAAACAG.........TTCCAAATAA 7511
WRC.GM.05.Pbt_05GM_X02        --G-C-G---GGAC-T-TG---T-TGG-C---G--AAG-----CA--CTT-TAA-A-CTC-TGGGT--ACTAT-CAGAG...GAGTGTAATAAGAAAGGG................AATAGCAGTGCATTTGCAACTTGTATTAAACAG.........TTTCAAATAA 6407
MND-2.CM.98.CM16              G-G-C-----GGA-TAG-CA-GC-TG-TC----A-TAG--------A--AT-GACA-CTCC-TATGG-G--ATG-AACA...AAGGGCTATTGCAAGGGGTTAAAAAGAAAG..................GAAAATAGAACGGGTTGTGCCCTTAAAACAATAAAGG 7476
MND-2.GA.x.M14                --G-C-----GGA-GAG----G--CGCAT--T-AC-TG------AGA--GT-AA--C--G--T-T-T-G-GGTT-TAATGCAAGTGGCCCTTGCAGGGGGATAAACAATAGA..............GGAGTGGTAAACAAAACAGGGTGTGCCTTGAAAACAATAGAAG 7402
MND-2.x.x.5440                --GG------GGAC-GG-----C-TAGTC--T-ATTAG-GC-----A-GAT-AA--AT-GC-T-T-C-A-GAC-A-CAC...AATATGACATGTAAAAGCAGTAACAACAAA..................AAGAACACAACAGGGTGTCATCTTAAAACTATAAGTA 7103
DRL.x.x.FAO                   --G-C-----GGA--GG-----GG-AGTC----C-AAG------A-TC-GT-AG----TCC-T-TGG-G--AT-CTACTAATAGCACATGCATAAGTCTCAAAGGCAAAAAGATC............CAGGACAGAAACGTGACAGGGTGCATGTTAAAGCAGCTTAAGG 7149
MON.CM.99.L1                  GA-CC-----G-G-CTG-CC--C-GGCAG-TCACGAGGC-G---A-ACTG-G-GG-A---G--A--C-ATGG--C-AAT..............GTGACCACAATAT 7262
MON.NG.x.NG1                  GA--C-G---G-CGCTG-C---C-TGCAGCCCA--TRGCTA-C-RGR-GG---GG-A--GC--A--T-AT-ATG--ACA..............GTGGCCAACATAA 5878
GSN.CM.99.CN166               T-AT-GG-------GGG-AA---G-G-AGCTATG-AAGC-A-G-AT-----T-GG-A------A-CTGGG.....................AGAAATGATACTCAAATAA 7251
GSN.CM.99.CN71                TA----G---G---TAG-AA---GG--AGCCATG-AGGCCA-G-AGA-C--T-GG-A---G-TC--C-ATGAT-C-ACC...............CCTAATAAGACTCAAATAA 7251
TAL.CM.00.266                 G-GC--T----C---AG-CA-GG--TC-GCCAA-GAAG--A--GC-A--AGG-ACAAGG-GC---TC-AC-AC-CAGAT.................AAAGTGAACATCA 7102
TAL.CM.01.8023                --G---T----CG-GGG-GA-GG--G-TG-GAA-GAG--GA-TGC-A--GCC-ACAAGG--C-C-T--AC-GC-GTAGT...............GGAACAGCACAAATAA 6649
MUS-1.CM.01.1085              GA--C-G---G--CCAG-A--G-G---AGCTATGCTAGCCA-G---A----T-GGCATGG--GATTGCA--A-CAGAACCACAGCTCGATAAACGAGACT.........AAAATAAAGATAA 7239
MUS-1.CM.01.CM1239            -A--C-G---G--GTAG-A----GG--AGCTATGTTAGC-A-G---A-G--T-GGCATGGCCGCTTGGAG-A-CAAGGAGAA...AATATTACAGAAGAA.........AAGATAAAAATCA 7208
MUS-2.CM.01.CM1246            -A-TC-----G--CCTG----G-G----GCTATG-TAGCTA-G-----C--T-GGCAC----GACTCCA--A-GCTAATAAG...ACAATAAACAAAACG.........GAAATAAAAATAA 7343
MUS-2.CM.01.CM2500            GA-CC-G---G-CCC-G-A----G---AGCTATG-AAGC-A-G----C---T-GG-CC...--A-TGGAC..............AACCGAAGTGAAAGT.........GATATAAAGATCA 7257
SYK.KE.x.KE51                 --G-G-----GGA--GTT---TG-G---G-CCA--ATG-GAGT-T-A-GGTG-GGAA----..........................GTGACCTCAACAT 6986
SYK.KE.x.SYK173               --G---T----CAG-ATT---T--G---G--CA-GAAC-GGCC-CGA--ACC-GGAA----..........................GTAACCAACACCA 7290
RCM.GA.x.GAB1                 --GG--T----CGG-TG-A-GG---G-AG--AAAGAG-G-G-G-A-AC-AC-AAA--GT--CGAGGT--T-GT-A-ACA.........GACAAAATAAAGA 7022
RCM.NG.x.NG411                T-G-G-T-----AGGTG-G-GG---G--G--AAAGAGGC-A-T-ACA-GAC-AAA---T----AGGT---CG-CAGCCT.........GAGAATATAACTA 7066
GRV.ET.x.GRI_677              --G-G-T-----AGGGG-A-GG---G-AG-CAGAGAAG--G---AGA--GTGAAAAATCT--CAG--GT--GC--AGAA...........AATATACATC 7297
SAB.SN.x.SAB1C                --GT--T----GAGGAG-C-GG--GG-AG-CAAAGAA-C-A--GTGAG--T-CCACC---A--GT-T-GTGGC-CAAT.........GACACAAACAAGA 7600
TAN.UG.x.TAN1                 G-GTG-T-----AGGAG-C-GG-GGG-AG-GAGA-AA-C-A--GTG--G-TTCCAAA-G-G--AT-TCG-GGG-CAAT.........AACACAAGGCAGA 7389
VER.DE.x.AGM3                 G-GC--C----GAGGAG-C-GG---G-AG--AAA-ATG--A--GT-A--TT-CCAAA-G---GAT-CCA-GG--C-AAT.........GATACTGAAGAGA 6927
VER.KE.x.9063                 --GC-GG----GAG-AG-C-GG--GG--G-TAAAGAAG--A---TCA--TT-CCAAA-G---GGT-TCGGGG--C-AAT.........GATACAGAGAAAA 7429
VER.KE.x.AGM155               G-GC--T-----AGGAG---GG--GG-AG--CAAGAGG--A--GT-A--TT-CCAAA-G-ACGGT-CCA-GGC-C-AAT.........GATACAAACAAAA 7409
VER.KE.x.TYO1_patent          --GC--T----GAGGTG-C-GGCGGG-AG--AAACAA---A--GT---GTT-CCAAA-G-C-GGT-T-A-GG--C-AAT.........AATACAGAACACA 7408
COL.CM.x.CGU1                 -ATG-CTG--C-G--T-AGAA---CA-A-C-CAATT-GC-A--GA-ATCA--AAGAATTTA-CAGCTTGGC--GAGCGCATTTCACGCAAAAATATA.........ACTATCACTC 6745
```

PLV Complete Genomes

```
                                                                                              V4 loop start
H1B.FR.83.HXB2              ..TTTAAGCAATCCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACT......................... 7448
Env                         . F  K  Q  S  S  G  G  D  P  E  I  V  T  H  S  F  N  C  G  G  E  F  F  Y  C  N  S  T  Q  L  F  N  S  T  W  F  N  S  T  W  S  T  E  G  S  N  N  T  .  .  .  .  .  .  .  .  .  .
H1A1.UG.85.U455             TC---GCTAGC--------------TAT-------AC---A--T--C--------------A-----------T--C--CA--T--GGC-----------C-T----AA-GGC--C-TG-CA-A-GACATG---C-----GG--.......................... 6893
H1B.US.90.WEAU160           TT-----A-------------------------------T----------------------------------------------------G-----------------CA-GC--A-GG-ACTT-G-AGA-TACTGA-GGGGCAGA-......................... 7458
H1C.ET.86.ETH2220           AA-------C-----------------T-------AC---A--T--C--------------A-----------T--C---A--T--A-C-----------C--AAAAC-AG-ACTGTT-AAT--C-G-AC-AAT-T---C............................ 6839
H1D.CD.84.84ZR085           TT-----A-C--------------T---------AC---A--------------------------------C---A--T--GG-------------G-A---AA--TATCAGGACAT-----A-GGTTAAATG---CA-TC......................... 6984
H1F1.BE.93.VI850            AA-----C-------------------T-------AC--T---T------------A----A------------CG--A--T--GG----------GAC--AG-C-CC----A-GGCACT............................................. 6743
H1G.SE.93.SE6165            CC-----CTC---TG------------T-------AC---A--T-----C-T----A----A--G---------T------A--T--GGC---------TCA-GCCT---G-GATCA-A-A-T---GAGA-T---A-T............................ 6869
H1H.CF.90.056               GC-------C-AA-------------ATG---G--AG---A--T-------------A----A-----------T--C---A--T--GGG-------------G----GAA-TGCA----AATTAC--ATC-AATGAC-CA--GGGA....................... 6773
H1J.SE.93.SE7887            AC----CATC-C--------------TAT----------G--A--T-----------------A--------A--T-----A--T--A-------C-------G----GA---G-A--GCATTGAAG--AC-AATGAC-CA-GTNAC....................... 6752
H1K.CM.96.MP535             TTAAACCA-C-AA-C-------A------------C---------TG---------C---A--------------C---A-----A-------------GAG--AG--GAG---G-----............................................... 6626
H101_AE.TH.90.CM240         TC---C-A-C-C-----------A--T-T-------AC--T---TCA------------A----------------T--C---A-----A-------------A------CC-AGGA-A-GAAACC-TGG-G-GGT-TAATG-C-CT........................ 7016
H102_AG.NG.x.IBNG           TC---GCTA-CC---T---------TGT-------AC---A--T-----------------A-----------T--C---A--T--A-------------C------GA------C-A-A-C-C-G-CA-CCACA--GGGTCA-A-......................... 6966
H103_AB.RU.97.KAL153_2      TC-----T-------------------------------T---------------------------------------A-----A-------------C------AA-GG--C-GAAGA-TTA-A-A-CACTGA-GGAG-T......................... 6637
H104_cpx.CY.94.CY032        ..---GCT-C-C--GT-----------T-------AC---A--T--C-------------AA---A-----------T--C---A-------C---------------ACACA-GC-A-A-GG-ACA-AC-T-AC-A-TA--G--TCT--A......................... 6830
H1N.CM.95.YBF30             GGGC-CGAG-GAGGAAT-A---A----T----G-GAC-CACTTA-TG--C-------A----A--G---------T-----CA-TT-CA--T-A------GAGGAA-TAC----CGAG--AG-TGAGC--......................................... 7011
H1O.BE.87.ANT70             CA--C--T--CAG-AGC--T--A--T-T---GG-AACCCATTTACAC-----C---CAT--A-----C--T--T-----CA--G-TA-GA--------TA---C-TT...TCAT---ACG-A-CC--CTGTA-TGTT-G---TGT-......................... 7487
H1O.CM.91.MVP5180           TA--C-GCAG-A-TAGT--T--A--TG-----G-AAGCCATTTACA------C---CAT--A-----C--T--T-----CA--T-TGGGA-------CTA-----TTA-C--CT----AAA-TCCGGATGCCA-GAG-TC--AGGG......................... 7498
H1P.FR.06.RBF168            TC---CG-A-CAGTA----G--A--------GG--TCTTTT-TGCAC--C-G----CAT--A-----C--T--T--C--CA-C-GCAGC----------TA--GC-ATAC-TG--A-GAGAA-...GG--T-T-C....G--T--A-AAT..............CATACA 6939
CPZ.CM.01.SIVcpzCAM13       CG---G---C-AG-AGT-----T-----------A-C--AT----TG--C--C--------A-----C--T-----CG--A-T-G-A--A---------.........................GAA-G----CCT-TTC-CG-G........................ 7041
CPZ.CM.05.SIVcpzEK505       CA----TA-CCAATC----C--A----T----G-AAC--ACTTA-TGA-C----------A-----C--T---------A-C-GT---T------CA-A-CAAAAT.............................--CAAC-CA-CTGGC....................... 6946
CPZ.CM.05.SIVcpzLB7         AC---GCAA-CAGT-----T--A--T------G-G-C--AC-TTCA-------------A-----------T--T-----CA-CT-GGCCT----C---GAC--C-T---ACGC-A--AAACCTTGCTCA-T-AAA-CTTGC---A-...................... 6980
CPZ.CM.05.SIVcpzMB66        CA---GGA-C-AG-A-G-----T---TT---GG-AAC--AC-TTCACC-C-TC---------A--G--C--T--T--C---A-CT-CAT---C-------CCT-CATC...............-TCTT-A-T-A-A-T--GG-TGA-...................... 7081
CPZ.CM.05.SIVcpzMT145       GC----T-GTCC-T--T--G--A--T------G-GACT-AC--TCAC-----C--C--G--A-----C-------------A---GTG-GA-CA-C----TC---AA-.........................AT-AACAA--C-G-G-A-...................... 6939
CPZ.CM.-.SIVcpzMB897        AC---G-AGGTG-A--C--C--A--T-T-----G-AACT--A--TTA---C----------A-----C--T---------A--T--A-CT-A--C----CAAT-CACTA--.............................................................. 6971
CPZ.CM.-.SIVcpzDP943        .....-CAG-TCAACA------A------G-----A----ACATG-TG--C---------G--A--G-----T--T-----CA-C--C-C-T----C----AC-G----ACA----CA---A-CTCAGA-AT-AATAGCTCA............................ 7005
CPZ.CM.98.CAM3              AC----CAGT-AGTC-------A----------G-AAC--ATATG-TG------------A-----C--T--T--C---A-T--TAC-T---------TA---G---AAA--C-AC-A-ATA-C--AG-G--ATAAC-CC-C-......................... 6873
CPZ.CM.98.CAM5              CC----CA-C-AA-C-C-----A--T-----------AC--ATATG-TG-----------A-----C--T--T------A-C--T-C---C-C----AGTT-AAT-A--C-T--GA-AACCAA-ACA-CT-CA-C-GA--T-AC.................ACAAAT 7158
CPZ.GA.88.GAB1              CAC-C--CAG-G----T-----A--T-----GG-CAC-CAT--T-TG-------------A-----------T------A-T-GC---A-A----C-GAC-ACATTACA---G-C-T-ATA............................................. 7485
CPZ.GA.88.GAB2              CA--C-CAA-GCATC-------A---------G-G-----ATTT--TG-----C--------------G--C--T--T------A--T--AG-T-AA---CCT---ACA-TAG-G-C-CA-G-GA-............................................... 6838
CPZ.US.85.CPZUS             AG--A-TT-C-AATG-------A---------G-G-----ATATG-TGC-----------A-----------T--T------A---TT-CC--A-----C-TG--C-ACAAC............-A---C--CAACA-C-C--T-........................ 7462
CPZ.CD.90.ANT               CAA-ATG-AC--T-CA--AT--A--T--T---G-AAA-GT---TT-G--------CCA---A-----C--T--T---G--AT-----CTTG---C---GCC--A-ACACG......GGAAACCTC-TCAC-AACGG-GCCCT-......................... 6901
CPZ.TZ.00.TAN1              CG--CC-ATGGCAGAGG-AT--A-----T--GG-CACT-GCTT-T-G--C-------CA---A-----C--T----------CTC---A-TTG-AC-----CCTGGACAGC-...-A--GAACC-A--A-ACTCATGGT-C-CTT......................... 7089
CPZ.TZ.2001.TAN2            AAAAGTTTG---GGAGTAA---A------T--GG--AGC-GCTT-T-G---------CA---A-----C--T--T-------CTT--TGGTTG-ACA---CCCTGGAACAAC-CA-ACCACA-A.............................................. 7100
CPZ.TZ.2002.TAN3_1          GG---G-CTGGCATAA------A----TT--GG-ACA--ACTT-T-G--C-------CA---A-----C--T--T-------CTT---ACTTG-ACAGG--ACTGGACACG---C--G--ACATCAG--AAGT-CCCTG--C--A......................... 7157
GOR.CM.2007.SIVgor2139_287  ........A-T--T--------A--T------G--ACCTTT-TGCA-----------CAT--A-----------------A-C--TA--A------GCTAC-A---CACA-CAGAGGG---C-A---CACT-AAA---GC............................. 6997
GOR.CM.2007.SIVgorCP2135con ........A-T--T--------A--T------G--ACCTTT-TGCA-----------CAT--A----------------G--A-C--TA--A-------CTAC-A---TACA--GGAG-GC--TGA-TGCAT-ACAGGG---TG-......................... 6981
GOR.CM.2004.SIVgorCP684con  ........A-G--T--------A--T-----GG--ACCT-T-TGCAC----------CAT-----------------------A-C--TA--A-------CTAC-AG--TACAGGGGAACAC--C-A-TG-AC--ACAAGG-GCTTTGCAATAAG................. 6931
H2A.DE.x.BEN                TTACG--A-CGGGAG----CTCA------G---G-G-C-TTTATGT-GAC---C--CA----A------C-----------CATG--TTGGT-CC-C---T-GGTAGAAGAC--G-ACCAAACAC-GCGC......................................... 7945
H2A.GW.x.ALI                TTACA--A-C-GGAAG---CTCA--TG------G-G---TATATGT-GAC---C--CA----A------C--C-T--C--CATG--TTGGT-CC-C---T-GGTGGAAAAC--A-CGGG-CA-GAACAG......................................... 7926
H2A.SN.x.ST                 TTA-AGC--C-GGAGA-C-CTCA----------G-G-C-TACATGT-GAC---C--CA----A------C-------C---ATG--TTGGT-CC-C---T-GGTAGAAAAC-GA-CG-A-CA--CACAG......................................... 7372
H2B.CI.x.EHO                TAGCAG----CG-GAG-A-CTCA--T-------G-AAG-TATATGT-GAC---C---A-G--A--G---C-----------ATG--TTTTT-C--A--CT-GGTAGAAAA--GA-C-GGGCTC-AG-GA......................................... 7904
H2B.GH.86.D205              TA-CAG-A--TGGAAA---TTCA--T--------CAAC-TATATGT-GACC-------A-----A------C--A----------ATG--CTGGT-C--A--CT-GGTAGA-AAC--A-CA-A-ACA-C-CGG......................................... 7918
H2G.CI.92.ABT96             TTACA-CA-C-GGAGA---TTCA--T-------G-AAAGTTCATGT-GAC-------CA----A------C--G--T------AATG--TTGGT-TC-A---T-G-TAGA-GAC-GA-A--TGACTCTCCTA......................................... 7276
H2U.FR.96.12034             TGGCAG-A-C-G-AGG-...--T--TT----GG-AAA-TTTATGT-GAC---C--CA----A------CC-----------GATG---TGGT-CC-C---T-GGTAGAAAAC-GG--C-TGGA-G-ACAA......................................... 7402
SMM.SL.92.SL92B             TAACGGCC-CGGGAGG-...------T--G---G-CAC-TTCATGT-GACA------CA----A------CC-T---------AATG-ATTGGT-CC-C---T-GGTAGA-GA--GG-A---GA-C--CC-A.........................AGGTGGACAACTCAA 7311
SMM.US.x.H9                 TAACAGCT-C-GGAGG-...--A--T------G--ACTTTCATGT-GACA------C-----A------C--A--T--C--AATG-ATTGGT-TC-----T-GGTAGAAGY--TA-AA-A-G-TTC-.........................AGATGGACAAGTCAA 7349
SMM.US.x.PGM53              TGACAGCT-C-GGAGG-...--A--T------G-CACTTTTATGT-GACA------A-----A--G---C----T--C--AATG-ATTGGT-CC-----T-GGTAGAAGA--GA-AC--A-CAG--T--.........................CTATGGACACAACAA 7809
SMM.US.x.SIVsmH635F_L3      TAACAGCT-C-G-AGG-...--A--T------G-CACTTTTATGT-GACA------C-----A------C--A------C--AATG-ATTGGT-TC-C---T-GGTAGA-GAC-GAGACCAAAAT----AC.........................AGATGGAAACAACAA 7882
                                                                          V5 loop start
MAC.US.x.239                TGACGGCT-CTGGAGG-...--A--T--G---G--ACCTTCATGT-GACA-----CA----A--G--CC----------AATG-ATTGGT-TC-A---T-GGTAGAAGA--GG-A---AGCT-AC.........................................CAG 7860
Env                         L  T  A  P  G  G  .  G  D  P  E  V  T  F  M  W  T  N  C  R  G  E  F  L  Y  C  K  M  N  W  F  L  N  W  V  E  D  R  N  T  A  N  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  Q
MAC.US.x.251_1A11           TGACGGCT-CTGGAGG-...--A--T--G---G--ACCTTCATGT-GACA-----CA----A--G---C----------AATG-ATTGGT-T--A---T-GGTAGAAGA--GG-A---AGCT-AC.........................................CAG 7856
MAC.US.x.251_BK28           TAACGGCT-CTGGAGG-...--A--T--G---G--ACCTTCATGT-GACA-----CA----A--G--CC----------AATG-ATTGGT-TC-A---T-GGTAGA-GA--GGGA-GTAACT-CC........................................CAG 7842
MAC.US.x.EMBL_3             TAACGGCT-CTGGAGG-...--A--T--G---G--ACCTTCATGT-GACA-----CA----A--G--CC----------AATG-ATTGGT-TC-A---T-GGTAGA-GA--GGGA-GTAACT-CCCAG............................................ 6991
STM.US.x.STM                TAG-GGCT-CTGGGGG-...--A--T-----GG-CACCTTCATGT-GACA-----CC-T--A------C-T--T--C--AATG-ATTGGT-C--A---T-G-TAGAAAAC-GA-----A-CAGAA.........................ATGAGAGATTGGAACAAA 7514
MNE.US.x.MNE027             TGACGGCT-CTCGAGG-...--A--T--G---G--ACCTTCATGT-GACA-----CA----A--G--C-----------AATG-ATTGGT-TC-A---T-GGTAGAAGA---G-A-CTGACTG-A..................................ACTACCCAG 7337
DEB.CM.99.CM40              TGACA-TA--C-A-C--AA---A--TAGG---G-ACA--GTGT-T-G---C-G---CAT--A-----C-----T-----CATTT-CA--GCCC-AG--CTCCTACTCC-AC-A-A--ACACT............................................ 7182
DEB.CM.99.CM5               TAACAGTA--C-ATC-------T--TAGG--GG-ACA--ACACGT-G---C-A---CCAT--A--G--C--T--T------GTTT-CA--GCA--.......-A-CT-C-C-TG-A--AAACT........................................... 7103
DEN.CD.x.CD1                ..GA-GTA-C-CGGGTG-----A-----------C-AGC-GTGCAC-G---C-G---A-G--A-----------T-----CCTC---AG-A----C--CTCCT--CAA-A-CCAGT--ACAAC............................................ 7484
LST.CD.88.447               CAAA--T-AC-AGAAG---G--AA-T-TT-C--CA-AG-T-TTG-TGCA--CA-------AA---A-G----T---C--GT----AG-A-A--CC-GGAATGGAATAA---A-ACT-AGAT-AATGG............................................ 6616
LST.CD.88.485               CAAA--T-AC-AGACGT--G-AA-GT-TT-C--CA-AG-T-TTG-TGCA--CA-------A----A-G--T-T---C--AT----AG-A-AC-CC-AGAATGGAATAA---A-ACT-AGAT-AATGG............................................ 6613
LST.CD.88.524               CAAA--T-ACGAGAAG-----TA--T-TC-C--CA-AG-T--T--TGCAC-C------T-A----A-G----TT------GTC---AG-A-C--CC-GGAATGGAACAA---A-ACT-AGAT-AATGG............................................ 6619
LST.KE.x.lho7               CTAA--T-ACTAGGCAT----CA---TT--C--CA-AG-T-TTA-TGCA--C--------A----A-G----TT------GT---TAGGA-A--CC-AGAATGGAACAA---A-A-T-AGAT-AATGG............................................ 7691
SUN.GA.98.L14               AACCA-T-G-GAAGCAT----CA--TG---C--CA-AG-T-TTG-TGA-G-C--------AA--GA-G----TT-------T----AG-A-T--C--AGTATGGAATGAC-CA-CCT-CAAT-AATGG............................................ 7759
MND-1.GA.x.MNDGB1           ATAA-T-TACTA-A---...--A--TAG--C-GCA-A--T-ATG-TGA-G-CA---CA---T--A-G----T-------GT----AG-A-CA-G-GGGCATGGAATGA-CC--A-GAGAA--AGTGG............................................ 7049
OLC.CI.97.97CI12            ATGAG--A-TTAAGAG------A--TGAT-CTGCA-A--ATATG-TGA-G-TG-----T--CT-CA-G-GG-TT-------ATC--TA--A-TC----A-CATGGAATAA-GCGTC-GA-CCT-AGTGG............................................ 7090
WRC.CI.97.97CI14            GAAACTTTAC-AGACA-----CC--TAG--CT-CA-A--ATATA-TGA-G-TC---------A---A-G----T---C---ATC---AG-A-AGCAG--GCATGGAACAA-CGC---GAGACC-AATGG............................................ 7594
WRC.CI.98.98CI04            GAAACTTCAC-AGAAA-----CT---AG--CT-CA-A--ATATA-TGA-G-T---------A---A-G----TT--C---ATT--TA-GA-TGC-G--GCATGGAACAAC--G-A-GAGA-C-AATGG............................................ 7639
WRC.GM.05.Pbt_05GM_X02      GAAACTTCAC-AAACA-----CT---AG--CC-CA-A--AT-TT-TGA-G-TA---------A---A-G----TT-----CATT---A--A-AGCAG--T-GTGGAATAA---A-CAGAAAA---ATGG............................................ 6535
MND-2.CM.98.CM16            TC-C-G-CT-CA--A-GAA---C--A----G-GCA-A---TATT-TGC-CCTC--------A----AC----TT-------GG---A--A-T-GG-GA-CATGGAATGA-C-A-A-T-A-CAGTATGG............................................ 7604
MND-2.GA.x.M14              TAAG---CT-CA--A-CATT--C--A----G-GCA-A---AATT-TGA-CCTC--------A----A-----TT-------GG---AG-A-T-GG--A-CCTGGAATAACC-A-ACT-AAATGT-TGG............................................ 7530
MND-2.x.x.5440              TAAG-G--TCCA--GTTAA---C--A----G-GC--A---TATT-TGC-CCTC--------A----A-----TT-------GG--TA--A-A-GG--GGCATGGAATAGC--ACAA-G--CAGT-TGG............................................ 7231
DRL.x.x.FAO                 TGGA---TT-CA--A-CAAG-----T-TG-C-TCA-AG-AC-T--TGA-GTTG--------A----AC----T-------GTTT-TA--A-A-GG-GA-CATGGAACAA--GA-CA-G-AATGTGTGG............................................ 7277
MON.CM.99.L1                CT-GG-GATTTCAAC-CCA-------AAG--GG-CCAG--T---T-G-----C--CCAG--A-----C--T--T------GTCT--GCTT-A--C-T--A--GAA-AACC--C-AA--GGACG-G......................ATATCTCCATTCGACGTGAATAAC 7411
MON.NG.x.NG1                CT-GG-GAT-CCAGC--AA---A---AAT---G-CCAG--C--TT-G--C-------CAG--A--G--C--T--T--C---CTGT-TAG---A--C-TYTCG-G-GACCCG-R--AG-A-AA--A-............G-A-CATT-GAATCCACCTTCAATGTTTCTACC 6036
GSN.CM.99.CN166             GA-GG-CCTC-GAGC-TAA---A--T-T---GG--CA---T--TT-G--CC-A---CA---A------------T--C----TGT--AT---C---C-ACT--A-AATAAC...-CC-T-AAC--C-GC...AATATTGGG---............................. 7388
GSN.CM.99.CN71              GG-GG-CCTCCGAAC-TAA---A--T-T---GG-CCA---T--TT-G--CC-A---CAG--A------------T-------TGT--GT---C---C-ATTC-ACAACAA--T--CGG-AGT-GA-GACAC-AACATT--C--T............................ 7394
TAL.CM.00.266               CA-GG-C-GC-GT-AGGA----A----T---GG-ACA-CTTTT-T-GCA-GCA---CAG--A-----C-----T--C--CAT--GTTCTA----CCTG-A-----CC-A-G-GTA--AC-AC-CAT-AA-CCAAGGC--ATTACACCAAGGGGTCATAGGAAAG...AAG 7269
TAL.CM.01.8023              GG-GGC-AGG-GT-AG-AAG--A----T----G-ATC-CTCTTTT-GCA-GCG--CCA---A--G--C-----T------ATTT-TAGCA----C-TG-AA--CAAC-A--CATAC-A--AC-CAG-A---AAAGG-GTACTGCACCAGGAGAAGATAGGCCGACAAGAG 6819
MUS-1.CM.01.1085            GA-GG-GATC-GAGC--AAG--A--T-A---GG-CCAG--A--TT-G-------------G--A-----C-----T------CTCT-TGTG-----CC-ATT-GT-AACCAC-CAGAG-TAAATGAA......................AGCAACATCCATACT 7379
MUS-1.CM.01.CM1239          GG-GG-CCTCCCAGC-CAA------TAAT--GG-CCA---A--TT-G----------C----------C--T-----C---CT-T--GTG-----C-ATT--ACAACAA--TC-CACAAATT-ACGAG...........................AGCAATATACATAAT 7351
MUS-2.CM.01.CM1246          GG-GG-CCTC-GAGC--AAG--A--TAGG--GG--CAG--A---T-G--------CCA---------C--------C--C-TGT--GT-T-A--CC-ATT-GAGAATAAC-CC--AGAAATT-A-GAC........................TCAAATATACATACA 7486
MUS-2.CM.01.CM2500          GG-GG-CCTCTGAGC--AA---A---AGG--GG-CCA---T--TT-G--------------A-----C--T--T--C--CCTT-GTGTC--A--CC-ATT-GAGAATGAC-CC--ACAAATTCAC-A-........................TCCAATATTGATACT 7400
SYK.KE.x.KE51               CA-GG-G-TCGCAAC--CA---A----TG---G-GAG---A--TT-G---C-A-----C--A--------T-----C---GTTT--A-GA-T---GCA-A-GTG-CAAA-GGCTA-G--AATCCA............................................ 7111
SYK.KE.x.SYK173             CA-GG-G-TCCCAGC----G--A--T-T---GG-ACG---A--TT-G--CC-A-----G--A--G--C-----------GT-T-TA-G--C---GC--AC-T-AC-AA-GGA-ACG-AA-T-AG-A-........................................ 7418
RCM.GA.x.GAB1               TAAGA-CAGT--ATGGT...--A--TGAT--GGCAAG-TATTT-T-GC-A------AAT--A--------A--T--C--G-T--ATTGGT-T--A---TTGTTAAATAA-G-G-CAGTAG--............................................. 7141
RCM.NG.x.NG411              TAAGATCAGTC-ATGG-...--A--TGAT--GGCAAA-TGTTTTT-G--A-------AT--A------C-T--T--C--AATG-ATTGGT-TC-A--CTA-CTAAATAA---A-CAGAAG-C-A--AA............................................ 7191
GRV.ET.x.GRI_677            TGAGA-G-AT--GG......--A--T------TCA-CG-ATTTTT-G--C------CA---T-----------T-----GATGGACTGGT-TA-C---TA-CTAAACAA-CGA-CAGAAGATGCAGAA.........GG--CT-A-AGGACCTGTGACAAAGGGAAGCCA 7452
SAB.SN.x.SAB1C              TT---CTA---AGACA-T-G--A--T------TC--A-TTCTT-TT-----------CA---T-----C-----T-----AATGGATTGGT-CC-C--CTA-TTGAATAA---GTCAGTAGACCCAGAC......CAT--C---TG-GCAAAAAAATAAT............ 7752
TAN.UG.x.TAN1               TT-GGCTAAGTAGACA-T-G--A--T------GCA-CT-ACATTT-GC-C-----CCA---A--G---------T----C-C-TGATTGGT-TG-----T-GCTGAATAA-G-GTC--A-A-TG-A-GA......AATGTAG-TGTAGAAGGTAATAATTGCACCACTGGA 7553
VER.DE.x.AGM3               TT-A-CT---GAGACT-TTT--A--T------GCA-C--ATTTAT-G----------CAG--------C-----T-----AATGGATTGGT-TC-A---TACCTGAATAA-CG--CAGTAGATCCGGAC......CAT---CCGTG-AATGGTACGAAGGGAAAAGGTAAG 7091
VER.KE.x.9063               TA-A-CT---GAGACAGTTT--A--T------GCA-C--ATTTGT-G---------CCAG--A--G--C---------C--GATGGACTGGT-TC-A---TA-TTGAATAA-C-ATCAGTAGATGCA-A-......CAT---C-ATG-AGTAATGAAACAAAAAAAGGGAAT 7593
VER.KE.x.AGM155             TC---TT----AGACA-TTT--A---------GCA-C--AT-TAT-G--C--C---CA----------C-------------AATGGACTGGT-T--A---TA-CTGAATAA-TTA-CAGTGGATGC-GA-......CAT---C-TTG-AAAAACAACGCAGGGAAAGGTCGA 7573
VER.KE.x.TYO1_patent        TA-ACCT----AGACA-T-G--A---------GCATCC-ACTTGT-G----------CA---A-----C--T--T-----GATGGATTGGT-T--A---TACTTAAATAA---A-CATGGGATGCAGAC......CAT---TTTTG-AGCAGCAAAAAG...AAAGGACAC 7569
COL.CM.x.CGU1               CAAGG--TGG-AATAG-AC-A-T--T------GGC-ACTTTCAC-TT-G-G-T----CATA--TT---C--T----------G--T-TAGTT-A-GG--ACA-GA--CC.............................................................. 6852
```

```
                              V4 loop end
H1B.FR.83.HXB2                .........GAAGGAAGTGACACAATCACCCTCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGAAAGTAGGAAAAGCAATGTATGCCCCTCCCATCAGTGGA.........CAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGATGGTGG 7600
Env                           _._._._._E_G_S_D_T_I_T_L_P_C_R_I_K_Q_I_I_N_M_W_Q_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
H1A1.UG.85.U455               .......................--A--T---A---T--------G-----------T---------G-------C----------------------CAA---.........GT---A--G---GA------C--------A--A-----------------G-- 7030
H1B.US.90.WEAU160             ................ACA-T...-----A-----------------------------G-------G-----------------------A------GAA---.........................T---------------T-A------------------- 7601
H1C.ET.86.ETH2220             .......................-----A---A--------------------------------GG-------CG----------------------TGAA---.......AT---A-TG---AG----------C-----A--A---C-G---C-T-----A-C 6976
H1D.CD.84.84ZR085             ...................--A---A---------------------------------------G---------------------------------TGAA--G.............AT---------------------A-----G-------------- 7121
H1F1.BE.93.VI850              .......................-----T--------TC-G------------G-----------GG-------CG----------A--T-------TGCA---.......A-C----CC---AAC--------------TT-A-----G-------------- 6880
H1G.SE.93.SE6165              .......................-----A--------T-AG------------G-G-GA----------GG--G---C---------------------GCA---.......A-C---GA---AAC----GC---------A--AA--------------------- 7006
H1H.CF.90.056                 ...............AC--A-AC--T--A--G-----------------------G-----------G-------CG----------------------CAA---.......A-C----TG---GT---------------A--AA----G----TT--C-AG-- 6919
H1J.SE.93.SE7887              ..............GCAACT...-----AA-A-----T-A-------------G-G-GA--------A-GGAC----C-------C--------------TGCA---.......A-C----CG--CA----------------AT-A-----G---C-T-----A-- 6895
H1K.CM.96.MP535               .......................-----A--------T-------G--G-----------------------G-----------A---------------TGCA---.......AGT----AC---AGC--------------AA--A----G-------------- 6763
H101_AE.TH.90.CM240           .......................-----A--T-------AG-----G------------------GG--C----C-----------T--------------------.......AG-----AT---GT---------------AA-A-----G-------------- 7153
H102_AG.NG.x.IBNG             ..............GACACT...--A--T----A-----------G------G----T------------------C----------------------CAA---.......AT---A------GAT-----C--------A--A--------------------A-- 7109
H103_AB.RU.97.KAL153_2        ....................T-G----A-------------------------------------G-------------G-------------------GCA---.............................................A----------------------------- 6777
H104_cpx.CY.94.CY032          ...............A-TCA--C-----A----A-------C--------T--G---GG----------G----G--GC-----------C---T-C-----TGCA--G.......AGC----AC--CAGC---G-----------AA-AA----G-------------- 6976
H1N.CM.95.YBF30               .......................-T--T--G--T--T--------G---G---G----TT-----ACA-GG-----------G---T-----A--A--A--TC-G---.......GTTC---AC---A-C--C--------T--A---G-TC--GA-TATAG------ 7148
H1O.BE.87.ANT70               .........AGTCA-G--A---ATGG---T--A--T----A-C-G-G---GG-GG---GGTCA---ATA-GG-G-CAGTCG-G-C-C-----A----------AA--T.......A-TC-A-C----ATG-----C--A--T--AA--A-CC--CA--TG---AACAC 7639
H1O.CM.91.MVP5180             ...AGCAAT--GACC-A-A-A-ATGGT--TA-A--T----AGT---G---GC-AG---GATCA---AT---G-G--AGTCGAG---C-----A----------CCC--C.......A-CT-A-C----CAT--C--C--A--T--AA--A-TC--CAGTT----CAACC 7656
H1P.FR.06.RBF168              GGTAATTACAC---GGAGA---T---A--A--G--T--T---C----G---G-AG-T--TTCA---AT--G---G---TC--GCT---T---G--A--T---GAG---.............T-GCAG--CCAC--C--C--------TT-AA-----GAC----CATCACC 7100
CPZ.CM.01.SIVcpzCAM13         ....................AC--G--AA-A-----------T-G-------AG-G--TTCA---AT--G---------GG-G---C-----A-----A--ACCA---........--T--A-CT---AAC--TCTA-----T-----CA----G---C-----.ACCA 7178
CPZ.CM.05.SIVcpzEK505         ....................AT-------T-A-A-----------G-------G-----TT-A---ACA-G-----------G---T-----T--G--A--T-AA---.........CT----AC---CTC--T--C--C--T----A-AA--C-TGAC............ 7075
CPZ.CM.05.SIVcpzLB7           ....................ACT-G--A--A--T--T--G--T-G-------G------C-A----TA-G----------G-G---A-T---T--A--T--T--A--T.......A-C--CC-C---AAT--C-C-------T----T--A-CC--GA-........... 7109
CPZ.CM.05.SIVcpzMB66          ....................AT-----TA-T--T-----------GG------G---GGC-A----A-G-----------G-G---A-T-CTT--C------T--A---.......AC---A-AC---ATC--G-----C-------A-AT-G--TG-G............ 7210
CPZ.CM.05.SIVcpzMT145         ....................TG-CA-TAA----T--T--------G------AG-G--TTCC---AT-CGT-----G-----G---T-T---A--A--A--T--A--T.......A-C----CC---CA-T--T-----C-------TA------C--GAG-T-C--AAGAA 7079
CPZ.CM.-.SIVcpzMB897          ...............ACA-T-TT-CACA-A-T--G----------G-------G----TCA------AGG-------------G---C-T---A------T--T--A---.......AC----CA----AAT--C-CA--C-----C--A--C---------------------AA 7117
CPZ.CM.-.SIVcpzDP943          ....................AT--A-TAA-AAAC--T--------GG------G-----TCC---AT--G---G--C--G-G---T--C-----A--A--T--A---.......ACT--A--T--CA------------T------C--GC-CGAGTT---CA-A-- 7145
CPZ.CM.98.CAM3                ........................TT-T-------T---------GG------G-----TCC---AT--G---G--G--G-G---T---------A--A--T--A---.......GTT--A--T--CA--------CG-----T---A-TA--C-TGAG-C--GGCA--- 7007
CPZ.CM.98.CAM5                GACACAAACTC--AC-A-AGT---C-A---AA-AT-----T-------GG------G-----TCC---AT--G---G--G---G---C--------A--A--A--A--G.......A-C--A-C---CAT---------A--T--A--C--CC-CGA-GC--GA--AAA 7319
CPZ.GA.88.GAB1                ..........................TA--G----------T-G---G--AG-T-GTTCA---AT--GG-------G-G---A-----A--C--A-----A---.......A-C--C-CT---CAAT--T--C------T--A--T--C------TTC----CAC-CC 7619
CPZ.GA.88.GAB2                .....................TA--TT----------A---T-GG---G-AG----TTCT---AT--G---G---G----GT--C-T-------G--T-GA--A---.......ACT--A-CC---AAT--C-CC-----------A--C--C----GA-GT-C-GAA-.. 6973
CPZ.US.85.CPZUS               ..........................A--TAAG--T-------G--------G----TCA---AT--G--------------G---C-T------A--A----AA--T.......GTGC-A--T---AAC-----------A--G--AA--A-TC-TGAC-T-AGCATAA- 7596
CPZ.CD.90.ANT                 ..........................TAGCA-AT---------T--G--G--AG-T--TCAT---GGC-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-CAAAT 7032
CPZ.TZ.00.TAN1                .........................GTTGCA----------C-G-GG--G--AG----TCAT---GGT-T---GTC------GGG-T--CCTT--C--A-GG--G---.......AC-G-A-A----CAC------C--C------A--TA-CA-G---GC---AAAA-A 7223
CPZ.TZ.2001.TAN2              ...............A--GA--TC-AGTAGCA-----T---C-T-G-------AG-T--TCAT---GG--T----TCT-R--G---T---CTT------A-GG--G---.......AC-T-A-AG--CAAT------------------A--TA-TA-G---GC---AAAA.. 7244
CPZ.TZ.2002.TAN3_1            ................ACACATC-ATTAGCA-A---T--GC-T-G---G--AG-T--TCAT---GGT-T----TCT----G---T---CCT-------A-GG-AG---.........T--A-AG--CGTT-----C--C-------A---A-TA-G---GC---A..... 7295
GOR.CM.2007.SIVgor2139_287    ......AATACCACTGAAACA--T-AATATA-T--C---CATC-------GG-AG-C-GGTCC---AT--G------CTC--GCC-A-T---T--G--T--T--G---.......ACTC-A------AT---T--C----------AA-AA-----CA---G----CACC 7152
GOR.CM.2007.SIVgorCP2135con   ..........................ATTATA----C---CATC-G------GG-AG-C-TGTCC---AT--G------CTC--GCC-A-T---T--G--T--T--G---.......ACTT-A------AA---T------------AA-AA-----CA---G----TACC 7121
GOR.CM.2004.SIVgorCP684con    ............ACTGCA--A-TGGAGTATA-T------C-CC-------GG-AG-C---TCC---AT--G------CTC--GCC-A-T---T--W--TM-T--G---.......ACTT-W------AT--YTW----------AA-AM-----CA-------ACACC 7080
H2A.DE.x.BEN                  ..........................A-TAT......---CAT-----G--G--A--T--T-CC-----T--------G---AATG-A---TTG-----T-GGGAA--G.......G-GT-GGCC---GA-----CAG-A--CA-CA-AA-TGCT-AC-TT--CATA-A 8076
H2A.GW.x.ALI                  ....................CAC-ATTATGCA--G---CAT-----G-----A--T--T--C-----C---C---G---AATG-A---TTG-----T-GGGAA---.......G-GT-G-CC--CAAC----CAG-A--CA-CT--A-TGCT-AC-TT--CACG-A 8066
H2A.SN.x.ST                   ....................CAC-ATTATG-G------CAT-----G-----A--T--T-CC-----C--G-----G---AATG-A---TTG-----T-GGGAA---.........--GT-A-CC--CAAC--T-CAG-G--CA-CA-AA-TGCT-AC-TT--C--A-- 7512
H2B.CI.x.EHO                  ..........................-ATTATGCGT-----CAC--C-G---G--AG-C----C------C--GA-T---G-AATG-----TTG-----A-GAGAG--T.......G--C-CTCC---AAT--C-C-G-----CA-C--CA--GCC-AC-TT--CT-GAT 8041
H2B.GH.86.D205                ........................G--ATTATGCG-----TCAT--C-G---G--A--C----C------C--G--T--G---AAT---A----TTG-----A-GAGAA--T.......G-GC-CTC---AAT--C-C-G-G---TA-C--CA--GCC-AC-TCA-CTCA-A 8058
H2G.CI.92.ABT96               AAGCCCCAG---C--CAAAGACGC-ATTATG-A--C---CAT--T-GG--G--A--T----CA-----C------------AATG----CTTG--C----GAGAA--T.......G-TC-G-CG---AAT----CAG-C--CA-C--AA--GCG-AT-T---A-AGAA 7437
H2U.FR.96.12034               ACGTCAAGG--GA--CA-A-A-GG-A-TATG-G--T---CAT--T-GG--GG-A--T----CA------T-------G-G-AATG-----TTG-----A-GAGAA---.......G-TC-C-CC---CAAT--T-CAG-A--TA-CA--A-TGC--AT-TT---C-GAT 7563
SMM.SL.92.SL92B               ACCAAAAAG---CAGCACA-A-G--ATTATG-A--T---CAC--C-GG-----A--T--T-C------C-GG----------AATG-C---TTG-----A-GGGAA---.......G-CT-A-CT---AAC--C-CAG-G--CA-C--AA--GC--AT-TT---T-GAT 7472
SMM.US.x.H9                   AACCAGAAA---C--TACC-G-A--ATTATG-G-----TCAT--TRG---G--A--C----C-------C--------C---AATG---R-TTG-----T-GGGAA---.......G-CC-G-C----AAC--C-C-G-A--TA-C--CA--GC-GA--T----T-GAT 7510
SMM.US.x.PGM53                ACAAAGGCA--C-A--AGA-G-G--A-TATG-G-----TCAT--T-G---G--A--C----C-------C-G------T---AATG-----TTG-----T-GGGAA---.......G-CC-G-C----AAT--C-C-G-A--TA-C--CA--GC-GAT-T----CT-GAT 7970
SMM.US.x.SIVsmH635F_L3        AAGAAGCGA--GCA-CAGA-G-G--ATTATG-G-----TCAT--T-G-------A--C----C-------C-----------C---AATG-A---TTG-----T-GGGAA---.......G-CC-G-C----AAT--C-C-G-A--TA-T--CA--GC-GAG-T----T-GAT 8043
                                                                                                                                                    V5 loop end
MAC.US.x.239                  AAGCCAAAG---CAGCA-A-A-GG-ATTA-G-G-----TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG---.........G-CC-C-CG---AAC--C-CAG-G--CA-T--CA--GC--AC-T----T-GAT 8021
Env                           _K_P_K_E_Q_H_K_R_N_Y_V_P_C_H_I_R_Q_I_I_N_T_W_H_K_V_G_K_N_V_Y_L_P_P_R_E_G_._._._._D_L_T_C_N_S_T_V_T_S_L_I_A_N_I_D_W_I
MAC.US.x.251_1A11             AAGCCAAAG---CA-CA-A-A-GG-ATTA-G-G-----TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG---.........G-CC-C-CG---ATC--C-CAG-G--CA-T--CA--GC--AC-T----T-GAC 8017
MAC.US.x.251_BK28             AGGCCAAAG---C-GCA-AGA-GG-ATTA-G-G--G--TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG---.........G-CC-C-CG---AAC--C-CAG-G--CA-T--CA--GC--AC-T----T-GAC 8003
MAC.US.x.EMBL_3               AGGCCAAAG---C-GCA-AGA-GG-ATTA-G-G--G--TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG---.........G-CC-C-CG---AAC--C-CAG-G--CA-T--CA--GC--AC-T----T-GAC 7152
STM.US.x.STM                  AACAAAAAG---CA-CAAA-G-GG-ATTATG-A------CAC--C-G----G-A--T--T-CA-----T--------G---AATG-T---TTG-----A-GACAA---.........G-CC---CT---AAT--C-C-G-G--CA-TA-AA--GC--AC-TT--CT-GAC 7675
MNE.US.x.MNE027               AAGCCACAG---CAGCA-A-A-GG-ATTA-G-G-----TCAT--T-G------A--C----CT-----T--------C---AATG-T---TTG-----A-GAGAG---.........G-CC-C-CG---AAC--C-CAG-A--CA-T--CA--GC--AC-T----T-GAT 7498
DEB.CM.99.CM40                AGAAATAGCACTT-GTC-----AGTGGTTGA-G--T--T--G-----T--GT-AG------CT---T-T-CG-----GC--CAT--C---CTA-----G-AAGAA---.........G-GT-G-AG--CAGC---C-C---T-T-CAT-TG-T--TGATGT----CACTA 7343
DEB.CM.99.CM5                 AATAACATAACCT-G---A-T-AC-CTTTGA-G--T--T-AG--T--T---T-GG------CA---TTC--G-------C--CATG-A---TTG-----A-AGGAA---.........G--T-G-AG---CAT--TTC-G--T-T-CA--TG-T--TGAGGT--C-CA-TA 7264
DEN.CD.x.CD1                  ...AGTAACAT-ATT-ACA--CAG-CATATA-GGTG--T-AC-----------T-AG-G----CA------C-C-------C-TG-AT--T--CTTA--A--AGCAGAA---.........G-------C---CA-C--C---G-GT----CA-TA-TG--GAC-CG----TCCA 7642
LST.CD.88.447                 ......................TATCC-TGGGC-AAT---CAC------TC-G-----G-TGAT---GCATCT-----G--GAA---T---CT------A-CATCA--CTTTAACAACAG---A------A-TCAT-GAG-A----AAA--TGG--TCA--TG--AAAGTT 6765
LST.CD.88.485                 ......................TATCC-TGGGC-AAT---CAC------TC-G-C---G-TGAT---GCA-CT-----G--GAA---C---AT------A-CATCA--CTTTAATAACAG---A------A-TCAT-GAG-A----AAA--TGG--TGA-CTG--AAAGT- 6762
LST.CD.88.524                 ......................TATCC-TGGGC-AAT---CAC------TC-------G-TGAT---GCA-CT--T--G--GAA---T---CTT-----A-CATCA---TTTAACAATAG---A------A-TCAC-GAG-A----AAA--TGG--TGA--TG--AAAAT- 6768
LST.KE.x.lho7                 ......................TATCCTTGGGC-AAC---CAC------AGC--C--TG-TGAT---GCT-C-A-------GAAG--T---CTA--A--T-CTTCA--TTTCAACAATAG---C------A-TCAT-GAG-A----AAA--T-C---TGAG-TG--AAAAT- 7840
SUN.GA.98.L14                 ......................TATCCTTGGGC-AAT---CAC------AGTG-G---G-TGAT---GC-TC---G--C--GAAG--T--CTTG--C--A-CATCA--GTTTAATAACAG---A------G--AAC-GAG-G----AAGCATGG--C---CTG--AA-G-T 7908
MND-1.GA.x.MNDGB1             ......................TATCCTTATGC-T----TCA---T-GG-----AG--G-TGAC---AT-C----------G-AAG--A---TTA--A--T-CATCA---TTTAATAAT--C--A--G--A--CAT-GGG-A----AAA--TAC--TGA--TGC-AAAGAT 7198
OLC.CI.97.97CI12              ......................TATCCATGGTCA-AG--T---C-G-G-----A-G--G-TGAT---AT-------------G-AG---T---CTT-----ACAACCA---TTCAATAATAG---CC-----..............---CAAAGAG--ACTGA--TGTT-TTG-A 7227
WRC.CI.97.97CI14              ......................TATCCATGGAGT-------C---T-G-A-T--AG--G-GGAC---GCAC-T--G----T-AA-G-A---TTG-----AG-ATCA--GTTCAATAAT--T--------C.............TCAAAGAG--AC-GA--T-TTGTT-AA 7731
WRC.CI.98.98CI04              ......................TACCCTTGGAGT-------CG--T-G-A-T--AG-GG-TGAT---GCAC-T-------GT-AA-G-A---TTG-----AG-ATCA---TTCAGGAAT--T--A------..............---CAAAGAG-GACTGA--TTCTATTCAA 7776
WRC.GM.05.Pbt_05GM_X02        ......................TATCCTTGGGCA--T----CT--C-G-TC---AG-GG-TGAT---GCATCTA-----G--GAA---A--CTTG-----GG-GTC----TTTAACAAT--C-----------............---CCAAAG-G---C-GA-GT-CTCTT-AA 6672
MND-2.CM.98.CM16              ......................TATCCATGGA-GT----T-AT--T-G------A---G-TGAT-----T----------GAAG--T---ATG-----TG-ATC----TTTAATAATG-G--A--------TAATG--G-A----AAA--T-C--TGAGGTTC-AAAGAC 7753
MND-2.GA.x.M14                ......................TATCCATATA-GT----T-A---C-G---G--AG-GG-TGAT-----T-----G----GGAAG--T---ATG-----TG-GTC----TTTAATAAT--T--A------CA-TAATG--G-C----A-A--T-T--TGAGGTCC-ACTA-T 7679
MND-2.x.x.5440                ......................TACCCTTA-A-GT-TC---AT--T-G------AG--GGTGAT-----T---------C-GAA---T---ATG-----TG-GTC---GTTTAATAATG----A------A-TAATG--G-G----AAA--T-C--TGAGGTCC-AAAAAC 7380
DRL.x.x.FAO                   ......................TATCCATATGCAT-C---CAC-----------A---G-TGAT---GCA-G------C-GGAAG--T---ATG-----TG-ATCA---TTTAATAATG----A------A--CA-G--G-C----AAA--T-C--TGAG-TCC-AAAATT 7426
MON.CM.99.L1                  AAGCCCAACACCACGTA-C-TGGTGGGTGG--AG----T-CC-----G-----GG-G-CACAA---GGAT-T--GTC---GAGT--C--CCTA-----ACGG-AA--C.........--TG-GCAG---A----T------C--T-CC--C--GA-T---G-G--GTTGTA 7572
MON.NG.x.NG1                  AATGCCTCCACT-CCTA-A--GG-GAGTGGT-AA-G---C--C-C-GG---C-AG---CTCA----GGAT-CA-CTCC--GAGTR-C----CTG--A--A-GACAA--G.........--TG-GCAG---A-C--T------C-----CC--CT-GG--GATG-G--GTTGTA 6197
GSN.CM.99.CN166               ...ATCACAAGCAA-TACA-AGG-CAATGG--GG-T----A---T-GG---T--G-C-C-CAA---GGAT-T--GTCT---T-----C--CTTA-----ACGGCAG--T.........--T--A-AT--CA----C--C---------T---T-GA--GATG-C-C-ATGTA 7546
GSN.CM.99.CN71                ...GTCACAAGCAA-TACA-AGG-CAATGGA-GG-T---------T-GG---T--G-C-CACAA---GGAT-T--GTCT-GGT-G--C--CTTA-----ACGGCAA--T.........--T--A-AT---A----T--C--C--T--C--C----A--GACG---CGATGTA 7552
TAL.CM.00.266                 CCGAGGTCGAGT-AT--GAC--AGT--TA-A-GT-----CA-C-G-GG------G--C-----G-----AGC-GG--G-AG--GTT-------CCT---A----GGGAA---.........--CC-GCAG--CAA-AGT--C--C-----AT--T-GG-GGAT-TCA-CCACCA 7430
TAL.CM.01.8023                CACAGGTATAGT-AC--AAC--A-T-TTT-A-GT-C---CAGT---GG-----G--C--T-G-----AGC-------AG--GCT-------CTG--G--A-GAGAG---.........--CC-G--T--CAA-AGC--CG-C-----TA-C--GG-GGAC-TCA-CCACTA 6980
MUS-1.CM.01.1085              GTAACAACTA-GTACCACA-GGATCAATGGA-GGTC--T-----T-G-----T--G-C---CA----GGAT-T--GTC---GT-T--C--CTTG-----ACG--AG--C.........TTG--C-----GAG--T-----A---------T--CA-TGACG---CCATGTA 7540
MUS-1.CM.01.CM1239            GTGACAACCA--TACCACC-AGATCAATGGA-GGTC--T--G--T-G---GT--G-C-CGCAA---GGCT-T--GTC----T-C--C--CTTG-----ACGAC-G---............--TG-GCAG---A-T--T---G-G--T------C--TA-TGATG---C-ATGTA 7512
MUS-2.CM.01.CM1246            GTAAAAACCA--TATC-CC-TGACCAATGGA-GGTC-----------GG--GT--G---C-CA----GGTT-C----TC---GT-C--A--CCTG--C--TCGA-AA--T.........--TG--CA----A----CC--G-A--------AT--A--GATG-G-CCATGTA 7647
MUS-2.CM.01.CM2500            GTAACAACCA--TACCACA-GGACCAGTGGA-GGTT--T--G----G----T--G-T-C-CA----GGTT-TACGTCG--GT-C--C----CTG--C--TCG-CAA--C.........--T--CCAG---A-G--C---G-G--T--T--C--CA--GATG---CCATGTA 7561
SYK.KE.x.KE51                 ......AGCA-TTATGCAA-A-ATT-ATAT---T-G--TGCC---CG---G---------TAT---GG-T-T---ACC--GCT--------TTG--A--T-CAGAA--G.........--C----AG--CA-T--T--------T-CTG-AT--AC-GAT-T---GTACTA 7266
SYK.KE.x.SYK173               .........A-TTATGCCAGT-ACC-TCGT--GT----TGCC--TCGC------A--T--TGAT---AGAT-T---A-----TTG--A--CCTG--C--T-CGGCA--G.........--T----A----A-T--C---G-C--T-CAG-AT-GACGGAC-T---ATACTA 7570
RCM.GA.x.GAB1                 ACAACAAAT--GAAG--AA-AG--CCTTTTG-A--------TC-C----ATG--AG-C--TGAT---T-T-C-----TCG-GGAAGG-A--CA-G--A--G-GGCCA-AT.........GCGT-A-AG--CAGTG--C-GG-AT-CTAT---T-GGC-GAC-T---CTA-.. 7300
RCM.NG.x.NG411                ......AAT----CT--GC-AG-C--GTTTG-T--T--T-TC-CT---ATGG--G-T--TGAT---T-T-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-AC 7346
GRV.ET.x.GRI_677              GGACCAGGACC-T-TGT-C-G-G--CTTATG-TG-C---CAT---CG-----G-AG----TGAT---T-C-CT--CTCT---AAGG-A-----T--A--A-GGGAA--T.........--TT-GGAG---AAC---TCAG-C--G-CA--ATACG-GG---T----TA-AA 7613
SAB.SN.x.SAB1C                ...ACAAAGCC-T-TT-GC-G-G--C-TATG-T--C---CAC----GG----G-AG----TGAT---T-C-CCT-GTCT---AAG-CA--C----------GAGAA--C.........--CT-GGAG---AAC--G-CAGCC----C----TATG--GA-TTGA-CTA-AA 7910
TAN.UG.x.TAN1                 AAGGATAAACGCT-CTACA-A-GG-CATATG----C---CAT--T-GGTC-----G-C---TGAC---T-C-CGC-CA-T--GAA--CC-----A--A--A-GGGAA---.........--CT-AGA---CA-------CAG-G---TCTA--A-GG--T--CTG---TA-AA 7714
VER.DE.x.AGM3                 GCACCAGGACCCT-TGCAC-A-G--CATATG-TG-T---CAT---CG-TCTG-C--T--TGAT---T-C-C-C--TC--GGAA--CC-----A--G--A-GAGAA--G.........--CT-GCA---CA----C-CGG-A--G--TA--TC-G-GGAGCT-A--TACAA 7252
VER.KE.x.9063                 GCCCCAGGACC-T-TGTAC-A-GG-CTTATG-TG-T---CAC--C-G-TCTG-C-----TGAT---T-T-CCA--TC---GAA-G-------T--A--A-GGGAA---.........--TC-GGA---CA----C-CGG-C--T--TA--AC-G--GA-CT-A--TA-AA 7754
VER.KE.x.AGM155               AGTCCAGGTCCCT-TGTAC-G-G--CTTATG-TG-C---CAT--CCG-TCTG-C-----TGAT---T-T-CTA--TC---GAA--CA-----T--A--A-GAGAA---.........--TT-GCAG--CA-G--C-CAG----T---A--AC-G--GAGCT-A-CTA-AA 7734
VER.KE.x.TYO1_patent          GCACCAGGACC-T-TGTAC-A-GG-CGTATG-TG-T---CAT--C-GGTCTG-A-----TGAT---T-T-CCC--TC---GAA--CT-----T--G--A-GAGAA---.........--TT-GCA---CAGG--C-CAG-C--T---A--AC-G-TGAGCTTA--TA-AA 7730
COL.CM.x.CGU1                 ........................CCAGT-A-GAAC--T-CC--T-G-A-GC-GG-G--T-GT---GTC-CTCAT-CC-GGATCC------GG--G--TCCAG-A--T.........--TC-ACA----AAC-GGG-G-AACAGCCAG-TA-TGCGTTC-TG-GCACCAT 6989
```

```
                            V5 loop start                                  V5 loop end                                                                                      Rev Responsive Element (RRE) region
H1B.FR.83.HXB2              TAAT.................AGCAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATAT..............AAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAG 7737
Env                         __N__.__.__.__.__.__._S__N__N__E__S__E__I__F__R__P__G__G__G__D__M__R__D__N__W__R__S__E__L__Y__K__Y__.__.__.__.__._K__V__V__K__I__E__P__L__G__V__A__P__T__K__A__K__R__R__
H1A1.UG.85.U455             --C-AAC.................--T-CAA-AAAT----C--------------------------------------A---------------G--..............-----------------C-----------------G------------ 7167
H1B.US.90.WEAU160           --G-AGT..............GAAG-G--CC--A-----------------------A------A---T----------------------------..............------------------------------------------------ 7741
H1C.ET.86.ETH2220           C--A..............GAGCCAC-T-GCACAAAA-----A---------A--------------------------------------------..............-----G---G-----A-G---C---------------T---C----A--G---- 7116
H1D.CD.84.84ZR085           -GC-.................-AT--T-C-C--AAT--C-C---T--------------------A-----C------------------------..............------------------C------------------------------- 7258
H1F1.BE.93.VI850            ----.................G-G-G-A-CATT----C----------AG------A------AA---------------C-------G---..............-----G------------------------------------A---CA-- 7014
H1G.SE.93.SE6165            G---AAT......AACAATACA-AT-CA-G----AGT-----------------------------------------------G---..............---AC---------CA--T--C----------------G-----G-------- 7155
H1H.CF.90.056               ---C.................GCGTCTGCA--AAATT-T-C----------------------------------G------------..............----------------C-G---A-------------A---G-------- 7056
H1J.SE.93.SE7887            A--CAGG......GGGAATGGC--TG-G----GCA-T----C----------AC-------A------A------------------------..............-------G-G----G---C--------------------------A------- 7044
H1K.CM.96.MP535             -....................-AT--T-C-C-TAAT----C---------------------C-----------------C-G-----------..............-------C-G---------C-----A-------------G---C-G-------- 6897
H101_AE.TH.90.CM240         -GT-AAT..................--T-CG--TAA-----C-------------------A-C--A-A-------------------------..............-------C------------C-----A-------------G---------------- 7290
H102_AG.NG.x.IBNG           G---.....................--T-G-ACAAAT--A-C------G--------------------------------G------------..............-----------------C----T------------CGT-----A-------- 7243
H103_AB.RU.97.KAL153_2      ---C.................CAG-G-----TTA--------T--T---A---------------------------------------------..............------T-------------------------------G------------- 6914
H104_cpx.CY.94.CY032        --C-.....................---A-TA-T---------------------C-----------------C-----------------------..............-----------G--------A--------------AT-------G----C--- 7107
H1N.CM.95.YBF30             GCC-...................G---CCA--GAAACA--AG-ATAT--CTC-------A-C---GTTA-TCTC------CAA--G--G-----G--C..............-------T-GC--A-----CA------------AGGT--A--T--A---C-CA 7282
H1O.BE.87.ANT70             ATGGAACAGC..............-G---CA-CAAT-TA-CA--T-----AAT-----G-----C-----AA--T-TA-------C-------G-TC--C--C..............---------GGG-AA----T--TA-T--G---------ACGTATTGCA--GCC-- 7779
H1O.CM.91.MVP5180           ATGG.................-ATTC--CA-GTGAAA-T-CAC-T-----A-T---G----------AA--T-TA-------C-A----G--C--C--C..............-------C-G--AA----T--TA-T---------T--A--AATGTCA---CC-A 7793
H1P.FR.06.RBF168            -T--AAT...........GCC-A--G--GCAGCAA-ACC-CT--A--C---AC---G------C-----AC-T-TC-----------G------CCC--C..............--G-----TC--G-AA--G-TC-G-CT--G-----T--A-GA-TCTCC---CCTA 7243
CPZ.CM.01.SIVcpzCAM13       -GT-.....................--T---ACTAATA-T-CA--T-----AAT---------C-----AAA---T------G------C-C------C..............--------G--GCGC--A---------TCT--G-----A------A-----A--G...C 7309
CPZ.CM.05.SIVcpzEK505       ............TATACAAAGTCTGGA-C----AAGT-C-C-A-ATAT--CAC---G-----C----CTA--CTC-----GCAG--G--G------G---..............--G--G--C-GC--A-----CA------------AGGA--A--T--A---CACA 7217
CPZ.CM.05.SIVcpzLB7         ..............AAACAC-CA----CA--CAATATC-C----C-T--AAT-------G-------CA----TC------------------C---..............---A----------------------------G--T----G---G-GA--G--GA 7248
CPZ.CM.05.SIVcpzMB66        ....CAGCAA...AAAACAGAC--G-TG--CA-A--A-CC--G----C-----TT--G-----A-----AA-T-------------------C-------..............---------------G-----------C-------------------------A 7358
CPZ.CM.05.SIVcpzMT145       CCGAGAAGATCAGGGAGAAGATCAGG-TC-GA-TAA-ACATA-G-TT-T-TAAC---G--TA-C----AA--T-T--------------C----C------C..............---A----TG-G--AC-----C----------------T--T--AAGT-GA---TAT- 7234
CPZ.CM.-.SIVcpzMB897        A---GAG.....................A-T--A-CT--T-------AC---------------C-----AA-T------------------C-----C..............--G------------------CC------T---------A-----------A--G---A 7245
CPZ.CM.-.SIVcpzDP943        ACTCCCGAGAAATGGGACT...-AT-C----A-CAAGTTA-CAC-ATAC---AC---------G------A--TTTG-----GCTG------C-------C..............-----------C-GC--A-----CA-----------G-----A-GT--A--T--A--GCACA 7297
CPZ.CM.98.CAM3              C-TC...............AAC-A--G--TCACCAA-ATC-CGC-ATAC---AC-------A-C---GTA---TTG-----GCT-------GC--------C..............---G-----T-GC--A-----CA-----------------A-GT--A--T--A--GCACA 7147
CPZ.CM.98.CAM5              ATGGGAA................-A-......A-TA-TATA-CAG-GTAC---TC---G-----A-----A----TTG-----GCT--------A-C-GG---..............---------T-GC--A-----CA----------------A--T-----T--A--GCACA 7453
CPZ.GA.88.GAB1              -GTGACT...............-A-----G--GTAA-CT--CA--T------AC--------A-C----AA--T-T------------GC-C------G---..............---------TCG---A--G--T---TC---G-----A--A--A----GA--GCATA 7759
CPZ.GA.88.GAB2              .......................G---CGGGGA-CAA-ACAGAAG-TTAC-TGTC---G------------A--T-T-------G-------------------..............---A---------------G---C-G--G--G-----A--A--A--------GTACA 7106
CPZ.US.85.CPZUS             CGCAGTC............AAT-A-G-T-G-AG-AATATA-CAG-G-TG---AC-------------C--CTTTA----A--A-------C-----G---..............--G--G--C-GC--A----TA---------G------AGGT-------C--A--GCATA 7739
CPZ.CD.90.ANT               AT--..........................----AA-T-TT-AAG-GTC-----CT-C-A---TAGCA...---C-A-------GCG--G--G-CC-GG--C..............C-G--G---G-G----RT--C--GTC-------C---A--A-CA-GN--A--GCC-- 7160
CPZ.TZ.00.TAN1              C--CAAT........................AATAGTTATAC-CCC-AATTTTCT-CT-TAG-AGAA---T------A-GTA------GCA-G-------..............--------G--GG-----C-G--C--GTC---G--T--A-GGCCA-G---A--GCCT- 7351
CPZ.TZ.2001.TAN2            ........................GGAGG-...AATA-C-CAG-TCCTA-ATTCTCC--TA-AG-AGAA---T------A-GTA---G---GCA-GG-----..............-----------G-G----C-G--C--GTC---G--T--T-GGCCA-G------GCCT- 7371
CPZ.TZ.2002.TAN3_1          ........................GGAGGAA-CAG-A-C-GAAC-CCCA-ATTCTCCA-TA-GG-AGAA---T------A-GTA------GCA-GG-----..............-------G--GG-G---C-G--C--GTC---G--T--T-GGCCA-G---A--GCCT- 7425
GOR.CM.2007.SIVgor2139_287  C.................CTA-AA-CTG-----AA-AGC-C---A-----AAT------T--C----CAA---T------G------C-G--CCCC--C..............--G--G--G-GGG-CA-GG---GACT--G-----A--A--A---------GCCT- 7289
GOR.CM.2007.SIVgorCP2135con C...CTGAATAAGACTAATGAG-AT-GT-C---T-A-A-C-C---A-----AAT-------G--C-----CAA---T------G-----G--G--CCCC--C..............-------G-----GG-CA-GG---GTCT--G-----A--A--A-----G-----CCT- 7273
GOR.CM.2004.SIVgorCP684con  CCYS...............GCTGAG--T---A-TAA-ACC-C---A-----ATT------T--A------AA---T------G------G--CCCC--C..............--G--G-----GG-CA-GG---GTCT--G-----A--A--A---------GCCT- 7220
H2A.DE.x.BEN                ---AAAT........................CG-A-TC-T-C-AA--TTA-CTTTA-T-C---AG--GCA--ACTG-ACC--TTA---C-GGGAG-C--C..............---T--A--G---AAC----A-T--CT-C------T--AG-TCAG-GA--GTACT 8207
H2A.GW.x.ALI                -GGC...........................AA-C---CAAAT-TTA-CTTTA-T-C---GG--GCA--ACTA-ACC--TTA-----GGGGG-T---..............---T-----G-G--AAC----A-T--CT-C------T--ATCA-A--G----TACT 8191
H2A.SN.x.ST                 AG-G...........................AA-C---CAAAT-TTA-CTTTA-T-C---GG--GCA--ACTA-ACC--TTA-----GGGGG-T---..............---T-GA--G--G-AAC----A-T--CT-T------T--ACCA-T---A---TACT 7637
H2B.CI.x.EHO                AG--AAG........................AA-CTT-CTAAT-TTA---TGA-T-C---AG--TCA--ACTG-AT-A-TTG---C-GGGGG-C--C..............---T-----G-G--AAC----A-T--CT-T------T--A-GTAT---A---TATT 8169
H2B.GH.86.D205              C............................A-CAGTACC-C-AAT-TTT---TGA-T-C---AG--TCA--ACTG-AT---TTG---C-GGGGG-C---..............---T-----G----AAC----A-T--CT-C------T--AG-T-T--GA---TATT 8183
H2G.CI.92.ABT96             --GGTCTAAC.....................AA-C-T-CAAAT-TCAT-TTCA-T-C---GG--GCA--ACTG-ATC--TTA-----GGGGG-T---..............---T-----G----AACT---A-T--CT-T-------------A--T-T-------TATT 7568
H2U.FR.96.12034             ---CAAA........................AATC-A-CAAAT-TCT-CATGA-C-C---AG--GCA--ACTG-AT---TTA--G--GGGGG-T---..............---T-G---G----AAG----A-T--CT-T--T-----ACCA-TG--A---TATT 7691
SMM.SL.92.SL92B             -G-CAAC........................AAT----C-AAT-TTA-CATGA-T-C---AG--GCA--ATTG-ATC--TTG-----GGGGG-T---..............---T-G---G-G--CACT---A-T--CA-G--T-----AC-T-TG--A--GTACA 7600
SMM.US.x.H9                 C---GGC........................AATA---C-AAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG--G--GGGAG-T--C..............---T-----G-G---ACT---A-T--CT-C--C-----A-GT-T---A--GTACA 7638
SMM.US.x.PGM53              -G--GGC........................AA-C---CTAA--TCA-CATGA-T-C---GG--GCA--ACTG-AC---TTG--G--GGGAG-T---..............---T-----G-G---ACT---A-T--CT-G--C-----A-GT-T---G--GTACA 8098
SMM.US.x.SIVsmH635F_L3      C---AAC........................AAT----C-AAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG--G--GGGAG-T--C..............---T-----G-G---ACT---A-T--CT-G--C-----AG-T-T--GA--GTACA 8171
                                                                                                                                                                               Rev Responsive Element (RRE) start
MAC.US.x.239                -G--GGA........................AA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T---..............---T-----G-G--CACT---A-T--CT-G--C-----AG-T-TG-----GTACA 8149
Env                         __D__G__.__.__.__.__.__.__.__._N__Q__T__N__I__T__M__S__A__E__V__A__E__L__Y__R__L__E__L__G__D__Y__.__.__.__.__._K__L__V__E__I__T__P__I__G__L__A__P__T__D__V__K__R__Y__
MAC.US.x.251_1A11           -G--GGA........................AA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T---..............---T-----G-G--CACT---A-T--CT-G--C-----A--T-TG-G---GTACA 8145
MAC.US.x.251_BK28           -G--GGA........................AA-C-A-CTAGT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG--G--GGGAG-T---..............---T-----G-G--CACT--GA-T--CT-G--C-----AG-T-TG-----GTACA 8131
MAC.US.x.EMBL_3             -G--GGA........................AA-C-A-CTAGT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTA--G--GGGAG-T---..............---T-----G-G--CACT--GA-T--CT-G--C-----AG-T-TG-----GTACA 7280
STM.US.x.STM                ---CAAT........................AAT----CTAA--TCA---C-A-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T--C..............---T-----G----AACT---A-T--CT-G--C--A--G--T-TG-------TACA 7803
MNE.US.x.MNE027             -G--GGA........................AA-C-A-CTAAT-TCA-CATGA-T-C---GG--GCA--ACTG-ATC--TTG-----GGGAG-T---..............---T-----G-G--CACT---A-T--CT-G--C-----A--T-TG-----GTACA 7626
DEB.CM.99.CM40              ----GGG...........................AGT--AAC-CTTA-ACC-TC--C---C--C----CAGTA------GCA--CC---T------..............---A-CA--G--G-AA----TA-----T--C--T---T--GCA-TG-GA--GTAC- 7468
DEB.CM.99.CM5               -GGGCAA...........................A-C--GAA--T-A-ACCCTCCTC------T-A--CAGT------GGCA--TC-----C--G---..............--GA-CA--G--G--A----CA-C--GA-G--C--T--A-GTAT--GA--GTAC- 7389
DEN.CD.x.CD1                -CCGAAC.....................AGAA-AC-T--AACATTGA-GCTCTC--CT---G-----C-AGTC------GC-------GGCA-G--GG..............---T-----G-GT-AAC----A-T--CC-T--T--T--TTCA-TG-GAC--TAT- 7773
LST.CD.88.447               -G-ACCAGAA......GAAAACCTGGGAGGAA-T-TGTCTG--AGATTCTTGCCTCCCTCCTGGGAA-CAA-TC-G-TTGTGGCAAC-GG-GC-C-C---..............---T--A----GGC--AG----A-T--CT-T--T--A--AG-T-TGC-TC-GTAC- 6914
LST.CD.88.485               GG-GCCAGAG......GAAAACCTGGGGGGAA-T-TGTCTG--AGATTCTTGCCTCCCTC-TGGGA--CAA-TC-A-TTGTGGCAACGGG-GC-C-T---..............---T--A----GG---AG----A-T--CT-T--T--A--AGCT-AGC-CCAGTAC- 6911
LST.CD.88.524               GG-ACCACAT......GAAGACTTAGGAGGAA-TCTGTCT--TAAGTTTTTGCCTCCCTC-TGGGA--CAA-TC-A-TTGTGGCAAC-GG-GC-C-T--C..............---T-GA-C-G-C-GAG----A-T--CT-T--T--A--ATCA-AGC-TC-GTAC- 6917
LST.KE.x.lho7               GG-ACCACAT......GAGGACTTAGGAGGCA-CCTAAGT--AAAGTTCTTGCCTCC-TC-TGGGAA-CTA-TC-A-TTGTGGCA---GGG-C-------..............---T-GA-----T-GA-T---A-T--GT-T--T--T--AG-T-AGC-C--GTAC- 7989
SUN.GA.98.L14               GG-AGACTGG...AAGGTAAATG-ATCA--CATC--A-TTG-AGC-TTC-AGCC-CCTACCA-C-CCTTAA-TC-A-TTGTG--CAC-GG-GCAC-T--C..............---C------G----AGG---A-T--GT-T--T--A--AG-T-AGC-CC-GTAC- 8060
MND-1.GA.x.MNDGB1           AG--AGTAATGAA...............CAA-AATGC-A--TAAATTCTTGCCTCCCA-T--A-CCTCCA-TC-A-TTGTTGC-T-TGG-GC-C-T---..............---T-----C-----AATG---A-T--CA-------T--AG-T-TG--A---CACA 7338
OLC.CI.97.97CI12            A-TACAAAAAGTAGAGAAC............-GAA-GC-AG-G-CATTG-T-CC-CC-TC---C-ACGT-A-TC-A-TTGTCTCA-TGGC-AG-G-T---..............---T-G---G-GT-G---TTGA-T--CT-T--T--T--AG--A--C-A--GTAT. 7369
WRC.CI.97.97CI14            CTGGGAA.....................AGAAATC-AG-GA-GTTC-AA-CTCCTTCCA-C--ACA-A--TC--TTGTGGCA-TGGG-C-A-------..............---T-G------T-AA--ATGA---GA-G-----AGGGG-T-G------GGTT- 7862
WRC.CI.98.98CI04            -TGGGAA.....................A-TAAGC-AG-GGAATTC-AA-CTCCTTCCA-C--ACA-A-TTCA-TTGTGGCA-TGGG-C-A--G---..............---C-G--G-G-C-AA--ATGA---GC-G--T--TGGTG-T-GG-GA---GTGA 7907
WRC.GM.05.Pbt_05GM_X02      -TGGGAA.....................A-CGGGC--G-GGAGTTC-AA-CTCCTTCCC-C---CAAA--TCC-TTGTGGCA-T-GG-C-A-------..............---C-G------C-AA--TTGA----------T-----TG-T-G------GGTGA 6803
MND-2.CM.98.CM16            AG-AGAGGGA......................-A-ATC---AAATTTGTGCC-CA--AT--GG-ACAAA--CGA-TT-C-GCG-T-GG-GCAC-C---..............C--C--------G-G--C---A-T--CT-T--T--T--AG-A-TGGCA---TATC 7884
MND-2.GA.x.M14              -G-AGGA...................--AAGA-A-CTA--AAAGTTT-TGCC-CA--AT--GG-ACAAA-TC-A-TC-C-GCA-TGGG-GCAC-T---..............---T-G------G-G--T--GA-----T-C--C--T--TG-T-T-C-C---TATC 7813
MND-2.x.x.5440              -G--GATGAC.................----GA-ATATC--AAAATTTATACC-CA--ATTGG--TCAAA-TC-A-AC-C-GCA-TGGGGGC-C-T--C..............---T-G--G--GG-G--T--GA----TT-T--C------AG-CAT-C-C---CATC 7517
DRL.x.x.FAO                 GG-GACAGAG......GAAACAG-AG-TTCCA-C-ATC----TAAATTTATACCGCA--AT--AG---CAAA-TC-G-TC-C-GCA-T-GG-GCAC-T--C..............---T-G--C--GG-G--G----A-T--CT-T--T--A--TG-T-T-C-T---TACA 7575
MON.CM.99.L1                -C-AAAC.........................A-CG-GAC-CTTGT-CCCTCG-C-C-GG----T---TCA-------TCA--GC-CAG-CG-----..............--G--G--GG-------C--C---TCTA-G-----T--A-CA--GC-------GA 7697
MON.NG.x.NG1                -GGCAGC.........................ACAG--ACGTTGAT-CC-TC--C-CGAG--TCT---GCA----C-CTCA--G--GAGC--G--C..............--G--G--CG-G--C--C--C--GTCT--G-----T--AGCA--CC-A--G---C 6322
GSN.CM.99.CN166             -G-AAGC.........................TCT---AA--TGA-CCCGTCG-CC----G--GCT--TGCG----ACTA-------G-CC-GG--C..............-------G--GG-G--A--T--GC-CTCGA-G--C--A--TCC---T--A--G-A-- 7671
GSN.CM.99.CN71              -GGACAT.........................TCC---AAT-TGA-CCC-TCG-CC----G-AGCC--TGCA----AGTA---G----CC-GG--C..............--G--G--GG-G--A--T---C-TTCGC-G--C--A---TC----TC-A--G---C 7677
TAL.CM.00.266               CCGCGGCTCT......................AGT--C--AACAGT-GAGCCCTCGTC-C-G--AC-T---TCG----C--GCA-C-G-T-CACGG--C..............--G--GA--G-G--C-TG--T--T--GTAT--G--G--A--TATC-GA--GTAC- 7561
TAL.CM.01.8023              ---GGGGAAT......................GA-ATC--AAGTGTGGAACCCTCTTC-C-A--C--A--TGC----C--GCA-CGG--GCACGG--C..............--G--GA--G-G--CAC----A-T--CT-T------A--G--TATC-GA--GTAT- 7111
MUS-1.CM.01.1085            --GAGAT.........................TCA---AAT-TGA--CC-TC--C-A--G--TT----GCC----C---TG--GCCTA-G--GGTACAA..............--G--G--GG-G--C--T---C-TTCCA-G---------GG-A-TGC-A------GC 7668
MUS-1.CM.01.CM1239          ---AGAC.........................ACAG--AA--TGA--CC-TCC-CGA--G---T-----GCC----C---TG---C--A-A--T--C..............--G-----GG-----A--C--GC-TTCCA-G--T------AG---TGC-----G--GC 7637
MUS-2.CM.01.CM1246          CGGCAAT.........................ACA--AAA--TGA--CCCCTC--CCA--G-C.........----CGC---CG-CTAGACTGAGAGAT..............-C---G---G-G--A--C--TC-TTCC--G-----G--AG-A-TGCGC--G---C 7766
MUS-2.CM.01.CM2500          -GGCGAC......................A-T-T--ATCGATTGA-CCC-TC--CCA--G--CTT---GC----C-C-TG------A-G-G---C..............-------G--GG----A--T---C-GTCC--G-----A--GG-A-TGC--------GC 7689
SYK.KE.x.KE51               -CCAAAT.......................--AC--T-GAATTTTG-ACCTACT-CTA--G--GAA--TGTG-------GC---T----TC--T---..............---T-GA-TCGG---A----TA-T--CT-T--G--T--A-GTCA--G-C-CTAT- 7394
SYK.KE.x.SYK173             CCCAGGA.......................--TAC-T-GAA-TTTA--CC-AC--C-A--G--GA---TGT-------GCA---TC-G-TC--T---..............---T-GA-TC-----A------A-T--CT-T--G------TG-TCA-CGTC-GTAT- 7698
RCM.GA.x.GAB1               .......................-TT-----CAGT----CAAA--TCA-CCTCTC--CG--AG--G-T--TT-----GC-GCA-----GGGG-G----..............--G-C-A--G----CAG-----A-T--CT-T------A--AG-AAT---A--GTACC 7430
RCM.NG.x.NG411              AG--AAT.........................ATGACA-CAAATGTCA-CTTGTCT-CT--AG-AG-A--T-TA---GC-GCA-----GGGC-G---C..............--------GG-G---ATG---A-T----TAT------A--A--T-TC-GA--GTAC- 7474
GRV.ET.x.GRI_677            C--CAAGTCT......................GG-CCA--AAATGTGA-CCT-A-TCCTC-GG-AC-CAG--TA---GCGTAC---C-GGGAG-C---..............---T-----G-G--AAC----A-T--CT-T--T--T--AG-T-T--GA---TATA 7744
SAB.SN.x.SAB1C              C-G-AAA.........................AA-AG--C-AATGTGA-GTT-TC-CCTC-G--TGAAAG--T----GCT-A------GGGAG-T--C..............---T-----GG----CA-G---A-T---T-T--T--A-----A-T--GA---TACA 8038
TAN.UG.x.TAN1               C-GCAAA.........................GAAAG--CTAATGTGA-ATTGAC--CTA--T--GA-A---TA---GCTTA------GGGA-G----..............--GC-CA--G----------A-C--TT-C--G--A--AG--AT--GA---TAT- 7842
VER.DE.x.AGM3               --G-AAG.........................AA-AG--CTAATGT-A-ATT-A-TCCCC-G--AGAAAC--TC---GC-GCA-----GGGC-GG--C..............---T-----G-----AC-----A-T--CT-C---------AG-A-T--GA--GTATA 7380
VER.KE.x.9063               C---AAC........................-ATAGA-C-AATGTGA-ATT-A-TCCCC-G--AGAA-GA-TC---GC-GCA-----GGGACG-----..............---T-G---G----CAC----A-C---T-T---------G--T-T--GAC-TTATA 7882
VER.KE.x.AGM155             ---CCAG.........................AA-AG--CAAATGT-A-ATTGA-TCCCC-G--AGAAAC--TC---GCGGCA-----GGGC-G---C..............---T-G---G-G---AC----A-T---T-T---------AG-A-TC-G-C--TACA 7862
VER.KE.x.TYO1_patent        --G-AAA.........................AA-AGA-C-AA-GTGA-ACT-A-TCCCC-G--AGAATCT-TC---GCGGC------GGGC-G---C..............---T----GG----CAC----A-T--CT-T---------AG-A-T--G-C-TTATA 7858
COL.CM.x.CGU1               AG-AGGA...........GAC-ATG-TGG-A-TGGGTGTGCT-ATCC-G-A-C-CC-AATTT--A-CAT-CATTAA-T-CTTTG-----GGGA-G---C..............---T-G--------GAGGA--ACTACTTAT-TC--A--AG-TAT---A---TC-- 7132
```

```
                              Env gp120 end Env gp41 start
H1B.FR.83.HXB2           TGGTGCAG...............AGAGAAAAAAGAGCAGTG......GGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAT 7886
Env                      V  V  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  Q  L  L  S  G  I  V  Q  Q  Q  N  N
H1A1.UG.85.U455          -----G--...............-----------------T.........---C--------A-C-------------A--------T-------A--------G--G----A----------------------------------------------A-------G---- 7316
H1B.US.90.WEAU160        --------...............----------------GGA...ATGC---------A-------------------------------------G-------------------------T--------------------A------------ 7893
H1C.ET.86.ETH2220        -----G--...............----------------CA.........---C------------------------------------------A----A-----------------------------------G---------A-----A-G---- 7262
H1D.CD.84.84ZR085        -----G-A...............---------------A-A.........---C--------A-------------------------G--------G--------------------------G--------------------A-----A------ 7407
H1F1.BE.93.VI850         --------...............----------------C-.........---C---------------------TC-T--G-ACAGCA-GGAGCA---------G--G--G--A----------------------------A---------A-----A------ 7163
H1G.SE.93.SE6165         -----G--...............-----------------T.........---C-G------G-C------------A-------------------G--G----A-------------T---------------------C--------A-----AGG---C 7304
H1H.CF.90.056            -----G--...............----------------...........--------G-------CT----------------------------G--G----A-----------------------G--------------------------A-G---- 7205
H1J.SE.93.SE7887         -----G--...............----------------...........-------------G---------------A----------T--------------G----A-------------T---------G--------C--------A-----A-G---- 7193
H1K.CM.96.MP535          --------...............-----------------A.........---T--------G-----T----A--------------------G--G----A--------------------------------------------A-----A-G---- 7046
H101_AE.TH.90.CM240      -----G--...............----------G-------.........------------A--A--T-----------A----------------G--G----A--------------------------------------------A-----A-G---- 7439
H102_AG.NG.x.IBNG        -----G--...............-------------G--T.........---C-G------G-C------------A-------------------GCGG----A-------------------------------C--C--------A--------G---- 7392
H103_AB.RU.97.KAL153_2   --------...............-----G-----------.........-------------G----------------------------------G--G----A--------------------------------------------A--------T--- 7063
H104_cpx.CY.94.CY032     --------...............---------T-------.........-------G--CA--------------------------------------G-------------------------------------C--C---------------------G---- 7256
H1N.CM.95.YBF30          CA---AGT...............----------------CCTTT......---C----T--GC---T--------TC-T----------G-----------G-----A----------------C-GAC------A-----G--------A--------T-T- 7434
H1O.BE.87.ANT70          -CA-AAGCACTAGAACTCAT...-----------------A.........---T-G---ATGC-A---T-G---G-TC-AA-T--------T----------------GG---CA------G--------A--CACACT--GC--AAG-----------A-------G----C 7940
H1O.CM.91.MVP5180        -AA-AA-CATTCACACCCCTCAC--G--------------A.........---T-G---ATGC-A---T-G---G-GC-AA-T--------T----------------GG---CAG----------G-A--CACAGTG--C--AAG-----------A-------G----C 7957
H1P.FR.06.RBF168         CAA-CAT-RCTCATGATGCTCAT--GA-------G-G--CA.........--CC-----ATGC-A------------A--A-T---------------------GG--G-T--------------------G---G-----CA---------------------------- 7407
CPZ.CM.01.SIVcpzCAM13    ATACAGTAGGA......GAAAGG---C-G--G-------CCTTT......--CC-G-----AC-------------TC----T-------------A---------GG--G-C-------------------C----G---C-A--G--A---------A--A---------C 7470
CPZ.CM.05.SIVcpzEK505    CA---ACT...............-------C---------CCTTT......---C----T--GC----T---------TC-T---------G-------------G-----A----------------C-GA-------A-----G----------------T---- 7369
CPZ.CM.05.SIVcpzLB7      CT---G-T...............-----------------CA.........--GC-G----TG--T---------------------------G-------------A---T-A----------------------------A-----C-------A-------GT---- 7397
CPZ.CM.05.SIVcpzMB66     C------T...............--G--------------CCGTA......---T-G---------------------------------------------------G---C-------------------C-GC------------------------G---- 7510
CPZ.CM.05.SIVcpzMT145    CA--AG-AAAACAACACCAT...--------G-----CC-T.........---C-------AC----T---------C---------------------------A---G-CGT------------A---C-G-----G--AA-G--C--------------T---- 7395
CPZ.CM.-.SIVcpzMB897     CA-----A...............-----------G--G-----A.........---C-G----TGA-----------A---A------------------------GG---C--------------------A-TC--C-C---------------A-----A-G---- 7394
CPZ.CM.-.SIVcpzDP943     CA---ACA...............-----------------CCTTT......---C----T--GC----T---------TC-A------------------------A----C-A------------------C-G--------------A--G--------A-----------C 7449
CPZ.CM.98.CAM3           CA--ATCA...............-----------------CCTTT......---C-G--T--GC----T-------------A---------------------A----G--GT------------------C-G--------------A--G--------A-----------C 7299
CPZ.CM.98.CAM5           C----GCA...............------G----------CCTTT......---C--------AC--------C-----TC-A------------------------A----G--GT------------------C-G--C--G------A--G--------A-----------C 7605
CPZ.GA.88.GAB1           CA---GCAAGACAG...AAAGAC---C-G----------CCTTC......--TC-G-----AC-------------C----T-----------------------GG--G-A--------------------------C----G---GC----G--------------A------T--- 7923
CPZ.GA.88.GAB2           CA---GCAAAAGCA...CTTGAT---TCC----------CCTTT......---C-G--G--GC----T----A---C-T---------------------------A--TG----T-------------A--TC--A-CC-GC-T--G--A-------------A---T--C 7270
CPZ.US.85.CPZUS          CA---A-A...............-----------------CCTTC......---C----T--AC------G-----TC-T----------------------------A---G-AGT--------------------C--------------A--G--------A-------T--- 7891
CPZ.CD.90.ANT            AAA-AA-ACAACACTCC.......--C---------GCA-T.........-------GCTG--C---T-G--TC-TC-CA-T-----T--C--T--A-----------G-----AG--------C----A----GA-T--GY-CCA-------T--A--A-------GC---- 7318
CPZ.TZ.00.TAN1           AAA-TA--GCCAATCAT...ACT--GTC--G------AT......GTG--C-------CTG-----T------A--TC-TA-T----------------T--A------G-----AG--------------C-------------GG----C-C--------T--A--------AC-A--C 7512
CPZ.TZ.2001.TAN2         AAA-TA--GCTAACCACACT...---TC--G---G-AT--T.........--C-------CTG-----T------A--TC-CA-T----------------T--A------G-----AG--------------C----------GGG---C-C-----A---T--A--------AC-A--C 7532
CPZ.TZ.2002.TAN3_1       AAA-TA--GCTAATCACACT...---TC--G---G-AT--T.........--C-------CTG-----T------A--TC-CA-T----------------T--A------G---C-AG--------------C----------GGG---C-C-----------T--A--------AC-A--C 7586
GOR.CM.2007.SIVgor2139_287 -TA-TGGCCATAAC.........--G-------C-T----CA.........--CC-G---ATGC----T-----------A--A-T------------G--------GG--G-A--------------------------G---G-C---CA-------C-----------------A-----C 7444
GOR.CM.2007.SIVgorCP2135con -TA-TAGCCATAAC.........--G-------C-T----CA.........--CC-G---ATGC----T-----------A--A-T------------G--------GG--G-A--------------------------G---G-C---CA-------C------------------AM----C 7428
GOR.CM.2004.SIVgorCP684con -T--TAGCCATAAC.........--G-------C-T----CA.........--CC-G---ATGC----T-----------A--A-T------------G--------GG----A--------------------------G---G-C---CA-------C-------A------A-----C 7375
H2A.DE.x.BEN             CCTCAACT.........CCAGTG--GA-C-------GT---...TTC-TGC----G..........-T-G--T--TC-C-CGA-------TTCTG-A---------GCGG--CC--------T-A-CC---T--C-GACT---C--G-C--G-----------A---C-AC-G 8359
H2A.GW.x.ALI             CCTCTACT.........CCAAGG--GA-T-------GT---...TTC-TGC----G..........-T-A--T--TC-C-CGA-------TTCTG-A------A-G--AG-TT-A--------T-T-CT---T-TC-GACT-------G-C--G-----------A---C-AC-G 8343
H2A.SN.x.ST              CCTCTGCT.........CCAGTG--GA-T-------GT--A...TTC-TGC----G..........-T-A--T--TC-C-CACGA-------GCTG-A---------G--G--CT----------T----CT---T-TC-GACT-------G-C--G-----------A---C-AC-G 7789
H2B.CI.x.EHO             CCTCAGT-.........ACACCG--GA-T-------...GTGTACTT-TGC----G.........TT-G--A---C-T-CGA-G-----TTCTG-A---------G--G--CT------T--T----CT---T--C-GACT---C--G---G-----------------C-AC-G 8321
H2B.GH.86.D205           CCTCAGTA.........AAACCG--GA-T-------GT--A...ATG-TGC--------.........-T-G--A---C-T-C-ATG-----TTCTG-A-----------A-G--TT---------T-A-CT---T--C-GACT---C--G-C--G-----------------C-AC-G 8335
H2G.CI.92.ABT96          CCTCTGT-.........ACACCG-A-A-T-------GT--A...TTT-TGC----G.........TT-G--A--TC-T-C-ATG-----TTCTG-A------A----G--GT---------T----CT---T--C-GACT--G---G---G-----------A---C-AC-G 7720
H2U.FR.96.12034          CCTCAGT-.........ACACCG--GA-T-----G-G---A...TTT-TGC-T--G..........-T-G--T--TC-C-CCA-G-----TTCTG-A-----T--G--G--GT---------T-A-CT---T--C-GACT---C--G---G-----------A---C-AC-G 7843
SMM.SL.92.SL92B          CCACAAGT.........ACCTCA-AGA-T-------G---C...TTT-TGC----G..........-T-G--T--TC-C-CGA-G-----TTCTG-A-----------G--GC-A--------T-CT---T--C-GACT-------G---G-----------A--A---C-AC-G 7752
SMM.US.x.H9              CCACAACTGGT......GCCTCA---A-T-------GG--C...TTT-TGC----G..........-T-G--T--TC-C-CGA-------TTCTG-A---------G--G--GT---------T---CT---T--C-GACT--G---G---G-----------A---C-AC-G 7793
SMM.US.x.PGM53           CAACTGGTGCC.........TCA---A-T--G----GG--C...TTT-TGC----G..........-T-G--T--TC-C-CGA-------TTCTG-A---------G--G--CG----T-G-T-C-CT---T--C-GACT--G---G---G-----------A---C-AC-G 8250
SMM.US.x.SIVsmH635F_L3   CCACAACTGGT......GCCTCA---A-T--G----GG--C...TTT-TGC----G..........-T-G--T--TC-C-CGA-------TTCTG-A---------G--G--GC---------T---CT---T--C-GACT--G---G---G-----------A---C-AC-G 8326

                              Env gp120 end Env gp41 start
MAC.US.x.239             CTACTGGTGGC......ACCTCA---A-T-------GG--C...TTT-TGC----G..........-T-G--T--TC-C-C-A-G-----TTCTG-A---------G--G--GT----------C-CT---T--C--ACT------G----G-----------A---C-AC-G 8304
Env                      T  T  G  G  .  .  .  T  S  R  N  K  R  G  V  .  F  V  L  G  .  .  .  F  L  G  F  L  A  T  A  G  S  A  M  G  A  A  S  L  T  L  T  A  Q  S  R  T  L  L  A  G  I  V  Q  Q  Q  Q  Q
MAC.US.x.251_1A11        CTACTGGTGGC......ACCTCA---A-T-------GG--C...TTT-TGC----G..........-T-G--T--TC-C-C-A-G-----TTCTG-A---------G--G--GT----------C-CT---T--C-GACT------G----G-----------A---C-ACTG 8300
MAC.US.x.251_BK28        CTACTGGTGGC......ACCTCA---A-T-------GG--C...TTT-TGC----G..........-T-G--T--TC-C-C-A-G-----TTCTG-A---------G--G--GT----------C-CT---T--C-GACT------G----G-----------A---C-AC-G 8286
MAC.US.x.EMBL_3          CTACTGGTGGC......ACCTCA---A-T-------GG--C...TTT-TGC----G..........-T-G--T--TC-C-C-A-G-----TTCTG-A---------G--G--GT-C-G------C-CT---T--C-GACT------G----G-----------A---C-AC-G 7435
STM.US.x.STM             CAACTAGT.........ACCTCA---ACT--G----GG--C...TTC-TGC----G..........-T-G--T--TC-C-CGA-------TTCTG-A------------G--GC---------------CT---T--C-GACT------A----G-----------A---C-AC-G 7955
MNE.US.x.MNE027          CTACTGGTGGC......ACCCCA---A-T-------GG--C...TTT-TGC----G..........-T-A--T--TC-C-C-A-G-----TTCTG-A---------G--G--GT----------C-CT---T--C-GACT------G----G-----------A---C-AC-G 7781
DEB.CM.99.CM40           A--GA-C-GAG......TCGGTA---C-C--G--G----CAGGAATT-CTT-T--T......C-GG-A-CA--TC-CA-TA-------TGCTG---------------G----CAG--------------T-TC-CTCT--GC--AG---G-----------A---C-AG-G 7626
DEB.CM.99.CM5            A--GA-C-GAG......TCTGCA--------G--G-C-GGAATT-CTT-T--T......C-AG-G-CT--TC--A-TA-------TGCTG---------------G----CAG--------------T--C-CTCT--GC--AG---G-----------A---C-AG-G 7547
DEN.CD.x.CD1             AA-GA-CAGAATCGGCAGAGTCTGTTTC-CGCC--AGGCG-GAAGTGTCTT---TGCT-GG-C--A-A--A--TC-CA-C----------CTGTC---------G-T-GG--CAGGCT-AG--------T--C-GG-T---C-TG-C--G-----------A---C-AG-G 7943
LST.CD.88.447            CACCC............AGTAGT-AGC-G--G--------CACCGCTG-CTC----T......GC-T-G--AC-TC-CA-T--T-----T-CTG-A------TT--TA--G-C--TA--A--T--C--A---CA-GT-G--A-ACAA--G---T--------------GC-G 7066
LST.CD.88.485            CACCC............AGTAGT-AGC-G--G--------CACCGCTG-CTC----T......GC-T-G--AC-TC-CA-T--T-----T-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GT-G--A-ACAA--G---T--------------GC-G 7063
LST.CD.88.524            CACCC............AGGGGA--GC---C---------CTCCACTA-CTC----T......GC-T-G--AC-TC-CA-T--T-----T-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GT-G--A-ACAA--G---T--------------GC-G 7069
LST.KE.x.lho7            CACCC............AGGGGA--GC---CC--------C-CCTCTG-CTT------------GC-T-G--AC-TC-CA-T--T-----C-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GC-G-C---CAA--G---T--------------GC-G 8141
SUN.GA.98.L14            CACCC...............AGG-AG------G--G----CACCTGTT-CCC-----------GCTT-G-CAC-TC-CA-T--T--G--T-CTG-A------TT--TA--G-C--TA--A--T--C------CA-GC-G-C---CAA--G---T--------------GC-G 8209
MND-1.GA.x.MNDGB1        CTT-A-CTGAA......CATCAT-A---G--G----G--CA...GTAAT-C-T--T......A-------TC-GC-CTCGCTG-------TC-G-G------T---TG--GG--G-A-----T--C--AT-TCAGTCT--GG--A-----G-------G-A--A--A--AC-G 7493
OLC.CI.97.97CI12         ................CAGCCT---A-T-------G---...GCATTG-TGC-G--G........--AT-G---G-TC-C--CCT---------TC-G-A------A---TA----C--TA-----T--C---T-TTCG--T--GC-TG-G--G-------G------------GC-C 7513
WRC.CI.97.97CI14         -ACCC-CTGAA......CATCGC------G-------GT...GCCATTATTC------------A-T--A--A---C-C--GCTG-------TCTG-A------A---TA----CAGTA-------A-C---T-TCAG--T--GC-TG----G---------------------A--G--C 8017
WRC.CI.98.98CI04         CACCAA-AGAA......CATCAT--------------G---...GCTATT-TCT-----------A-T--A--A---C-C--GCTG-------TCTG-A------A---TA----CAGTA-------A-C---T-TCAG--T--GC-TG----G---T-------------A--G--C 8062
WRC.GM.05.Pbt_05GM_X02   CACCT-CTGAC......CATCAT---------C---G---...GCAATT-TTT-----------C-T--A--C---C-C--CCTG-------TCTG-A------A-G-TA----CAGTA-------A-C---T-TCAG--T--GC-TG----G---------------------------G--C 6958
MND-2.CM.98.CM16         ATC---CAGAA.........GCT--GC-G--G--G-G--C-...GTCTTGC-T--G......A-GT-C--CC--C-A--TTTG-----TTC---G-------T---TGG-GG--G-C------T--C---T-TCAGGCT--G---AA---G--T---G---------------GGT- 8036
MND-2.GA.x.M14           ATC-A-CAGAT.........GCA-AGC-G--G----G--CA...GTCTTGC-T--------.A-G--C--CC---------CTTG-----TTC-G-G------T---TGG-GG--G-C------T--C---T-CAGGCT--G---AA---G--T---G---------------G-T- 7965
MND-2.x.x.5440           ATT---CTAAT.........ACA--GC--------G--C-...GTCTT-C-T--------.A-G--C--CC--C-A--TTTG-----TTC-G-G------T---TGG-GG--G-C------T--C---T-CAGACT---C--AA---G--T---G---------------GGT- 7669
DRL.x.x.FAO              ACC---C-GGA.........CAC--GC-G--G----GT-CA...GTC-TGC-T--G......A-------CC------A-CTTG-----TTC-G-A---T--T---TGG-G-C-G-CA----T--C---T-CAGGCT--GC---A----G--G----G--------A--GC-A 7727
MON.CM.99.L1             CT-G-GTA.............CAC--G------G-----......ATTACTT-G--TATGGCC--T--G--C---C-CA-CA-G------G----A--------------GG-G-CAG-C------------T-TC-GTC-----C---G----G-----------------G-G--C 7849
MON.NG.x.NG1             AT-G--TA................TGAG-G-----CG--CT...ATCAC-T-G---ATGGCC--T--G--C--TC-CA-CA-G------G----G--------------GG-G-CAG-C------------T--C-GTCT--G---G----G-----------------G-G--C 6474
GSN.CM.99.CN166          AACAC-CC.............GCTGTG--G--------C.....CTTTC-T-G---ATA-CC---T-GAC--TC-CA-T--G-----T-C---G--------------GG---CAG-----------A----T--C-GTCT--G---G---G---A--------------G-G--C 7823
GSN.CM.99.CN71           CACAT-CAGGGGTG......CATGTGA----G--G--C.....ATTTC-T-G--CATA-CC---T-GAC--TC-CA-T--G-----TGC---G--------------GG---CAG--------------AT--C-GTC---G---G---G------------------G-G--C 7835
TAL.CM.00.266            AT-G--CC.........GCTAAC-AGC-----C----G-CCCCGCTA-CCT----C.........--A-A--A--TC-CA-CA-------C-CTG-C-----------TGG-G-CAG-C---------G---T-TC-CTCT--GC--AG---G------------------G-AC-C 7716
TAL.CM.01.8023           AT-G--C-.........GCAACA-A---G-------C-C-CCTCTG-CTT-G---------.--A-A--A--TC-CA-CA-------C-CTG-C-----------TGG-G-CAG-C---------G---T-TC-TTCT--GC--AGC--G------------------G-GC-C 7266
MUS-1.CM.01.1085         CT--C-CA.............CATGTGC-TGC------C......ATCTC-T-G--CATAACC--TT-A--A--TC-CA-C----------G----A--------------GG-G-CAG-C------------T-TC-GTCT--G---G-C--G------------------GCA--- 7820
MUS-1.CM.01.CM1239        CT--T-CC.............CATGT-A--GC---------...ATATCTC----CATAACC--T--G--A--TC-CA-C--G------G----G--------------GG-G-CAG-C------------T-TC-GTC----C--G-C--G------------------GCA--- 7789
MUS-2.CM.01.CM1246        CT----CC.............CACGTCA-GGCT--G--------...ATCA-CT-G---ATAACC---T-A--A---C-CA-T--T--G---G----A--------------GG-G-CAG-C------------T-TC-GTCT--G---G---G-----------A---GCT--- 7918
MUS-2.CM.01.CM2500        CT----C-.............CATGTCC--GCT--G--------...ATAACCC-----ATAACC--TT-G--A--TC-CA-T--T-----G----A--------------GG-G-CAG-------------T--C-GTCT--G---G---G-----------A---GCT--- 7841
SYK.KE.x.KE51            A-C---CT.............ACG-A-C-G---------CACCATTG-C-C-T--G.............G--A--AC-CTCG-----G--T-CTG-A-----GAGC--TG-G-CAG----T--AC-C---T--CAGACT--G---G---G-------------------C-A--G 7546
SYK.KE.x.SYK173          A-T---CT.........AATACC--G--------G-----CACCATTG-C-C-T--G.............G--A--AC-CTCG--T--G--C-CTG-A-----G-GC--TG-G-CAG----A--AC-C---T-TCAGACT--G---G---G-------------------C-A--G 7853
RCM.GA.x.GAB1            A-AC-...................-A-C-G-----G-T-CCT...TTG-TGC-G--T.........-T--A--T---C-CTC----------T-CTG-A------------GG-G-CAG-C-----T--C---T--C-G--T--GC-TG-A--G---T--------------A--G--C 7573
RCM.NG.x.NG411           A-ACA....................-A-C----G-----T-CCT...TTG-TGC-G--T.........-T--A--T---C-CTCG---------T-CTG-A------------GG-G-CAG-C-----A--C---T--C-G--T--GC-TG----G---T----------------G--G 7617
GRV.ET.x.GRI_677         CT-GC-CCACA..............-----------G-TGCCATTC...-TGC----G.........-T--A--C-------------T--T----CTG-A------------GG---CA---------A--C---T-TC-G--T--GC-TG----G---T----------------G--C 7893
SAB.SN.x.SAB1C           CT-GC.............CCCGAA---C--------G-TG...CCATTC-TGC----G..............A--T---------------T---GCTG-A-----T-----GG-G-CAG-C--------C---T-TCAG--G---C-CG----G---C--------------A--- 8187
TAN.UG.x.TAN1            -T-G--CAACT..............C-------G--G-TGCCCTTC...-TGT-G--G.............A--C--T-----G-----T---GCTG-A-----T---A-AG-G-CAG--T-------C---T--CAG------C-TG-A--G---T--------------G--- 7991
VER.DE.x.AGM3            C--GAGGT.........CATGAC---AC---GC---TC...CCGTTC-TGC----G.............A--C-----A-----T--T--G-CTG-A-----A-----GG-G-CAG-C--------C---T-TCAG--T---C-TG----G---C--------------G--- 7532
VER.KE.x.9063            C--GAGGA.........CAAGAC---TC----C---TC...CCGTTT-TGC----G.............A--C-----A-----T--T-----CTG-A-----A-----GG-G-CAG-C--------C---T-TCAG--T---C-TG----G---A--------------A--- 8034
VER.KE.x.AGM155          C--GAGGC.........CAAGAG--GC-----C---TC...CCGTTC-TGC----G.............A--C-----------T--T--G-CTG-A-----A-----GG-G-CAG-C--------C---T-TCAG--T---C-TG----G---T--------------G--- 8014
VER.KE.x.TYO1_patent     C--GAGGA.........CATGAG--GC----G----TC...CCGTTT-TGC----G..........-T-A--C---------G--T--T---CTG-A-----A-----GG-G-CAG-C--------C---T-TCAG--T--GC-TG----G---C--------------G--- 8010
COL.CM.x.CGU1            --AATGTAAATTGGCATCATGGG---C----G---.........GGCATCT-T-C-TTC-CCA--T-A-CCC-----A-T-GG--C--TGCTG-A------T----G--TG--G--------A-C-----TCA-TCCC-TAATGGAA-GGCTAGTGCCAGCAGC---CGA 7293
```

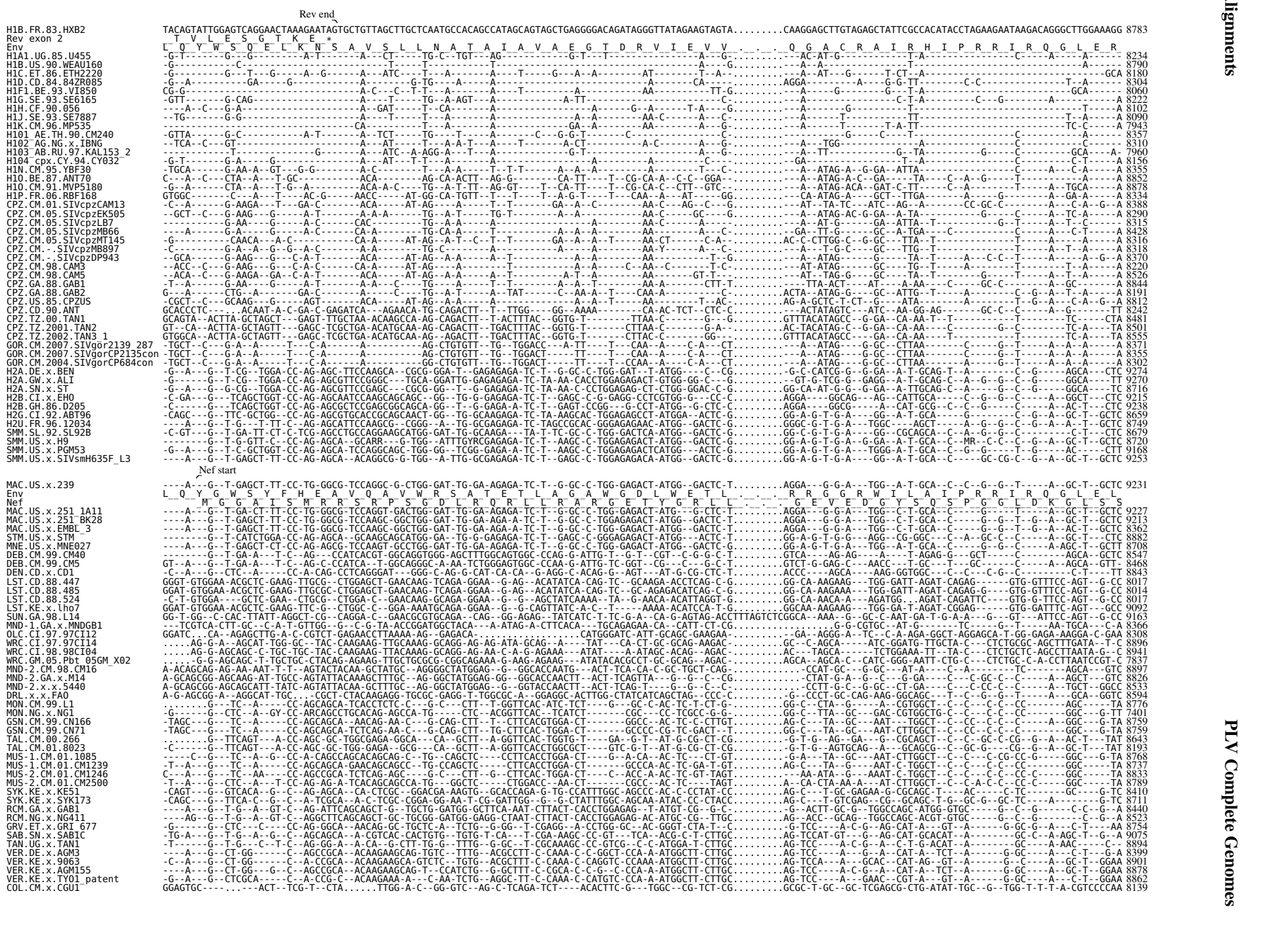
```
H1B.FR.83.HXB2               TTGCTG...AGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTG 8053
Env                           _L__L__._R__A__I__E__A__Q__Q__H__L__L__Q__L__T__V__W__G__I__K__Q__L__Q__A__R__I__L__A__V__E__R__Y__L__K__D__Q__Q__L__L__G__I__W__G__C__S__G__K__L__I__C__T__T__A__V__P__W
H1A1.UG.85.U455              C-----...--------A-----T-------------A-------T----------T--A--------------G-----------------------C---------------A--A--------C-----------G-----C---------A-------C-- 7483
H1B.US.90.WEAU160            C-----...-----------------------------CG--------------------A---------------G-------------------------------G---------------------------------------C---------A---------- 8060
H1C.ET.86.ETH2220            ------...-AA-----A---------------A-------G-----G-----------T-----------A-----G-------A-A------C------G----------------A---------------C-------------------C------------------ 7429
H1D.CD.84.84ZR085            ------...--------A-------------------------G-----------T--A-----------------------A-----------G-----------------A--A--------------------A-----------A----A--C-- 7574
H1F1.BE.93.VI850             ------...-----------A------------------G---------------T--A-----------------G-------------------------------------A-----C-----C-------------------C---------AA------C-- 7330
H1G.SE.93.SE6165             ------...--------A-------G-----------------------------T--A-----------------C----------A--------------------------A---------------C-------------------C---------AA------C-- 7471
H1H.CF.90.056                ------...--A-----AC------G-------A-------G-----G----------T--A-----------------G----C----------------------G---------------A---------------C-------------------C---------AA--------- 7372
H1J.SE.93.SE7887             ------...-A------A-------------------A-------T----------T----------------------G------------------------------------A-----A---------------C-------------------C---------AA------C-- 7360
H1K.CM.96.MP535              ------...--A-----A--A---------A---------A--G-----G----------T--A-------G------------------A-------------------------A---------------C-------------------C---------AA------C-- 7213
H101_AE.TH.90.CM240          ------...--------A-------G---------------------------------T--A-----------------G--------------------------------A--T---A--AC--------C-------------A-----C----------------C-- 7606
H102_AG.NG.x.IBNG            ------...-AA-----A-----T-----A---------A-------G-----------T--A-----------------G--------C---------------GA----------A--A--------C-------------------C---------A----A--C-- 7559
H103_AB.RU.97.KAL153_2       C-----...--------------------A-----------------------------------------G---G------------G---------------------------------------------------------------------------------- 7230
H104_cpx.CY.94.CY032         ------...--------A-----T-----A---------AG-------G-----------T--A-----------------G---------C-------C-----------------------A--A--------C-------------------C---------AA---A------ 7423
H1N.CM.95.YBF30              C--T--...--A--A--A-----------A---T-----------T--A-----------T--A-------------A-G-----T---A-A-----------T-G------G--AA----AA-TC-A-----C-----A------ACA--A---TAT--CA---------- 7601
H1O.BE.87.ANT70              C----A...--A--A--AC----C--G-----AT--C--AGG--AT-TR-A-----T----GA--A----GA--TC-CC-G--A--CT-A----CC-TA---C--A-----G--A-----AA-CC-A-----C--TAAA-----G--AG-C---TA---AT-A--AAAA-- 8107
H1O.CM.91.MVP5180            C-----...--A--G--AC----C--G--A--CT--C--AGGT-AT-T--A-----T--T-GA--A----GA--TC-CC-G-AA--CT-A----CCCTTA--C--A-----G--A-G---AAACC-A-----C--TAAA---------A--C--TTA---AT-A--AAAA-- 8124
H1P.FR.06.RBF168             A----A...C-A-----A-----A--G--AG-AT--C--AG-----T-T--G----------A-GA-------GA--TC-CC-G--T--CA-T----CC---T---G------G--A------A--AC-G-----A------A---C--A-AG-A--TTAT----AA-------A-- 7574
CPZ.CM.01.SIVcpzCAM13        C--T--...--A--A--A-----C------------A---------A-----------AG-A--A--A-----A--C---T-A--T-----A--------------C-------G---A-T--A--CC-------C-------------GTCT--C--TTAT--CA---------A-- 7637
CPZ.CM.05.SIVcpzEK505        C-----...--A--A--A------------A--CT--C--------T-----T------------T--A--------A------G-----T--A-A-------------G-G---C--G--AA----A--CC-A-----C------A--------TCTG-G--TTA---CAA--------- 7536
CPZ.CM.05.SIVcpzLB7          ------...--A-----A--------G--A--CT--C--------T-------------------A-----------A------G-----T--A-C--------------------------A----A--A--A-----------A--------------C------------T-------A-- 7564
CPZ.CM.05.SIVcpzMB66         C-----...--A--C--A-------------------------T-------------------T--A-----T------C-GG-T--T----------G-------------------T-A--AC-A-----C--------------------------------T--------- 7677
CPZ.CM.05.SIVcpzMT145        C----A...--A--A--A--A--T-----A-----------G--TT-T---------A-----A----------G---G-T--T------------G--------G--------G--A-----A--TC-A-----C--TA-A------ACT-----TC-----------A-GC-- 7562
CPZ.CM.-.SIVcpzMB897         ------...-AA-----A--------------------------T-------------------T--A------T-----GC--G-T--T-----------G--T-------------------A-----A--AC-A-----C-------------------------T---AA--------- 7561
CPZ.CM.-.SIVcpzDP943         C-----...--A--A--A-----------------T------------T----A---------T--A----------A------G-----T----------------T--A------G--AA----A--TC-G-----C------A------TC----C--TTAT---A---------- 7616
CPZ.CM.98.CAM3               C-----...C-A--A--A-----------------T------------T----T---------T--A----------A------G-T--T-----A-----------C-GA-----G--AA----A--CC-A-----C------A------GC------TTAT---A---------- 7466
CPZ.CM.98.CAM5               C-----...C-A--A--A-----------A---T------------T----T-------------A----------A------G-T--T-----A--------------------G---A----AA-TC-A-----C------A------GC----C--TTAT---CA-G-------- 7772
CPZ.GA.88.GAB1               C--T--...-AA--A--A-----T-------C--A---------AT--A---------AG-A--A--A--A--A--C---T-G--T-----A------G-----GC--------G---A-T-----CC-G-----C--------------GGCTG-C--TTAT--CA-G-------- 8090
CPZ.GA.88.GAB2               ------...C-A--C--A--A--T--G--A--C---C----G--AT----------------A--A------------G--GG----T---------------T------------G--AG-----A-CAC-A-----C--------------ACTG-G---TA-T--A----A--C-- 7437
CPZ.US.85.CPZUS              C----A...--A--A--A-----C--------T-A---------AT----T---------T--A-----T--A--C---G-A--T------------------T----------G---A-----A--TC-G-----C-----A-------ACA------TTAT--CA---------- 8058
CPZ.CD.90.ANT                C-----...CAA--C--A---A-A--G--A------C-A--G---T-G----------AG-A--A--A-------------------G--T--A--C--G-AG-------GA---------A--AT--A-CC-C--------TG---AC--GG-G-CC--TCA----A-G-------- 7485
CPZ.TZ.00.TAN1               C----T...CA---C--A--A-----------A--CT--------G---T-T--A-----------T----------------C-----G--T--A--A--G-A-----A---GA--C-------------AA-CC-C-----A--TG--AAC---T-GG-G--TCA--G-AG------A-- 7679
CPZ.TZ.2001.TAN2             C-----...CA---C--A--------------A-------------G---T-T--G---------------A---------------C-----G--T--A--A--G-AR---A---GA----------TA---AA-C---------A--TG-CAAC---T-GG-A--TCA--G-AG------A-- 7699
CPZ.TZ.2002.TAN3_1           C-----...CA---C--A--------------A-------------G---T-T--G---------------A---------A--C-----G--T--A--A--G-A----A---GA----------TA-T-AA-C---------A--TG-CAAC---T-GG-G---CA--G-AG------A-- 7753
GOR.CM.2007.SIVgor2139_287   A-----...--A-----A-CA--T--G--AG-AT--C--AG-----T-T--T----------A-GA-------GA--TC-CC-G--A--CA-A-----CC-------GA--C--G---------A--AC-G-----A---A-A--G-G-T-A-----TTA----AA-------A-- 7611
GOR.CM.2007.SIVgorCP2135con  A-----...--A-----A-CA--T--G--AG-AT--C--AG-----T-T--T----------A-GA-------GA--TC-CC-G--A--CA-A-----CT-------GA--C--G---------A--AC-G-----A---A-A--G-G-T-A-----TTA----AA-------A-- 7595
GOR.CM.2004.SIVgorCP684con   A-----...--A-----A-CA--T--G--AG-AT--C--AG-----T-T--T----------A-GA-------GA--TC-CC-G--A--CA-A-----CT-------GA--C--G---------A--AC-G-----A---A-A--G---T-G-----TTA----AAC------A-- 7542
H2A.DE.x.BEN                 C--T--...GAC-TAG-CA--AGA-----AG-AA------G---G--C---------A-CG--AA-C------------G--ACT---A-C--G-AG----------C----GGCA-AG--AAATTCA-----A--TG-GTTT-G--AGG-C---CA----A----A-G-- 8526
H2A.GW.x.ALI                 C--T--...GAC-TGG-CA--AGA------G-AA------G---G--C---------A-CG--AA-T------------G--ACT---A-C--G-A----T--------C--GGC--GG--AAATTCA-----A--TG-ATTT-G--AAG-C---CA----A----A-A-- 8510
H2A.SN.x.ST                  C--T--...GAC-TGG-CA--AGA-----AG-AA------G---G--C---------A-CA--AA-T------------G--ACT---A-C--G-A----T--------C--GGC--AA--AAATTCA-----A--TG-GTTT-G--AAG-C---CA----A----A-A-- 7956
H2B.CI.x.EHO                 C-CG--...GAC-TGG-CA--AAGA----AG-A-------GG--G--C---------G-CG--AA-C------------G--ACT--CA-C--G-A---T--C--A--C--GGCA-AA--AAATTCA-----A--TG--TTC-G--AGG-C--TCA---GA----A-A-- 8488
H2B.GH.86.D205               CCTG--...GAC-TGG-CA--AAGA-----G-A-------GG--G--C---------G-CG--AA-C------------G--ACT--CA-C--G-A------------------GGCA-AA--AAATTCA-----A--TG--TTC-G--AGG-C--TCA---GA----A-A-- 8502
H2G.CI.92.ABT96              C--T--...GAT-TGG-CA--AGA------G-AA------G---G--C---------G-CG--AA-C------A-T---G--ACT--CA-C--G-A---T---------C--GGCA-GG--AAATTCA-----A--TG-GTTT-G--AGG-C---CA----A----A------ 7887
H2U.FR.96.12034              C--T--...GAC-TGG-CA--AAGA----AG-AA------GG--G--C---------A-CT--AA-C------A-T---G--ACT--CA-C--G-A------------------GGCAAGT--AAATGCA-----A--TG-CTTT-G--AGG-C---CAT---A--------A-- 8010
SMM.SL.92.SL92B              C--T--...GAC-TGG-CA--AGA-----AG-AT------G---G--C---------A-CT--AA-T--T---A-T---G--ACC--CA-C--G-A------------------GGCA-AG--AAATTCA-----A--TG-ATTT-GG-AAG-C---CA----A----A-TA-- 7919
SMM.US.x.H9                  C--T--...GAT-TGG-CA--AGA-----AG-AT------G---G--C---------A-CT---A-C------A-T---G--ACT---A-C--G-AG-----------------GGC--CRN-AAATTCA-----A--TG--TTT-GG-AGG-C---CA----A-K--A-A-- 7960
SMM.US.x.PGM53               C--T--...GAT-TGG-CA--AAGA----AG-AT------G-T-G--C---------A-CT---A-C------A-T---G--ACT---A-C--G-AG----G-------------GGC--AG--AAATTCA-----A--TG--TTT-GG-AGG-C---CA----A----A-A-- 8417
SMM.US.x.SIVsmH635F_L3       C--T--...GAT-TGG-CA--AGA-----TG-AT------G---G--C---------A-CT---A-C------A-T---G--ACT---A-C--G-AG------G-----------GGC--AG--AAATTCA-----A--TG--TTT-GG-AAG-C--TCA----A----A-A-- 8493

MAC.US.x.239                 C--T--...GAC-TGG-CA--AGA-----AG-AT------G---G--C---------A-CA---A-C------A-T--GG--ACT--CA-C--G-AG---T--------C--GGC--AG---AATGC------A--TG-GTTT-G--AAG-C---CA----A----A-A-- 8471
Env                           _L__L__._D__V__V__K__R__Q__Q__E__L__L__R__L__T__V__W__G__T__K__N__L__Q__T__R__V__T__A__I__E__K__Y__L__K__D__Q__A__Q__L__N__A__W__G__C__A__F__R__Q__V__C__H__T__T__V__P__W
MAC.US.x.251_1A11            C--T--...GAC-TGG-CA--AGA-----AG-AT------G---G--C---------A-CA---A-C------A-T--GG--ACT--CA-C--G-A----T--------C--GGC--AG---AATGC------A--TG-GTTT-G--AAG-C---CA----A----A-A-- 8467
MAC.US.x.251_BK28            C--T--...GAC-TGG-CA--AGA-----AG-AT------G---G--C---------A-CA---A-C------A-T--GG--ACT--CA-C--G-AG---T--------C--GGC--AG--AAATGC------A--TG-GTTT-G--AAG-C---CA----A----A-A-- 8453
MAC.US.x.EMBL_3              C--T--...G-C-TGG-CA--AGA-----AG-AT------G-T-G--C---------A-CA---A-C------A-T--GG--ACT--CA-C--G-AG---T--G-----C--GGC--AG---AATGC------A--TG-GTTT-G--AAG-C---CA----A----A-A-- 7602
STM.US.x.STM                 C--T--...GAC-TGG-CA--AGA-----AG-AT------G---G--C---------A-C----A-C------A-T---G--ACT---A-C--G-A------------------GGCA-AG--AAATTC------A--TG-GTTT-G--AGG-C---CA---CA----A-A-- 8122
MNE.US.x.MNE027              C--T--...GAC-TGG-CA--AGA-----AG-AT------G---G--C---------A-CA---A-C------A-T---G--ACT--CA-C--G-AG---T--------C--GGC--AG--AAATGC------A--TG-ATTT-G--AAG-C---CAT---A----A-A-- 7948
DEB.CM.99.CM40               C--T--...-A----G-G-----T--CGG---G-----AAC-T-A----C------AG-G-GAA-C---A-CA-CC-CC-GACT---A-A--G-AG---T----A--C---GCAAAAT-AAATGAA-----C---G-CTTC----AG--C--TCAT--CA-G--C------ 7793
DEB.CM.99.CM5                C--T--...-A----G-G-----T--CGG-----------TC---A----CA-----AG-G-GAA-C---A-CA-CC-CC-GACT---A-A----AG---------A-----GAGCAAGT-AAATGAA-----A--TG-CTTT----AA--C--TCA----A-G--C------ 7714
DEN.CD.x.CD1                 C--T--...--A---G----A-GT--CAGTGCA--AC----GT-GT----A--------T--AA-C---A-C--TC-CC-GACT--AA-A--G-A---T--T--A--C----GCA--G--AAATGAA-----A--TG-CT-G--G-AA-----TCA---CT-A--AGAA-- 8110
LST.CD.88.447                ------...GT-CT-G-----AAA--G--AG-G--C--AAG-------CA-------AG-G---A-T-----A---C-CC-GACA---A-T--GGAG-------------C---ACT--G--T-CTTCC-----G--TCAAT-G----AAG-A--TCAT---AA---AGAA-- 7233
LST.CD.88.485                ------...GT-CT-G-----AAA--G--AG-G--C--AAG-------CA-A-----AG-G---A-T-----A---C-CC-GACA---A-T--GGAG-------------C---GCT--G--T-CTTCC-----G--TCAAT-G----AA--A-TCAT---AA-----GAA-- 7230
LST.CD.88.524                ------...GT-CT-G-----AAA--G--AG-G---C-AAG-------CA-------AG-G---A-T-----A---C-CC-GACA---A-T--GGAG--T-G--A--C---GCA--A--T-CCTC-----G--TCAGT-G----AAG-G--TCAT---AA---A--A-- 7236
LST.KE.x.lho7                ------...GT-CT-G-----AAA--G--AG-A-----AAG-------CA-A-----AG-G---A-T-----A---C-CC-GACA--CC-T--GGAG--TG-G--AC-----GC------T-CCTC-----G--TCAAT-G----AGG----TCAT---AA----GAG-- 8308
SUN.GA.98.L14                ------...GT-CT-G-----AA---G--AG-G--AC-AAG-------CA-A-----AG-G---A-T-----A---C-CC-GACA--CC-A--GGAG--TG-GC-A--C---AGC--TT-A-CTTC-----G--TCAAT-G----AAG-A---CAT---AA------C-- 8376
MND-1.GA.x.MNDGB1            ---T--...-A-CTC--A---CAA--GTCTG-A--C--AA--------CA-A-----AG-A---A-TT-A---A-TC-CC-GACCAG-T----G-AT--TA-C------C---GCTT-G---TCTCAA-----G--T--AT-GGC--AGG-G--TCAT---T----AGAG-- 7660
OLC.CI.97.97CI12             A-----...GA-ATG------AA---G--AG-AT--C--A-G-----TA--------AG-G---A-C--------TC-CC-G-----GA--A--G-A-------C--A-----GGCA--A--AAAACAA-----A--TGAATTT-TG-AAG-G---CAT---A-A-----A-- 7680
WRC.CI.97.97CI14             C-----...GA-CT-G-A---AA------AG-G-------G-------CA-------GG-G--AA-C--------TC-CC-GTCA--CA-T---G---T-T--G-A--C--GTT-AAAT--AAACAA-----A--TGAATTG-C--AAG-A---CA---C-----T------ 8184
WRC.CI.98.98CI04             C-----...GA-CT-G-----AA------AG-AT------G-------CA-------GG-G--AA-C--------TC-CC-GTCA--CA-T---G---T-T--G-A--C--GTTAAAAT--AAACAA---------TGAATTG-C--AGG-A---CAT--CAA---A------ 8229
WRC.GM.05.Pbt_05GM_X02       ------...GA-CT-G-A---AA------AG-G-------G-------CA-------GG-T--AA-C--------TC-CC-GTCA--A--C---G-------T--G-A--C--GGT-AAAT--AAACAA-----A--TGAATTG-C--AAG-C--TCA---A--A--C------ 7125
MND-2.CM.98.CM16             C-----...--CCTG--A--TCA---CTCCG-GT-A--AA----A--TA-------TG-A---A-T--T-----CC-CC-TACA--CT----GGAC--TG--GCT------GCAAGA--ATCC--G------C------ATTCGCC-AGG-A---CA---ACA---A--A-- 8203
MND-2.GA.x.M14               C-----...--CCTG--A--TCA---CTCCG-GT-A--AA----A--TA-------TG-A---AA-T--T-----CC-CC--ACA--CT----GGAC--TG--GC---C--TCAAGA--A-CAG-A-----A------ATTCTCC-AAG-G--TCAT---AA---T--G-- 8132
MND-2.x.x.5440               C-----...--CCTG--A--TCA---CTCCG-GT-A--AA----A--TA-------TG-A---A-T--T----TCC-CC--ACA--CT----GGA----G--GC---C--TCAAGG--CTC-G-A-----A--T--ATTCTCC-AAG----TCAT---AG---AAAA-- 7836
DRL.x.x.FAO                  C-----...C-CCTGG-G--ACA-------G-GT-AC-CA-GT-A--TA--------G-A---A-C---------CC-CC--ACA---C-T--GGA----A--GGA--C---GCAA-GT--A-CC-C-----A------CTTTGC--AAG-A---CAT--AAA----GTA-- 7894
MON.CM.99.L1                 C-----...------CG-AACA--C------AGCT-AC-A--G--T--T--T-----AG--------------A------CC-CC-GAC------A--G-AG-T-A-------C---ACAT-G--AAATGCA-----A--TG-CAAT----GC-G-G--TCA---AA-A------G-- 8016
MON.NG.x.NG1                 C-----...------CG-AACT--C--G---AGCT-A--A--G---T-T-----------------A--AT-A-----CC-CC-GAC-------A----AG-T-A-------C---ACAT-GT-AAATTCA-----A--TG-CAATCGCGCAG-R---CA---ACAA----TA-- 6641
GSN.CM.99.CN166              C-----...-------G-A-----C------AGCT-A--A--G-CAT-T--A----------A--A------A--AA-CC-CC-GTCAT-AC-A--G-A---TT-G-G-------ACCA-T--CCAAGCA-----C--TG-GAAC-GG-CA--C--TCA---CATA--A------ 7990
GSN.CM.99.CN71               C-----...-------G-A-----C------AGCT-A--A--G--AT-T--A----------A--A------A--A--CC-CC-GTCTT-AC-A--G-A-------G-G-------ACAA-T--CCAAGC------C--TG-GAATC-G-CA--C--TCA---CATA--A------ 8002
TAL.CM.00.266                C-----...------A------CAT--G--A--C-----A---T-A-----G--------------A-T---A-T--CC-CC-GACC--GC-T--G-A-----T--G----C--GGCTAAG--AAATTCA-----G---G-AT-G----AG--C---TA---CT-G--A--G-- 7883
TAL.CM.01.8023               C-----...------A------CAT--G--A--C-----A---T-A-----G--------------A-T---A-T--CC-CC-GACT--GC-T--G-A-----T--G----C--GCCAGGT-AAATTCA-----A--TG--T-G----AG--C---TA---CT----A--G-- 7433
MUS-1.CM.01.1085             C-----...--A--CG-A--A--C------A-----A--A--GT-GT-T--T--------T--A-----A-----TC-CC-GACA--GC-T--G-AG-T-A--------C---GCC--T---AATC-C-----A--TG-AAAC-GG-AG-----TCA----AGA------A-- 7987
MUS-1.CM.01.CM1239           C-----...------CG-A--A--C--G--A----------A--GT-GT-T--------T--A--A-----A-----CC-CC-GACA--GC-T--G-AG-T-A------------GCCT-GT--AACC-------A--TG-GAAC-GG-AG--C--TCA----AGG------A-- 7956
MUS-2.CM.01.CM1246           C--T--...--A--CG-A--A--C------A-----AC-A--GT-GT-T--T--------A--A--A--A-----TC-CC-GACA--AC----G-A--T-G-G------C---GCCA-A--AAATC-G-----G---G-CAAT-GG-AA--C---CAT--CT-G------A-- 8085
MUS-2.CM.01.CM2500           C--T--...------CG-A--AA-C------A------------A--GT-GT-T--T--------A--A-----A-----CC-CC-GACA--AC----G-A--T-A------------GCCT-GT-AAATC-G-----A--TG--AAC-GG-AG--C--TCA---CAGG------A-- 8008
SYK.KE.x.KE51                ------...GA----G-A-----A--------C--C--AGGCT-A--G--------TG-G--AA-C--TA-C--CC--C--ACT---C-A--G-CG--T--C-GA--C---GCCA-TA--TCA-A------A---G-ATTT--G-AA-----TCA---A--A--AA-G-- 7713
SYK.KE.x.SYK173              ------...GA---AG-------A--G--A--C--C--AGGCT-A--T--------TG----AA-C--TA-C--CC-GC--ACA---T-A--G-CG--T--C-GA--C---GGCCA--T--TCA-A------A---G-ATTC--G-AG--C---CAT--A--A---A-G-- 8020
RCM.GA.x.GAB1                C-----...GACATAG--A--CG---G---A-----C-AA-G-----C--------A-CT--AA-T--------GC-TG--ACT---A-T--G-A--------GCA--C---TCT--AT--AAT-CA-TT--G--TG-CT-G-G--AAG-C---CAT--A-TG------G-- 7740
RCM.NG.x.NG411               ------...GACATAG----ACA---G---G-A---C-AA-G-----C--------G-CT--AA-C--------GC-TG--ACT---A-T--G-A--------GC---C---TCT--AT-AAAT-CA-TT--C--TG-AT-G-G--AAG-G--TCA---AA-G---GAG-- 7784
GRV.ET.x.GRI_677             ------...GC----G-G--ACA-------A--GT-----A-G--G--CA-T----TG-G--AA-C---A-T--CC-CG--ACA---C-C--G-AG-------G-------GGCA-GG--AAATTCA-----A--TG-GT-G----AAG-A--TCA---CA-A------A-- 8060
SAB.SN.x.SAB1C               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-A--------G-------GGCA-GGT--AAC---------A--TG--TTC-GG-AGG-G--TCA---GA-G--C-T--- 8354
TAN.UG.x.TAN1                C-----...GC---AG-----CA---G-----GA--C-AA-G-----CA-T----TG-G--AA-C---A-T--CC-CG--ACT--CC-T----A--------G-------GACA-GG--AAATT-G-----A--TG-ATTC----AGG-G--TCA----A-G------G-- 8158
VER.DE.x.AGM3                C-----...GC----G-G-----T--------GA-----A-G--G--CA-T----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG-------G----C--GGC--GGT--AATGC-----G---G-AT-G--G-AAG-C--TCAT--AA-C--A--G-- 7699
VER.KE.x.9063                C-----...GC----G-G-----T--------AA-----A-G--G--CA-T----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG-------G-------GGCA-GAT--AACG-C-----G---G-AT-G----AGG-C--TCAT--CA-A--A--A-- 8201
VER.KE.x.AGM155              C-----...GC----G-G-GA--T--------GA-----A-G--G--CA-T----TG-G--AA-C---A-T--CC-CG--ACA---C-T--G-AG-----GGC------GGCA-GGT-AAACGC-----G---G-GT-G----AAG-A--TCAT--AA-A--A--C-- 8181
VER.KE.x.TYO1_patent         C-----...GC----G-G-----T--------GA-----A-G--G--CA-T----TG-T--AA-C---A-T--CC-CG--ACA--CC-T--G-AG-------G-------GGCA-GA--AAACTCC-----G---G-AT-G----AAG-A--TCAT--CA-A----GAG-- 8177
COL.CM.x.CGU1                A--T--CTG-AACTGG-C---A----GTCTGCA---C-A--G-----G---------GG-T--AA-C-------T-C--G-TGCCA--A-A---G-T--TT-GG----A---GCAAAG--T-CC-GCATT--G--TG-AAAC-TG-AA-----T-GA---AT------C-- 7463
```

```
                                  RRE end
H1B.FR.83.HXB2               GAATGCTAGTTGGAGTAATAAATCT.............................CTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATG 8193
Env                          _N__A__S__W__S__N__K__S__._._._._._._._._._._._._._._L__E__Q__I__W__N__H__T__T__W__M__E__W__D__R__E__I__N__N__Y__T__S__L__I__H__S__L__I__E__E__S__Q__N__Q__Q__E__K__N__
H1A1.UG.85.U455              ---CT----------------------.............................-A---GG-C--A-----A---T-------C--C-A----G-A--------GT-G-------G--A----T-TCAAC-------G--------G-------G---------- 7623
H1B.US.90.WEAU160            -------------------G-------.............................-A---TT-C--------A---T------------------G-----------------G-------T--AA-----------------G-----A--------A---- 8200
H1C.ET.86.ETH2220            ---CT-------------G--------.............................-AA--GG-------G--A---T----------C------T---------GT-----T---GA-A----T--AAT--GC------TC-------------G--C--A---- 7569
H1D.CD.84.84ZR085            ---CT--------------G-------.............................G-A--GT-T------GGGA---T---------C------A---------G---------GG-------T--AA--G-----------------T------------ 7714
H1F1.BE.93.VI850             ---CT----------------------.............................-A----G----------A---T------------------A-AG--G---G-------T---A-A----T--AAA--G--------G--A-----------G---------- 7470
H1G.SE.93.SE6165             ---C-T---------------------.............................TATA-TG-A------G--A---T---T-----A--A-----A--G-----C-----------TA-CA----T--AG-C--C-----------A--G--------G---------- 7611
H1H.CF.90.056                ---CT------------------A...............................-A-AGTG-A--C---G-CA---T---T--------A-----T-A-C------G---------GAGGA----T--AGG--GC------TC-------C------G---------- 7512
H1J.SE.93.SE7887             ---------------------------.............................TAT--GG-C------G-GA---T---------AC-A-----A--G-----------------G-AA-----T--AGTC-----------G-A-----------G-----T---- 7500
H1K.CM.96.MP535              ---CT----C-----------.............................TCTTG----G----------CA---T-----------------A-A---G---GG-------T--GA-AC----T-TAAG-------------A----C------G---------- 7353
H101_AE.TH.90.CM240          ---CT-C-C-----------G------.............................T-T---G----------CA---T--A-----A--A-----G-----------G-----------A-CA----T-TGAGA--C--AC---------G-----A--G--C-G----- 7746
H102_AG.NG.x.IBNG            ---CT----C------------A----.............................T-TA-TG-C--A---G--A---T---------AC-A-----G-A---------G-----------GA-A-----T-TAATC---------------G----G---G---------- 7699
H103_AB.RU.97.KAL153_2       ----A-----------------C----.............................--A--TG------A---A---T------------------A-----------T-----------G-T------T--AAT-----A--------------G-----A----------- 7370
H104_cpx.CY.94.CY032         ---CT----------------------.............................TATA-TG-T--A---G-CA-T-T-------T--C-A-----T-A-------------------CAAA-----T-TGGG---C------------A--G---------G---------- 7563
H1N.CM.95.YBF30              -----AG-C------C--C--TA-CTCT............................TAT--TACA--C------A--TTTA------CAAC-A-----TGAGA--G-A-GA--C--TT--G-TG-C--TTTTGGAC-T--A---C-GG-A---G-A--A--GA-C-CA---- 7744
H1O.BE.87.ANT70              ----AGA-CA----TAGGA.....................................AAC---AGC-------G-CAC-TTA--A---CA---A-----TC-GC-G--A-G---CATA-GCTC-ACC---T-TGAGGA---AC--A-GG-A---GTA-----G---C-A---- 8241
H1O.CM.91.MVP5180            ---CA-ATCA---TCAGGA.......................AGATATAATGAT--CAGT-------G-CA--CTT--A---CA-C-A------CA-C-C--A------GTA-GCTC-A-T---T-TGATGA---AC---C-G-A---G----A--G----------- 8267
H1P.FR.06.RBF168             ----AGG--C-----CC--C---AG-.................GAGACTGAAT-A--TGGCR------C-A-TCTA--A---CA---A-------A-CTGG-GG-T--------TGA-AC---TT-T-TAGA---AC--CG-G-A---G-T------A-RGCT---- 7723
CPZ.CM.01.SIVcpzCAM13        ---CAAA-CA---TCAGGC.............................AAGTCTA--AGTG-C---------CA-TTTA--A---CAAC-A-------A-TTG--C-CT---------G-GAC---TTTTGGT---T-A-----GG-A---TCA--A-----G--A---- 7777
CPZ.CM.05.SIVcpzEK505        ----A-C-C------C--C.......................AATAACTCCTAT--CACA--A---GGAA-T-T---A---CAAA-T-----TGAGC--G-A-GA------TT--G-TG-C--TTTTGGGC-GC-A--GC--G-A---G-G--A---AGC-TA---- 7679
CPZ.CM.05.SIVcpzLB7          ----AGA-C----------C............................AAAACATATA-TG-A------G-CA---T---A---------A--------G---G-A-GA---------GAAA-T--TT-TGGG-----A--GC--G-A---G---------G------T---- 7704
CPZ.CM.05.SIVcpzMB66         ---CA---CC-----C----............................AAATCTTAT--TG-T------T--A---T---A-------C-A------A----G---G-----------GATG-T--CT-TAATC--T-A---A-GG-A---C--------G------T---- 7817
CPZ.CM.05.SIVcpzMT145        ---CAAG-CA---G-A----TT..........................AGTGACTATC-GGTT---------A-TTAC--A---CAAC-A-----T------G-A--T----------G-AC-C--CT--A-A--GC-A--G--GG-TA-T-CA-----G---G--A---- 7705
CPZ.CM.-.SIVcpzMB897         ----AA--C---------A---....................TTAGGAACAAACCT-TCCT-----G--A---T--A------C-A-----G-A---G---G-T--C------GAAACT--CT-TGAAC-TC-AAC-CG---T--G--T--A--G---GTC---- 7707
CPZ.CM.-.SIVcpzDP943         ----AG--C----G--GC---T..........................ACCTCCTAT--TG-A--C------A--CTA--A---CAA--C-------AGAG-G-A--A-----TT--G-AG-C--CTTTGA-C-T--A--GC--G-A---G-G--A---RA-C-CA---- 7759
CPZ.CM.98.CAM3               ----AA--C-----------GC---T......................ACATCCT-T--TG----C------A-TTTA--A---CAA--C-------A-AGGG-A--A-----TT--G-TG-C--TTTTAG-C-T--A---C--G-A---G-G--A---A-C-CA---- 7609
CPZ.CM.98.CAM5               --------C-----------GC---C......................ACATCCTAT--TG----C------A-TTTA--A---CAA--T-----T-A-A-GG-A--A--------T--G-AG-C--TTTTAG-C-T--A--GC--G-A---G-G--A---GA-C-CA---- 7915
CPZ.GA.88.GAB1               ---CAACTC----CC-GGG...AGC.......................AATTCCACA--TG-C------GGGA-TCTA--A---CA-C-A-----T-A-TT-G-GTCT--C-------G-GAA---TTTTGGTC-GT-A--G---G-A---TCA--A-------G--A---- 8233
CPZ.GA.88.GAB2               ----A--TCC----AC-G---CAAA..........................TCTTAT--GG-C-------G-A-TCT-------CA-C-A-----T-AGTT-G-AG-A---C-T---G-AACC---TTTTCAA---T-AC-GAG-G-A---TG-G--A---GA-T-GC---- 7580
CPZ.US.85.CPZUS              -----A--CC--------C--C...........................CTCTCCTAT--TGCT------GGCA-TCTA--T---CAA--A---------A--G-A-GA--C--TT--G-TACT---TTTTAGTC-T--A---C--G-A---G-A-----GA-T-CA---- 8201
CPZ.CD.90.ANT                ----AA-TCC----GTA--CTTCA-GCAAACATGT...GCAAAGAACAGCAGTGATA-AC--TGT------G-AA-T-T---A---CAA--A---------TT-G-AC-G----CA---G-ACAG---T-TAATA-CT-AC--AT-G-A---TG-G--A-------G-GA---A 7652
CPZ.TZ.00.TAN1               ---CCTC-CC---GC-G-AG-T---ACAAAGTGCAATCACAGTGATGCAAAGTACTAT--CTGT--A------CA-TTT---T---CA---A-------TC--TT-G-AG-A--C-CT---G-AACC---T------C-GT-A--GA--G-A---CA--A--G--G--A---CA 7849
CPZ.TZ.2001.TAN2             ---CCTCTC----GC-G-AG-C--CCAAAAGTGCAACAGTACTAATACAAAGTACTAT--GTGC--A------A-TTTA------CA---A-------T---CT-G-GC-A--CGCTT--GAAACT---T-------GT-A---AT-G-A----CA--------G--A--CA 7869
CPZ.TZ.2002.TAN3_1           ---CCTCTC----GC-GGA-CC---TCAAAATGCAACTATACTGGTGCAAAGTACTAT--GTGC--A------A-TTTA------CA---A-------T-AGCT-G-AC-A----C-T--GAAACT---T-------GT-A---AC-G-A----CA--------G-GA--CA 7923
GOR.CM.2007.SIVgor2139_287   ---CAAG-CC----CCGGG---....................AATGACACTGAA--A--TG-C------GGAA---T---T---CAAC-A--------A-TTGG-AG----C-----TGA-AC----TTT-TAGA---AC--AGGG-A--GG-A-----G--GGCT---- 7760
GOR.CM.2007.SIVgorCP2135con  ---CAAG-CC----CCGGG---....................AATGAAACTGAA--A--TG-C------GGAA---T---T---CAAC-A--------A-TTGG-AG----C-----TGA-AC----TTT-TAGA---AC--A-GG-A--GG-A-----G--GGTT---- 7744
GOR.CM.2004.SIVgorCP684con   ---CA---CC----CCGGG---....................AGTGATAGTGAA--A--TGGC------GGAA---TT---T---CAAC-A--------A-TTGG-AG----C-----TGA-AC----TTT-TAGA---AC--A-GG-A---G-A-----G--GGCT---- 7691
H2A.DE.x.BEN                 -GTAAA-GAC-CT...........................................T-ATCG-CTGAC-----AA-T-T---A---CA---------G-A-C--G-CCG-T-CCTAGAGGCAAAT--CAGTCAAAGTT-A------G-C----TT--A--------------A 8654
H2A.GW.x.ALI                 -GTAAA--AC-CC...........................................T-AA---CTGA-----G-CA---T---G---CAA--------ACA-C--G-CCGTT-CCTAGAGGCAAAT--CAGTGAACAGT-A----CGGG-A----TT-----------------A 8638
H2A.SN.x.ST                  -GTAAA-GACACC...........................................T-AACG-CTGA------CA---T---A---CA---A-----GCA-CG---CCG---CCTAGAGGCAAAT--CAGTGAAAGTT-A---C-GG-A----T----------------CA 8084
H2B.CI.x.EHO                 -GTAAA-GAA-CC...........................................--TA-G-CAGAC------A---T---A---CAAC-A-----G--GC--G-CCG-TTC-TGGATGCAAAT--ACAAAA----C-A-----GG-A--G-TA--A-------------CA 8616
H2B.GH.86.D205               -CCAAA-GAAACC...........................................--CAC--CA-A------A---T---A---CAAC-------A-AGC--G-CC-TTTC-TGGATGCAAAT--ACGG-TC-GT-A-----GG-T--G-TA--A--G-----------CA 8630
H2G.CI.92.ABT96              -G----CCT-G-TGC-------A-A...............................T----G-CACAA------A-T-T---A---CAA--A-----A-AGC-G--C---TT--TGGAGGA-AAC--CACACGGC--T-A------G-A----TT-------------A---A 8027
H2U.FR.96.12034              --TAAACGACACA...........................................T-AAC--CT-A----G--A-T-T---G---CAA--A-----GGA-A-GG-G---T-C-TGGA-GAAAAT--CACACAGC--T-A----C-G-T----TA--A-----G--A--CA 8138
SMM.SL.92.SL92B              -CCAAA-GACA-T...........................................--T-TC-CAGAC-----CA-T-T---A---CAA--A-----G-A-A-GG--G-ATTCCTAGAGGCAAAT--ACTCAAA-GT-G--G---G-A-G-CTA-----------G--A--CA 8047
SMM.US.x.H9                  -CCAAA-GACACA...........................................T--AC--CT-RC-----CA-T-T--YT---CAA--------A-A-C-GG-G---TTCCTAGAGGCAAAT--ACTSAA-Y-T-G------G-A----TT---------M-----A 8088
SMM.US.x.PGM53               -CCAAA-GCC-CA...........................................T---TG-CT-A------CA-T-T---T---CAA--A-----A---C-GG--G--G-CCTAGAGGCAAAT--ACTCAGGC-T-G------G-A----TT-----------------A 8545
SMM.US.x.SIVsmH635F_L3       -CCAAA-GAC-CA...........................................T---TG-CT-A----G-CA-T-T---T---CAA--------A---A-GG--G--TTCCTAGAGGCAAAT--ACTCAA---T-A------G-A----TT-----------------A 8621
                              RRE end

MAC.US.x.239                 -CCAAA-GCAA-T...........................................--AAC--CA-AG-----CA-TGA--T---CAA--------GC--A-GG--G--TTC-TGGA-GAAAAT--ACAG--C-CC-A--G--GG-A----TT--A-----G-----CA 8599
Env                          __P__N__A__S__._._._._._._._._._._._._._._._._._._._._._L__T__P__K__W__N__N__E__T__W__Q__E__W__E__R__K__V__D__F__L__E__E__N__I__T__A__L__L__E__E__A__Q__I__Q__Q__E__K__N__
MAC.US.x.251_1A11            -CCAAA-GCAA-T...........................................--AAC--CA-AG-----CA-TGA--T---CAA--------GC--A-GG--G--TTC-TGGAGGAAAAT--ACAG--C-CC-A--G--GG-A----TT--A-----G-----CA 8595
MAC.US.x.251_BK28            -CCAAA-GCAA-T...........................................--AAC--CAGAC-----CA-TGAT--T---CAA--------GC--A-GG--G--TTC-TGGAGGAAAAT--ACAG--C-CC-A------GG-A----TT--A-----G-----CA 8581
MAC.US.x.EMBL_3              -CCAAA-GCAA-T...........................................--AAC--CAGAC-----CA-TGAT--T---CAA--------GC--A-GG--G--TTC-TGGAGGAAAAT--ACAG--C-CC-A------GG-A----TT--A-----G-----CA 7730
STM.US.x.STM                 -CCAAA-GA--CT...........................................T---T--C-GA------CA-T-T---A---CA---A-----G---A-GG--G--TT-CTTGAGGCAAAT--ACACAAC-GT-A------G-T---GTT--A--G---------CA 8250
MNE.US.x.MNE027              -CCAAA-GCAA-T...........................................--AAC--CA-A------CA-TGA--T---CAA--------GC--A-GG--G--TTC-TGGAGGAAAAT--ACGG--C-TC-A------GG-A----TT--A-------------CA 8076
DEB.CM.99.CM40               ----AACTC-CT-...........................................GA---T-CAGAC---G-CA-T-T------CA---------A-TGA-GG-AGCA-----T---GATGAGTGGG-AGGAGCTT-AC-GAG-G-A---G-G-----G--G-G----- 7921
DEB.CM.99.CM5                ---CCA--CA---...........................................GGA----CTGA------A-T-T------CAA--------A--GA--G-AGCA--C--T---GATGA-TGGG-AGGAGCCT-AC-GAG-G-A--GG-A--A-----G-GA---- 7842
DEN.CD.x.CD1                 --GCAA-CAA.........................TCAATACCTGGATTGGCATTATACC-G-CTGAC-TCC-GA---T---A---CAAC-A-----G-C---C--AGCAGC-CTATTTG--AATG-GTCAGA-GCCT-GA-GA--G-A--GGCA-----G----GA---- 8256
LST.CD.88.447                ---CTA--A-..............................................A-AACT-CT-A-----CAAGGGAT--------T-----------C--G--GGAGTA-TAGA-GCAAAT--CTC-A--C-GT-GC-----G-CT-T-C-ACAG-GTT-G-A---A 7358
LST.CD.88.485                ----TA--A-..............................................A-AACT-CT-A-----CAAGGGAT--------T-----------CG-G--GGAGTA-TAGA-GCAAAT--CTC-A--C-GT-GC-----G-CT-T-C-ACAG-GTT-G-A---A 7355
LST.CD.88.524                ----TA--A-..............................................G-TACT-CT-A-----CAAGGGAT--------A--A-----A---C--G--GGATC--TAGA-GCAAAC--CAC-A-AC-GT-GC-----GG-TT-C-CAACAG-GTT-G-----A 7361
LST.KE.x.lho7                --CATA--A-..............................................A-CAC--CC-A-----CCA-GGAT-------GA--A-----ATC-A-GG-AGCA-T---TGAC-AGAAC--TACT--TC-GT-AC-----G-AT-C-C-ACAG-GCT-G-----C 8433
SUN.GA.98.L14                ----TA--A-..............................................A-CACG-CC-A-----C-A-GGAC-----------A-----T--GC--G-A--A-TG--TGATGATAAT---ACAG-TC-TC-CC-G---GG-AT-TGTAACAG-GTT-G-----C 8501
MND-1.GA.x.MNDGB1            --C-AA--CAA-C...........................................A-CACT-CA-A-----CATCAGAA--T----A---A-----G-C-AG--C-G-TT--CTGCA-CAAAAC--TACAGAAA-GT-AA--C-GG-AT-TG-T-GAG-GC---GA--CA 7788
OLC.CI.97.97CI12             ---CTAC-C-..............................................-A-TC--CAGA-----GGCA-TGAT--T---T--A-----C-A--GA--G-AG-T----TA---CAGGA----G-TGA-AG-T-GAC-AT-G-AT-TC-AG----GC---G----C 7805
WRC.CI.97.97CI14             ----CAGTC-...............................TGGAATGCAACT-CA-A----G-AA-TGAA--T---CAAC-T---T-T-A-AG--CAG-TG-ACTGCAGGCAAAT--AGTG-AA--T-A-T-ATTG-C--TG-T--AG-GC--GCA---A 8315
WRC.CI.98.98CI04             -----AGTC-...............................TGGAATGCTACT-CA-A-------A-TGA-------CAAC-T---T-T--GAG--CAG-TG-A-TGCAGGCAAAT--AGTGTAC-CC-A-T---GG-T--T---ACAG--C-------CA 8360
WRC.GM.05.Pbt_05GM_X02       -G-G-AGGCA...............................TGGAATA--AC--CTGC-------A-TGAA--A---CAAC-A---C-TTT-AG--CAG---CC-TG-TG--TAAT--CTCAGATA-GT-A-TGC--G-T--TGCT--AG--C--TCA---A 7256
MND-2.CM.98.CM16             -CCAAA-GAC-CT...........................................A-AAC--CA-A-----CCTCAGA---A---C----A-----T-A-AG-G-A-CAGC-CTA--TGA-AAT--GACAGTAAACT-AC-GA-GG-AT-TG-ATT-G-GC--------A 8331
MND-2.GA.x.M14               -CCAAA-GAAA-T...........................................A-AACT-CT-A-----CCTCAGAA--A---C----A-----T---AG-G-G-CAGC-AT---T-A-AAT--GACAAT-GATT-AC-GAGGG-AT-TG-ATT-G--C--------A 8260
MND-2.x.x.5440               -CCCAA--A-A-T...........................................A-A-TG-CC-A-----CCTCAGA---A---C----A------------AG-G-A----GCATAGTG-CAAAT--GAC-ATAGACT-AC-GAGGG-AT-TG-ATTAG--C-G-G---CA 7964
DRL.x.x.FAO                  -CCCAA-GAG-CA...........................................G--ACT-CA-A-----CCTCAGA---A-----------C-A-A-AG-G-AG-T-G-ATAT---A-AAC--TACTCTTGATT-AC-GA-GG-AT-TG-A--AG--C--------A 8022
MON.CM.99.L1                 ---CAA-TC-...............................TGGGCAAAAGGT-ACTTC-CTGAG---G-CA-T-T---A---CAAC-----AGTGAGTT-G-AG-T---G------TGACC--T--GCAGC-CT-G----CTG-----G-G-------GC--A--CC 8153
MON.NG.x.NG1                 ---CAA--CG...............................TGGGCTAAAGGC-ACTTC-CCGAG---G-CA---T---A---CAAC-A---AGT-TGCT-G--G-----G-T--GGCAC-T--C--GCADC-CT-G--G--GG-T--GG-G--A----G---A---- 6778
GSN.CM.99.CN166              ---CA-CTCA...............................TGGGCTAATGGCTC-CTT-CAGAC---G-GA---T-------CAAA-----AGT-TGTT-G--G-G---G-----TATAC------GCAG--GT-A--GC--G-AA-CC-G--A----C-TCA--CC 8127
GSN.CM.99.CN71               -----ACTCA...............................TGGGCTAAGAACTC-ACG-CAGA----G-G----T-------CA-------AGT-AGTT----G-A---G-----TATACC-----GCAA---T-A---A-TG-AA-CC-T--A---TC-------CA 8139
TAL.CM.00.266                ---CAAA-CA...............................TGGTCGAATTACACA--T-CTCAA---C-GA-T-T---A---CAA--A-----G-TGA--G-AG-----C-----G-GC-C---TCACAG--GT-AC-----G-A---G-G--A--G----GA---- 8020
TAL.CM.01.8023               ----AAG-CA...............................TGGACAAACTCCACAA-C-CTGA----C-GA---T---A---CAA--------G-AGTTGG-AG-----GC-T--GA-ACT--CACAGTA--GT-AC-----G-A---G-A--A--G----GA---- 7570
MUS-1.CM.01.1085             -----A-TCA...............................TGGGCAAATCACACCC-G-CAGG----G-AA-T-T---A---CAAC-----AG---GTTGG-AG-----G-T--T-CAACT-----GGAG---T-G---TTGG-A----GA-----G---G-----C 8124
MUS-1.CM.01.CM1239           ----A-ATCA...............................TGGGCCAATCATACCTT--CAGGC---G-AA---T-------CA-C-----AG--AGTTGG-GG-T---G-T--T-ATACC-----GGG-C-CC-A----C-G-A------G--A--G---G-A--CC 8093
MUS-2.CM.01.CM1246           -----A-TC-...............................TGGGCAAACCACACCC---CAGGA---G-CA---T-------CAAC-----AG----TT-G-GG-----G-----GCAG-C--C--TGGTC--T-G---ATGG-T---C-A--A--G---G-A--CC 8222
MUS-2.CM.01.CM2500           ----AACTC-...............................TGGGCTAATCATACCC-G-CAGAA---G--A---T-------CA-------AG---GTTGG-GG-----G-----TGCAACC--C--AGGTC--T-A--GTTGG-T--G--A-----G--GG-A--CC 8145
SYK.KE.x.KE51                -C-GCAAGCA--TG-C-----C---.....................CGTTGC-CAAC--CCCAA---G-AA---T---A---CATACA-----AC-GC-GG-AG-T----TG---GA-CAC---G-TAAT---C-GAG-----G-T--GG-G---------G-------- 7859
SYK.KE.x.SYK173              -G-GAAAGCA--TG-A--C--C--CAAC...................TTCTGT-CCA---CACAG----AA-T-T---A---CACAGA-----ACA---GG-AG------TG--TGA-CAT--TG-TGGA--GC-CAG----G-A----G-A--A--G--G-GA---- 8169
RCM.GA.x.GAB1                --CATTC.................................AACAAAACG-CTGAG---C-GA-AGAAT-A---T--C--------A---A-T--CTCTT---TAGAGGCTAAC--TACAATAGC-T-AC-G--GG-C--GG-T--A--T--G--A---- 7865
RCM.NG.x.NG411               --TATA----CAA...........................................AC--CAGAG-----CA-ACA---A---C-T---------A---A-T---TCA-GA-TAGA-G-AAAC--CTCAGTAGCTT-AC-G--TG-A---G-A-----T---CGT---- 7909
GRV.ET.x.GRI_677             ---GTA--A-AAC...........................................ACT-CT-AG---G-CA-T-T---T---T----------G---C-------TGCC-TGGA-G--AAC---ACTCAAC--T-G-------G-A-----T---G--TC-------C 8185
SAB.SN.x.SAB1C               ---GTA--A-AAC...........................................ACT-CCGAC---G-AA---T---G---CAA--A-----G--GC----AG-A--G--TGA-GCTAAC--CAGTAGGA-TT-A--GC--G-A--TG-A---G--C--------T 8479
TAN.UG.x.TAN1                --C-TTC-ACAAT...........................................AC--CAGAC---G-CA-T-T------CA---A-----G--CC----A-CTGCC-TGGA-G-AAAT--TAGTA-TACTC---TCA--G-AT-TG-G--AG-GC-G--A---A 8283
VER.DE.x.AGM3                -C-GTGG-A-AAT.........................................AG-ACC-CTGA-------A-T-T---T---C----A-----A---C-G--ATCGT---TGGA-G-TAAC---ACAA-ACA-T-A--G---G-CAG-GCA---G-G--G------T 7827
VER.KE.x.9063                -C-ATGGCAAAAC.........................................A--ACG-CC-A----C-AA---T-------T----------A---C-G---GGAG-G-TGGA-G--AAC---ACAGAACAGT-A-T-A--G-AAG-G-A--AG----G--A---T 8329
VER.KE.x.AGM155              --CGTGG-A-AAT...........................................AC--CAGAG------A-T-T-------T----------A-A-C-G--AG-AGGA-TGGAGG--AAC---ACAAAACA-T-G---C-GG-AAGGG-A--AG-G---------T 8306
VER.KE.x.TYO1_patent         -CCCTGG-CAAAT.........................................-G-ACT-C-GA----C-AA-T-T---T---T----------A---C----AGCTG---TGGA----AAC--TACGAGACA-T-A-TGA-GG-TAG-G-A--AG-G---------C 8305
COL.CM.x.CGU1                ---CAAG-CC---....................................GGAGA---GG-TCCA---C-GA---T--------AAC-----C-TGAGCG-G-A-GA---------GATA-----G-GG-AGACT-A-T---GG-AT-TG--TTA-----GG-A---- 7594
```

```
H1B.FR.83.HXB2            AACAAGAATTATTGGAATTAGATAAATGGGCAAGT...TTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATAT 8360
Env                       E  Q  E  L  L  E  L  D  K  W  A  S  .  L  W  N  W  F  N  I  T  N  W  L  W  Y  I  K  L  F  I  M  I  V  G  G  L  V  G  L  R  I  V  F  A  V  L  S  I  V  N  R  V  R  Q  G  Y
H1A1.UG.85.U455           ---T---C-------C---G--C--G-------A-...C----A-------------CT-------------------T-G-C---TG----A---------------AA----A---------------A----G---------A-C--------------------C 7790
H1B.US.90.WEAU160         -G-----------------G-----------------...-------C-------G-----T-----C------------------A-C-------------------------A---------------------A-----------C-----A--------------------C 8367
H1C.ET.86.ETH2220         --A-----T---C-A-C---G--C--------A--A-...C------------------------------------------A------------------G-----TG--A----------------A----------G----------------------------------C 7736
H1D.CD.84.84ZR085         --A------------A-----G--C--G-----G---...--------------G--------C-A------------------A-------------------------A------------------A-----------------A---------------------G--C 7881
H1F1.BE.93.VI850          ---------------C------C--------------...C--------------G-----T----C------------------A------------------G--------A----C-----------------------------A---------------A--------C 7637
H1G.SE.93.SE6165          -------C--------CC--G--CC-G----------...C--------------GG--------GA-----A---------------A----T-------------------T--AA---------------------G------------A---------------------C 7778
H1H.CF.90.056             ----G--C-----A-C---G--C--------------...C-------C-------G-----T--C------------------A------------------------T--AA--------------A-----------G------------A---------------------C 7679
H1J.SE.93.SE7887          --A----C-------C---G--C--G---A---A---...-------------------T-----C-----------C--C---A----------------A-------T--AA--------------A-----------G---G---------A---------------------C 7667
H1K.CM.96.MP535           -------T-------C---G--C--------------...C--------------G------------A------A--------------A------------------A------------A-------------C----------G-------G----A----------C----------C 7520
H101_AE.TH.90.CM240       --A-G--T--G--A-----G-----------------...C--------------G------------------------------------A----T--------------------T--AA--------------A-----------G-----------A---------------------C 7913
H102_AG.NG.x.IBNG         -------C--G----C---G--C--G-----------...C-A------------G-----------------------A----------G-A----T--------------------T--AA--------------A-----------G---A-----A-A---------------------C 7866
H103_AB.RU.97.KAL153_2    --------A------C---------------------...---------------G-----T----A-------------------------A------------------------------------------A-----------------------------------G-------------- 7537
H104_cpx.CY.94.CY032      -G-----C-------CC--G--C--G--------C..C----------------G-------------A------A---------------A----T--------------------A------C----------A----------G-------------A---------------------C 7730
H1N.CM.95.YBF30           -GA--TC-C-C--------G---C------AC---...C------GC------GGT--T-----A-------------------A--GCT-------------C-----A-T-----CA-------CA-AAG-ATA---A-AA-----A-AGCA------------------ 7911
H1O.BE.87.ANT70           -GA--A-G--GC----G------G--------CTC-...A-T---------C--G--------T--A---T---------------A--GCA-----C----------CAC-A-----GG-G---G-TA-CA-GATA------AA----------A-ACA--------------- 8408
H1O.CM.91.MVP5180         T-A---C---G-----GC-----G--------CTC-...C-T--------------G--------T--A---T--------------A--GCT-----C------G-----CAC-AA-----A-----G-TA--A-GATAA-----AA-C--------G-ACA--------------- 8434
H1P.FR.06.RBF168          --A--A-G------A--G-----CC----------CAA...---------C-----GG-------RC-A---T----------C--------GA-T------T-------------AA-AA----AC-------CT-G--A---R--R--AA-G-----C-GG---A-C-------------C 7890
CPZ.CM.01.SIVcpzCAM13     --A-G--C---------G--G---C---------TC-...---------C--------G-----------C--C------------C-------A-T---C-T---GC----------GA-TA--------------TA-CA-GT-A--TG-C--AG-GA-C-GG-----C---------T--C 7944
CPZ.CM.05.SIVcpzEK505     --A-GTC-C-C--------G---C-----T------...---------C-------G-----T------A--------------C--------A-T----------G-----C------A-T------CA-----------A-AAGCATAA--A-G-----G--AGCA------------------ 7846
CPZ.CM.05.SIVcpzLB7       -GA--A--C-T--A---C-G---C-T---A----C..C--------------G------GTC-C----------------C--------A---------T-------A-------------AA---TG-GT--G---A-------------G--CG-------A----------C------------C 7871
CPZ.CM.05.SIVcpzMB66      --A------C-CC-A-----------C--G-----G--C..C-C--------------G-----------TCA-------------------A----T--C--A-------------A--AA---------------------------T-G-----A-----A------------------------ 7984
CPZ.CM.05.SIVcpzMT145     --A-G--G------------G----TC------T-A-...--A----G-------G-------------T-------C------GA-G---T----AG-----------A-TA----C-----------TGC-----AA--GGA--AC--A-C-----G-------A--------C 7872
CPZ.CM.-.SIVcpzMB897      -G--------G--A-CTC-G----------T----C..C-------------CG-T-------C-G-----------------------TG-C--------------------AA----CA--------------A-G------------C---------------------------C 7874
CPZ.CM.-.SIVcpzDP943      -GAG-AGC------------G---C-----T-----...-----------------G-T--TT-T-----------------C-------A----TC------GC------C------AA-----CA-----------AGGA-T-A-CA-G--CT--A-AGCA-AG---------------- 7926
CPZ.CM.98.CAM3            -GA--AGTC------------G---C-----T-----...--A-------------G-T---T--CCGC------------C------------G--T------------C-------A------CA------------AGGG---A-CA-A--AT-T--AGCA-AG---------------C 7776
CPZ.CM.98.CAM5            --A-----TC-T---------G---C-----T-----...--A----G------C--T-----TC-G---T------------------A----TT-G--AG------C----T--AA--------T---C--A-AGGAAT---TA-G--AG-TA-AGCA-AG---------------C 8082
CPZ.GA.88.GAB1            --AGG--CC------------G---C---------C--C..----------------G-------------A-------------------A-T---C-T---GC------------AA-CA------G-----------TA--A-GA-A--TT-C--AG-----C-GG--------------C--C 8400
CPZ.GA.88.GAB2            --A-----GC-----------G---C-----T-CTCA...-A----------------G-------------C---------------------A-G--T------CT-----G------GG-TA------------T--AA-A-G----G-AAA-----A-A-GA---AG------------C 7747
CPZ.US.85.CPZUS           --A--TC-C-C----------G---C-----T-----...C-A------C-------G-T--T--C--C---------------------A-----TT-----AG--------C-A-----A-------AA-C------T--AGG--TGA--T-----AC-----AGCA-A--------------T--- 8368
CPZ.CD.90.ANT             --A-G--------AT---C----C------AGCTCA...-A------------------G-----------C-A-------A----------A------T--T-------------CTA-T------AC--------TT-GC---TAT--G--AG-TGCT-A-GA-AG------------------ 7819
CPZ.TZ.00.TAN1            -------G--G--A---------C------AGC----...C-T---G----------G-T---------C-A-------------------A--GCT------C-------------C-----A--A------AC-T------TC-CA-GTT-A--G--AA-G-----T--GCA------------T--- 8016
CPZ.TZ.2001.TAN2          --------------A--G------C------AGC---...C-T---G----------G-T---------C-A------------------GG---GCT------C-------------C-----G--A------AC-T------TC--A-GTT-A--G--AA-G-----T--G-A-------------T--- 8036
CPZ.TZ.2002.TAN3_1        ---------C----A--G--G----------AGC---...C-T---G----------G-T---------C-A--------------------A--GCT------C-------------C-----G--A------AC-T------TC--A-GTT-A--G--AA-G-----T--G-AG------------T--- 8090
GOR.CM.2007.SIVgor2139_287 --A-G-CCC--C-A--G--G----------------GA-...C-T---G----C--G-----------C-A------A--------------A------------A---A---C----GC-A------AC-------TC-AA-G---A--A-AAA---GTGTC-C--G-------------T--C 7927
GOR.CM.2007.SIVgorCP2135con --A-G-CC---C-A--G--G----------------GA-...C-T---G----C--G-----------C-A------A--------------A-------------A---A---C----GC-A------AC-------TC-AA-G---A--A-AAA---GTGTC-C--G-------------T--C 7911
GOR.CM.2004.SIVgorCP684con --A-G-CC---C-A--G--G----------------GA-...C-T---G----C--G-----------C-A---A-A------C------G-A--------------AG---A---C----GC-A------AC-G-A---TC-AA-G---A--A-AAA----TGTC-C--G-------------C--C 7858
H2A.DE.x.BEN              TGT-T------CAAA------A---GC----AT-T-...C-TG-C--C-------G--T-----CTCC---G-CAA-------TC---ATGGAG-GCAT--------TG--AA-AA---C------------CAA-CTA-----G-GCAAT-GT-A-G-----T-----AA-----C--- 8821
H2A.GW.x.ALI              CGT-T---C--CAAA------A---GC----ATGT-...--TACC--C---C--G--T-----CGCC---G-CAA-------TC---ATGGAG-TTAT---A---T---AA-A----C-C-T-----------AA-ATA-----G-GCAA--GT-A-G-----C-C---A-----C--- 8805
H2A.SN.x.ST               TGT-T---C--CAAA------A---GC----ATGT-...--TG-C--C-------G-TT-----CTCC---A-CAAA------TC-G-ATGGAG-TTAT---------T---AA-AA----T---------------AA-ATA-----G-ACAA--GT-A-G-----C-------AA-----C--- 8251
H2B.CI.x.EHO              TGT-T--G---CA-A------A--C------AT-T-...--CA-T--------G--T-C--CTCC---A--GCA--C--C-GG---GGAT--TAT--------AT----AA-A-----TA-----------CAA-ATACA-TA-ACAG--GC-AGCA--GC-------A------C--- 8783
H2B.GH.86.D205            TGT-T------CA-A--A--A----GC----ATGT-...--TG-T--------G--C-C--CTCC---A-AAAA-------CC-T----GGAC--TAT---------C-----G--A-----TG-------------AG-ATACA-TG-GCAG--GC-AGCA--C-------A------C--- 8797
H2G.CI.92.ABT96           TGT-T--G---CA-A------A---GC----ATGTA...--TG-C--------G--C----CTCC---G-TAAA---CG-TT-C----GG-C-TTATG-----G-C-----AG-AA----TG-------------CA-ATA-----G-ACAA---C-AGGA---C---------A--G--- 8194
H2U.FR.96.12034           TGT-T------CA-A------A----C----AT-TC...--TG-T--C--------G-TC-C--CTCA---G-TAAA----G--T-C----GGAC-GTATG--------C-----AA-AA---ATAC-T----------GA-ATA-----G-ACAGC-GT-AGG----TC-------A-A--T--- 8305
SMM.SL.92.SL92B           TGT-T------CA-A-----GA---GC----ATGTG...--TG-A--------G--C-T--CTCC---G-AA-A--C----C-G-ATGGAG-CT-TC-----CAT----AA-A----TTG-----------------A-CTA-----G-ACAA--GT-A-G-----GT-A----------T--- 8214
SMM.US.x.H9               TGT-T-----GCAAA------A---GSK---ATNT-...--YG-C-------K--G--C-T--TTY-----A-AAAA-------C----ATGGTG--C-------TCT-----AG-AA-----A-------------GA-ATA-----G-GCAG--GT-AGC----GT-A---A-------T--- 8255
SMM.US.x.PGM53            TGT-T-----GCAAA------A---GC----AT-T-...--TG-C--------G--C-T--TTC-----A-AAAA-------C-G-ATGGTG--C-T------TCT-----AG-A-----G--------------GA-ATA-----G-GCAA--GA-AGC----GT-A-----A--T--- 8712
SMM.US.x.SIVsmH635F_L3    TGT-T-----GCAAA---C--A---GC----AT-TC...--TG-C--------G--C-T--TTC-----A-AA-A-------C----ATGGTG--C-A------TTT-----AG-A-----G--------------AA-ATA-----G-GCAA--GC-AGC----GT-A--A-------T--- 8788

MAC.US.x.239              TGT-T------CAAA-G--GA---GC----ATGTG...--TG-C---------G--C-TG-TTC-----A-AAA-------C---ATGGAG-TTAT------T-T----AG-AA--CTG-----------GA-CTA-A--G-ACAA--GC-AGC--AGT-A----------G--- 8766
Env                       M  Y  E  L  Q  K  L  N  S  W  D  V  .  F  G  N  W  F  D  L  A  S  W  I  K  Y  I  Q  Y  G  V  Y  I  V  V  G  V  I  L  L  R  I  V  I  Y  I  V  Q  M  L  A  K  L  R  Q  G  Y
MAC.US.x.251_1A11         TGT-T------CAAA-G---A---GC----ATGTG...--TG-C---------G--C-TG-TTC-----A-AAA-------C---ATGGAG-TTAT-----T-T----AG-AA--CTG-----------AA-CTA-A--G-ACAA--GC-AGC--AGT-A----------G--- 8762
MAC.US.x.251_BK28         TGT-T------CAAA-G--GA---GC----ATGTG...--TG-C---------G--C-TG-TTC-----A-AAA-------C---ATGGA--TTATG-----T-T----AG-AA--CTG-----------GA-CTA-A--G-ACAA--GC-AGC--AGT-A----------G--- 8748
MAC.US.x.EMBL_3           TGT-T------CAAA-G--GA---GC----ATGTG...--TG-C---------G--C-TG-TTC-----A-AAA-------C---ATGGA--TTATG-----T-T----AG-AA--CTG-----------GA-CTA-A--G-ACAA--GC-AGC--AGT-A----------G--- 7897
STM.US.x.STM              TGT-T--G--GCAAA--C--A---GC----ATGTA...--TG-C---------G--C-T--CTC----G-AA-A--C---C---ATGGAG--TATT-----TAT----A--A---ATG------G---CAA-CTA-A--A-GCAAT-GT-AGC---GC-----A-A--T--- 8417
MNE.US.x.MNE027           TGT-T------CAAA-G--GA---GC----ATGTG...--TG-C---------G--C-TG-TTC-----A-AA--------C---ACGGAG-TTAT------T-T----AG-AA--CTG-----------CA-CTA-A--G-ACAA--GC-AGC--AGT-A-----A--G--- 8243
DEB.CM.99.CM40            T---T-CCC--CAATC-C--C-AG-T----ACTC-...--A--------------G--C-GT-C-GA---T-C----GG--T-G-----G-AG--TAC---A---CT-CT--AA--TTGC----------C-A-GTT--G-G-GAA---T-GA-G--AGT-GT-T-GA------- 8088
DEB.CM.99.CM5             TG--T-CC---CA-TC-C-GAC-G-T----ATTCA...--A--------------G--C--T-C-GA---T-C----GG--T-G----G-AG-GTAT---A---C---CT---A--TTGC-T--G-----CAA-GTT--GGG--AA---C-G--C--AGT-GT-T-GA------- 8009
DEN.CD.x.CD1              TG--C---G--CAATCC---T--G-T----AT-A-...-----------------G-TT---GC--A---T-C----GG--T-G---GG-AG-GTATG--A---C---G--A---ATA----------A-AA-GTT-A-CA--AA-G-GT-A-G---GC--T-TAGA--C--C 8423
LST.CD.88.447             G-A-T-CC--CAA-A-GC--C-GG-G-TCAAT.....--T-------------AG-----CTT-TCA---T-CCAA------T--G-ATGCAG--C-A---A--AT---TA-TA---T--------G---AAGCTT-A-TG-GCAA-ATA-AGTA-A-A-GT-TAG-------- 7522
LST.CD.88.485             G-A-T-CC--CAA-A-GC--C-GG-G-TTAAC.....--T----G---------GG-T---TTT-TCA---T-CCAA------T--G-ATGCAG--C-A---A--AT---TA-TA---TG---------------AAGCTT-A-TG-GCAG-ATA-AGTA-A-A-GT-TAG-------- 7519
LST.CD.88.524             GGA-TA-C--TAAAA-G---C-AG-T-TTAAT.....--T----G----A-GG--C-C--T-C----T-CCAA------T--G-ATGCAG-GC-T---A-CAT---TA-AA---ATAC--------TC-AAGCTT-A-CA-ACAA-GT--AGTG-AGA-GT-TAG---C--- 7525
LST.KE.x.lho7             -GA-TA----TAA-A-G---C-AG---TTAAC.....--T----G------GG-T--CT--C-C---T-TACT----G-T--G-ATGCAG-GT-A--TA--CTT-TAA-CA-----G------G--T-GAGCTT-A--A-ACAA-AT--AGTA-A-A-GT-TAG---G--- 8597
SUN.GA.98.L14             --A-CA-G--CAAAC-GC--C-AG---TTAAC.....--C----G------GG--C-GT-TC-A---T-TCTA--C---------ATGCAG-CC----TA-----CATTA-AA---CAGCT------TC--AG-TT-A-CA-ACAACAGA-CT-----GA-GT-T---------- 8665
MND-1.GA.x.MNDGB1         C-T-T------CA-A-G----GAG-CCTTA--TC-...-G-GCA-G-------G--T-T--TTGG---G-TCAA--CT-------GGGGAG-TT-CT-----GTT----AA-TA----A-----------TT-G--A--CT--TGGAA--C-A-A-G---GT----------C--- 7955
OLC.CI.97.97CI12          --T-T--G--GCA-AGG--TC-AG---TAAGTGA-...-G-ACA--A-----C--TT-GG-T-------A--ACT----G------A--GGAC-TT--G-----GTT--CA-AT--G--AC-C-AT-----AC-A-GAT--G--AGG-AGA-T-GG-TTA-GC---------T--- 7972
WRC.CI.97.97CI14          TGTTT--G--GCA-A-G---TC-G-T---A-TGC-...-G-G-T-GC---C-CC-ATGGTTT---ATCTG-AAC-GGT-------GGGGAT-GT-AG-TA------ATAG--------A----------TT-AGCATG-T--A-GAA--GTA-T-GG-T--C----------T--C 8482
WRC.CI.98.98CI04          TGTTT------CAAA-G--GC-AG-T---A----CA...-G-G-C-G---C-CC-GTGGTTT---ATTTG-AAC-GGT-------GGGGAT-GC-T--TA-----ATAG-A-----CA---------T-AGGATGGC--A-GA---GCA-T-GG-T--C-----------T--C 8527
WRC.GM.05.Pbt_05GM_X02    TGTTT--G---CA-A------AC-G-T---A-TCAA...-G-G-TTCA---C-CTCATGGTTT---ATTTG-AAC-GGT-------GGG-TC-TT-AG--A------TTAG-A------AA------G-TT-AGGATGCT--T--A---GTC-T-GG-T--C-----------T--- 7423
MND-2.CM.98.CM16          T-T-T--GC--GAAA------GAG-T---A-TTCC...-G-GCT-GC------CG--T-T--GTGG------TAAA----G-T--GA--GGTT--C-A-------AT--TGA-TA----TG------G--TT-AGC-TGCT--TGG--AG-CT-AGGA-AGT-------------T--- 8498
MND-2.GA.x.M14            TGT-T------CAAA-----G-GAG-T-T-A-TTCA...-G-GCT-GC------CG--C-C--GTGG------CAAA----G-T--GA--GGA---T------A--ATG-TAG-AA-----AC-T------TT-AGC-TG-T--TGG-----C-A-AGG---GT-------------T--- 8427
MND-2.x.x.5440            TCTTT--G--GCAAA------GGG-C-TAAACTTCCATGG-CTA-C-G-----CG--C-T---TGG------TAA-----G-T--GA--GGAC-GT-AG------T--TTA-CA-----A----------GT-AGC-TGCT--TGG--AG--T-AGGG-AGT--------------- 8134
DRL.x.x.FAO               TCTTT--G----CAAA------GGG-T-TAA--TCA...-G-GCT--C-------G-TT-C--TTGG---TCTAAA------T--GA--GGAT-TT-C-------ATG-CAA-AA-----A----------TT-AGCA---T--TGGAGC-C---AGG---GT-------------T--- 8189
MON.CM.99.L1              ----T--G---A--A-GCCG-GAC-------ACTTC...C----------------G-----CT-C--A---T-------C--C---A---------T--AG-------C--CT---A-----C-----------C--A-GTTCA----AGGAG-TA-C-G---GT-AG---------C 8320
MON.NG.x.NG1              C---T--GC-CA--A----G-GAC------ACT-G...--A--------------G-T--CT-C--A-----T--------T--GA-C---------AG-G-----C--CT----------CC----GG-GC-CA-GTTCA-C---AGGG--CA-TCG----G--AG---------T--C 6945
GSN.CM.99.CN166           TGA-T--GC-GA--A-G--GTCC--------ACTCC...C-T----G--------G-T--TT-CG-------A-A----C--T--GA-C--TG----AG-----T-C--C---AA---C----------------AA-GTT-A-C---AA---GC-T-GA--GA----------C--C 8294
GSN.CM.99.CN71            TGA-T--C------AA-G---TCC--------ACTCC...C-T----G--------G-T--TT-C-----------C-----C--T--GA-T----------G-----T-C--CT--A----C---G----------A-AA-GTT---T---AA---GC-T-GA--G----------C--C 8306
TAL.CM.00.266             TC--T--G--GCAAA-G---A--G-C----AC----...C-C----GC-----C--TT-GT-TGCC---T-TA-A-GGC-T-GGA--GCAG----AG-------CTA-TC-TA-CTTAC-T--G--T--AA-GTACA--G--AA-G-GC-C--GTTGC-C----------C--C 8187
TAL.CM.01.8023            TC--T--G--GCAAA-G---A--G-C----AC--C...C-C----GC-----C--TT-GT-TGC----T-TA-A-GGC-T-G-A--GCAG----AG-------CTA----AA-TTTGC-T-----T--AA-GTACA-TA-CAA-G--T-A--GTTGC-C----------C--C 7737
MUS-1.CM.01.1085          -G--TA------CAAA----GTTAG-G----ACTCC...C-C---G-A-------G-T--CT-C--A-----C---------C---A-C---TGC----G-----T-C----TC-A--GCT----T--GC----CA-GTT----GT-AGGA---C-A-GG--------------C--C 8291
MUS-1.CM.01.CM1239        ----CA-----GCAAA-----TTAG------ACTC-...--A---G-A-------G----CT-C--G-----T-----------T--GA-C--TTGT---G-------C----T--A--GCT----T-------T-CA-GTT---G--AGGC---C-T-GG--GA-C-----------C--- 8260
MUS-2.CM.01.CM1246        ----CA-G--GCA-A-GC-TTTAG-------ATTCC...C-T----G-A-------G----TT-C--A-----C---------G-C--GA-----TTGC---G-------C------AA---C--------------A-GTT-A--T-AGGA--GC-T-GA--GA----------C--C 8389
MUS-2.CM.01.CM2500        -G--TA-GC-GCAAA-GC-CTTGG------ATGCC...C-C---G-A-------G-T--TT-C--A-----C-TC----G-T--GA----TTGCT-AG-------C------AA---CA-----------TA-AA-GTT-C-G--AGGG---C-C-GA-AG----------T--- 8312
SYK.KE.x.KE51             TG--T--TC-TACTA------C-AG------ATTCC...C-------GC------G-TC-GT-G--A---T-CCA---CC-T--GA--GGAT-CT-TGCCA----C--CTA-A--GATA-------G--C--AG-TT--CTTGGGGC------A-GA-ATA-GTT-GGA--C--- 8026
SYK.KE.x.SYK173           TG--T--TC--ACAA-G---C-GG-------ACTCG...---------GC------G--T-GT-G--A---T-C-TT---C-----GA--GGAT-TTATG--A---------CT--A--GCT-------GC------AG-TT-AGTG-AGGA--CA-A--G-ATT-GTT-GGA--C--- 8336
RCM.GA.x.GAB1             TG--T------GGA-A------AG----T----G-GA-...GCA-TC-G-----C-G--TC-TGACTGG---A--CAA-------------A--GG-T-CT-T-------AT----TA-CA-----A----------G----C---GGCTGT--TGGAA-TGTC-T-G--ATC--------A--G--- 8032
RCM.NG.x.NG411            TG--T--C---GAAA-----GA--TC-----G-GAC...A--CTC-G-----C-A------GGATTGG---T-AAA--------GGA--GGA--TT-C--TA-TTT---AA-AA-----C-----------A-C---TTAC-CTGGAGCTGT--T-G---AGC------GCA--T--C 8076
GRV.ET.x.GRI_677          TGG-TCTG-ACCA-A-----G---G-T---T--G-G...--C---TCA------CTCAC-GT----C-----T-AG-C----G-T---A--GGAT-TT-AG-GA---TGATTA-TC----A---------T-T-CA-GG-TAT--TGGGGATGTA-C-GA-ATA--------------- 8352
SAB.SN.x.SAB1C            TGG-TAGT-ATCAAA-----G-TGTCT---T--GAC...--T---TCG-------G--C-C-------A---T-TG-A-GG--G--GA--GC----C---G-TA----C------A-CA-----TAGCT----G-TC-GC-A-TAA--A-AGGA--TC-T-GG-AGT-----AA-A-G--- 8646
TAN.UG.x.TAN1             TGG-TACT-ATCAAA-G----G-G-T---A-TTC-...-G-------CATC----G--G-CT---TCA----T-C-----GG--T-----GGGGAT-TTAT-------AT----AC-TA---TTG--T--G--G-C---GGCT-A-TTGGGGATGCA-AGC----G-------------T--C 8450
VER.DE.x.AGM3             TGG-T-C--ACCAAA-----AG--GT---T--GA-...--C---TC-------CG-TT-CT-----G-------AACAT-C------A---GGAT-TT---GAT----CT----TA-TA-----A-------T-GC---A-A-A---TA------TGCA-AGC----G-------------T--C 7994
VER.KE.x.9063             T-G-T-CC-ATCAAAG---GACATCT---T---AC...--C---TC-------CG-TT-CT-----A-------AAACAT-C-T----A---GGAT-CT-AG--------T---TA-CA-----AC-T--GT--C-G-A-A-AA--TA-AG-TGCA-CGG----G-------------C--- 8496
VER.KE.x.AGM155           TGG-T-CT-ATCAAA-G--GTCAG-C---T-G----...--T---TC-------CG-TT-TT-----A------AACAT-T-----GA--GG-T-TT--GC-----AT----G-TA-----G-------T-GC---ACA-AT--TA-A--TGCA-AGC----G-------------T--C 8473
VER.KE.x.TYO1_patent      T-G-T-CC-ATCA-A-G---AC--GT---T--GA-...--C---TC-------CG-TT-CT-----A-------TAACAT-T-----A-GGGAT-TT-AG--A---T---AA-AA-----G-------T--C----ACA-A----TA-GGATGTA-AGTG--G------------- 8472
COL.CM.x.CGU1             --A-GA----GGCA-------GGG-T---A-T-A-...-G--TTTCAG-C---GG-C--TTC--CATCT-TAA-----G--CTC-ATGCTGC-TATG-T---------T---AA------AC-T------TA--A-G-TA----A-AG--TGCA-A-GAG---C-TTCAGA-TTA-G 7761
```

```
                          Tat Rev intron end Tat Rev exon 2 start                      Tat premature stop in HXB2                    Tat end
H1B.FR.83.HXB2            TCACCATTATCGTTTCAGACCCACCTCCCAACC.....................CCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGAC...AGATCCATTCGATTAGTGAACGGATCCTTGGCACTTATCTGGGACGATCTGCG 8506
Rev exon 2                                  N P P P N . . . . . . . . P E G T R Q A R R N R R R R W R E R Q R . Q I H S I S E R I L G T Y L G R S A
Tat exon 2                                   P T S Q P . . . . . . . . R G D P T G P K E * K K K V E R E T E T . D P F D *
Env                       S P L S F Q T H L P T . . . . . . . . P R G P D R P E G I E E E G G E R D R D . R S I R L V N G S L A L I W D D L R
H1A1.UG.85.U455           -----TC-G-------------TTGC-----T-.....................--AGA---T-T--G------A-G-A---C-----------------CA--G--A---...-----G----C------G----T--A--GA--GC--------C----- 7936
H1B.US.90.WEAU160         --------G--A------------------G--.....................-------------------------------C------------------------...------GGAA-------G-T---T--AA-----------T---C--A-- 8513
H1C.ET.86.ETH2220         ----------------------TTA-----CA-.....................-------------------T--G-----C-----------------CA--G------...--------A--C-----------T--A--A--T-T-------C----- 7882
H1D.CD.84.84ZR085         -----T----------------T-------G--.....................--------------------A--------------------------CA-------GT...-C-----------G---------T--CA----------------------A- 8027
H1F1.BE.93.VI850          -----T--G--A--A-------TTA------G-.....................-----------------------------C-------G---------CA--G--A---...------G-G--C------C----T--A--T---GC------------A- 7783
H1G.SE.93.SE6165          -----T--G--A----------TTAC--ACCA-.....................-A-----A---------------------C-------G-----C--CA--G----G-...----AG-G---------G----T--AC-G-----------C----- 7924
H1H.CF.90.056             -----T--G--T----------TTG----G-A-.....................--AC-------------------------C--------G-----C--CA----------...------G-GA---------------T--AC--G--G---------C--C-- 7825
H1J.SE.93.SE7887          -----T--G--A----------TTA------A-.....................--A-C--A-G-----------G------C--------G---------CA--G---AC-...-----G-----------------T-C-A--T---GC--------C----- 7813
H1K.CM.96.MP535           -----T--G--A----------TTA--------.....................T--------G-A-----A-----------C-----------------CA----A-A--...-----AG-----------G---C-T--A--G--GC--------------- 7666
H101_AE.TH.90.CM240       -----T--G--T--C-------CTTC--ATCAT.....................-A--A--A-----------A-----------C-------G-----C--CA--G-------...-----AG-G---------G----T--A-----GC-----------A-- 8059
H102_AG.NG.x.IBNG         -----T--G-----C-------TTAC--ACCA-.....................-A-----A---------A-----A---C-------G---------CA----A----...------G-G----------G----T--A-----GC--------------- 8012
H103_AB.RU.97.KAL153_2    -----------A----------G-------G--.....................-A----------------------------------------A----------------------...-C-----------------A--T--A------------------C---A- 7683
H104_cpx.CY.94.CY032      -----T--G--T--G-------TTA-------AACC..................-AAC------T-----------G-----C----------------C---CA-------AG-...-------------C-------------T---C-----------------C----- 7879
H1N.CM.95.YBF30           --T--CC-T-----G-------TTA-------A.....................G-A--------A--------A---A--C----G--GC-T------CA-------G-...-----G-G---------G----T--CA--T---G------G--C--C-- 8057
H1O.BE.87.ANT70           CA---CC-C-----A----T--C-AA--ATCA-.....................-AAGA--A-G-A-GA-C---A-G-A---C--G--G--G---------AGA--G---GCC-...--G-GG--A-CC-CGCC-C-A--G-T----C----GT-G-ACACG--C--CA- 8554
H1O.CM.91.MVP5180         CA---CC-C-----G----T--CTG-----CA-.....................-G-CA--A-G-A--A-C---A-G-A---C--G--------------AG-------GCC-...-AG-GG-CAGCC--GCCACCA---T----CA--AGT-G-ACACG-----CA- 8580
H1P.FR.06.RBF168          -----TG-------A-------TTGG--TG-A-.....................GG-GATCC-G-A-GA-TAG-ACC-----C-A-C-------------AGC--G--AC-G-...-----G--CA---GC--G------T----C-T--CG-----------C---AA 8036
CPZ.CM.01.SIVcpzCAM13     -----CC-C-----G-------TTA-----G--.....................-A-----A---A-GA-C--T--G--A--C--GC-G-----------CC--G---C-G-...------G-GA-------A-CA--G-G--G-C-G--C-------------C--C-- 8090
CPZ.CM.05.SIVcpzEK505     --T--CC-C-----G-------TTA-------G.....................A-TC-------A-------A---A---C---------C-C------CT----AC-G-...------G-G--C------G----T--A--T---GCG----G--CT-CA- 7992
CPZ.CM.05.SIVcpzLB7       -----TC-G--A----------TTTC---C-T-.....................----------T--CTCA----G-----C-----------------CA--G--C---...-----AG-GA---GC--G------T--A-----GCT--A----C---A- 8017
CPZ.CM.05.SIVcpzMB66      -----TC-G--A----------TTA-----G--.....................-G-----AT-GA--------A---A--C-------G----------CC----AC-T-...-----A----C-------G----T--A-----GC---A----C--A-- 8130
CPZ.CM.05.SIVcpzMT145     -----T--G--A--G-------TTA-----G--.....................AATCA------A---G---T--G--A-C----A--G-----------AC-T----G-...-----A--CA---G----G---G-T----C-T--G-------A-----T-- 8018
CPZ.CM.-.SIVcpzMB897      -----T--G-----C-------CTT-----GT-.....................-G-----AC--A--------A--------C-------G--------A-AC---AC-T-...-----AG-G--C-------G----T--A-----GC---------CT--A- 8020
CPZ.CM.-.SIVcpzDP943      --T--CC-C-----G-------TTA----G--A.....................A-A--------A--------A-G-----C-----------------CC--G----G-...-C----CG--C-------G----T-----T---GC---------------C-- 8072
CPZ.CM.98.CAM3            --T--CC-C-----G-------TTA----G--A.....................A-A----A---A--------A-G----C---------C-T-------CC--G--A--G-...-----------C-------G----T--A--T---------A----C-- 7922
CPZ.CM.98.CAM5            --T--CC-C-----G-------TTA----G--A.....................A-A----A---A-------GA-G----C---------C---------CC--G----G-...-------CGA---G--C-G---G-T--A--T--G-------A-----T-- 8228
CPZ.GA.88.GAB1            -----TC-C-----G-------TTA-----GT-.....................-A-----A--AA-GA----T--G--A---C--C----G------C---CA-------AG-...------AG-GA--------G-A--G-G---C----------------C--C-- 8546
CPZ.GA.88.GAB2            -----T----------------TTA-----G-A.....................A-AG-------A--AGT---A--------C-G--G--G----------CA-------G-...-----TG-GA------C-G-----TT--A--T--CTG---------CA-C-- 7893
CPZ.US.85.CPZUS           --T--CC-C-----A-------TTT-------A.....................A-TC---A---A-------A-------C--------GC-C---A-A-C----AC-T-...------A-CGA-------C-G-----T--A-------G-------A-----T-- 8514
CPZ.CD.90.ANT             CAT--TC-G--A-----......A---T---CAAAAC.................-A-CA--AT--A--GCA---A---A---AG-------------AGA-A------GAT-...--G-GG-GGGCC--GCA-C----G-T---C-----CT-G----TG--C---AC 7965
CPZ.TZ.00.TAN1            A----CC---TT-CA----T--CTAC--A-G-G.....................GA-CA--AT--A--ACA---A-G------C-C--G--G--------GC------ACAT-...--G-GG-CG-CC-CGCCAGCA---T---CAGTA-CG-------G--C--CA- 8162
CPZ.TZ.2001.TAN2          ATG--C--G-TT-CA-------CTAC--A-G-G.....................-A-CA--AT--A--ACA-----G------C-C----G--------GC---G--ACAT-...--G-GG-CG-CC-CGCCA-CA---T---CAGTA-CG-------G--C--CA- 8182
CPZ.TZ.2002.TAN3_1        ATG--C--G-TT-CA-------CTAC--A-G-G.....................GA-CA--AT--A--GCA---A-G----CC-C--G--G-------GCG---G--ACTT-...--G-GG-C--C--CGCCA-CA---T---CAGTA-CG------G--C--CA- 8236
GOR.CM.2007.SIVgor2139_287 -----TC-G--A-------T--TTGG--GG-A-.....................-A-GTCCC-G-A-GA-TAGTACC--A-----G-------------AG-GC-----AC----...-----G--CA---GC--GCA---T---AC-T--C------------A-CA- 8073
GOR.CM.2007.SIVgorCP2135con -----TC-G--A-------T--TTGC--GG-A-.....................-A-GTCCC-G-A-GA-TAGTACC--A-----G-------------AG-GC------C----...-----G--CA---GC--GCA---T----C-G--C----------------CA- 8057
GOR.CM.2004.SIVgorCP684con -----T--G--A-------T--TTGG--GG-A-.....................-A-GTCCC-G-A-GA-TAGTACC--A-----G-------------AG-GC------C----...-----G--CA---GC--GCA---T----C-T--C----------------CA- 8004
H2A.DE.x.BEN              AGG--TG-T-TC-C-TCCC---C-GGTTATCT-CAACAGATCCATATCCACAAG......-AC-GG-GACA---A-CCAACGA--G-AC-----AA-AC-TCG---G-GACAGT...G-T-A-GACTTG-GGCCTTGGCC-ATAAACTATG-GCAG-TCCTGATC-AC-T 8982
H2A.GW.x.ALI              AGG--TG-T-TC-CCTCCC-T-C-GGTTAC-T-CAACAGATCCATATCCACAAG......-AC-AG--ACA---AACCA--GG---AC-----AA-AC-TTG------AC-TT...G-GGA--GATTG-GGCCCTGGCCGAT-GCATATT-ACAT-TCCTGATC-AC-T 8966
H2A.SN.x.ST               AGG--TG-T-TC-C-TCCC---C-GCTTACTT-CAACAGATCCATATCCACAAG......-AC-GG--ACA---A-CCA--GA---AC----AA-AC-TTG--A----C-TT...G--GA--A-T-G-GGCCCTGGCCGATAAGATATA-ACAT-TCCTGATC-GC-A 8412
H2B.CI.x.EHO              AGG---G---TC-CCTCCC-T-C-TCTTAT--TCAACAGATCCCTATCCGCAAG......-AC-GG-GACA---A-CCAACGA---AC-----AA----GTG--A---AC--G...G-C-A--GATCT-GGCCCTGGCAGAT-GAATACA-CCA--TCCCGAT--GC-A 8944
H2B.GH.86.D205            AGG---G---TC-CCTCCC-T-C-TCTTAT--TCAACAGATCCCTATCCGCAAG......-AC-GG-GACA---A-CCAACGA---AC-----AA----GTG--A--GAC-GG...GAC-A--GATCC-GGCCCTGGCAGATAGAATACA--CAT-TTCTACTG-GC-A 8958
H2G.CI.92.ABT96           AGG--TG-T-YT-C-TCCC-T-C-TCTTATGTTCAGCAGATCCATATCCGCAAG......-AC-AG--ACW---AACCAA-GA---AT------A-A-AGTG--A---A--GT...G-C-A--GATCT-GGCCTTGGCAGATAGAATATA--CAT-TCCTGMTC-GC-A 8355
H2U.FR.96.12034           AGG---G-T-TC-C-TCCCA--C-TCTTATGTTGAACAGATCCCTATCCAACGG......-AC-AG--ACA---AACCAA-GG---AT-----AA--C-CTG-------C-GT...G-C-A--GATCT-GGCCTTGGCAGATAAGATATA-ACAT-TCCTGATC-AC-A 8466
SMM.SL.92.SL92B           AGG--TG-T-TC-CCTCCC---CATCTTATCATCAGCAGATCCATATCCAGCGG......-AC-AG--ACT---A-CCAA-GA--G-AC------A-A--GTG--A--G---GT...G---A--GATTG-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-A 8375
SMM.US.x.H9               AGG---G-G-TC-C-TCCC-T-C-GCTTRTGTTYAGCAGATCCCTATCYARACG......--C-AG--ACT---AACCAA-GA--G---------A-AC-GTG---R-----GT...G-CAA--GATCT-GGCCTTRGCAGATAGAATATA--CAT-TCCTGATC-GC-A 8416
SMM.US.x.PGM53            AGG---G-G-TC-C-TCCC-T-C-G-TTATGTTCAGCAGATCCCTATCCACAAG......--C-AG--ACA---AACCAA-GA--G---------A-AC-GTG---------GT...G-CAA--GATCT-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-A 8873
SMM.US.x.SIVsmH635F_L3    AGG---G-G-TC-C--CCC-T-C-GCTTATGTTCAGCAGATCCCTATCCACAAG......-AC-AG--ACC---AACCAA-GA--G---------A-A--GTG---------GT...G-CAG--GATCT-GGCCTTGGCAGATAGAATATA--CAT-TCCTAATC-GC-A 8949
                                           Tat Rev intron end Tat Rev exon 2 start                                                                 Tat end

MAC.US.x.239              AGG---G-G-TC-C-TCCC-A-C-TCTTATTT-CAGCAGACCCATATCCAACAG......-AC--G-CACT---AACCA--GA--GCA----A-A-AC-GTG---AG-C-GT...G-CAA--GCTCC-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-A 8927
Env                       R P V F S S P P S Y F Q Q T H I Q Q . . D P A L P T R E G K E R D G G E G G . G N S S W P W Q I E Y I H F L I R Q
Tat exon 2                                         P I S N R . . T R H C Q P E K A K K E T V E K A V . A T A P G L G R *
Rev exon 2                                        N P Y P T . . G P G T A N Q R R Q R K R R W R R R . W Q Q L L A L A D R I Y S F P D P P
MAC.US.x.251_1A11         AGG---G-G-TC-C-TCCC-A-C-TCTTATTT-CAGTAGACCCATATCCAACAG......-AC-AG-CACT---AACCA--GA--GCA------A-AC-GTG---AG-C-GT...G-CAA--GCTCC-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-A 8923
MAC.US.x.251_BK28         AGG---G-G-TC-C-TCCC-A-C-TCTTATTT-CAGTAGACTCATACCCAACAG......-AC--G-CACT---AACCA--GA--GCA------A-AC-GTG---AG-C-GT...G-CAA--GCTCC-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-A 8909
MAC.US.x.EMBL_3           AGG---G-G-TC-C-TCCC-A-C-TCTTATTT-CAGNNNACTCATACCCAACAG......-AC--G-CACT---AACCA--GA--GCA------A-AC-GTG---AG-C-GT...G-CAA--GCTCC-GGCCTTGGCAGATAGAATATA--CAT-TCCTGATC-GC-A 8058
STM.US.x.STM              CGG--TG-G-TC-C-TCCC-T-C-TCTTGTCGTCAGCAGATCCCTATCCACAAG......--C-AG--ACA---AACCAA-GA--G-AC-----AA---GTG---------GT...G--AT--ACTCC-GGCCTTGGCA-ATAGAATATAC-CAT-TCCTGATC-GC-A 8578
MNE.US.x.MNE027           AGG---G-G-TC-C-TCCC-A-CTTCTTATTT-CAGCAGACCCATATCCGACAG......-AC-AA-CACT---AACCAA-GA--G-AC------A-AC-GTG-------C-GT...G-CAA--G-TCC-GGCCTTGGCAGATAGA-TATA--CAT-TCCTGATC-GC-A 8404
DEB.CM.99.CM40            --C--T--GCA-ACC-TT----CAAAG........................GA-GA--A-GTGCTTC-A-T--C--A-GA--G--GCAG-A------GAGAC-G---G-C-...--G-GG-CAGC---GCC--GA----T---CAGT--C--------G-----C-A 8228
DEB.CM.99.CM5             --C--T--GCA-ACC-TT----CAAAG........................GA-GT--A-GTGCTTCAA-T--G----GA--G---CA--A------AG-C-T--AGAG-...--G-GG-CGAC--CGCCC-GA----T----AGG--C-----------CA-CTC 8149
DEN.CD.x.CD1              --T--TC-T-T-CAGATCC-TAT--GA........................--CCA-AAGG-A-CA-AT-AGCC--A-GA--G-----G--AAA-T-GAGAC-G--ACAGGCTC-AG-GGTGG-A--ATCC--GA----T----AGTA-C--T-----A-----CTC 8566
LST.CD.88.447             AGGGTTC-TG-CCCCTCTG-TT-TG-TGA-CAGGACTGCAAGTGGGAGAAAAGA...GA-AAC-AA--ACA---A--CAA-GA--G---GAT--AGAA---C-A---A-ACAT-TAC-TCAA-T-AGA-CAGTAC--GAA-GAA-CTT-CAC-CCACC-TGGA-CG--GA 7689
LST.CD.88.485             AGGGTTC-TG-CCCCTCTG-TT-TG-TGA-CAGGACTACAAGTGGGAGAAAGGA...GA-AAC--A--ACA---A--CAA-GA--G---GAT--AGAA---C-A---AGACAT-TAC-TCAA-T-AGA-CAGTGC--GAA-GAA-CTT-CA--CCACC-TGGA-CG--GA 7686
LST.CD.88.524             AGGGTTC-CG-CCCCTCTG-TT-TG-TGA-CAGGACTTCAAGTGGGAGAACGGA...GA-A-C-AA--ACA---A--CAA-GA--G---GAC--AGAA---C-----A-ACAT-TAC-TCAA-T-AGAGCAGTGC--GAA-GAA-CTTTCAGGCCACC-TGGA-CG--GA 7692
LST.KE.x.lho7             AGGGTGC-C--CCCCTCTGTTT-TA-TGA-CAGGACTACAAGTGGGAGAAAGAA...GA-AAC-AA--ACA---A--CA--GA-----GA----A-C---T-C---G-CTAT-TAC-TCAA-T-AGAGCAGTGC--GAA-GAA-CTT-CAGGCCACT-TGGA-CG--GA 8764
SUN.GA.98.L14             AGGGTG--G--CCC-TCTG-TT-TG-TGA-CAGGATTGGCTCCAGGAAACCTGC...CCCAA----ACAGACAAA----A-GA-----GAC--AGAA---A---AT-......TAC-TCAA-T-AGAGCAGTCC--GAA-GAA-CTTTG-CACCACCTTGGACAG-AGA 8826
MND-1.GA.x.MNDGB1         CG---TG-C-TT-CA.....................CAGGACTGCCAGCAGAAC...CTATAC-G-A-AC----A--CAACGG---------A-CAACAGCTT---ACT--GA...GAGCA--ACTCCGAGAACTTG.........AAGGAAGAA-CCTTAA-CAGATC 8089
OLC.CI.97.97CI12          GTGT-T--GCA-C-C-CA..................TTTTCAGATACCCTC......CA--ACAAG--ACA---A--CAACGG----G-----ACA-C---TTCTT-GAGAG-......A--T-CAACGAGTG----T-TA--ACA-AGTG--CAG--ACATTGAGAGT 8112
WRC.CI.97.97CI14          --TAAGC-TATA-C----................CAGAACCACCTCAACAAC...TT--A-GAG-GACA--AA--CA-CGA-----G-AT--AA-A---TTACA----TATAGAA-TCAA-GA-T-GCAGTAC-GGA--GAA-CT---A-AG---TCTTGA-CAGCTT 8631
WRC.CI.98.98CI04          --TAAGC-TATA-C----................CAGAGCCACCTCAACAAC...TT--A-GAG-GACT---A--CA--GA-------C--AA-A---C-ACAGT-CTTTAGACGTCAG-GA-T-GCAGT-C--GAA-GAA-CT---A-A-C--T-TTGA-CAGCTT 8676
WRC.GM.05.Pbt_05GM_X02    --TAAGC-TGT--C----................CAGAGTCATATCCTGAAC...CT--A-GAG-GACT---A--CAACGA----AC-----AA-A---TTTCATAGCCATAGAAGA---AGACT-GCAGT-C--GAA-GAA-C-CTGA-AC---TCTGGA-CAGCTT 7572
MND-2.CM.98.CM16          CGC--TC-T.........C-TT-TG--TTC-AGGGAGACTATCTCCGTCCCCAC.....AAC-T-A-ACA---A-GCAA-GA-AG-------AGCCC--CTC---A-A-C-G...--TA---AATC-GACAGCTC-AA-GTAGAATTTGGG-AGCCTTGGAGCAAAGA 8650
MND-2.GA.x.M14            CGC--TC-T...CC-T-TGT-TTAAAGGG-GA-TATCTCCGACCTCGC...........AAC-T-AGACA---A--C-A-GA-AG----G--AGCTCA-CTT---AGA-C-G...-ACAT--AGTC-GAGAGCTC-AAGAAAGAATTTGGC-GACCTTGGAGG-CAGA 8579
MND-2.x.x.5440            CGC--TC-T...CC-T-TGT-TT-AAGGG-GA-TATCTCCGGCCCCAC...........AAC-T-A-AC----A--CA--GA--G--G-----AGCC---CTT---A-A-C-G...-ACAT--AGTC-GAGAGCTC-A--CAAGAAT-TAGG-AGCCTTGGA-G-CAGA 8286
DRL.x.x.FAO               CGC--TC-T...CC-T-T-T-TTTAAGGACG-ATACCGCCCAGAGACG......A-C-ACTAC-A-AGACA---A--CA--GA-A---G-----AGCA---C----G-GAGTGG...-ACAT--GATC-GA-CCCTC-AAGC-AGGAT-CTCC-AAGCTTGGAGCGAAGA 8347
MON.CM.99.L1              --T-TTC-T--T-C-----T--CTA-----T--......................-ACGC-----AGCCA-C------C----C--GC-C--GC-----C--CG-CAG--ACTC-...-----GCCAGC--ATC----A----TT--TA-CA----------A-----CA- 8466
MON.NG.x.NG1              --T-TCC-C--TCCA---.........ATCC-TACCCAAGTCCCTCGG...........-A---A-GAGC---A----A--C--G-----GC---------CA--G---GA--AGGTCRA-AG-CTATGCGAGAGGATTT-TAACCATTA--TGGACA---TGC-GAGA 7094
GSN.CM.99.CN166           -----TC-C--CCCG----T--TTA-----CAA....................G-AGT--A--G--GTCC---A--C-C--CC-------GC-T---C-C-CA--G--AG-T-...--G---G-GA---ATC-A-CA----T--C-AGT--CT-G---------C---A- 8440
GSN.CM.99.CN71            -----TC-C--CCCG----T--TTAC----CAG....................-AAGT--A--G---TCC---A--C----C-AG-A----C-----C-C-CA--G--AC-T-...--G---G-CA---ATC-C-CA----T--C-AGT--CT-G---------C---A- 8452
TAL.CM.00.266             -T---TC-TC-----A-----CTA----G-G-...................AAC----AT--A--AGC---A-G----GAC-CTAC---------A--CA--G---CT--...-AG-GG-ACA---CGCAAGCA----T---AGCA----T-----A------AC 8333
TAL.CM.01.8023            -T---TC--C-C---A-----CTA----G-G-...................AAC----A---A-GAGT---A----A-GA--GTAC---------A--CCC-----GAG-...--G-GG-GAAC--CGCCTC-A----T---T--CA---------G--C--CAA 7883
MUS-1.CM.01.1085          --T--CC-T--TCC-----T--TTA---G-C-G..................-A--C------A--AGC---A-------C-----G---C-----AGCG--AG-------...-------G-GA---ATC---G-----T---CT-T--CT---------C--C-- 8437
MUS-1.CM.01.CM1239        --T--CC-T--TCCA---T---TAC--A-C-T..................--AGA--C-G-A--GCC---A-------C-----G--GC-T---CAGC--C-G---GAC-...-------G-G----ATC---------T----T-G--GT-T--------C--C-- 8406
MUS-2.CM.01.CM1246        --T--CC-C--TCCA---T---TAC--G-..............................-AGTC--C-GA-C----GC--CAA-A-A-A-C---A-A--G-...GCTGGA-GAA-CGGCAG-C--A--------T-G--GT-T-CA--A--C--C-A 8514
MUS-2.CM.01.CM2500        --T--TC-TC-TCCA----T---TAC--G-G--...................---GA-----AAGA-C-------A--C-------G-----C-CC----G---T-T-...-------G-G--C-ATC---------T----T-T-------------C--C-- 8458
SYK.KE.x.KE51             --T--TC-T-T-CAGA--C-G--G-C---T---.....................CTCTAC-A-A-GGCC--A-C---GCC--------G-ATA-A-ACAGATACAGACA-T...TACAAGCCCA-CG-CTCT-CGTTCAG-AGA-A-T--T-GA-AC-ATCG--CGC 8169
SYK.KE.x.SYK173           GTT--TA-T-T-CAGA-CC-TAC-............................CA---C-GGA-GGAC--A-GCAA-CCC-C---C----AA-A--GC--T-GAGACAGA...GA-GGAC-CAACG-CTCT-C-TTCAG-AGA-A-TC-T-GA--C-ATC-T--GA 8470
RCM.GA.x.GAB1             AGG--TC-C......TCCC-A-C-TCTTATGTTCAGCAGATCCATATCCACAAC...-C---GGAACCGCAAA-TCC----GA-A--AG-----AC--T-GAGA---AG-T-G-......AA--AGTACAACAATTGGCTGAGAGAA................TATT- 8169
RCM.NG.x.NG411            A-G--T--GCTC-C--CCC-A-CTTGTTATCATCAGCAGATCCCTATCCGCAAT.....C-GGAG--ACA---AATCAA-GA--G----A----A--CAGTG----AG-C-GT...G--CA--AATAT-ACAACTGGCAGAGAGGA................TATTT 8219
GRV.ET.x.GRI_677          AAT--TC-CC-C.....................CAGATCCATATCCACAGT.....TC-G-G--AC----A--CAACGG--G--GGC---ACA---GTG---A--CAG-AGC--CAAAT-GAT-AG-T--C-G-A-GAG-CCT--ACACCT-C-AGGATCAACAA 8492
SAB.SN.x.SAB1C            G----TC-C--C-CC-TCC--TCTTCT-ATTAA...CAGATCCATATCCACCTG.....-ACAAG-GACA---A--CAA-GA------GC----CA-C-GTG-GCTA-ACAG-GAC--CAGGTCATACACTTG-C-GA--GAA--CTTG.................... 8787
TAN.UG.x.TAN1             -TT--T--G--TCC-....................CAGATCAATATCCGCCTG......--GAGG--ACA---A--CAACGC--G--G-----ACAA--TTCCAG---CAG-...----GA--AGTCGCCGCCCTCA-TGAAAGAAT-TT-AT-GCCCA--AGAGGAG- 8590
VER.DE.x.AGM3             --T--TC-T--TCCA..................CAGATCCATATCCACCCG.....T-GAAG-GACA---A--CAACGC-----GGCC---A-A--GTG-----AGCG-AAG-ACAG-TCCGAGCCTTG-C-GAA-GAA-CT-GCACAGCAGA-TGGA-GAGCAA 8137
VER.KE.x.9063             --T--C-----TCCA..................CAGATCCATATCCACCCG.....T-GAAG-GAC----A-GCAACGC--G---GCC--AA-A--GTG---G-----AACGACAG-TCAGAT-CTTG-C-GAA-GGA-CT-G-ACG-GACAAA-GAGAGGCA- 8639
VER.KE.x.AGM155           --T--T-----TCC-..................CAGATCCATATCCATCCG.....T-GAAG-GACA---A--CAACGC--G---GCC--AA-A--GTG--AGA-C--G-AAA--CAAATC-ACGCATTA-C-GAA-GAA--T-GGGGACGAGACA-GAGGACCA- 8616
VER.KE.x.TYO1_patent      GTT--TC----TCCA..................CAGATCCATATCCACCAA...GT---GAAG-GAC----A--CAACGCC--C--GCC---A-A--GTG-----ACAG-AGG-TCAAAT-AGAG-CTTG-T-GAA-GA--CCAAGAGC-GG--TATGC-GT--AC 8618
COL.CM.x.CGU1             GG-TTTCAGCA-A-AGGC-GGACTAATGTGT-TTCA.................-A--TCC-TGG-C-AGAC-GG--GT-GGA-C-GCC---CAAC-C---AGA-ATACACAGA...GAGCAAGAAAACAGCAGC-G-A---TT-G-..................... 7887
```

```
                                                                                                                          Rev end (TAA) in some lineages
H1B.FR.83.HXB2            GAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACTTCTGGGA....................CGCAGGGGGTGGGAAGCCCTCAAATATTGG................................TGGAATCTCC 8622
Rev exon 2                 E  P  V  P  L  Q  L  P  P  L  E  R  L  T  L  D  C  N  E  D  C  G  T  S  G  . . . . . . . . . . T  Q  G  V  G  S  P  Q  I  L  . . . . . . . . . . . . . . . . V  E  S  P
Env                         S  L  C  L  F  S  Y  H  R  L  R  D  L  L  L  I  V  T  R  I  V  E  L  L  G  . . . . . . . . R  R  G  W  E  A  L  K  Y  W  . . . . . . . . . . . . . . . . .  W  N  L
H1A1.UG.85.U455           --A-----------------------A----------CGC---------G----AGC---------------CGCAGCAGCCTCAAGGGACTGA-ACT---------G-G---------CT-.................................--------T- 8073
H1B.US.90.WEAU160         ----------------CT--------------TC-------------C--A-------------------------...................................................T--------T...........................---------- 8629
H1C.ET.86.ETH2220         ------------------T--------------------A--------C-G------CA---------------CGCAGCAGTCTCAAGGGACTA-AG------------A----T------CT-...............................G-A-GC--TG 8019
H1D.CD.84.84ZR085         --A-------------------------------A---A----A----C-G-----------------------.......................-------------------------CT-.................................--------- 8143
H1F1.BE.93.VI850          --A-----------------G--A-----------CA-A--A----C-G------------CAGGGGCCTG....................A-G---------------T---------CT-...............................G--------A 7899
H1G.SE.93.SE6165          -----------------------------------CCA----------G------CA---------------CGCAGCAGCCTCAAGGGACTGA-ACT-----------GT--T-------T-.................................--------T- 8061
H1H.CF.90.056             --------CA------------G--T-----------------A------GT-----C--------------G.....................A-A------A-----------------CTC...............................--------T- 7941
H1J.SE.93.SE7887          --A--------------T--------A-----------CG--------C-G------C-----G-AC-------.....................-T------------GAT---------CT-...............................GT---C--TG 7929
H1K.CM.96.MP535           --A-----------------G--AA------A-----A---A-----G--------CC-----AGGGGACTG....................A-GG----T----------------CT-.................................-----C--TG 7782
H101_AE.TH.90.CM240       --------------------------G-------------AC--------C-G------CG----------------CACAGCAGTCTCAAGGGACTGA-AC-----------G-----------CT-...............................G--------T- 8196
H102_AG.NG.x.IBNG         --------------------------A------------A----------C-G------C-----------------CACAACTGTCTCAAGGGACTGA-ACT--------G-------------CT-.................................--------TA 8149
H103_AB.RU.97.KAL153_2    ---------T------T-------A-----------------------C-G-----------------------.......................................................-.................................---------- 7799
H104_cpx.CY.94.CY032      --A-------------------G--A------A-----------A-----G------C---------------G.....................ATA-------------------G---CT-...............................-----CT--- 7995
H1N.CM.95.YBF30           --A-----TGA-----CT--------------C-----C---------AC-G-G-----C-C-------C------CAGAGTCTCAGCAGGGGACTG-AACTACT-AAT---CT-AGA-C-C-C-T-..........................---GGAA-A- 8194
H1O.BE.87.ANT70           A-CAA-AATAT-G-GG-CT-------T------CA-----GCA-CAGGGA-CCA--A-G-GA-CAGCTA----A-GCTTGGACTGTGGATCCTAGGG-AG-A-ATAATTA-T-TTTG--G-AT---T..........................GCAGC-G-AA 8691
H1O.CM.91.MVP5180         --CAA-AAT-T-G-GG-CT-------T------CA-----A-A-CAGGGA-CCG----C-GA-C--CTAC------CTGGGACTGTGGATCCTGGGA-AA-A-ACAATT-----TTGT-G-CT---T..........................G-AGC-G-AA 8717
H1P.FR.06.RBF168          --A----GTGG-GCAG-T------GAT---AGTC-G--GCA-A---GGGA-C-A-GACC-CC-AAC-A-C---T-GATACACCTGGGCCAACTGCTGAC---A---CTTA-CTG---GCG-G-C--C...........................TTGC-GCGT 8173
CPZ.CM.01.SIVcpzCAM13     C-A----GT-A---GG-------GGA------C-AG----GC--GC-----TTG---ACAGT-----AAGT-----CAACTGTTAATCAACATCCTCA-ACT-ATACA---GAAA--G-CTCTACT-..........................A--GGGA-AA 8227
CPZ.CM.05.SIVcpzEK505     --A---TCTA------CT-------------C-----GT----CA--CC-G-G-----CCT-----------A--CAGAACATCCACAAGGGACTG-AACTACTTAAT---CT-AGA-T-----T-..........................---GGAA-AA 8129
CPZ.CM.05.SIVcpzLB7       -GA-------------CTT--------A----C-------G-----A--A-CG-G-----CAT----GA--G-----CAATACACTCTGAAGGGACTGA-ACT--T------------ACT----CT-..........................CA-GGAA-T- 8154
CPZ.CM.05.SIVcpzMB66      -GA------T------CTT--------A-----------------A--GCT-----CAT-------GG-----CAGACCCTCCTCAAGGGACTGA-AC----AC----------A-T-C-CCT-..........................A--GGCA-T- 8267
CPZ.CM.05.SIVcpzMT145     --A---CCT-AG----CT---------AA----------GTGCA-CTC-GA-TTG---AC-AC----GA-A------CAATACAGCCTCAGGGGAGTA-AACAAATAG-TACTCT---AC-CC-ACT-..........................A-AGGGAC-- 8155
CPZ.CM.-.SIVcpzMB897      --A-----------------------------------G---A--G-G-----CGC----GA--G-----CAGAAGTTCCTYGCAGGACTGA-ACT--T-------A----A-C---C-T-..........................AA--GCA-T- 8157
CPZ.CM.-.SIVcpzDP943      --A---CCTA------CTT--------A----T-----------CA--CC-G-G-----CAT-----A--------CAGAACATCAACAAAGGACTG-AGT--CT-A----CAT-AG--G-C-CCT-..........................---GGAG-GA 8209
CPZ.CM.98.CAM3            --A---CCTG------T-T--------A---------------A-----AC-GGG-----CAT-----AAC------CAGAGCCTCAACAAGGGACTG-AACA-CT-A--A-CTT-AG--G---CCT-..........................---GGAG-GA 8059
CPZ.CM.98.CAM5            --A---CCTGA-----TT---------A----T-----------CA--CC-GTG-----CA---C--A---------CAGAGCATCAACAAGGGACTG-AACA-TT-AA---CTTTAGT-C---CCTA...........................AA-GGAG-A- 8365
CPZ.GA.88.GAB1            A-A----G-GA-T-GG--------GA------C-AG----GC--GC-A---GTG---ACAAC-CA-GACGT-----CATCTAATACTTCACAGCCTCA-ACT-CTAC----GAGA--GTGCCTGCT-..........................G--GGAA-TA 8683
CPZ.GA.88.GAB2            --A----ACGA-----CT------G--T---------T-GT----C-G-CC-TTG---CC-GC------AAT-----CAACAGGTCCTGAGGAGGACTGA-ACTACTCA----GCTT--G-GCC-GCT-..........................AAAGGGA-AG 8030
CPZ.US.85.CPZUS           --A---CCTGA-----CT----------A---GAG--------A--A---C-G-G-----CA--AC--A-C------CAGAACATCAACAAGGGACTG-AACT-TT-AAT---CT-AGAGC-CGC--C..........................----GGGG-AA 8651
CPZ.CD.90.ANT             ----A-AAT--AG-GG-T------GAT--GC---AC--GT-------ACC-TTG-GCAG-CC-CC---AC--CT-CAGA...................ATT-CTTTCA-ACTGTG-AA-C-TCTGGA-..........................AAC------A 8084
CPZ.TZ.00.TAN1            --A---CCT-A---GG-TA-----GAC---TCA-A----CA---G---CC-CTG--TC-GCC--C--GCA---AA-CAGGGGATAATCAGCTTGGCA-A---CCTAGTAAT--TG-AT-G-AC-ATCATAGTAGGAGTTAGACAGATC...........ATTG-GTGGA 8320
CPZ.TZ.2001.TAN2          --A---CCT-A---GGCT------GAC--GTC--A----CA---G-G-CC-GTG--C---CC--C--GCA---AA-CAGGGGACAATCAGCCTAGCAAA--ACCTAGTAAT--TG-AT-G----ATAGTAGTAAAAATTAGACAAATT...........ATTG-GTGGT 8340
CPZ.TZ.2002.TAN3_1        --A---CCT-A---GG-T------GAC--GTCA-A----CA---G-G-CC-GTG--CA--CC--C--GCG---AA-CAGGGGACAATCAGTCTAGCA-A--ACCTAGTAAT--TG-AT-GG---ATAATAGTAAGAGTTAGACAGACT...........ATTG-GTGGT 8394
GOR.CM.2007.SIVgor2139_287 A-C----GTTG-G-GG-T-----GGACA--AGTCAT------A-C-GGGA-C-A-GA---AACACTCAGC---ATCGAGTACCTGACCAGACTACTGA-A--A-T-AATA-TTTGACA-G-G-C--C.............................-TTGC-T-TA 8210
GOR.CM.2007.SIVgorCP2135con A-C----ATTG-G-GG-T-----GGACA--AGTCAT------A-C-GGGA-C-A-GAA--AACACTCAGC---ATCGAGCACCTGACCAGACTACTGA-A--A-T-AATA-TTTGACA-G-G-C--C.............................-TTGC-T-TA 8194
GOR.CM.2004.SIVgorCP684con A-C----GTTG-G-GG-T-----GGACA--AGTCAT------A-C-GGGA-C-A-GA---AACACTCAGC---ATCGAGYACCTGACCAGACTACTGARA--A-T-AATA-TTTGACA-G-G-C--C.............................-TTGC-T-TA 8141
H2A.DE.x.BEN              ACTGACTC-------G-T-GGG-TATA-AAC-TCTG-AG-GA---AC-ATCC-A--ACTCCCC-...........ACCCGCCGACTGATCTCCCAGA-TCTAACAGCAATCAGGGA-TGGCTGA-ACTTAAG..........................GC-GCC-AA- 9113
H2A.GW.x.ALI              -CTAGCTC-------G-T-GGG-TGTA-AGC-TCTG-AGGGA---AC-ATCC-G--TCTCCCC-ATC--C-AACCGATCTTCCGGAGTCTTCAGAGAGCGCT-ACAACAATCAGGGA-TGGCTGA-ACTTAAA..........................GCAGCCTA-- 9109
H2A.SN.x.ST               -CTGA-TC-------G-A-AGA-TATA-AAC-TCTG-AGGGA---AC-ATCC-G---CT-CCA-ACC--C-AACT-ATCTCCCAGAGTCTTCGGAGAGCATT-ACAGCA-TCAGAGA-TGGCTGA-ATTTAAC..........................ACAGCCTA-- 8555
H2B.CI.x.EHO              ACTGAG-GA------G-TT-GG-TATA-AGCG-CTG-AG-AC---AT-ATC--A--CCT-CCAAACC--C-AACC-GTACTCCAACCGCTCAGGCTT-CACCT-CA.............................................TA-- 9054
H2B.GH.86.D205            -CTGAG-AA------A-TT-GG-TATA-AACG-CTG-AG-AC---AC-GT-G-A--CCT-CCAA.......................ATCCTCCAC-AA-TCTCAACCA-TCT--AACC-CTCA--CTCCCG.......................GTTGCGTA-- 9077
H2G.CI.92.ABT96           -CTAGG-AA------G-CT-GG-TATA-AGC-ACTG-AG-GC---AC-ATCC-G--TCY-CCA-ACT--C-AACCGCTGTTCCAAAGAATCTCCAGGAC-CTCCAAGCAATTAGAGAAC-TCTCC--CTAGAA..........................GCAGC-TATT 8498
H2U.FR.96.12034          ACTGA-TC----G--G-CT-GG-TATA-AGC--CTG-AGGGA----C-ATCC-G--TTTGCCA-AGC--C-A-CCGTTGTTC.....................CA-A-CATCAGAGAACGTCTGCACCTTGAA..........................ATAGC-TA-- 8588
SMM.SL.92.SL92B           -CTGA-AC-GA-A--G-CT-GG-TATA-AAC-ATTTGACT-G---AGCATCC-GAGCCTACCA-A-C--C-AACA-CTGTGCCAGAGACTCTCAGAGAT---TCA-CC-ATTAGAGAGCTTGTCA-AAGAGAAGCAGGC.....................TATA 8518
SMM.US.x.H9               -CTGA-AC-------G-CT-GG-TATT-AGC--TTG-AGGGAT-G-C-AT-G-G--NCTGCCAAATC--C-A-CC-GTGCTCCAGAGCCTCTCAAGGACGCT-CA-A-A-CCCGTGAAGTCATCA-AGTTGAGATAACC.....................TA-- 8559
SMM.US.x.PGM53            -CTGA-TC-------G-CT-GG-TGTT-AAC-ACTG-AGGAC--ATC-CT-G-GA-TATACCA-ATC--C-AACC-GTGTTCCAGAGGCTC.........TTACA-A-AATCCGAGAAGTCGTCA--ACTGAG..........................ATATCCTA-- 9007
SMM.US.x.SIVsmH635F_L3    ACTGA-AC-------G-CT-GG-TATT-AGC--CTG-AGGGAT-G-C-AT-G-G--TCTACCA-ATC--C-AACC-GTGCTCCAGAGACTCTCAAGGACGCT-CAAA-A-TTCGTGAAGTCATCA-AATTGAA..........................ATA-CCTA-- 9092
                                                                                                                             Rev end
MAC.US.x.239              ACTGA-AC-------G-CT-GG-TATT-AGC-ACTG-AG-AC----C-ATCG-GAGTATACCA-ATC--C-AACC-ATACTCCAGAGGCTCTCTGCGAC-CTACA-A--ATTCGAGAAGTCCTCA--ACTGAA..........................CT--CCTA-- 9070
Env                         L  I  R  L  L  T  W  L  F  S  N  C  R  T  L  L  S  R  V  Y  Q  I  L  Q  P  I  L  Q  R  L  S  A  T  L  Q  R  I  R  E  V  L  R  T  E  . . . . . . . . . . . . . . L  T  Y
Rev exon 2                 T  D  T  P  L  D  L  A  I  Q  Q  L  Q  N  L  A  I  E  S  I  P  D  P  P  T  N  T  P  E  A  L  C  D  P  T  E  D  S  R  S  P  Q  D  *
MAC.US.x.251_1A11         ACTGA-AC-------G-CT-GG-TATT-AGC-ACTG-AG-AC----C-ATCG-GAGTATACCA-ATC--C-AACC-ATACTCCAGAGGCTCTCTGCGAC-CTAC-AA--ATTCGAGAAGTCCTCA--ACTGAA..........................CT--CCTA-- 9066
MAC.US.x.251_BK28         ACTGA-AC-------G-CT-GG-TATT-AGC-ACTG-AG-AC----C-ATCG-GAGCATACCA-ATC--C-AACC-ATACTCCAGAGGCTCTCTGCGAC-CTAC-AA---TTCGAGAAGTCCTCA--ACTGAA..........................CT--CCTA-- 9052
MAC.US.x.EMBL_3           ACTGA-AC-------G-CT-GG-TATT-AGC-ACTG-AG-AC----C-ATCG-GAGCATACCA-ATC--C-AACC-ATACTCCAGAGGCTCTCTGCGAC-CTAC-AA--ATTCGAGAAGTCCTCA--ACTGAACTGACC...................TA-- 8201
STM.US.x.STM              -CTAG-AC-------G-CT-GG-TGTA-AAC-ACTT-AGGGCT-GTC-ATCC-G--TCTACCA-ACC--C-ACCC-ACCTTCCAAAGGATCTCCAGGATACT-CA-A-AATTAGAGAAGTCGTCA-ACTAGGA..........................GCAGCCTATT 8721
MNE.US.x.MNE027           ACTGA-AC-------G-CT-GG-TATT-AGC-ACTG-AG-AC----C-ATCG-GAGCATACCA-ATC--C-AACC-ATATTCCAGAGATTCTCCACGAC-CTACA-A-A-TCCGAGAAGTCCTCA--ACTGAA..........................CTA-CCTA-- 8547
DEB.CM.99.CM40            ACAA---CTGACG-GGCT------GAC---TC-CA---GCA-T-G-C-GA-T-GA--AGGCT-T---T-A---CA-GAAGGAATATATAGACTGTCATTTGCTCTT---C--AGA--G-G-CTC-T-GGGGAACGCTTTAGAGTGCTT...........G--GGGTA-T 8386
DEB.CM.99.CM5             ACAG--TCTGACG-GGCT------AAC------GA---GCGC--G-C-GA-T-GA--AGGCT-T--G--C---CA-GAAACAATATATAGACTTTCATTTGCATTT---C--AGA--G-G-CTC-T-GGGGAACATTTAAGAGTGGTC...............---GG-TATT 8307
DEN.CD.x.CD1              ACAG---CTGACG-GG--T----TGAC-------A-----G-G--TC-GA-T-G--CCGCCTGTAGT-GGA-TT-GGACTTTGTA................................AT-GGC-CA-CTTAATTTTGCTTGGC.......................TA-T 8682
LST.CD.88.447             CT-GAGCGAG-CT--ACAAG--TCA-TG--TGTGACTC-C--GAA-TGGA-C-A-GCCT---GAATT-----CCTCAGCCTTGTGTGGCAGTTCCTTTCGT--CTAG-CC-TTTAG-G-TCCT--TCTTCCAA...CATGGGCAACATCTTCGGAAAGCC--C-GCAGA- 7856
LST.CD.88.485             CT-GAGCGAG-CT--ACAGG--TCA-TG--TGTGACTC-C--GAA-TGGA-C-A-GCAT---GAATT-----CCTCAGCCTTGTGTGGCAGTTCCTTTCGT--CTAG-CC-CTTAG-G-TCCT--TCTTCCAA...CATGGGCAACATCTTTGGAAAGCCA-C-GCGGAT 7853
LST.CD.88.524             CT-GAACGAG-CT--ACAGG--TCA-TG--TGTGACTC-C--GAA-TGGC-G-A-GC-TGG-GAAT---G--CCTCACCCTCGTGTGGCAATTCCTCACTT--CT-G-AC-TATAG-G-TCCT--CTTACCAC...CATGGGCAACATCTTTGGCAAACCCTC-GCAGAT 7859
LST.KE.x.lho7             CT-GAACGAG-CT--ACAAG--TCG-TG--AGTGACTC-C--GAA-TGGC-C-A-GAAGGG-GAAT---C--CCT-TCACTAGTGTGGCAGAGCCTCA-TT--CT----C-TTTGT-G-TCCT--TCTTCCAA...AATGGGCAACGCCTTTGGCAGACCA-C-GCAGAT 8931
SUN.GA.98.L14             CT-GGACGAG-CT--G--AG--TCT-TG--AGTGACTC-T---AA-TGGC-G-AAGCTTGG-GAAT-G-C--A-C-CAAAACATCTACCACCTGCTAA-TTTCTT----C-CCTA--T-CGAC--CCTTCCAT...CATGGGCAACGCCTTTGGCAGACCCTC-GAGGTT 8993
MND-1.GA.x.MNDGB1         CTTGA-AGAGGA-C-G-C-AG-TTTGC-AG-GAGTGACC-A-T-G-C--TGGCT--AA-A-C--AG-GCAGCAATTGAATATGGGTTC..............................CTCGCAGTCCAAGAAGCG.........ATCAGAAGC-T-GG-TCG- 8214
OLC.CI.97.97CI12          A-ATTG-A-AGA---ACTA-CA----TG-G--AG-GAG--G-GAGAG-CACTCGT..................................................................................................................T 8168
WRC.CI.97.97CI14          -CTAACC-CA--A------ATTTG--AACACCTCCTTCACA-TG--TGGAGTCT-CTTTCGTGCCTC--GTG-T-GATAAACGATGGGTGG.....................AGTCTTCTCAAAGACATC.............................................. 8740
WRC.CI.98.98CI04          --TAAAC-T-T-A-CTGCAGT-TG--AAACCCTCCTTCA-A-AG--TGGAGTCTCCTTGCGTGCCTCT-CTG-T-GATCCAGCATGGGTGG.....................AATCTTCTCAAAAGAATC.............................................. 8785
WRC.GM.05.Pbt_05GM_X02    -GCAACC-TA--T-A----ATTTG--TGAGCCTCCTGCAG--AAGCTACAGCCTTCT--GCTATCTGTGGA-A---ATACAGAATGGGTGG.....................AGTCTTCTCAAAAGACTC.............................................. 7681
MND-2.CM.98.CM16          -.........AAA----AG-TTGG-TGAA--CTTCAAGGGG--ACG-GTGGCT--A--A-C-AC--GCAG-AATTGAATATGGGTGGCAAGAG................................................................................---A 8739
MND-2.GA.x.M14            -.........AAA-A--GA-TTGGTTGAA-GAGTCAAC-A-T-AC---TGGCT--A--A-C--C-GGCAG-AATTGAATATGGGTGGCAAGAG................................................................................---A 8668
MND-2.x.x.5440            -.........AAG----GAG-TGG-TGAAA--GTCGACG--T-AC---TGGCT--A--ACC-CC--GCAG-AATTGAATATGGGTGGCAAGAG................................................................................---A 8375
DRL.x.x.FAO               A......ACAG-AGGG-C--GGTTGAAGGAATCGAGAGG-TA-A-CTGGC-G-A--ACG-GAAA-C-G----T-A-TATGGGTGGTCAGAA...................................................................................---- 8436
MON.CM.99.L1              --A---TGT------GG-CA-----GAT-----A-----C-G-A--AG-GA-CTACC-----CC-CC-GAGAG--TCCCAGAGGCTGCCACCCCTCCTG-A--TAC-CCT-CTCCAA--GTGGG-AA-CCTTAGACGCTTGCTAGCCTACTGCCAGTAT............. 8623
MON.NG.x.NG1              CCTAA-AGT-TGGAC-TT-C-G-TATTGAG-GACTC-GGCT-GAC---CTACCG---CC-CCA-AG-G-G-GCA-CCACCTGATCCCTCTTCTCCGCGA-CT-T-CC-AC-GCTT-GGG--GCCA-TAGCCGCCTTCTT......................GCCTA-- 7240
GSN.CM.99.CN166           --A----GTGA-T-GG-T------GAT----GCCAG----GC--G--C-C-CCGTG----C-G-A-GT-C---T-GGAACAGCTGCAGAAAGCAATC-AG-AA-CAATCC-GCAAGCG-G-C-GCTCAGGGAAGTAGCAGCCAGAGTG...........ATAGCCTA-A 8598
GSN.CM.99.CN71            ---T---GTAA-T-GG-T------GAT----GCCAC----G---G--C-C-CCA------C-GCA-GT-C---T-GGAGCAG...........CTG-AG-A--C--TCC--CAAGCG-G-C-GCT-GGGGAAGTGGCAGCCAGAGCA...........GCAGCCTATA 8598
TAL.CM.00.266             AG-T---AT-AGA-A--T------GAC---AGCCA-----GCT-G-C-GA-TTG-GACC---C-A--GC-GC-AA-GCCAAGAGCCAACAGGTAGCG-AAGCCACA-AC-CCTT---ACG-CTCCACCTGACAGCAGCAGCAGCGACTCTCCAGTAT............ 8490
TAL.CM.01.8023            C-----CCT-AG--A---T-----GAC-------A-----G-T-G-C-GAGTTG-GACC-CC--A-G-GC---AA-GCGACCACCCAACAGGCAGCG-AAGCCACA-AC-CCTA---G-G-AT-AAATTCACAGCAGCAGTA...................GCAACAA 8032
MUS-1.CM.01.1085          C-A---AGTGA-T-GG--------GAC---A-T-AG--C-GC--CC-CCC-TTG-T--C-AC-TC-GAGAGCA-TCCACCTGATACAGAAGGGGTCAATA--ATT----TTCA-G----GGGCACT-ATTGAGCCTCTCAGGGAAGCCTTGCAAAGAGCCCTCGC-TG-T 8607
MUS-1.CM.01.CM1239        C-A---AGTGA---GGTTT-----GACA--AGTCAG--C-G-A-CA-CCC-TCGCT----AC-CC-GAGGGCA-CTCACCTGATACAGAAGTGGTCAGTAC-ACT----ATC-TGA-T-GGG-ACAAGCCCGCCGCTGCAGTGAACTCTTAGCGCGCCTC-TAGCATA-A 8576
MUS-2.CM.01.CM1246        C-A---AATGA-T-GG-CT-----GAT-----CCAG--CCGC--C--ACC-CTG-T--T-GACTC--AGAA-CTACCACCACCTGACAAAGGGGCTAGTA-ATCC----ACTCTG--T----CC-TTAGGGAAGCCTTAGGGCGCTTCATCGCACAC............G 8672
MUS-2.CM.01.CM2500        C-A----ATAG-G-GG--------GAT-------AG--CT----CA-ACC-TTG-T--C-GA-CC-G--GA-A-TCCATCACCTCCAGAAGTGGAGCAT-TCTCT----CTCAGGGCG-G-GCAACCAGCACCCACCTTCTGGAAGTCCTTAGACGCTTGCTTGCATAT- 8628
SYK.KE.x.KE51             AGAAGC-CAGAAAC-TT-GA--AGAGTT-C-GAGAG--GCAGA-CAC-GA-T-GAG-AC----AATTGCCTG----.............................................................................................T--A 8249
SYK.KE.x.SYK173           AGCAGGACAG-AGC-TT-GAGGACAGT--GC--CAG---CAGA-CTC-GA-T-GACA-C--ACTATCACCTG----..............................................................................................T--A 8550
RCM.GA.x.GAB1             CT-GA-TCAA--GA--CA-CCGTTGA--AG--TTTGGACG-AGC-ATCGCAG-TTT-C-GAAGCTGCAGCTCAAT-ATCTTCCAGAGCCTCCGGTGGATTTTA-CT...............................................................AAGA 8279
RCM.NG.x.NG411            TTATA-CCTGA-AC--C--CCGATAGAGA-CCTTTG-AGGAGGC-CTA-CA--TCT-C-GCAACTC---CTCA-TGATCTACCAGAGCCTCCAGTCAAT-CCTTCTC-AATCCTT--TC-T-TGC--TGGATAGGAGCT.................................AAG- 8362
GRV.ET.x.GRI_677          CT-GTG-CT-AA-----AGAG-TG-A--------ATAAGGACT-G-TG-TAC-AC-TCTGCC--ACC--C--CAT-TTCATCAGGACAGCAGTGGGA........................................................................TA-- 8593
SAB.SN.x.SAB1C            .................AC---TGT-TCAC---TTGA-CACT-G-C-CAGG-AT...T-GACCAGTTGGT-CTCGACAATCTTCAGCAACCTCCAT---T-CCTCCA---CAT--AAC-GAGAACCAGACAGCTAACAGCT..........................-AT- 8914
TAN.UG.x.TAN1             --T--AAGCGGAGGAG--AGCTTGG--GCA-CATTTGACCAAT-G-TGCT-G--A-TC-GCAGTTGGT-AT--A--CTTTACCAGATCCTCCGCAGGA--CTCAC-ACTCTTCTT-AACTGCT-A-ACAGGAG...........................--CC-ATA-A 8733
VER.DE.x.AGM3             CT-GTGCAAG-GA--G-C-A-TTGGT---C--TCAG-AGCA---G-C-ATAC-AC--TTGCC--ACC--C--A-TTCATCTTAGGAGCGCTTTCCAG.......................................................................TA-A 8238
VER.KE.x.9063             TT-GAGCAAA-G---G-C-A-TTGG-T--C--CGTT-AGCAC--G-C-ATAC-AC-CCT-CC--ACC--C--AATTCATCTTAGAGGCGCTTGGCAG.......................................................................TA-A 8740
VER.KE.x.AGM155           TT-GTGCA-G-AA--G-C-A-TTGGTT-----C-CT-AGCA---G-T-ACAC-AC--CTGCC--ACC--C--TCTCAAGCTTAGAAGCGCCTGGCAG.......................................................................TA-T 8717
VER.KE.x.TYO1_patent      CGC-TG-CTGACGAGGCT-A--..................ACT-G-C-ATAC-AC--TTGCC--ACC--C--CATTCAGCTTAGGAAAGCTTTTCAA.......................................................................TA-- 8701
COL.CM.x.CGU1             .....................-GATTGA--GA---G-GGTGGAGC-GCT-GTT-C-TCA-TGCAGCAA-GG-ATCCTTACT...................................GTCCTGCCTCAACATCCGCCAACAACCTCA............GAC--GAGGA 7987
```



PLV Complete Genomes

```
                            Env gp41, gp160 end                        Nef start
H1B.FR.83.HXB2              ......ATTTTGCTATAAG...........................ATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGGAAAGAATGAGACGA..................................GCTGAGCCAGCAGCAGATAG 8882
Nef                                                                       _M__G__G__K__W__S__K__S__S__V__I__G__W__P__T__V__R__E__R__M__R__R__.__.__.__.__.__.__.__.__.__.__.__A__E__P__A__A__D__R
Env                         _.__.__I__L__L__*_
H1A1.UG.85.U455             ......GC---------C............................---------------------AG--CAGAG-G-A-------GAG--T---A-----------GA-...................................A--CCTG--------A-AG- 8333
H1B.US.90.WEAU160           ......GC----------............................-----------TA---------C------GT-G----------G--A-------------A------------------------------------------------AG- 8889
H1C.ET.86.ETH2220           ......GC-----A----A...........................-----G----C-AT-------T-----CCAG-A---------G--A----A---------A-----.....................................----CT-----------GG- 8279
H1D.CD.84.84ZR085           ......GC----------A...........................----------A-------------A-AG-----------G--A--------------A---AA-...........ACTGATCCAAGGGAAAGAAGAAGAC----------------G- 8427
H1F1.BE.93.VI850            ......GC------G---A...........................-----G----------------C---A-AG-----------G-----G------------A-...................................A-CCCTA----------AG- 8159
H1G.SE.93.SE6165            ......GC----------A...........................-----A--------------------CA-AG----------CGAG--------------A---AAC.................................A-CCCTA----------AG- 8321
H1H.CF.90.056               ......-GC---------TA..........................-----A-----A-------------GA--GG---G---T-----A---------------G---...................................-----A----T------AG- 8202
H1J.SE.93.SE7887            ......GC----------A...........................-----GAA---A---------......................CAG-----------------G---...............................GCTC-C-CT-----------G- 8180
H1K.CM.96.MP535             ......GC---A------A...........................---------------------A-AG-----------G--A---------------G---..........GCACGACCAGCAGCAGACAGGGTGGGAA-AC-A-----------CG- 8069
H101_AE.TH.90.CM240         ......-C----------C...........................-----AA-T----------------CA-AG-G---------CAG--C-------A---A-AG-A-...............................A--CCT------A----AG- 8456
H102_AG.NG.x.IBNG           ......GC----------C...........................-------------------C--CA-AG-G---------AG--T-T-A-----------A-...................................A-CCCTA--------ACAG- 8409
H103_AB.RU.97.KAL153_2      ......GC-----A----C...........................-----G-----A-------------CA-AG-G---------CAG--T-----------A-----..............................GCTC---CT-----G---AGAG- 8062
H104_cpx.CY.94.CY032        ......GC------T---A...........................-----A-----A-------------CA-AG-----------GAGA----------------G---GCTCGAGCTGAGCCAGAAAGAATGAGGCGAGCTCAA---------------CAG- 8291
H1N.CM.95.YBF30             ......GCAC--A------...........................-----AAAG-TT--------G--C--CC-AG-A--------AGAAA-C--A------------A-....................CAAACGCAAGAACCA--A-TAGAGC------TAG- 8469
H1O.BE.87.ANT70             ......-G---AT-G---C...........................-----AAA-GCA-T-AG----G---AAT-TGAG------G-AG-A-----A-----------A-.................ACTAGAACTTTCCCTGAGT----A---TGC---CC-G- 8969
H1O.CM.91.MVP5180           ......--C--AG-G---C...........................-----GAATGCA---AGC-----C-AAT-TGCA------T-AGAA-----A--T-----------........................TCCTCCTCTGATCCTCAAC-A---TGT---CC-G- 8992
H1P.FR.06.RBF168            ......-GC---------............................-----GAATGCA---AAG--------CT-AG-G--C-----AG-A--C--A----------AT-AG...................ACTGCCC-A--AA--C-Y---CC-G- 8437
CPZ.CM.01.SIVcpzCAM13       ......-C---A--T---ATCTTGCTTTAAAATAGTCTTGCTTTAAG.-----A-CTC--------T--------CC-AG-A--------GAAA-C--AA-------A-A-GAG..............................---CCAAT---------AG- 8515
CPZ.CM.05.SIVcpzEK505       ......GCAC-TT-----A...........................-----AAAAGTT--------G------CC-TG-G--G-----ACAA--C-TA--T----------A-...............................CAGC-AGATC--------AG- 8389
CPZ.CM.05.SIVcpzLB7         ......GC---AT-----A...........................-----AA----A--------------CC--G-A---------GAA--T--AAG------T-----A-...............................-AAACTGT---------AG- 8414
CPZ.CM.05.SIVcpzMB66        ......GCA--AT-----A...........................-----A----------------C---C--G-G--C-----AGAG-----A--C---C-T--GA-G................................A-AC-AA-T--------AG- 8527
CPZ.CM.05.SIVcpzMT145       ......GC------T---............................-----A-C-----------------CC-AG-A---------AGAAA-----A--G-----A-A-A-...............................---CCAA-----------G- 8415
CPZ.CM.-.SIVcpzMB897        ......GC----------............................-----AAA---A--------------CC--G-A----------G-A---------C---C-T---A-G..............................A-ACCT-----------AG- 8417
CPZ.CM.-.SIVcpzDP943        ......-GC---------C...........................-----CAA---A----------C--CC--G-A--------AGAA-----A-CT-----A----A-...............................---C-A-----T---CCAG- 8469
CPZ.CM.98.CAM3              ......-GC---------C...........................-----CAA--G----------------CC--G-A--G-----AGA---C--A-------C-A--CGA-...............................CTCA-C-TGCTT---CCAG- 8319
CPZ.CM.98.CAM5              ......-G----T-----C...........................-----AA---------------------C--G-G------G-AG----T--AC-----G--A----A-...............................A--CCA----------TCAG- 8625
CPZ.GA.88.GAB1              ......GC---A--T---CACTGCATTAAG................-----AAC---A-----T-----------C--G-A----------GAG--C--AAG------A--GGA-..............................---CCAA----------GG- 8955
CPZ.GA.88.GAB2              ......GC---A---G--TAAA........................-----CAA---A-----T------C--CA-TG-A--------GCAG--C--A--G----C-T--G--G...............................A--C--GA----------AG- 8293
CPZ.US.85.CPZUS             ......-GC--AT-----............................-----AA----------T--G-----CA-AG-A---------AGAA--C--AA-C---T------A-.......................ACTCAGA-AACAG-----------AG- 8915
CPZ.CD.90.ANT               ......GCA--AAAT---AC..........................----TCTGCA-----T----TC-AGTG-G-----GCA.........AC--GC---T--GAA-...........................ATACA---AA--AATC-----GA 8336
CPZ.TZ.00.TAN1              ......GCC--AAAT---A...........................-----AAA--TA-TTGGT-G-T-GCC-.......-GGCC-GG-AA-CC-TC---GATC-T...........................CA-A-CA-CT--AGT--GCC 8568
CPZ.TZ.2001.TAN2            ......GCC--AAAT---A...........................-----AAA--TC-TTGGT-G-T-GCC--GAGC-C--A-AG---T--A-GATCTT...........................CATAAAAGCT--CAT-AGCC-.. 8589
CPZ.TZ.2002.TAN3_1          ......GCC--AAAT---A...........................-----GAA--TC-TTGG---GT-GCC--GAGC-CAGA-AG---T--A-GAACTC...........................CATAAAAGCT--CAT-AGCC-.. 8643
GOR.CM.2007.SIVgor2139_287  ......-GCC-AT-----A...........................-----AAA--CA---AA---G-----CT-C--A---------AG-A-------------------A-ACC.......................AACCCT-AGCCAAAGC-G--TCCAG- 8479
GOR.CM.2007.SIVgorCP2135con ......-GCC-AT-----A...........................-----AAA--CA---AA---G-----CT-C--A---------AG-A-------------------A-ACC.......................AGCCCT-AGCCAAAGC-G--TCCGG- 8463
GOR.CM.2004.SIVgorCP684con  ......-GCC-AT-----............................-----AAA-GCA---AA---G-----CT-TG-A---------AG-A-----------------G-A-ACC.......................AGCCCT-AGCCAGATC-G--TCCAG- 8408
H2A.DE.x.BEN                ......GCCC-C--G-G-..............................................................................................GGGACAGCAGT-T-AGCAG-G---GTA 9313
H2A.GW.x.ALI                ......GCCC-C--G-G-..............................................................................................AGGACAGCGGT-T-AGCAG-G----TT 9309
H2A.SN.x.ST                 ......GCCC-C--G-G-..............................................................................................GGGACGGCGGT-T-A-CAG-G----TT 8755
H2B.CI.x.EHO                ......GCCC-CT-G-G-.......................................................................................................................... 9227
H2B.GH.86.D205              ......-CCC-CT-G-G-.......................................................................................................................... 9250
H2G.CI.92.ABT96             ......-C-C-CT-G-G-..............................................................................................AGGGCAACAAT-CA-GCAG-GTC-ATT 8698
H2U.FR.96.12034             ......GCGC-TT-G-G-..........................................................................................AGCTCAGGGCCAA-GCT-TAGT-A--G---GTT 8794
SMM.SL.92.SL92B             ......--GC-TAAC--G............................................................................................GGATGGATATAAAT-CAAT-T--GG--ATT 8721
SMM.US.x.H9                 ......-C-C-CT-G-G-..............................................................................................GCCTCAGAAAT-TTGT-AG-GTC-ATT 8759
SMM.US.x.PGM53              ......GCAC-C--T-G-..............................................................................................GGGACAG-GAT-CAGT-AG-G-C-ATT 9207
SMM.US.x.SIVsmH635F_L3      ......-C-C-CT-G-G-..............................................................................................GCCTCAGAAAT-TAGT-AG-G-C-ATA 9292
                            Env gp41, gp160 end
MAC.US.x.239                ......-C-C-CT-G-G-..............................................................................................GGGACAGAAAT-CAATCAG-G-C-GTA 9270
Env                         _.__.__T__L__L__*_
Nef                         __.__.__L__S__C__E##__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__.__G__Q__K__Y__N__Q__G__Q__Y
MAC.US.x.251_1A11           ......-C-C-CT-G-G-..............................................................................................GGGACAGAAAT-CAATCAG-A-C-GTA 9266
MAC.US.x.251_BK28           ......-CGC-CT-G-G-..............................................................................................GGGACAGAAAT-CAATCAG-GGC-GTA 9252
MAC.US.x.EMBL_3             ......-CGC-CT-G---..............................................................................................GGGACAGAAAT-CAATCAG-GGC-GTA 8401
STM.US.x.STM                ......-CAC-CT-G-G-..............................................................................................GCCTCAGAGAT-TAAT-A--GTC-ATT 8921
MNE.US.x.MNE027             ......-C-C-CT-G-G-..............................................................................................GGGACAGAAAT-CAATCAG-G-C-GTT 8747
DEB.CM.99.CM40              ......--CC-CAAC-G-.......................................................................................................................... 8559
DEB.CM.99.CM5               ......T-AC-CAAC-G-.......................................................................................................................... 8480
DEN.CD.x.CD1                ......G-CC-CAAC--G..............................................................................................GTTACTG-AGC-TGAT-AG-AC--AGA 8882
LST.CD.88.447               ACAGAG-CCGCT--G--G.......................................................................................................................... 8035
LST.CD.88.485               ACAGAG-CCGCT--G--G.......................................................................................................................... 8032
LST.CD.88.524               ACAGAG-CCACT--G--G.......................................................................................................................... 8035
LST.KE.x.lho7               ACAGAG-CCGCTT-G--G.......................................................................................................................... 9110
SUN.GA.98.L14               ACAGAG-CCACTT-G---.......................................................................................................................... 9181
MND-1.GA.x.MNDGB1           ......CCCC-TAAC---.......................................................................................................................... 8378
OLC.CI.97.97CI12            ...ACA-C-GGAG-C-G-.......................................................................................................................... 8323
WRC.CI.97.97CI14            AGAAGA-GC--CAGTG--AGAAGACAG................................................................................................................. 8923
WRC.CI.98.98CI04            AGAAGA-GAC-CT-C---.......................................................................................................................... 8959
WRC.GM.05.Pbt_05GM_X02      AGAAGAGGACA--AG---.......................................................................................................................... 7855
MND-2.CM.98.CM16            ......C-CC-CAAC-G-.......................................................................................................................... 8909
MND-2.GA.x.M14              ......C-GC-CAAC---.......................................................................................................................... 8838
MND-2.x.x.5440              ......C-AC-CAAC-G-.......................................................................................................................... 8545
DRL.x.x.FAO                 ......C-C--CAAC--G.......................................................................................................................... 8606
MON.CM.99.L1                ......T-CC-TAAC-G-.......................................................................................................................... 8788
MON.NG.x.NG1                ......C-CC-TAAC-G-.......................................................................................................................... 7413
GSN.CM.99.CN166             ......C-A--TAAC-G-.......................................................................................................................... 8771
GSN.CM.99.CN71              ......C-AC-TAAC-G-.......................................................................................................................... 8771
TAL.CM.00.266               ......TGC--TACT-G-.........................................................................................................T-AA--ATGGCA----ACAG- 8676
TAL.CM.01.8023              ......TGCC-TGCT-G-.........................................................................................................T--A-GATGGCA----TCCGA 8226
MUS-1.CM.01.1085            ......C-CC-TAAC-G-.......................................................................................................................... 8780
MUS-1.CM.01.CM1239          ......C-CC-TAAC-G-.......................................................................................................................... 8749
MUS-2.CM.01.CM1246          ......C-CC-TAAC-G-.......................................................................................................................... 8845
MUS-2.CM.01.CM2500          ......C-CC-TAAC---.......................................................................................................................... 8801
SYK.KE.x.KE51               ......TACC-TAAC-G-.....................................................................................................................-C-CATCTTGGA 8434
SYK.KE.x.SYK173             ......TACC-TAAC-G-.............................................................................................................GCCCATT-ACC--C-CG- 8741
RCM.GA.x.GAB1               ......C---GTAAT--G.......................................................................................................................... 8452
RCM.NG.x.NG411              ......T--C-TAAC--G.......................................................................................................................... 8535
GRV.ET.x.GRI_677            ......G-CC--GGG---.......................................................................................................................AG- 8769
SAB.SN.x.SAB1C              ......TGG--TAAT--G......................................................................................................................GCA--G---CTC 9099
TAN.UG.x.TAN1               ......TGG--CAAC--G.............................................................................................................................GAGGG- 8912
VER.DE.x.AGM3               ......--CC-TAAT--G......................................................................................................................GAAGAG-A--G- 8423
VER.KE.x.9063               ......--CC-TAAT--G......................................................................................................................GAAGAGGA--G- 8925
VER.KE.x.AGM155             ......G-CC-TAAT--G......................................................................................................................GAAGAG-A--G- 8902
VER.KE.x.TYO1_patent        ......--CC-TAAT--G......................................................................................................................GAA-CA-A--G- 8886
COL.CM.x.CGU1               ......GG-CCTGCC---.......................................................................................................................... 8151
```

```
H1B.FR.83.HXB2              GGTGGGAGCAGCATCTCGAGACCTGGAAAAACATGGAGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCACAA......GAGGAGGAGGAGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGCAG 9046
Nef                          V  G  A  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  V  G  F  P  V  T  P  Q  V  P  L  R  P  M  T  Y  K  A
H1A1.UG.85.U455             A--A-------T-----A---TT-A--T---T--------G-------C--T-----T--T---T----CAG----------G----G---.....-----A-GA--C--A--C--C-----T-GG--A--------------------------T------- 8497
H1B.US.90.WEAU160           -----------T---------T---C------------------------GAG......-AT-------A---C-T----T--A---------.....-----T--------------------T-G---------------------------------------- 9047
H1C.ET.86.ETH2220           A--A--------G---------T-A--C---T----G---C-T-----C-----C---C-C--C-AT---C---A-----------G-----G---GAG...-----A--A-----A--C----------G--------------------------T------- 8446
H1D.CD.84.84ZR085           -----------T---------------G------G-----------------------A----T-------C---------G----------GGAA...-----C-----------C----------G---------------------------------------- 8594
H1F1.BE.93.VI850            -----------TG---------T-A--C-G--GG--------T-----C---------TAG-A----T---C---AC-TG---------G----------.....-------------A--A--C----------G------------G------------C--T------- 8323
H1G.SE.93.SE6165            A--A-------T-----A---TT-A--T-GG----------------C---------------C-AT---C---A-------------G---------GAG...-----CTCA--A--A--C----------G---A--A--------G------------TT----GT- 8488
H1H.CF.90.056               ---A-------TG--------TT----T-G--GC--G---G------T--AT----------T----T---CG--A-GCC---------G---------GAG...--C-G------A--A--C----------GG---------------------------C--T----G-- 8369
H1J.SE.93.SE7887            A----------TG----A---TT---CT--G-----G-----------C------------------A-------AC-----------G----------.....ACA-----------A--A-----------A-------A----A----------------A-------G-- 8344
H1K.CM.96.MP535             -----------T-----A---------CT-G------G---G-------C---------TT--CAC-AT---C---A------------G----------G.....-----A-----A--A-----------G------------A------------------T-------- 8233
H101_AE.TH.90.CM240         A--A-------T-----A---T--A--T--------------A---------T--......ATAGAT-------A-----T------GAG--------.....-----C---------A--C----------TG--G----------C------------------T-----G-- 8614
H102_AG.NG.x.IBNG           A--A-------------A---TT-A--T-G-------------------C--T-------------CAA--T--CC---A------------G-----------.....-----T---A-T--A--C----------GG--A-----G-----G-----------------T-----G-- 8573
H103_AB.RU.97.KAL153_2      A--A---C---T-----A---TT----T--GT--------G---------C--T--------------AT-------A---C---------G--G--G---.....-----A-----------A--C----------GG--G------------------G-----G-----A--T-----G-- 8226
H104_cpx.CY.94.CY032        A--A-------TG----A---T----C---------G--------T--A----------------T---C---ACAAAA-------G---------GAA...-----A--A-----A--C----------GG--A------------------G----------C-TT--A-G-- 8458
H1N.CM.95.YBF30             A-CA--------T----A---T--A-CT--T-GA--G--C-----C-TA--A-----TAG--AC-AT----AAAG-ATA--T--------------GAA...--A--A-----A--A--C----------ACGC--------------A-----G-----A--C--T--ACAG- 8636
H1O.BE.87.ANT70             A--A---CAGATC--CA-G--AT-A-C-GCTAGA----GG--AC------TC-C----TC-TCAA-A-------A--CCT---A-TC-------AGT--C......C-A------A--A--A-----------AG--------A--G---C----G---------C--T--A-G-- 9133
H1O.CM.91.MVP5180           A--A-----T-TC--CA-G--GT-A-C--CTAGA--G-G---AT------TC-C-C--TC-TCAA-A-------A--CCT---A-TC------CAGC--C......A-A--T------T--A--C--C-----A-G-------A--G---C----G---------C-TT--A---- 9156
H1P.FR.06.RBF168            ---T--G-AGATT--CAA---AT-A-C-C--GGAAA--G---ACGT-----AAT---ATT--AAG-A-------A--A-TA------TCT--------T--T......-A--C--A--A--A--A--C-----Y-GG-------A-----CR--------C-----A----------- 8601
CPZ.CM.01.SIVcpzCAM13       A--A----AG-TT---AAG--T--A------T--------T-----T-----AC------C--AG---T---CAAA---TA--A---T-----T-A-ATGAAACATC-T--A--A--A-----C----------T-G---A-----G--A--------------------A-G-- 8685
CPZ.CM.05.SIVcpzEK505       ---A--------TT-----A---T--A-CT--T-GA--G--T--T--C-TA--A-----TAAG-AG-AT---CAA--CCTA--T------------AG---......A-A------A-CA--A--C----------ATGT-------------A---------------------A--A-- 8553
CPZ.CM.05.SIVcpzLB7         A--A---C----C---A-A------A-CT-GG------G-----T--------T------CC--CAA--T----AGA-CAT-------T----G-AG---......A-------A--AA-AT-G----------A-A----------------A------G---------------TT--A-AG- 8578
CPZ.CM.05.SIVcpzMB66        A--A---C---T---A--G-----A-C--G-------------A--C-----A--C---T-CCAG--A----AGA-CCT-------T-------A-GT-......C-AA-T--A--A--A--A--C----------A-GG-----A------C-------------------C-TT-----G-- 8691
CPZ.CM.05.SIVcpzMT145       A--A--G-A--T---A-A----A-A-C--GC-G--------G------T---A-----CC--CAA-------AGA---TA--T----------A-ATG......C----T--A--A--A--A--------------A-------------A--A--G-C---------T--------C-TT-----GT- 8579
CPZ.CM.-.SIVcpzMB897        A--A--GC---TC--G--------A--G-G------G--T--A-----------G-----TC--CAG--A----AGA--CT-------T-----G-AG---......AG---C--A--A--A--A--------------A-G---C--A--G-----------G--C-----C--T--A--G- 8581
CPZ.CM.-.SIVcpzDP943        ---A---T---TC--A------TT-A--GGC-------------A-C-C---A-----CC--CAA-A----CAAA--CTA--A-------------A-ATGCAGGGAC-T--A----------A--C----------T-G---A--A-----CC-T------T------A-TT--A-G-- 8639
CPZ.CM.98.CAM3              A--A---C---TC--A-A----T-A--GGC----------G----C-C-----GC-C---C--CAG-AT---CAGA--CTG--------T--------A-ATGCAAGAAC-CA-T--A--A--A--C----------T-G-------A---------C-T--------T------C-TT--A----- 8489
CPZ.CM.98.CAM5              A--A---C---TC--A-A----T--A--GGC--G------------C-C---A----CAC...A-A-AAT--CAGA-CCTG------T---------A-ATGGAAAATC-T--A--A--A--A--C----------T-G---C------------CC--------------------A--------CA-- 8792
CPZ.GA.88.GAB1              A--A----A--TT---GAAG-----A-----G---C-------T------T------G--C--CC---AG--T---CAAA--CTA--T----------G-A-ATG......--CA-T--A--A--A--A--C------A-G----A------T--AACG----------------------T--A---- 9119
CPZ.GA.88.GAB2              A--A----A--TG-----AG--TT-A-C--GG-G----------A--TTCC--AC-----CC-TCAG--T---CAAA--CT---A-------------A-ATG.......AC--------T--A---G-----------A-G------------------------------------A--T--A---- 8457
CPZ.US.85.CPZUS             A--A---C-T-TC-----A----T-A-C-G-------G--T------T-C---A--------C--CAG-A---CCAAA--CTG--A---T-----T-AGATGACTAATC-T--AAGT--A--A--C--C--T--TCGC------A------A------G--C-------------ACA-- 9085
CPZ.CD.90.ANT               TA-A--GC-CTGTGGAAAC--AT---CC-GTAGA--G---C-A--CTCCTCT-C--T---GTA-AGAA--A-A--TAATCA---ATTC------AT--T......ACA-----A-G-ACA--C------T--G-G-------A--C--CA-G------C-------AG-A--ATT-- 8500
CPZ.TZ.00.TAN1              T--A---CAG--C--A-A------CC-G--TA-A----GTC-----T-C--A--CCCT--GTA-CT-AGCA-A--TG-TA-AA-ACTCT-C--AC--T......ACT--A--A--A--A----------------G---AGCA------CA-G------C-------AG-G---CT-- 8732
CPZ.TZ.2001.TAN2            .A-A---CAG--C--AAC-------CC-----TAGA--G-GCT-A--C-AC-A--CC-T--GTA--T-AGCA-A--TAGTA-AG-ATTCT-C--AC......CATACT------A--A--A--G--------T-G---AGCA------CA-G------C-------AG-A--ACT-- 8752
CPZ.TZ.2002.TAN3_1          .--A---CAG--C--AAA-------CC-----TAGA-----GCT-A--C-CC-A--CC-T--GCA--C-AG-A-A--TA-TA-AACACTCT-C--AC......CATACT----------A--A--G--------T-G---AGCA------CA-G------C-------AG-A---CT-- 8806
GOR.CM.2007.SIVgor2139_287  ------T-A--TG---A---CAT-A-C-C--AGA----GG--AC-TGCA-AAT-----TC--CAA-A-----CAA-----TA--A-TC---------T......CAT--A--T--A-----A--A-------------G------A--G--G------------------C-TT--A-GG- 8643
GOR.CM.2007.SIVgorCP2135con ------T-A--TG---A---TGT-A-C-C--AGG----GG--AC-TGCA-AAT-----TC--CGA-A-----CAA-----TA--G-TC---------T......CAT--A--T-----------A--A----------GY-------A--G----------G--------------C-TT--A-GG- 8627
GOR.CM.2004.SIVgorCP684con  ------T-A--TG---AT--CAT-A-C-C--AGA----GG--AC-TGCA-AAT-----TC--AAA-A-----CAA-----TA--A-TC---------T......CAT--AA-T-------A--A--C----------GT------A--G----------G--------------C-TT--A-GG- 8572
H2A.DE.x.BEN                CA--AACAGCC---GGA--A---CA-C--C-G-AA--CAG-AAGATTTGTAT-GGCAGCA-AA--TGG---A--TA...-AT-CTGAT..............--T--T--CCTAA-A--AG----T--T-----AAGA-----ACGG---GA------C--T--ATTG- 9465
H2A.GW.x.ALI                TA--AATA-CC---GGA--AC--CA-C--C-G-AA--AG-AAGA-TTGTA---GCA-CAGAA--TGG---A--TA...-AT-TAGAT--T............--T--TAGCCTA--A--GG-CT-T---------AAGA-----AA------A-------A-----ATTG- 9464
H2A.SN.x.ST                 TA--AATA-CC---GGA---C--CA-C-...G-A--G-AG-AAGGCTCGTA---GCA-CA-AA--TGG---A--TA...-AT-CAGAT..............--T--T--CCTA--A--GG-C--T---------AAGA-----A------GA------A--T-G-TTG- 8904
H2B.CI.x.EHO                .......................T-G-C--A-GACTTTGGGAG--GAAG-GGGAGG-AA-CAAGATTCA-A-......-AGGATGAT..............-----CA-T--A--A--GG-C-GT--A-G---CGG---C--AC-G--G---------A-T---ACT-- 9352
H2B.GH.86.D205              .......................-G-C--A---CCCGGGGGG--GAAG-GGGAGGGCA-CAAGATGCA-A--AA...AGTGATGAG..............--C--T--A-T--GA-CCA-ATGTAAG--C--AATA------AC-G-------------A-----ACT-- 9378
H2G.CI.92.ABT96             TA--AATA-CC---GGA--A---CA-C--C-G-AA-------AGTT-GCATA--GGYAGCA-AA--ATG---A--TA...-ATA-TGAT..............--TA-T--CTTA--A--AG-A-------TAC--CA-A--C--GC-----GT------G------ATTG- 8850
H2U.FR.96.12034             CA--AATA--C-T-GGA--AC--CA-C--C-GGGAA--AT-AATT-CAATAT--GCA-CA-AAC-TGG---A--TA...-ATGAAGA-..............--TA-T--TTTA--A--AG-AG----A-G---AAGA---G-AC-G--A------------------ATTG- 8946
SMM.SL.92.SL92B             TA--AACA-CC---GGA--A---CT-CT-C-G-AA--CA--AGCA-CAGTAT-GGCAGCAGTG--TGG---A-ATA...-ATGA-GAT..............--T--T---CT--CT--AG-AGTT---TGG--AA-A------CC------G------T-A-----ATTG- 8873
SMM.US.x.H9                 CA--AATA-CC-T-GGA-GA---CA-C--C-G-A-------AATT-G--TAT-GACA-CA-AAC-TGG---A--TG...-ATAATGA-..............--T......TTA--A-----G------G-GC--AAGA--C--CG------ATC------A-----ATTG- 8905
SMM.US.x.PGM53              CA--AACA-CC---GGA--A-T-CA-CT-G-G-GA---A--AGTT--AATAT-GACA-CA-AAC-TGG---A--TA...-ATGATGAT..............--T--T---CT-A-A--AG-AT----ACAC--AA-----G--A------G-C---T-A-----ATTG- 9359
SMM.US.x.SIVsmH635F_L3      TA--AATA-CC-C-GGA-GA---CAAC-GC-G-AAA-----AATT-G--TAT-GACA-CA-AAC-TGG---A--TG...-ATGATGA-..............--T--T--CTTA--A---G-CT----GCAC--AA-A--C--C------GG-C------A-----ATTG- 9444
                                                                                                           Premature stop in SMM239

MAC.US.x.239                TA--AATA-TC---GGA--A---CA-CTG--G-GA---A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-TA-..............--T--T--CTT---A--GG-AT----G-GG--AA-A--T--CC-----A------G------ATTG- 9422
Nef                           M  N  T  P  W  R  N  P  A  E  E  R  E  K  L  A  Y  R  K  Q  N  M  D  D  I  .  D  E  *  .  .  .  .  .  D  D  D  L  V  G  V  S  V  R  P  K  V  P  L  R  T  M  S  Y  K  L
MAC.US.x.251_1A11           CA--AATA-TC---GGA--A---CA-CTG--G-GA---A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-GA-..............--T--T--TTT---A--GG-AT----G-GG--AA-A--T--CC-----A------G------ATTG- 9418
MAC.US.x.251_BK28           TA--AATA-TC---GGA--A---CA-CTG--G-AAA--A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-GA-..............--T--T--CTT---A--GG-AT----G-GG--AA-A--T--CC-----G-------------ATTG- 9404
MAC.US.x.EMBL_3             TA--AATA-TC---GGA--A---CA-CTG--G-AAAG-A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-GA-..............--T--T--CTT---A--GG-AT----G-GG--AA-A--T--CC-----G-------------ATTG- 8553
STM.US.x.STM                TA--AATA-CC-T-GGAA-A---C--C-GC-G-AA-C--T-AGCT-GAATAT-GACAGCA-AAC-TGG---A--TA...-ATGA-GA-..............--T--TA-TCTA--A--AG-AG----ACAT--AAGA--C--AC-T--GGA-------------TTG- 9073
MNE.US.x.MNE027             TA--AATA-TC---GGAA-A---CA-CTGG-G-GA-G-A--AATT-GCATA--GA-A-CA-AA--TGG---A-ATA...-ATGA-GA-..............--T--T-GCTT---A--GG-A-----G-GG--A-GA--T--C-------G-------------ATTG- 8899
DEB.CM.99.CM40              .......................-TTCTGGGGAA-TC-T-GA-G-GAA-TA--GTGCA-CAA--AGC-ACAACAAGACTTAGAT................C--ACA--A-CA--A--C--C-----ACGC--C------T-CA-CC--GACCCC--A--T---TT-A 8687
DEB.CM.99.CM5               ..................G--A-CACTCT-TGCA-ACC-CCA--G-GAG-TA--ATGCA-TCAG-GAGCCCAG-AG---TATGAT................------AC-T-G-A-A--C---G-T----G---AAG---C--AA-T--GGA-CCC--------ATTGA 8614
DEN.CD.x.CD1                AACCTTT-T--ATG-AGA-...GAAATC-TC--ATC-T-TCCAG--CTGGA-CC-CA-CATC-A-GA.........ATACA-CAATAT-TG-T-G-CCATGAT--T--A--A--A--A--C---------ACGG--A-GACG---AC-GCACG--CC---C--T--A-ATA 9040
LST.CD.88.447               ........................................................................................................................................................................CCCT--C--T---A-TC 8052
LST.CD.88.485               ........................................................................................................................................................................CCCT-----T---A-TC 8049
LST.CD.88.524               ........................................................................................................................................................................CTCT--C--T---AA-CT 8052
LST.KE.x.lho7               ........................................................................................................................................................................CCCT--C--T---A-TC 9127
SUN.GA.98.L14               ........................................................................................................................................................................GCCA--A--T---CAGC 9198
MND-1.GA.x.MNDGB1           ...........................................AGATG-G-A-CTT-ACTTA-CACAGCAGGA-C-AGAAGAG..............-----A--A...--T--A-----T--GTGT-GC--A--TT-C------GTGCCAT-A-----A-ATC 8482
OLC.CI.97.97CI12            ......G-ACCAGCA-.......-AAAC---T-TG-C-AT-GAGT-TTTA-TG-C-GGG-AC....................ACTGAC..............AC---T--TCT--AA--A---------G-AG--------G--AG-GC---A-GCC-----T--A-TC 8445
WRC.CI.97.97CI14            ..............................................................................AGTGATGA-..............--T--A-GA...--A----------A--T-----A--T--CA-------A-CCA----TT--ACTGC 8997
WRC.CI.98.98CI04            ....................................................................A-AAGA-AGT-CTGAT..............--T--A--A...--A--C--------TCAT--------T--AA-------A-GCT--A-TT--ACTGC 9039
WRC.GM.05.Pbt_05GM_X02      ....................................................................A-AAGACT-A-CTGAT..............--T--T-CA...A-A--A--C--T----G---A--A--T----------AGGCA--A-T----ATCT 7935
MND-2.CM.98.CM16            ..............................................................................-GAAGAG..............--A--AAC-.......--C---------ATAT--A---TGC---AG-----GA-CCC--C-------ACT 8977
MND-2.GA.x.M14              .......................................................................AGTA-ATGATGAGC--.............A-A--A--A...AG---C---------TTAT--CACACAG---G-GC-TGA-GCA--C-------AAACC 8918
MND-2.x.x.5440              .....................-CGGG-T-TAA-GCACCAGCAGGA-GTA-C-CTCC-CT-AGAA--A-AGCAACCAAGTGATGA-..............-----A---...--A--C---------TAC--AAC-CAG---AG-GCCTGA-GCA--C-------ACT 8673
DRL.x.x.FAO                 .................CAG-T-A--T-CTAC-CATATGCAAG--GCAGCAGCAGAGAA--A-GT--CCCT-C-CACA-AAGAAGA-C--C-CTC-GAAGAA--A--A--A--A--A-----------ATGT------ACGC---------GAGCCA--C-----A-ATC 8758
MON.CM.99.L1                ...........................................GG-C--A---GGCAA-CAACG-CGGCAAC--..........-G--C-AC...GACGAC--C--C--C-GT--C----G------A-GG--C-GA--T--GC-TC-CGACCCT----GG---CTCA 8901
MON.NG.x.NG1                ...........................................GG-GC-G---CCACCGCT-A-A-CGCCGCC-AACCA-TGTCG-T--T-ACGCCGACGAG--CC-A--TTT-CCT--C-G----G--TCGC---AT---G--AC-T---GACCCT--A-GG--ATCA 7538
GSN.CM.99.CN166             .........................................................-AGC-A--AC-CAAGA-A-GAA-C--C...............-------T-T---ACCATCA-GC-----A-G---A-G------AC-T---GACCCC--C-----A-TCA 8866
GSN.CM.99.CN71              .........................................................-AAT-ACCCT-CAGGCAAAAT--CA-C...............--TT-T---ACCATCA-GC-----G-G---A-----G--GC-T---GACCCC--C-----ACTTA 8863
TAL.CM.00.266               A-C--ACTACATG...................--CAAGGCATT--CACAAGGAGAG-TTTGGTATG-AAAA-AACAA-AAGAAAAT...............--CAT---ACTA-----C--C--T--G-GG---A-A--T--G-G---A-CTGT-----T--ACT-- 8810
TAL.CM.01.8023              C-CAAATTGGATG...................--GAAGGCACTCTCGGAAGGAGA--TCTGGTATG-AAAG-AGCAA-AAGAAAAT...............--CAT----CTT--A--C---------A-GG-----A--T--A--G--G--CTGT-A------ACTG- 8360
MUS-1.CM.01.1085            .........................................-G-AGC--G--A-AG-C--GTG-TA...........................GAA---C-T---...CGTACC-GC--G--GCGT--GAG---G--AA-CC-TGACCCC--C------CTGA 8872
MUS-1.CM.01.CM1239          .........................................-G-A------GAGCGGC-ACT--TC...........................GAA--AC-T---...CGTTCA-GC----GCGG--G-GA--G--CC--C-GGATCCC--A------ATGA 8841
MUS-2.CM.01.CM1246          .........................................--AC-AGGGCCGCAGGCAA-T--TA...........................GAA--AC-T---...TGCTCC-GC-----ACGC--GA-A--C--GA-TC-TGATG-C-AC------ACTCA 8937
MUS-2.CM.01.CM2500          .........................................GG-A--T--C-A-AA---A-T--TA...........................AAAA-AC-TA-A...CG-TCG-G------ACGC--G--ACAC--AA-CCAGAACCCC--C--T---TT-A 8893
SYK.KE.x.KE51               TCCCAA-AG--AC.........T-T--CTCTGGG--GAA--AATGG-A-GCGGGAGACATCTG-...C-.........-ATGA-GG-..............--T-----TCTA--A--C-----T--TTT---A--A-----C--G------CT-A-------ACTT- 8568
SYK.KE.x.SYK173             -CC--ACAG--AC-GG......GGAC-TTCGGGA---CA--AATTCTCACCTGGAGATAT--TG...C-G.........-ATGAAGG-..............--CACA-GTCTC--A-----C--T--GTGT-----AAC---AC-C---A--T-A--C-----ATTG- 8878
RCM.GA.x.GAB1               ........................A-G-GGCAT-AC----GAA--GCAAGCTTCA-T--AT-AC...-T--ACTGG-A--AAGATACT..............--T--CACCTT---A--A---------G-A------A-----AC-T-----------G-------CT-- 8580
RCM.NG.x.NG411              ........................................AGG-G-GTTT-CA-T--A-A--CAGC-GT-A-TGGA---AATAGATTGGTATG--GATACA--T--TACCTT---A--G---------G-G----A--A--C--CA----G-------T-------ATT-- 8663
GRV.ET.x.GRI_677            TAAAATGA-TC--..................G-C---CCGCCG-CTGCAAGAAGGAGACA-CTT-...G---AGTGG...-ATGATGAT..............--A--A--A...--A--C--C--T--GCA-----GA--C--C--------AG-----C--T--ATT-- 8897
SAB.SN.x.SAB1C              CTCTA--A---A-..................G-A--T-T--AA-TG-AACATC-AGG-AG-CAG...CCCT-GTGG-A--ATGA-GAT..............--A--A--A...--A--C---------A-GG---TGCT----AC--C--G-------A--T---TTG- 9230
TAN.UG.x.TAN1               AAA-A-CA-CC-T..................G-A---AG--AATTGGCAGCTG--GACA--TGG...G---ACTGG...-AACCAGAG..............--A--A--A...------C----------A-G---A-GA--G----------A------A--T---TTG- 9040
VER.DE.x.AGM3               TAAAATGA-TC--..................G-A---AG--AATT-CAAGAAGGAGATAA-TGG...G---AATGG...T-TGATGA-..............--A--T--A...--A--A---------A-G---AAGA--G--GC------A------C--T--ATT-- 8551
VER.KE.x.9063               CAACATGA-TC--..................G------AG-CG-CTGCAAGAAGGAGAC-A-TGG...G---AATGG...T-AGATGAT..............--A--T--A...------A-----T--A-A---AAGA-----A--G-G-A------------ATT-- 9053
VER.KE.x.AGM155             CAACATGA-TC--..................G-A---AG-CGTCT-CAGGA-GGGGACCA-TGG...G---AATGG...T-AGATGA-..............--A--T--A...------A---------A-G---AAGA--G--AC------A---A--A-----ACTT- 9030
VER.KE.x.TYO1_patent        CAACATGA-G-A-..................G-G---AGG-GGCTTCAAGAAGGAGACA-CTGG...G-A-AGTGG...T-GGATGAT..............-----A--A...------A---------G-G---AAGA-----C-----G-A----------T--ACTT- 9014
COL.CM.x.CGU1               ...........T-GC-CTTG--C-CTTTGCT--A--CATCCAGATGTATATT-CCACAAC.........C-A-AACCA-AGGACG--...............--A--T-TA--A-CA--A-G-TAT--G-TGT---GAC-G--AG--C---AGGC---CATT---CT-C 8285
```

```
                                        3' LTR U3 start                                              Nef premature end in HXB2
H1B.FR.83.HXB2                CTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGG....................GTC 9195
Nef                           A  V  D  L  S  H  F  L  K  E  K  G  G  L  E  G  L  I  H  S  Q  R  R  Q  D  I  L  D  L  W  I  Y  H  T  Q  G  Y  F  P  D  *  Q  N  Y  T  P  G  P  G  . . . . . . .  V
H1A1.UG.85.U455               --T-T-----C---TT--------------------------T--T---------A-G-A--------A-------------G----T---------A-T---------G------T---------------A.................A-- 8646
H1B.US.90.WEAU160             -GCAT--------------.----------------------------T---------A--------------------------G-----------------T---------G-------------------------.................ACT 9195
H1C.ET.86.ETH2220             -AT-C-----C----T--------------------------T-----T-----A-G-A---G--G--G--T------T-----G----TA-----------T---------G---A------------A------.................--- 8595
H1D.CD.84.84ZR085             ------------------------G-A----------------G--T------G-A---------------------T--G-------------------T---------G---A-------------------.................A-- 8743
H1F1.BE.93.VI850              ----------C--T--------------------------T-----T-----A-G-A---GG-----C------------G-T--------------------------G----------------------.................A-- 8472
H1G.SE.93.SE6165              --T-T-----C---TT--------------------------T--------T-----A-G-A--------G--------C------G----A-T--------A--------A---G-----------G----------.................ACT 8637
H1H.CF.90.056                 --T-T-----C-----T-------------------------T--T-----T-----A-GCA------G--C---------T-A---G----TA--------------------C-G--------------------.................-AG 8518
H1J.SE.93.SE7887              ----T---------TT--------------------------T--T-----T-----A-G-A--------G---------------G-TC-TA---------T-----------C-G---A----------------.................A-- 8493
H1K.CM.96.MP535               --T-T-----CG--TT--------------------------T--------T-----A-G-A--------A---------------G-T--T-----------T-T---------G---A----------------.................A-- 8382
H101_AE.TH.90.CM240           --T-T---------TT--------------------------T--------T-----A-G-A--------G--------CT-A---G----TA-T-----------T---------G---A----------------.................A-- 8763
H102_AG.NG.x.IBNG             ----C-----C-------------------------------T--T-----T-----A-G-A--------A---------------G----TA--------A-T-----A---G------------------------.................ACT 8722
H103_AB.RU.97.KAL153_2        --T-C-----C-------------------------------T--T-----T-----A-G-A--------A---------------G----T-----------------------G--..........-----ATTA-A.................. 8363
H104_cpx.CY.94.CY032          --T-------C-------------------------------T--------T-----A-G-A--------G---------------G----T-----------T-T---------G-G-T--------------A----.................-AG 8607
H1N.CM.95.YBF30               --T-T------TC-TT-----------T--------------G--TGG---AG--A---G--------T--A--C--C-----G--T-----------AT-C-------C-G---T--------------------A.................A-T 8785
H1O.BE.87.ANT70               -AT-T--C--C---TT------------------A---------T-----T--------T-A----GC---A-----G-----T---G-G--TA----T--G--A-T-------------------------A-----A.................AC- 9282
H1O.CM.91.MVP5180             -CT-T--C--C---TT------------------A-------T---T-----T-------T-AG---GC---A-----G-----C------A--T-----T--G--A-T----------G-----TGT-------G--A-----A.................CCT 9305
H1P.FR.06.RBF168              -AT-T--C--C---TT--------------------------R--T---T-----...................................................................................... 8654
CPZ.CM.01.SIVcpzCAM13         -CT-T--C---TC----------G--G---------------T--G--T-------G-A---G-----A--T-A-----C---G-T--T-T----------T---------C-G-----------C-----A-----A.................-AG 8834
CPZ.CM.05.SIVcpzEK505         --T-T------TC-TT-------G------------------G--TGG--AACC-AG-------A--T--A--C--C---G-T--T--T--T-------AAT----------C-G-------------T-----A-----C.................--- 8702
CPZ.CM.05.SIVcpzLB7           --T-T------G--T---------------------------T--G--T-----TA---AG-------A-----------C--CG--G-T--T-----------AAT----------C-G-----T-----C---------------.................A-T 8727
CPZ.CM.05.SIVcpzMB66          --T-T-----C-------------------------------G--T-----A---AG-------G-----------T---G----T--T---------A-T-----------G---A-------------------.................A-- 8840
CPZ.CM.05.SIVcpzMT145         -CT-T-----CTCT----------------------------T----CTGG--AAGG-A------G--G---------T-A---G-T--T-----------T-----------C-G---A-------C--G-------A.................--G 8728
CPZ.CM.-.SIVcpzMB897          -CT-T-------------------------------------CT-------------G------G-----------C--G----T--T-----------T-----------G---A----------A-----------.................--- 8730
CPZ.CM.-.SIVcpzDP943          --T-T---T-AG--T---------------------------T-----AG----------G--T-G-----C--G----T--T---------A-T-----------G------T----------A-----A.................--- 8788
CPZ.CM.98.CAM3                --T-T---T-A-------------A-----------------T-----AT--NNNNNNNNNNNNNNN-G---ACG--TA---T--------G----T----------------G------T---------A-----A.................--- 8638
CPZ.CM.98.CAM5                --T-T---T-A-------------------------------T--G--T-----AC--A---G-----G-----A-----C--G-T--T---------G-AATT--------C-G------T------------------A.................--- 8941
CPZ.GA.88.GAB1                --T-T------TCA----------------------------T--G--T-----AGG---------G---------C--C---G----T-------------------T---------C-G------------A---A-----A.................ACA 9268
CPZ.GA.88.GAB2                -CT-T-----C-------------------------------T--G--T-T----A-G----------G-----G--C--C---G-T--T--T--C-----A-------------------G-A-T-----T--T--T--A-----A.................--G 8606
CPZ.US.85.CPZUS               -AT-C---T-AG--TT--------------------------T--G--T-T---AG----------G-----A-----C--G-T-----T--------AAT----------C-G------T-----C---------T--A.................--A 9234
CPZ.CD.90.ANT                 -------C--GTCATGN-----------------------TN--------T-NN-A-TT-A---GC--C-----A---A------G-TTA---C--G--AAT---------------G------T------GA--A------A.................--T 8649
CPZ.TZ.00.TAN1                -AA-------GTCATGG--C----------------------T------T--TT---T-C--A---GC--CC-----A--CACC-----G--TA-T-----G--TGT--T--A--C-G-----------C--T--A-----A.................A-- 8881
CPZ.TZ.2001.TAN2              -A--------GTCATGG--C----------------------T---T--TA-TT---TAC--AG---GC--CC-----A--CACC-----G-TTA-T--------AATA-----A--C-G-----------C--T--A-----A.................--- 8901
CPZ.TZ.2002.TAN3_1            -A--------GTCATGG--C----------------------T------T--T-----T----A----GC--CC-----A---ACC-----G-TTA------------AAT-----A--C-G-----------C--T--A-----A.................--- 8955
GOR.CM.2007.SIVgor2139_287    -AT-T--C--C---TT--------T-----------------T--G--TT---T--GGCC---AGT-------T-A--C--C---G-T--T--T--T--G--A-----------C-G------T-------------------.................A-A 8792
GOR.CM.2007.SIVgorCP2135con   -AT-T--C--C---TT--------------------------T--G--TT---T--GGCC---AGT--G--T-A-------C---G-T--T--T--T--G--A-----------C-G------T--T----------------.................A-A 8776
GOR.CM.2004.SIVgorCP684con    -AT-T--C--C---TT------------------A-------T--G--TT---T--GGCC---AGT-------T-A--C--C---G-T--T-----T--G--A-----------C-G------T-------------------.................--A 8721
H2A.DE.x.BEN                  -AA-----A-GTCA--T---A---------A---------C-----A-GT--T--AGTAGG--G-----TAGA-----A--CA-A-ACC-AG-AA-AGAGG---GATAA-A--A---G-----T--T--T-AT--------A.................--A 9614
H2A.GW.x.ALI                  -A------A-GTCA--T--AA---------G-----------A-GT--T--AGTG-G--------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G---GATAA-T--A---G----------T--T-AT--------A.................A-A 9613
H2A.SN.x.ST                   -AAG----A-GTCA--T--GA-----------------------GTA-T--AGTG-T--G----GTAGAG----A--CA-A-ACT-AG-AA-GGA-G-G--AATAA-TGGA--C-G-----------T--T-AT--A-----A.................--A 9053
H2B.CI.x.EHO                  -A-----CA-GTCT--T-------------------A---------A-TT-CT-TAGTG-G-----G--TA-A--A--A--CACA-ACT-AG-AA-TGA-G-----ATTG-GT---GA-G-------------AT--------A.................--A 9501
H2B.GH.86.D205                -A-----CA-GTCT--T---A--------C----------------A-GTA-T-TAGTG-G---------CAGA--AT-A--CACA-ATT-TG-GA-TGA-G-----ATTG-GT---GC-G---A-----T--T-AT--------.................A-A 9527
H2G.CI.92.ABT96               -AA----CA-GTCA--T---A-------------------------A-TTA-T--AGTG---------TAGA-----A--CA-A-AC--GG-AA-GGAG-----TATAA-T--A---G------T------G----A-----T.................A-A 8999
H2U.FR.96.12034               -A------A-GTCA--T---A----------A----------T--A--T--T-T--T--G---------TAGA---T-A----A-C-ACT-AG-AA-TGAGG-G--AATAA----A---G------T------CT----A-----A.................ACT 9095
SMM.SL.92.SL92B               -AA----CA-GTCA--T---A---------------------T--A-TTA-T--AGTG-G--G-----TAGA------A---A-A-ACC-AG-AA-AGAGG----TAT-A----A---G------T--T---AGT---------A.................A-A 9022
SMM.US.x.H9                   -AA----CA-GTCT--T---A-------------------------G-TTA-T--AGTATT--------TA-A--AT-A-----R-ACT-AG-AA-RGA-G----AATAA-A--A---G---A--T------G---------------A.................A-- 9054
SMM.US.x.PGM53                -GA----CA-GTCT--T---A-------------------------A-TTA-T--AGTG-G--------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-AC---A---G------T-------T----A-----A.................A-- 9508
SMM.US.x.SIVsmH635F_L3        -AA----CA-GTCT--T---A-------------------------A-TTA-T--AATG-G-A------TAGA--AT-A---A---AC--GG-AA-GGA-G----AATAA-A--A---G---A--T------TT---------A.................ACT 9593
                                        3' LTR U3 start
MAC.US.x.239                  -AA----CA-GTCT--T---A-------------------------A-TTA-T--AGTGC---------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A---G----G-T-----CT----A-----A.................A-T 9571
Nef                           A  I  D  M  S  H  F  I  K  E  K  G  G  L  E  G  I  Y  Y  S  A  R  R  H  R  I  L  D  I  Y  L  E  K  E  E  G  I  I  P  D  W  Q  D  Y  T  S  G  P  G  . . . . . . .  I
MAC.US.x.251_1A11             -AA----CA-GTCT--T---A-------------------------A-TTA-T-TAGTGC---------TAGA---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-A-A--A---G----G-T-----CT----A-----A.................A-T 9567
MAC.US.x.251_BK28             -AA-----A-GTCT--T---A-------------------------A-TTA-T--AGTGC---------TAGA---T-A--CA---ACT-AG-AA-GGA-G-----AT-A-A--A---G----G-T-----CT----A-----A.................A-T 9553
MAC.US.x.EMBL_3               -AA-----A-GTCT--T---A-------------------------A-TTA-T--AGTGC---------TAGA---T-A--CA---ACT-AG-A..........AT-A-A--A---G----G-T-----CT----A-----A.................A-T 8693
STM.US.x.STM                  -AA----C--GTCA--T---A-----AGC-----------------A-TTA-T--AGTG-G--------TAGA--AT-G--CA---ACT-AG-AA-AGA-G---GATAG-T--A---G------T------G----A--------.................A-A 9222
MNE.US.x.MNE027               -A------A-GTCT--T---A-------------------------A-TTA-T--AGTG----------TA-A---T-A--CA-A-ACT-AG-AA-GGA-G-----AT-G-A--A---G----G-T-----CT-G--A-----A.................CCT 9048
DEB.CM.99.CM40                TGA----CTACTCT--T-----G-----------T--------G-ATG-TT--T-----TGCC-----G--T-C---TT-A--G--TCATGC-C-GA-TGA-TGG--AATTA-A--A-GC-G--TAC-G------AAG--A--T--A.................--T 8836
DEB.CM.99.CM5                 TGA----CTACTCT--T-A---------------T--------G-ATA-TT--T-----TGCT----------T-C---AT-A--AA-CCATGC-C----TGA-TGG--AATAA-A--A-G--G-TT---G------GA---C-----A.................CCA 8763
DEN.CD.x.CD1                  TGA-----A-GTCT--T-----G-----------A------C-G-ATA-TT--T--AGT-C-------G--T-C---TT-A--C--C-ATGC-C-GA-TGA---T-----T-A-AA-A-GA-G-----C------GAT--------C.................A-T 9189
LST.CD.88.447                 TGA----C--CTCA--T---A-------------T---------------TA-TGG-----G---------G--C---T-ATG---T-ACTGTG-GA-TGA-TGG--TCT-A-TGGA--C-TCAT--CT------GAT-----------.................ACG 8201
LST.CD.88.485                 TGA----C--CTCA--T---A-------------T---------------TA-TGG-----G---------G--C---T-GTG---T-ACTGTG-GA-TGA-TGG--TCT-A-TGGA--C-TCAT--C-------GAT--A--------.................ACA 8198
LST.CD.88.524                 TGA----C--CTCA--T---A-------------T---------------TA-TGG-----G-----GA-T--C---T-AACA--C-ACTGTG-GA-TGA-TGG--TCT-A-TGGA--C-TCAT----------GAT--A-----A.................ACT 8201
LST.KE.x.lho7                 T-A----C--CTCA--T---A-------------T---------------TA-TGG-----G---------G--C---T-AAT---C-ACTGTG-GA-TGAGTGG--TCT-A-TGGA--C-TCAT----------GAT--A-----A.................ACA 9276
SUN.GA.98.L14                 TAA-------CTCA--T---A-------------T--------------CTGGT-----AG--CC---G----A---T-G---T---ATGCAG-AA-TGAGTGG----T-A-AA---GC-G---AG-T-----CAA------T--A.................--G 9347
MND-1.GA.x.MNDGB1             TGA----CT-CTCT--T---A--------------------G---A--TA-T-TAG-AGG------G-----A-----A-----C-ATGCAG-GA-TGAGTGG--A-TTGAA----GA-G---AC-G--T--GA----T-----A.................AC- 8631
OLC.CI.97.97CI12              TAA-T--G--CTCG--T-------------------------G---A-GTA-TGG--A-CT-----GG-GA--T-GA--T------ATGCT-TGA-TTGTTGG--AATAA-AGG-A-C-TC---GCAACAGAGGAG-TAGA--A-..............GGCACGT-A 8600
WRC.CI.97.97CI14              TA-----C--CTCG--T--A--T-----------A--------G-ATA--T--TGG--AGCG------G----A--AG-GA-C--C-ATGCAC--A-TGA-TGG--GCTGA-AGA--GC-G---AGCT------GA---A--------.................A-A 9146
WRC.CI.98.98CI04              TA--------CTCG--T--A--T-----------A--------GTATA--T--TGG--T-CT------G----A--AG-GA-C--T-ATGC--T-A-TGA-TGG--ACTGA-AGAA-G--G----GCA------GAG--A-----C.................A-A 9188
WRC.GM.05.Pbt_05GM_X02        TA-----C--CTCT--T--A--T-----------A--------G-ATA--T--TGGAGT-CT------G----A--AG-GA-CT-A-ATGC---T--TGA-TGG--TATAG-GGAG-GA-G---AGC------CGAT--A--------.................--- 8084
MND-2.CM.98.CM16              TG--T---A-GCC---------------------------------A-TTGG--T--TA-------G----G---T-G--CT-A-ATGCTC-AA-TGA-TGG--A-T-A-T---ACC-G---ATCT--T--TGAT--C--------.................A-A 9126
MND-2.GA.x.M14                TGA-----A-GTC---------------------------------A-TTGGTT---TAG------G----G---T-AA-C----ATGCTC-GA-TGA-TGG--A-T-A-A--A---G---AG-A-----TT------------.................A-A 9067
MND-2.x.x.5440                TGA----CA-GTC---------------------------------A-TTGGTGG--TA-------G----G---T-G--C--T-ATGCAC-AA--GAGTGG--ACTAA-A---A---G------------AAG--A--G---.................A-A 8822
DRL.x.x.FAO                   TA-----CT--TC-----------G-------------A-----A-TTGGTGG--T-TC------G--ACA---T-G-----C-ATGCTC-AA-TGA-TGG----T-A-AAAA-GC-G---A-GT-----CAAG--A--T--C.................A-- 8907
MON.CM.99.L1                  TGA-G--C--CTCT--T-A---G-----------A--------G--A-A-GT--T---GTG--GAC-----TAGA-AGA-A--G-AA-ATGC---T-TGGA-TGG----TAA-----A-GA-G--TAC-A--T--GGA-----------.................--- 9050
MON.NG.x.NG1                  TGA-G--C---TCT--T-----------------A--------TAA---GT--T---GTGTWGACC-Y--TC-A-AGT-A--G-AA-ATGC---TGT--TGTGG--ACTTG-G----GA-G--TAC-G-TT--CY-C--A-----C.................--G 7687
GSN.CM.99.CN166               TG--G-----CTCT--------G-----------A--------G--CA-GT--T---GTG-GGAC--G--CC-G-AGT-G--A-AA-ATGC---T-TGGAGTGG--CTGG-G----GA-G--TCTCA-T---T-----C-----A.................AC- 9015
GSN.CM.99.CN71                TG--G-----CTCG--------G-----------A--------G-CCA-GT--T-T-GTG--GAT-----CC-A-AG--A--GAGC-ATTG---TT-TGAGTGG--ATTG-G----GA-G--TCC-G-GG--T-----A-----A.................A-- 9012
TAL.CM.00.266                 G-A----CT-CTCGTT------G-----------A------A-G--A--T--T--AA-A--C-T--G--T-CA---T-GA-----ATGC-C--A-TGA-TGG--ATAC-G--A--C-G------------GAG--A-----A.................ACA 8959
TAL.CM.01.8023                G-A----CT-CTCGTT------G--------A----------A-G--A--T--T--AA-A-C-----G--T-CA---T-AA-C----ATGC-C-TA-TGAGTGG--ATAC-G--A---G----TCA------GA---------A.................ACA 8509
MUS-1.CM.01.1085              TG--G---A-GTCG--T--C--------------A-----------A-GT-CGT-AGTAC-GA---G-TCC-G-AGT-G--GACC---GC---TGTGGAGTGG--AATAA-A--A-GA-G--T--TG--T---GAT--C------.................ACT 9021
MUS-1.CM.01.CM1239            TGA-T--C--CTCT--T-A---------------A-----------A-GT--T-T-GTG-GGA------CC-A-AAT-G--GACT-AT-C---TGTGGACTGG--AATTG-A--A-GA-G--TAC-G-TT---GAG--C--G---.................AC- 8990
MUS-2.CM.01.CM1246            TGA----C--CTCT---A---G---A----A-A-----A-----AA-GT--GTG-G-AT-GAC----TGC-A-AA--A--GACT-ATGC---TGTA-A-TC-AAAATAG----AAGA-CA-TA-TA----GGAT--C-----C.................AC- 9086
MUS-2.CM.01.CM2500            TAA--A-C--CTCT--T-----G-----------A-----------A-GT--T--AGTG-GGA------CC-A-AAT-G--GACC-ATGCA--TGTGGAGTGG--TCTTA------GA-G--T--TG-----GA---C--T--T.................--T 9042
SYK.KE.x.KE51                 -CA-------CTC---T---A-----AG------A------C-G--CA-GT--TT--GT-----AG-----T--G--T--AC-G----ATT-AA-GA-TGAG--T--TGTGA-AGA---C..ATC-C---T--GAGT--------A.................ACA 8714
SYK.KE.x.SYK173               -AA----C--CTCT--T---A-----A-C-----A------C-G--CA-G-A-T-T-GTG---AG---G-T--A--T---C----C-ACC-AC-GA-TGAG--T--AAT-A-A...-CAG-ATC--T-----CT-----------A.................AC- 9024
RCM.GA.x.GAB1                 -AA----CA-GTCT--------------------------------A-TTA-TGGAGTATC---------AGA--AT-G---A---ACC-GG-AA-TGAG--T---ATAA-A-------G---A---------T------------A.................A-A 8729
RCM.NG.x.NG411                -AA-------ATCT--------------------------------A-TTA-TT---AG----G---AGGCGC--TT-A--C--C-AC--GG-GA-TGA---T--AATAA-A--A---G------T-----T--------------.................A-A 8812
GRV.ET.x.GRI_677              -A--G--CT--TC------------TC--------------T--A--TA-T----TG--------G--A-G-----GA--T---ATGC--TGA--GAGTGG--AATAA-AGA-----G---AGCT---T-------C--G---.................A-A 9046
SAB.SN.x.SAB1C                -AA----CT--G----------------------------------A-TTA-T--AGTG---G---A-GA-G---T-G-----C-ATGCTCTTA-TGAGTGG--ATAG-AGA--GA-G------T--T---GAT--------T.................AC- 9379
TAN.UG.x.TAN1                 -A------T-CTC---------------------A--------G--CA--TA-TGG--A-C--AG---G-GC-A-----GA----C-ATGC-CTAA-TGA-TGG--ATTA-AGA-----G---AGCA--TT-C--T--------C.................AC- 9189
VER.DE.x.AGM3                 -G--G--CT--TCG--------------------------T--A-TTA-T-----G-C--G--GA-TC-G-----AA-C----ACGC-CT-A-TGAGTGG--AAT-A-TGA-----G-A-TGCT-GGT--GA---A-----A.................A-- 8700
VER.KE.x.9063                 -A--G--CT--TCG--------------------A-------T--A-TTA-T-----G----G--GA-TAGA-----CA-C----ATGCTCTTA-TGA-TGG--AAT-A-TGA-----G-A-CGC--GGT--A----A-----C.................--A 9202
VER.KE.x.AGM155               -A------T--TCG--------------------A-------T--A-TTA-T-----G-T--G---A-TA-G-----A------ATGC-CT-A-TGA-TGG--AAT-A-TGA-----G-A-CGCA-GGT--AA---A--T----.................A-A 9179
VER.KE.x.TYO1_patent          -A--G---T--TCG--------------------A-------T--A-TTA-T-----G-T--G---A-TA-G-----GA------ATGCTCTTA-TGA-TGG--GATAA-TGA-----G-A-TGC--GGT-GAAG--A-----A.................A-A 9163
COL.CM.x.CGU1                 TG--G--CA-CTCTTG------G-----------T--------G---A-T---AGAA---CTGAG---G-T--CC-GA-A--G-A--ATGC----ATTGAGTGG----G---AAAA-G--G-TT-G-G--TGA-GAT-A--TG--AGAGGATGGAGCATTAAAAGAG-A- 8455
```

```
H1B.FR.83.HXB2              AGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAGAGGCC...AATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGGATGGATGAC.........CCGGAGAGAGAAGTGTTAGAGTGGAGGTTTGA 9353
Nef                          R  Y  P  L  T  F  G  W  C  Y  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  R  E  V  L  E  W  R  F  D
H1A1.UG.85.U455             -----C-----A--A-----------------------------T---C-G-AG---G--A--T...C-GG----------A---TC-AC--------A-AT---AA----AG-A-----T.........GAA---A---------AT-----A------ 8804
H1B.US.90.WEAU160           -----C------TG-------------T--------------G--A-----G---G-------A--T...---G-----------A----------------A------------------------.........-AT---A------A---AT-----A------ 9353
H1C.ET.86.ETH2220           ---------------------------T-----------------C---AG-G-AG-G----AAT-...---G-----------A-T-----C-G-------C-----------A-----G--T.........GAA--C--------A---A------A------ 8753
H1D.CD.84.84ZR085           ---------T-A--------G-------T-G----------------T-----GGT-G-----A-----...-C-G-------AG---A-T-----C---------A-CT---A------A-----G---.........GA-----AG----------T------A---A- 8901
H1F1.BE.93.VI850            -----------------G---------T-----------------T-----AG--G-G--A-----...---G-----------A-T-------------CA-----AA----A-----G--T.........GAA--C--------AC-GAG-----A------ 8630
H1G.SE.93.SE6165            ----TC--C------------------T---A---------A-G--T----CAG--G----G--A---...------------------A---TC-A---------CA-CT---A------A-----G---.........GAA--C---------C-G-T------A----- 8795
H1H.CF.90.056               ----T---C----------G-------T---------------AA-T---C-GG--G-----C-----...---G-----------A------C--------CA---------------A-----G--T.........GAC-G---------C-GAT-----AA----- 8676
H1J.SE.93.SE7887            -----C---T----A------------------------------T---AGCG-AG---G--A--T...---G--------A----A-T--C----------CA-AT---A------A--A--G--T.........GAA--A-----------C------A------ 8651
H1K.CM.96.MP535             -----C-------------G-------------------------T----CAG-AG---------A--...-CAG--------G---A-T--C----------CA-A-A--A------A-----G--T.........GAAC-C------A-A---AT-----A------ 8540
H101_AE.TH.90.CM240         ----TC------TGT------------T-----------------C-A-AGAG-AG---G----A-...--C----------A----A-T--C----------CA-------A------A--A--G---.........GAA--A---------C-GAT-----A------ 8921
H102_AG.NG.x.IBNG           ----TC-------------G-------T---A---------A-G--T----CAG--G----G--A---...---G------------A----T--A---------CA-CT-T-AA----A-----G---.........GAC--T---------C-GATC-----A----- 8880
H103_AB.RU.97.KAL153_2      ...-TC-----A--A-----G----------------------C--C----C-G-AG-----------...-CAG-G----------A----C-A---------A-AT---AA--C--A--------T.........GA---A-A----------AT-----A------ 8518
H104_cpx.CY.94.CY032        ----TC------TG-------------T---------------A--T---C-GG--G-G---------...-C-G----------------T-T---C-G------A-A----A------A-----G--T.........GAA-----------A---A------A------ 8765
H1N.CM.95.YBF30             -----C--CG-A---------------T---A---------T-GTCAG-T--AG-AG---------T...---G-------C----ATGC-C-C--------CA-AT-T-AA-----AGCA-----T.........GATC-T-A---------G-T----C-A----- 8943
H1O.BE.87.ANT70             --G-TC---------A----------TG-TT--A-----------GTCAGA---AG--GC----AGACTAGGA----C-T-T----GGG-T-ATC-CC-G--T--A-CAT-TGCC-----AT-T--A--T.........A-AC-T-A----A-AC-GAT-----A------ 9443
H1O.CM.91.MVP5180           ----TC---------A----------TG-TT--A--G---------GTCAG----AG--GC---GAGACTGGGT----C-AAT--AG-TG-T--TC-TC----T--A-CTT-TAAT-----AGCT--G--T.........G-AC-CG-G--GA-AC--A-A---CA------ 9466
H1P.FR.06.RBF168            ............................................................................................................................................................................ 8654
CPZ.CM.01.SIVcpzCAM13       -----C----CCTG--------------TT--A---------TT-GACAA-T--GC-AG-G----CA---...---G-G-----T----A-T----AC-G------A-AT-T-A------A-----G--T.........GAAAGT-A------C---ATA---C-C----- 8992
CPZ.CM.05.SIVcpzEK505       -----C--C------------------T------------C-GTCA--T--GG--G-------A--T...---C-------C----ATGT-C-------T--CA--T---AA-----A-----A--T.........--A--C-A-------C-G-T----C-C----- 8860
CPZ.CM.05.SIVcpzLB7         ---------A-A--------CG------T------------C-AACC-----AG--G-G--G-----T...-------------------ATCTAC-AC-G-----AA-AT---A------A-----A--T.........GAA--C-A-------C-GATA----AA-A--- 8885
CPZ.CM.05.SIVcpzMB66        --------------------C-------T---A---------T-A--A-----AG-AG-G---AGA--A...---G-------C----A--TA---------T--CA-CT---A---C--ACAA--G--T.........GAA-CA---------C---TC----CA----- 8998
CPZ.CM.05.SIVcpzMT145       C-C--C-----A-----C--G-------TT--A--G--C--TT-GCCA--T--GG--G-G--GA-A---...---G-------A----A-T--C--C-----T--CA--T---A------A-----G--T.........GA---A---------AC-G-TA----CT----- 8886
CPZ.CM.-.SIVcpzMB897        -----C--GG-------C--C-------T---------G--T--A--T---ACAG-TG-G--G--A--A...---G-------C----ATGTGC-A-----T---A--T---A---C--ACAA--A--T.........GAAC-C----------C-GAT----CA------ 8888
CPZ.CM.-.SIVcpzDP943        -----------A--A-----G-------TT--A--G--T---C-GACA-----AG-AG-G---A-A---...---G-------T-C--A--TAC--C-T-----CA-CT---A-----AA--A--A--T.........GAAC-TG--------C--G-TT-----A-A--- 8946
CPZ.CM.98.CAM3              -----C-----A--A-A---GG------TT--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-TG--------C--G-TC-----A----- 8796
CPZ.CM.98.CAM5              -----C-----A--A-------------T------G-----TC-GACA-----GG-AG-G--GA-A---...-CG-------C-C--AT-TAC--C-T-----CA-AT---AA-----A-----G--T.........GAAC-TG--------C----TC-----A----- 9099
CPZ.GA.88.GAB1              ----TC-----CTGT------------T---A-----G--CC-GACAGAG--GC--G-----C-A---...---G-------T----A-T--C----G--T--CA-AT-T-A---------------A--T.........GA---C-A---G--C--G-TC---C-C----- 9426
CPZ.GA.88.GAB2              C-T-----GT------A-----------T---------G-----ATCA-----GGCAG--------A---...-C--G------G-T-ATCT-C-A-----T--CC--T--ACT-----AT-T--G--T.........GAA----AG--G--TC-GATA----A------ 8764
CPZ.US.85.CPZUS             --------C--A--A-A---G-------TT---T----C--TC-CACAGA---GG-AG---GC-----...----------------C--A--TAC--C-G-----CA--T---A------A-----A--T.........GAAC-TG-C-----C---ATC----CA------ 9392
CPZ.CD.90.ANT               -----------CTGTAGA--------T-TT---T----------A--AT--CCACCTGAT.................-AT----GG-A--TAC-CC-C-------CTT-TACA-----AGAT-GA---..........--AC-T-A---GA-C---AGA---GAA----- 8792
CPZ.TZ.00.TAN1              -----C------TGTAGG-------TA-TT---T-G-----G--A--C...............CCAC-A...G-AG-T-AT-----G-A--T----C----T--A-CCT-TAGC-----A-CTACC--T..........--A--TG----GACTC-GATC---C-C----- 9024
CPZ.TZ.2001.TAN2            -----C--G--A---AGG------CTG-TT-----G--------A--C---CCAG-AGAT.................-AC---C-G-A--T----C---G---A-CCT--AGC-----A-GTA-C---..........--A--AG----GACTC-GATC---C-C----- 9044
CPZ.TZ.2002.TAN3_1          -----C------GT-AGG------CTG-TT-----G--------A--C---CCAG-AGAT.................-AT-----G-A--T---------T--A-CCT-TAGC-----A-GTA-C---..........--A--AG----GACTC-GATC---C-C----- 9098
GOR.CM.2007.SIVgor2139_287  --------C----------------CTG-TT---T----G-----CTCAGAG-CAG-AGCG--G---ATGGGA-CATC-AAT-----GG---TGC--C-G--T--A-C-T--AGC-------GA--C---..........--AT-TG----GA-CC-T-TC-----A--C-- 8953
GOR.CM.2007.SIVgorCP2135con --------C----------------CTG-TT---T----G-----CTCAGAG-CAG-AGCT--G---ATGGGA-CAGC-A-T----GAG-T-TAC--C-G--T--A-C-T--AGC-------GA--C---..........--AC-TG----GA-CC-T-TC-----A--C-- 8937
GOR.CM.2004.SIVgorCP684con  --------C-------A-------CTG-TT---T----G-----CTCAGAG-CAG-AGCT--G---ATGGGA-CAGC-A-T----GAG-T-TA---C-G--T--A-C-T--AGC-------GA--C---..........--AC-TG----GACCC-T-TC--------C-- 8882
H2A.DE.x.BEN                --G--C---A--TA---C--G---CTG-GG---------T----A--A...CTCTCACA---G-CA-AG...G-AG-T...---GC--A-T----AG-------A-CACAAACAAGCA-ACAT-----T.........GA-C-TG-G--GACA----T----CA------ 9766
H2A.GW.x.ALI                --G--C--GA--TT------G---CTG-GG---------------A--T...-TCCCACA-----G--AG...G-C-CT...----CTCA-T--C--C-------A--ACAAACAAGCA--CAT--------.........A--C-TG-G--GACA----TT------A------ 9765
H2A.SN.x.ST                 --G------AA-TT------G----TA-GG---------------A--T...-TCCCACA---G-GA-AT...G-C-GT...----CTCA-T----AG-G--T--A-CACAAACAAGCA--T-T--------..........---C-TG-----ACA----TT------------ 9205
H2B.CI.x.EHO                --------CAA-TT------C---CTA-GG--------G---A-AA-T...ATG-TAGC----CCA-AG...G-CG-G...--A-C-CATT-TC--G-G--T--A-CACAGACCTCC-CATG--------..........--CC-CGA---GACCC-T-TC---CA------ 9653
H2B.GH.86.D205              --G-----CAAATA------T---CTG-GG-----G---------A---...-TGCCAGC--CGACCCGA...G-GG-G-AG--A-C-CATT--C-AA-G-----G-CACAGA-CTCCTCATG--------.........ATCC-TG-G--GACTC-TATC---CA------ 9682
H2G.CI.92.ABT96             ---------R----G-----C---CTA-GG--A-----------CAC-...-TCTCTGAT--G-CACAG...G-GG-C...----CTCATT-----G----T--A-CACAAACTAGCCCATG--------...........-------CC--------G-C-----A------ 9151
H2U.FR.96.12034             ----GG---A--TG----------CTT-GG-------A-------A--T...-CA--TGATA---CACAG...G-GG-T...----GGCTGTATC--G-GGGTT-A-CACAGACCAGCT-CGA---G----..........-ACTG-G----G-CC----TA----A------ 9247
SMM.SL.92.SL92B             --G--C---AA-TT------G---CT--GGC----G--G-----A--C...-TCTC-GAT--G-CAAA-...---G-T...----C-CATT-----G----T--A-CCCAGACAT--CA-CATTCA---..........--TTG-G--A-G--CC---CC----A----A- 9174
SMM.US.x.H9                 -----C--TA--TT----------CT--GG--AT---WC--G--AA-T...-TCTCAGAT----CTCAG...G-AG-C...----CACATT-----A-G--T--A-CACAAACTTC-CA-TG--------..........--CTG-G----G--AC-G-CA----A------ 9206
SMM.US.x.PGM53              -----C---A--TT------C---CTA-GG--AT-G--C-----AAGT...-CCTCAGAT----CTCAA...G-AG-C...----CACATTA--AG-G--T--A-CTCAGA-ATCCCA-TG------T..........--CTG-G----G--TC-G-CA----A------ 9660
SMM.US.x.SIVsmH635F_L3      -----C--TA--TA------G---CT--GG--AT----C-----A--T...-TCTCAGAT----CCCAG...G-AG-C...----CACATT--C--G-G--T--A-CACAGACT---CA-TG--------..........--CTG-G----G--AC-G-CA----A------ 9745

MAC.US.x.239                -----C---AA---A-----C---CTA-GG--AT----C--T--AA-T...-TATCAGAT--G-CACAG...G-GG-T...---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCA-TG--------..........--TTG-G----G--TC---CA----A------ 9723
Nef                          R  Y  P  K  T  F  G  W  L  W  K  L  V  P  V  N  .  V  S  D  E  A  Q  .  E  D  .  E  E  H  Y  L  M  H  P  A  Q  T  S  Q  W  D  D  .  .  .  P  W  G  E  V  L  A  W  K  F  D
MAC.US.x.251_1A11           -----C---AA---A-----C---CTA-GG--AT----C--T--AA-T...-TATCAGAT--G-CACAG...G-GG-T...---G-GCATTAT--AA-G--T--A-CTCAAACTTCCCA-TG--------..........--TTG-G----G--TCC--CA----A------ 9719
MAC.US.x.251_BK28           -----C---AA---A-----C---CTA-GG--AT----C--T--AA-T...-TATCAGAT--G-CACAG...G-GG-T...----GGCATTAT--AA-G--G--A-CTCAAACTTCCAA-TG--------..........--TTG-G----G--TC---C-----A------ 9705
MAC.US.x.EMBL_3             -----C---AA---A-----C---CTA-GG--AT----C--T--AA-T...-TATCAGAT--G-CACAG...G-GG-T...----GGCATTAT--AA-G--G--A-CTCAAACTTCCAA-TG--------..........--TTG-.----G--TC-GAC-----A------ 8844
STM.US.x.STM                -----C---AA-CAG---------CTA-GG--A-----------A--T..ATGTCA-AT--G-CACAG...G-GG-T-AT-G--CACATTATC--G-G--T--A-CACAGACA---CA-TG--------..........--TTG-G-------AC-G-TT----AA----- 9377
MNE.US.x.MNE027             -----C---AA---A-----C---CTA-GG--AT----C--T--AA-T...-TATCAGAT--G-CACAG...G-GGGT...---G-G-ATTAT--AC-G--T--A-CTCAAACTTCCCA-TG--------..........--TTG-G----G--TC---TA----A------ 9200
DEB.CM.99.CM40              --G------A-TTA---C--G-TCCT--TT--A--G--G-----A--A...ATAGCAGACCC---TTAT...G-G--T-AT----GG-A--T---------TGA--CCCA--A-GGGAT------A---..........--CC-T-----GAG---G-TT----A------ 8991
DEB.CM.99.CM5               --G------ACATA------G-TCCTA-TT--------G-----A---...ATAGCAGATCCT--CTAT...G-G--T-AT----GA-AT-T-C-C-----TGA--CCCA--AAGGACAAGCA--A---..........--CC-T-AG--GAG-C---TA----A------ 8918
DEN.CD.x.CD1                C-G--C--G--AGAG-----G-TCCT--GG--AT-G--C--G--GACT...ATAG--GAG---T-T-A-...------G......G-A-ATT-TC--C-G--TGACAGATATGA-GG-CAACAA-C----..........--TTG------GACCC-G-TA-----A------ 9341
LST.CD.88.447               C-T-TC---T-AG-A----------TG-GGC----G-A------G-CC...TG-G-TGAGT-CACA-AT...CCCTC--ATT-C-GACAGT--C--C-G---T-AAGTCAG---GGG-TCCA---G---..........--CTG-G-----AG-C-TATC---CAC---A- 8356
LST.CD.88.485               C-T--C---T-A--A----------TG-GGC----G-A------G-CC...TG-G-TGAGT-CACA-AT...CCCTC--AT-CC-G-CAAT--C--C-G---T-AAGCCAG---GGG-TCCA---G---..........--CTG-G----GAG-C-TATC---CAC--CAG 8353
LST.CD.88.524               C-T--C-----CG-----------CTG-GGC--T-G-AG-----A-CA...TG-G-TGAGTTTA-A-AT...CCCTC--AT---GACAAT-T---C----TT-AAGCCAGT--GGG-TCC-A--A---..........--CTG-G-----AG-C-CATC---CAC---A- 8356
LST.KE.x.lho7               -----C--T-----A----------TG-GGC--T---A------G-CT...TG-G-TGAGT-TA-A-AT...CCCTC--AT----CACAAT--C-T--G--TT-AAGCCAG---GGG-TCC----A---..........--CTG-G----GAG-C-TATC---CAC--CA- 9431
SUN.GA.98.L14               ---------AA-G-------C---CTA-GG-------C------CACC...AT-G-TGA---CAG--AT...CC---TCATCCTTGTCAAGC-C-AC-G--TT-CAGCCAG-AAGGG-TT-AT--A---..........--TTG-G----GAGAC-TATC----CA----- 9502
MND-1.GA.x.MNDGB1           --------TAA---A-------TCCTG-TT-----G-A------GAGCAG--C--TAGG---T---TAT...GCAGCT.......----ATCATC--------T-CTCCCAGT-ATG-CCTCA---A--T..........--A--AG----GACCC-CAT----TCTGGGAC 8783
OLC.CI.97.97CI12            --GAGA--TAA---GCCA--G----TA-GGG-T--TAGGAG---G--A---TTGCCTGAGACCCTT-AAGAG-GAGGG.................................................................................... 8678
WRC.CI.97.97CI14            ----GG---AA---A-----G-------GG--AT-G--T-----AC-C...TTC-GTGA----TCA-AT...----C-...--CCCATA-TATGA-AG-A-TAAGT-ACTG-AT-CA-CATGT-C---GCCAGAGGAT--CTG-G-----CATC-G-TC----A------ 9307
WRC.CI.98.98CI04            C-C-GG---AA---A-----C-----T-GG--AT----G--T--G--A...TT--GTGA--GTCTAAT...G-C-GC...TCTC-GTA-TA-GA-AAGA--AAGC--TTA-AT-CA--CTGCC-A--ATCAGAGGAT--TTG-G-----CAC--G-TC----A------ 9349
WRC.GM.05.Pbt_05GM_X02      C-C--C---AA---A-----C-------GG--AT-G--T-----GC-C...TTC-GTGA--GTCTC-T...--CTC-...--CC-AGG-TTTGA-AAGA---TGC--CTA-AT-CA-CAG-A-C---GTCAGAAGAT--ATG--AG---TAC--G-TC----A------ 8245
MND-2.CM.98.CM16            -----C--GAA---A---A-G-TC-TA-T------TTGC--G--G-CA...-TGCCTCC---CC---AG...--C--T...----TGT-A--AA---C-C--GT-CTCTCAG---GGAATCCA---A--A........T--TTG-G----GAG-C---T-----A------ 9279
MND-2.GA.x.M14              -----C--GAA--GA-----G-TC-TA-T----T--A-C-----AC-A...-TACCACC---CC---AG...--CC-G...----TGT-A---GC--C-TA--T-CTCTCAG--AGGAATCCA---A--A........T--ATG-G-G--GAG-C-GAT-----A------ 9220
MND-2.x.x.5440              --G-----GAA---T--C--G-TC-TG-T------G--C--G--G-CA...-TGTC-CCCCCCCT--AG...G-GG-T...--ATG--A---GC--C--A--T--TCTCAAA--GGAATCCA---A---..........--A---G----GAG---GAT-----A------ 8974
DRL.x.x.FAO                 --G--C---AA--TG-----G-TCCTG-T------G--C--T--G---...-TAG-TGA---TCT--TA...---C-G...C--TGT-A---GC----GA--T-CAGCCAGA--GG-CCT--T-------..........--A--AG-----AGAC-CAT----CA------ 9059
MON.CM.99.L1                --G--C--GACT-TGCCC--G-TCCTG-GGTGCT-GCGC--G--A-CC...ACC-C-GAG--TTCC-AG...G-GGG--AT---G--TT-CT-C-AACC--T--G-CCTAT-AAGGCC-------G---..........--CC-T---C-GT-C--G-TC-TCTCA--CTG 9205
MON.NG.x.NG1                C-G--C---ACC-T-CC---G-TCCTG-GGTGC--GCGT--G--A-CC...ATCC--GA---CTC--AG...G-CGG--AT--YG-GTT-CT-C--ACC-----G-CCTA--A-GG-C--GAT--A---..........--CC-T-AGC-GT-C--G-TC-TYTCC--CTG 7842
GSN.CM.99.CN166             C-C-----TAC--TACCA----TC---AT-TGT--GCGG-----A-CC...ACA-CTGAG--CTCT-AG...CC-GGG-AT--TG-ATATCTGC--ACC-------CCTAT-A-GGAC---GT--A---..........-A-C-T-AG--GT-C--G-TC-TCTCC--CTG 9170
GSN.CM.99.CN71              C-C--C---ACT-TGCCA--C-TC----GGTGCT--CGG-----G-CC...ATG-CTGAG--CTCT-AG...CCCGGG-AT--TC-ATATCTG---AAC--T--C-CCTAT-A-GGACA-CAA--G---..........-ACC-T--G--GA-C--G-TC-TCTCC--CTG 9167
TAL.CM.00.266               -----C---T--TG------TATCCTG-GG--A--CTGC-----A---...ATCC-TGAG--T--CA-G...G-GG-C...-GAC-T...CT-C-CCCG--T--A-CATATGATGGGCA-GCA--A---..........--ATG-G-----TCCC-G-TC---GTC----- 9108
TAL.CM.01.8023              --G--C---T-ATG---C--TATC-TG-GG-----GTGC--C--G---...ATCC-CGAG--C--CA-G...G-GG-C...-GAC----...CT-C-CA-G-----A-CATATGACGG-CA-CA---A---..........--CTG-G-C----CCC-G-TC---GTC----- 8658
MUS-1.CM.01.1085            C-G-----CAA--TACCA--C-TCCTA-GGTGT--GAG-GA---G-CC...ATG-C-GAG-GCTCT-AG...G-GGG--AT--CC-ATA-CT-C-AGAC-------CCTTT-A-GGGA-AGAA--A---..........-ACC-T-AGC-GT-CC-G-TC-TCTCC--CTG 9176
MUS-1.CM.01.CM1239          C-G--C---ACACTGCCA--G-TCCTG-GGTGC---AGGGA---C-CC...ATA-CTGAG--CTCT-AG...G-GGGG-AT---G-ATATCT-C-CACC--TG-A-CCTA--AAGGCAA-GA---G---..........--CC-T-AGC-GT-C--G-TC-TCTCA--CTG 9145
MUS-2.CM.01.CM1246          C-T--C----G-CAACCA--G-TCCTG-GGTGCT-GAG---G--A-CC...ATG-CAGAG--GTC--AG...CCAGG--AT--CC-GTA-CT-C--AC------G-CCTATG--GG-C--GAT--A---..........-ACC-C-A---GT-CC-G-TC-TCTCC---TG 9241
MUS-2.CM.01.CM2500          C-C--C--G--T-TGCCA--TGTC-TG-T-TGCT--AGG--T--GC-C...ATA-CTGAGA-CTCA-AA...CTGGGG-AT--TG-ACATCT-C---GC--T--C-CCTT-A-TGGGA--GAA--A---..........--CC-T-AG--GT-CC-G-TC-TCTCC--CTG 9197
SYK.KE.x.KE51               -----C--G----T---C--G----TA-GGG-AT-G-CC--TAA---A...ATAG-AGGGT-CTT--TA...G--G-G...--AG----ACTGA--A-G-----A-CAGCAGGAGTG---GCTTCA--AGAC.......--AC-T--G--GAATC-CAT-----AC--CA- 8869
SYK.KE.x.SYK173             C-G--C--GT-A-T-----------TA-GGG-A-----T---AA----...ATAG--GG-TGCTT--AG...T--G-G...---C-T--ACT-C-CC-C--T--A-CCTCTGGA--G----GTTCGAG-ATGGGAGAA--TC-TGTG---C-TCA-CCCCCACCTGGCT- 9185
RCM.GA.x.GAB1               ---------ACACTG---------CT--GGC-AT-G--G-----A--T...-TATCTGAT----CAAGA...G-AG-T...--AG-GCAT--T---C----T--A-CAGAAACAAG---------A---..........--ATG-G-G--G--C--G-CC----A----A- 8881
RCM.NG.x.NG411              ---------ACA-TG-----C---CTA-GGC-AT-----------A--T...-TC-GTGAT----CAA-A...G-G--T...----GGCATT-TC-CC-G--T--A-CAGAAACTAG---------C---..........--CTG-G-G--G--A----CA----A----A- 8964
GRV.ET.x.GRI_677            --G--C--GAGAGT------C-TC----TT--------C-----G--C...CTGC-TGAG--G-CACG-...--CTGT...----GACA-T-TC--A-G--T--A-CACAGA--GGG-AAGATCC---TGGAATA...GATC-TG-------C--G-TC----A------ 9204
SAB.SN.x.SABIC              --G--C---AA-TG-------------TT--A--G-----T--G--C...CTGTCAGAG--G-CTAAG...----GT...------CA-T--C--C-G--T----CTCAAG--GCCTATGA--------..........G-CTG--A---GACCC-G-T-----A------ 9531
TAN.UG.x.TAN1               --GA-G--C-GATG---C--C-TT----TTG--T----G--G--A--T...-TCTCCCAG---CCCAG...G--...........----GACA-T-T---C-C-------CCCAGA-TG-AT--GAAAGC---..........--TTG--A---GACA--G-T-----AA---A- 9338
VER.DE.x.AGM3               -----C---AGATG---C--C-TC----TT--AT-G--------A--C...CTGC-TGAG---CA-AG...-C-TGT...----GACATT--C--G-G--T--A-CACAAG--AGG-AAGACCC---TGGAATC...AACC-TG-------C--G-TC----A------ 8858
VER.KE.x.9063               --------C-GCTG------C-TT----T-----------G--A---...ATGC-TGA--G-CA-AG...-CCTGT...----GACATTA----GCG--T----CACAAA-AA-A-A-GACCCC--TGGCATC...AGTC-TG----GACC--G-TC----A------ 9360
VER.KE.x.AGM155             -----C--GAG-TG------C-TC----T-----------G----CC...CTGC-TGAG---CA-AA...-CATGT...--A-GGCATT-----G-------A-CACAA-------AAGACCC---TGGTATA...AATC-TG-----A-A--G-CA----A------ 9337
VER.KE.x.TYO1_patent        ----TC--TAAATG------G-TC----TT--------G-----G--C...TTAC-TGAG---CACAA...-CATGT...--A-GACATT--C-AG-C--T--A-C-CAGA--GGA-AAGATCCA--TGGTATC...AGCC-TG----GA-C--G-T-----A------ 9321
COL.CM.x.CGU1               ---A-G--T---GTTGC-------CTT-GG--AT----CTACA-A---...C-ACTT....................................................................................................... 8509
```

```
                                                                      TCF-1 alpha binding
                                                                                 Nef end
H1B.FR.83.HXB2             CAGCCGCCTAGCATTTCATCAC...GTGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGC...TGA.................................................................................... 9417
Nef                        _S_R_L_A_F_H_H_._V_A_R_E_L_H_P_E_Y_F_K_N_C_._*_
H1A1.UG.85.U455            ---TAC---G-----AA-A---..AGA--TTAT---------------T--AT--AG-----..---..........................................................................AAGA 8869
H1B.US.90.WEAU160          ----AAA-----------------..--A-----------------------T--AG-----..---.............................................................................. 9417
H1C.ET.86.ETH2220          ---T-A-------CGCAGA---..A-------C-----A-----------T-A---AG-----..---.............................................................................. 8817
H1D.CD.84.84ZR085          ----A-A---------G-G---..AA----AA---AAA-T--------------------------..---.............................................................................. 8965
H1F1.BE.93.VI850           ----A----G---C-GAGA---..A-A---A-----AGA----------T--A-C-AG--..---..............................................................GACTGCTGACACAGAGA 8708
H1G.SE.93.SE6165           ---TA--------CGGAGA---..A-A------------------------A---AG-----..---.............................................................................. 8859
H1H.CF.90.056              ---T--A--------GACA---..T---------TAAA-------------..-AG-----..---.............................................................................. 8737
H1J.SE.93.SE7887           ----TC-------CGGAGA---..A-A--------------------TT-A---AG-----..---.............................................................................. 8715
H1K.CM.96.MP535            ---TTC-------CGAAGA---..--A------------------C-----A---AG-----..---.............................................................................. 8604
H101_AE.TH.90.CM240        ---TGC-------CGAA-A---..--A--------A-A------A-------AT--AG-----..---.............................................................................. 8985
H102_AG.NG.x.IBNG          ---TA-A--G-----CAGA---..ACA--TA--------------------A---AG-----..---.............................................................................. 8944
H103_AB.RU.97.KAL153_2     ---T-----G--TC-AACA---..AGA----------------------T--A---AG-----..---.......................................................................CACAG 8587
H104_cpx.CY.94.CY032       ---T-----G----ACA-G---..--A----------------------TT-A---AG-----..---.............................................................................. 8829
H1N.CM.95.YBF30            ----TC-------AGAAGA--T..--A--AA------------------TT-A-------------..---.............................................................................. 9007
H1O.BE.87.ANT70            T--ATCT----GCAACACC--T..--T--TATGATAACT--C--A---CT----C----GGA-...-A-..........................................................................AA 9509
H1O.CM.91.MVP5180          T--ATCAT---GC--AACA--T..A-A----TGC-AAA---C--A---CT----CCC--G.....-A-.............................................................................. 9527
H1P.FR.06.RBF168           .................................................................................................................................................. 8654
CPZ.CM.01.SIVcpzCAM13      ----A-A--G--T--AAGA--T..A-A--AA-----AAA-----------T-A---AG--..---.......................................................................CTGGC 9058
CPZ.CM.05.SIVcpzEK505      ----TC-------AGAGTA---..AGA--AA------------------T--A-C--------..---.............................................................................. 8924
CPZ.CM.05.SIVcpzLB7        -----AG--C--CC-AAGA---..A-A--T------------T---------A-----------..---.............................................................................. 8949
CPZ.CM.05.SIVcpzMB66       ------G--T---C-GA-G---..AGA--T--------A-------------A---AG-T---..---.............................................................................. 9062
CPZ.CM.05.SIVcpzMT145      T---AAG--G--TC-CAGA--T..C-A--AA-----AAA--------A--T-A--GAG--..---....................................................................CCACAGC 8954
CPZ.CM.-.SIVcpzMB897       -----A-T-G--CC-GACA---..AGA------------------------A---AG-----..---.............................................................................. 8952
CPZ.CM.-.SIVcpzDP943       ----A-G-----T--GAGA---..A-T--TA----A-AA-------------A-C--G--..---.........................................................................AGC 9010
CPZ.CM.98.CAM3             ----A-G--G--C--AAGA---..A-A---A----A-AA----------T-A-C------..---.......................................................................TGCCC 8862
CPZ.CM.98.CAM5             ----A-G--G--TC-GA-A---..C-A---A----A-AA----------T-ATGGAG--..---....................................................................CACAACCC 9165
CPZ.GA.88.GAB1             ----A-G--G--C--AAGA--T..A-T---A----A-AA-------------A---AG--..---...................................................................CTTTAGACT 9496
CPZ.GA.88.GAB2             ---T-AA-----CC-G-GG--T..C-T--AA----AAAA-----T-----T-A--GAG--..---..................................................................AACCGCAACT 8835
CPZ.US.85.CPZUS            --CTGAA--G--TCGGAGA---..AGA--TAA-----------A---------CG-----..---..................................................................TGGCTGTAAC 9463
CPZ.CD.90.ANT              TGCAA----GATGAGAAGA---..A-T--AA----AAGA--C--T--A--T----GAG--......-A-.............................................GACTGCTAACAGGCGGATATAAAACC 8879
CPZ.TZ.00.TAN1             ----A--------AGAAGG---..A-A---A-----AAGAT--------------AT-A.........................................................GGACTTCCGGGTGCCATGACT 9103
CPZ.TZ.2001.TAN2           ----A-------GAGA-GG---..A-A---A------G-T---------------A.........-A-.........................................................................GGA 9105
CPZ.TZ.2002.TAN3_1         ----A--------AGAAGG---..A-A---A------G-TT----------A---A.........-A-.........................................................................GGA 9159
GOR.CM.2007.SIVgor2139_287 T--AGCAT-G-G-ACAAC---T..--T---TT-C--AA---C--A--ACTG-------GAGA-...-A-.........................................................................AAC 9020
GOR.CM.2007.SIVgorCP2135con T--AGCAT-G-G-ACAAC---T..--T--ATT-C--AA---C--A--ACTG-------GAGAA...-A-.........................................................................AGA 9004
GOR.CM.2004.SIVgorCP684con T--AGCAT-G-G-ACAAC---T..--T----T-C--AA---C-----ACTG--TC---GAGA-...-A-.........................................................................AAC 8949
H2A.DE.x.BEN               -TC-ATG--G--C-A-A-CT---..AA----TTCACT---T-C--A--AG-G--TGG-C--AAGTCAG--TTGCCAGAGAAAGAATGGAAGGCAAAA......CTGAAAGCAAGAGGGATACCATATAGTGAATAA................CAGGAACAACCATACTTG 9911
H2A.GW.x.ALI               -CCTAAG--G--TCA-G--T---..AAA---TTTATT--A--C--A---G-A--TGG-T--AAGTCAG-CCTGCCAGAAGATGAGTGGAAGGCAAGA......CTGAAAGCAAGAGGGATACCATTTAGTAAGAACAGGAACAGCTGA..................TTTG 9908
H2A.SN.x.ST                -CC-ACG-----T---AGCT---..-A----TTTATT-GAT-C--A---G-G--TGG-T--AAGTCAG-CCTGCCAGAGGATGAATGGAAGGCAAGA......CTGAAAGCAAGAGGGATACCGTTTAGCTAA...............AAACAGGAACAGCTATACTTG 9350
H2B.CI.x.EHO               -TC--T--------A-G-CT-T..------TTCAGCAG-TTC--A---G-G--TGG-T-TCAGTCAG--ATGCCAGAGAAAGAGTGGAAGGCTAAA......CTGAGAGCAAGAGGGAATACCTACAGAGTAG...........ACAGGAAATAACAGCACAAAGAGGAA 9803
H2B.GH.86.D205             TTC--T---G----A-G--T-T..------TTTCA-TAG-TT---A--AG-G---GG-T--CAGTCAG-GCTACCAGAGGAGGAGTGGAAGGCTAGA......CTAAAAGCAAGAGGGATACCTACAGATTAG...........GCAAGAGACAGCAGCATAAACAGGAA 9832
H2G.CI.92.ABT96            -CCTACAT----C-A-G-CT---..AG---ATTTATAT-------T---G-G----G-TGGAAGTCAG-GCTACCAGAAGCAGTATGGAAAGAGAAA......CTAAAACAGAGAGGACTGCCTATAGAATAA............................................ 9275
H2U.FR.96.12034            TTC-TCT--G----A-AGCT-T..CA---ATTTATTAA-TGC--A--AG-G--TGG--GTAAGTCAG-TCTGTCAGAGGAAGAGGTTAAGAGGAGG......CTAACCGCAAGAGGCCTCCCAGTAAAGAATTGCTGA.............................. 9377
SMM.SL.92.SL92B            TCCT-A------C-ACACAT-T...-AA--ATTT-TAAG---C--A--AG-G--TGG-TGGAAGTCAG-CTTGACAGAGGAAGAGGTTGAGAGAAGGTTAGCTAACAAGCCAAAACCGCAGAAGAAGATGGCGGATAAGAAGGAAACAAGCTGA................ 9325
SMM.US.x.H9                TCCARAGT----T-A-R-CT-T...AA---ATTT-TTGA---C--A--AG-G--TGGT-GTCAGTCAG-CTTGTCAAAGGAAGAGGTGCAGAGAAGG......CTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGAAAACAAGCTAG................ 9351
SMM.US.x.PGM53             TTCA-AA-----C-ACAGAT-T...-A---TTTCATTAGA--C--A--AG-G--TGGC-GTAAGTCAG-CTTGTCAGAGGAAGAGGTAAAGAGAAGG......CTAACCGCAAGAGGCCTTTTAAAAATGGCTGACAAGAAGGAAACAAGCTGA................ 9805
SMM.US.x.SIVsmH635F_L3     TCCAGAAT-G--T-A-AGCT-T...AA---ATTTATTAA-T-C--A--AG-G--TGGT-GTAAGTCAG-CTTGTCAGAGGAAGAGGTAAAGAGAAGG......CTAACCGCAAGAGGCATTTATAAAATGGCTGACAAGAGGGAAACAAGCTGA................ 9890
                                                                      TCF-1 alpha binding                                                                  Nef end
MAC.US.x.239               TCCAACT--G--C-ACAC-T-T...-A---ATAT-TTAGAT-C--A--AG-G--TGGA-G-AAGTCAG-CCTGTCAGAGGAAGAGGTTAGAAGAAGG......CTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTCGCTGA................ 9868
Nef                        _P_T_L_A_Y_T_Y_._E_A_Y_V_R_Y_P_E_E_F_G_S_K_S_G_L_S_E_E_E_V_R_R_R_._._._L_T_A_R_G_L_L_N_M_A_D_K_K_E_T_R_*
MAC.US.x.251_1A11          -CCAACT--G--C-ACAC-T-T...-A---ATAT-TTAGAT-C--A--AG-G--TGGA-G-AAGTCAG-CCTGTCAGAGGAAGAGGTTAGAAGAAGG......CTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTAGATGA................ 9864
MAC.US.x.251_BK28          TCCAACT-----C-ACAC-T-T...-A---ATAT-CTAGAT-C--A--AG-G.-TGGA-G-AAGTCAG-CCTGTCAGAGGAAGAGGTTAGAAGAAGG......CTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAGGGAAACTCGCTGA................ 9849
MAC.US.x.EMBL_3            TCCAACT-----C-ACAC-T-T...-A---ATAT-TTAGAT-C--A--AG-G--TGGA-G-AAGTCAG-CCTGTCAGAGGAAGAGGTTAGAAGAAGG......CTAACCGCAAGAGGCCTTCTTAACATGGCTGACAAGAGGGAAACTCGCTGA...G..ACAGCA.... 8996
STM.US.x.STM               TCCA-TAT----CCA-AC-T-T...-A----TTT-TTAGA--C--A--AG-G--TGGTTCTAAGTCAG-CTTGCCAAAGGAAGAGGTTGAGAGAAGG......CTAACCGCAAGAGGCCTTCTAAAGATGGCTGACAAGAAGGAAACAAGCTGA................ 9522
MNE.US.x.MNE027            TCCAACT-----C-ACAC-T-T...-A---ATATATTAGAT-C--A--AG-G--TGGA-G-AAGTCAG-CCTGTCAGAGGAAGAGGTTAGAAGAAGG......CTAACCGCAAGAGGCCTCTTAAAAATGGCTGACAAGAAGGAAACTAGCTGA................ 9345
DEB.CM.99.CM40             ----ACT--G--C-A-TG-T-T...AA---AG-CC-TGCTG-A-AAAGAG-GCAT-CT-GAA-ATGTATGTTCCCTAAGAGGAAGTAG...................................................CTAGCTTAACCGCAAACCA 9095
DEB.CM.99.CM5              TTC-TC---T--C-A-TGCT-T...AAA--AG--C-TGAAGCC-ATA-ACCA...G--C-TAA-AGAA--TGCATGTTCCCTAAGAGGAAGTAG...................................................CTAGCTTAACCGCAAACCG 9025
DEN.CD.x.CD1               -CCTGAAT----C-GGTG-T-T...AAA...GC--GC-ATA-GAA-CTTC-AACTTCAC-GCATAATAACTACTGCCAGAGAGCTAAGCCCAGGAAG......TAG............CCTAACCGCAAACCACATCCTACTGCAGAACTTGCTGAGAGGCTCA...... 9481
LST.CD.88.447              TCCTATG------G-GG-CT-T...--A----TCAGAAAA--A--TA-AAG-A-AC-AGCA-CAGCTATTGGCTTG......CAGAGGAAGTAG.........CCAAACCGCAGACCGCAACACATCCTCTTGCAGCCTGGTTGCCAGGGCAACCAGGCTAGCGCATGCG 8508
LST.CD.88.485              TCCTATG------G-GG-CTTT...--A----TCAGAAAA--A--TA--AG-A-AC-AGCA-CATCTCTTAGCTTGAACTGTAAGAGGAAGTAG.........CCAAACCGCAGACCGCAGCACATCCTCTTGCAGCCTGGTTGCCAGGGCAACCAGGCTAGCGCATGCG 8511
LST.CD.88.524              TCC-ATGT-----G-GG-CT-T...--A----TGAGAAA---A--A-CAAGTG-TC-AGCAACTGCTATGAGATTTAACTGCGAGAGGAAGTAG.........CCAAACCGCAAACCGCAGAACATCCTCTTGCAGCCCGGTTGCCCTAGCAACCGGGCTAGCGCATGCG 8514
LST.KE.x.lho7              TCCAATGT-G---G-GG-CTTT...A-A----TCA-AAA---G--T-C-A-GA-AC-A--TACTGCT-TCGCTTTTGACTGCAAGAGGAAGTAG.........CCAAACCGCAAACCGCAGCACATCCTCTTGCAGCCCGGTTGCTAAGGCAAC...CGGGCTAGCGCAT 9586
SUN.GA.98.L14              TCC-ACT--G--C-A-G--TTT...AG----AT-C-AAA---C--T--AG-G------C-TGTTACTA-TCTGCAGTGGGAGGTGTAG.............CTTAACCGCAAACCGCA.....TCCTCTTGCATCGCC..TAGGCAACGGGGCTAGCGCATGCGCGCT 9647
MND-1.GA.x.MNDGB1          -CT-AT-T-GC-TA-GAC-TTG...CA-CATTA-C-TA-CACC-T-AGTGTT-CA-TA-GGCTAAGAGT-TTGAACATCTGCCATTTTGGAAGAGGA......AGTAG..........................................................CCTAACCG 8890
OLC.CI.97.97CI12           .................................................................................................................................................. 8678
WRC.CI.97.97CI14           -CCTAAG--G--C-G-G--TTTGTA-CA-GAA--CTC-CTGGA-A-TCAAT-GC--CTGGGAAAGAGGCCATTGAGAGTCTCTATAAG...AACAAG......AAGCAGTGCAAGTAG.................................................. 9416
WRC.CI.98.98CI04           -CCT--A-----C-G-G-CTTTGTG-CAAGTA--CTC-CTGGA-ATATAATAGC-C-AGGAGAAGATGCTATTGAGACCCTCCATAAAGAGAAGCTG......AAGCGGCACAAGTAG.................................................. 9461
WRC.GM.05.Pbt_05GM_X02     TCCTA-G--G--T-G-G--TTTGTG-CT-GGA--CT-GCCG-CGGTC--ATAGC--CTGGTATTGAGACCCTC......TATGAGATG...AAGAAG......AGAAACAACAAGTAG.................................................. 8348
MND-2.CM.98.CM16           -TCTG----T--T-ACACCTTT...TAT----CCATAAA-AGA--A-GTG-A------C-TGTG...CA-AGCATCAGCCATATTGCTTATGCTAAG......GAGCATAAGCCAGAGTGCTGCAAGAGGAAGTGGTGGCAGTTCTAG...............CCGACCA 9422
MND-2.GA.x.M14             -CCTGAAT-G--T-A-ACATTT...TAT----CCATAAAATGC--AA--G-A-A--G-C-TGTT...................ACAAGCCTCTCCTAT......ACTGCCTACCAGGAG......AAATCAGACTGCTGCAAGAGGAAGTGGTGGCAGTTCTAGCTAACCA 9357
MND-2.x.x.5440             TTCAG-G--T--T-A-ACGTT-...TAT--T-CTATAA-TAGG--A---G-A-A----TGTGTG..................ACAAGTCTGAGCTAT......GAAGCCTACAAGAAGGAGGAAAAGCCTGACTGCTGCAAGAGGAAGTGGTGGCAGTTCTAGCTAACCA 9117
DRL.x.x.FAO                TCCAG----G----A-ACATTT...CA---TT-TATT-AC-----A---G-G--TGGAC-TGTTAAAA-TATAAAGTCAGTAGAAGAGCCAGGT..............................TGTTGCAAGAGGAAGTGGTGGCAGTTCAAGCCAACCA 9187
MON.CM.99.L1               -TC-AAG--G---G-CA-GAG-GGCAGACAG-T--CA-A--T--A-C-AG-GGAGCG---GAAGCGC-T-GCCGCAAACCGCATCCTCTAG.................................................................... 9296
MON.NG.x.NG1               TTCTAAG--G-G-G-CA-GAG-GGCCCTCAG-T---TGA-TTA-A-C-AG-AGAACG---GA-GCGC-T-ACCGCAAACCGCATCCTATAG..............................................AACTCTGCCC.......... 7943
GSN.CM.99.CN166            -TC-AAG--G--CA-AA-GTCAGGCA-ACAG-T---C-A--TG-A-C--G-GGAGCG---GATGCGCCT-ACCGCAAACCGCTTCCTATAG.................................................................... 9261
GSN.CM.99.CN71             -TCTA-G-----C--GA-ATCAGGCTG-CAGATGA-C-A--TG-A-C--G-GGAGCG---GAAGCGCCT-ACCGCAAACCGCTTCCTATAG.................................................................... 9258
TAL.CM.00.266              TGAGAAG--G--C-ACACC-C-GGA-C-AAGATG-CCGA-TGGGACAG--TGGAG-GAG-GAAGCGGATGCTGCTAGCTCCACCGCAAACC...................................................................... 9199
TAL.CM.01.8023             TGAGAAG--G--C-ACACC-CAGGA-C-AAAATG-CTGA-T--GACAG--TGGAG-G---GAAGAAGGAGCTGCTAGCTCCACCGCAAACCGCATCC......TCTTGA............................................................ 8761
MUS-1.CM.01.1085           -TC----T-G--CC-GA-GTCAGGCT-ACAA-T---C-AAATG-A-C--G-GGAGCG---GAAGCGC-T-GCCACAAACCACATCCTATAG.................................................................... 9267
MUS-1.CM.01.CM1239         -TCT-----C--CA-GA-GTCAGGCAG-CAATT---T-AAATA-A-C-AG-AGAACG---GA-GCGC-T-ACCGCAAACCGCATCCTATAG.................................................................... 9236
MUS-2.CM.01.CM1246         -TCTA-G--G--CC-AA-GTCAGGCCC-CAG-TTA-T-AAATA-AAC-AA-AGAGCG---GA-ACGC-T-ACCGCAAACCGCATCCTCTAGAACTCT......GCTAGAGACCTTCCA.............................................. 9353
MUS-2.CM.01.CM2500         -TC------G--TC-CA-GTCAGGCCGCCAGTT-A-C-A-ATG-A-C--G-GGAGCGA--AC-GCGC-T-GCCGCAAACCGCATCCTCTCGCAGTCT......CCATAA............................................................ 9300
SYK.KE.x.KE51              TCCA-A-T----C-A-ACA-CAGGCTG--AAATG-CT-GT--G-AATTAG-GAGAC-A-CAG-AAAGC-CTAGCTCACCGCAGAAACCACAGGGTAC......ATCTCCTAG................GAGACACCTAGGAGACAGAGACCGCCCAGGGGAGGAGTCG 9015
SYK.KE.x.SYK173            T-CA-CAGGTTGGGAGATGGC-...CGCTTG-AGTTAGA-AGA-AAACAGGAAAACCAC-AGAG.........CTGCAGTCAGCACTTAGTAAGAAC......ATCTCCTAG................GAGACTCCATGGTGACAAGCTCGGCCCAGGGGA....... 9312
RCM.GA.x.GAB1              TCCTATG--G---G-AG--T--...A-A-G-TATAGA--------A----T---TGG-G-AA-GAAGAACAAGACCCAGTAA.......................................................CCACATCCTCTG 8972
RCM.NG.x.NG411             ACC--T-T-----G-AG--T--...-CA-GGTATAGA--------T----T---TGGAG-GA-A...AAGAACAAGACCAAGTAA..................................................CCACTTCCTCTGGAGTTGCCTT 9065
GRV.ET.x.GRI_677           -CCGAAGT-G--GG-GG-GT--.....................GC-----CATG--T---G--ATGCACGA-CATGCAAAGCGCTAG....................................................TGTCAGCACTTTGCGGTT... 9289
SAB.SN.x.SAB1C             -CCA-TA------G-GG-CT-T...--T--TT-GAG------C--A---C-GG-GCC-TCAGCA...CA-GGCTAA........................................................CCACAAAACCACATCCTA 9619
TAN.UG.x.TAN1              TCCA-T-------G----GT--..................A-C-----C......-GCTTTAAGGACATGCATGGTCTTGTTAAGAGGAAGTAG...........................CCTAACCGCAGGCTTGCGGTTAAGCTAACTA 9439
VER.DE.x.AGM3              TCC-ATGT-----G-C--AT-T...-ACC-TAACAGAAAAT---TCACTG--A-GC-TG-TCTT...G-CAAGAGGAAGTAG........................................................CT.AACCGCA. 8943
VER.KE.x.9063              -CCTATGT-----G-G--GT-T..................G-C--AA-CAGACAGT-TTTTGAAGACATGCATGCAATAGTGAAGAGGAAGTAG..........................CTAACCGCAGGCTTGTGGTTAAGCCGTTGC 9466
VER.KE.x.AGM155            TCCAATGT-G--TG----GT--..................G-C--CTCAAGGGAGT-CTTTACAGAC-T-TATTCAACAGTTGGTACAGGAAACTAG.......................CCGACCACAGGCTTGCGGTTTCCTGGTTGC 9446
VER.KE.x.TYO1_patent       TCCTATGT-G---A-A--GT--..................G-C--CA-TCGGGAGT-CTTTACTGACATGCATGGGCTGGTGAAGAGGAAGTAG..........................CCAGACCGCAAGCCTGCGGTTAGAACATCAC 9428
COL.CM.x.CGU1              .G-AGAGTAT----ACAGCT-TGACT--AGT-TGTT-AGCGT-A-AAGCAGGAAG-----GAAACCACA-CAGGTTGCCATAGAGATGGTGGACTGA................................................................ 8605
```

```
                                                                                         NF-κ-B-II                      NF-κ-B-I                    Sp1-III                                   Sp1-II     Sp1-I
H1B.FR.83.HXB2                .......CATCGAGCTTGCTACAA..........................GGGACTTTCCGCTG..................GGGACTTTCCAGGGAGGCGTG................................GCCTGGGCGGGACT.GGGG.AGTGGC.GA 9495

H1A1.UG.85.U455               CTGCTGA--CA--AG-----GAC-..........................----------A---..................----------GA--GA-G.--T.............................G-TT-------AGT-.----.G------.T- 8955
H1B.US.90.WEAU160             ........-------T--T-------........................----------A---..................----------G---------..............................AT-----------A.----.------.-- 9495
H1C.ET.86.ETH2220             ........--CA--CGGGA--TTCCGCC......................----------A---..................---.-G-------AG-AG-G...............................-T-----------.----.------.C- 8895
H1D.CD.84.84ZR085             ..........---------C..............................................................................................................................................  8975
H1F1.BE.93.VI850              TTGCTGA--CA--AGAAT---A--..........................----------A---..................--------A-G-CG--CC.............................AGAG--.--------.----.------.TC 8793
H1G.SE.93.SE6165              ........--CA--AG-----GAC-AG.......................----------A-CT..................----------G-------C-G...............................------A--AGT-.----.------.T- 8939
H1H.CF.90.056                 ........--CA--AGA-.--GA--.........................--------------..................----------G---------A...............................T---------GA-.----.------.C- 8813
H1J.SE.93.SE7887              ........--AA--AG--T---GCG.........................--------------..................------------GA-GAGTG.................................-----------GT-.----.------.T- 8794
H1K.CM.96.MP535               ..................................................................................................................................................................  8604
H101_AE.TH.90.CM240           ........--AG--AG--T--.ACT.........................A-A---.------..................------------GA-GTGTG............................GC-G.------AGT-.----.---A--.T- 9062
H102_AG.NG.x.IBNG             ........--CA--AG-----GAC-.........................----------A---..................---------.---GA-GTGT.............................G-TT----A--AGT-.----.------.T- 9022
H103_AB.RU.97.KAL153_2        ACTGCTGACAT--AG-----GAC-..........................------------G...................------------GA-GTGT...............................GG-TT-G-C-GAGTT---G------TA- 8673
H104_cpx.CY.94.CY032          ........--CA--AG-----GAC-AA.......................------------CC..................------------GA-GCGCG..............................GCCT-G-A-GGGTT---G------TA- 8911
H1N.CM.95.YBF30               ........--AG-GA---TACTGCTGACAA....................--------ATA--TG....GGGACTTTCCGCCA.--------------.A-GTGT..........................GGTT..GGG--AGT--CT-.-..........  9096
H10.BE.87.ANT70               ACTGCTGACCT--AGA-TGCTG-CACTGT.....................--A-------AGCAAAGACTGCTGACACTGCG..------------T-G-AG-GA..............................-A---G-C-GT-C.---G------TA- 9617
H10.CM.91.MVP5180             .............CTGACACTGC...........................----------AGAC.....TGCTGACACTGCG..------------C-T--GAG-GA............................TAA--G-C-GT-C.---G------TA- 9613
H1P.FR.06.RBF168              ..................................................................................................................................................................  8654
CPZ.CM.01.SIVcpzCAM13         GCATGCG-C-AAGAAG--..CTG-CTGTGC....................----------AA....................----------G-------G-...........................TC.-AG---A-T-GT-.----.-----TTTC 9147
CPZ.CM.05.SIVcpzEK505         ........--A.......................................---------TA--TG....GGGACTTTCCAGCA.----------G------...............................T-T-GT-T----.----.-----TTG 8994
CPZ.CM.05.SIVcpzLB7           ........--A.......................................-A------T-TGACCA................-------------CCA---G...................CTTCCGGGGAGGCGT-T-AA--GA--GT-T----.----.---TTT- 9029
CPZ.CM.05.SIVcpzMB66          .....CAA-GGACTTG--..CTG...........................----------A...CA.............---GA-----G--------...............................GT-A-CAT-G-C-GGGGA-A-G------TT. 9139
CPZ.CM.05.SIVcpzMT145         GGCGCATGCG-ACT.AA-AACTGCTGATCAACATTTGCTGACAA......----------A.................A-------------G--TTG-GT....................TTGGGGCGTAACT--G-C-GAACT.--G------TTG 9068
CPZ.CM.-.SIVcpzMB897          ..........................................CAAA-------GTAA.......................A------------CAAG--GT--T....................CCGGGAGGCGTG-CTA-G-C-GGGGTA--G------TT. 9028
CPZ.CM.-.SIVcpzDP943          CCAGAGC-CAGCGC--GCGC-GT-GAAACTGCTGACAA............----------TAA...................-----------A--G--GTG-T..................TCA.......................--GC----TTTT 9107
CPZ.CM.98.CAM3                AGCGCCTGCG-AGTAGAAA.CTGCTGACAA....................----------TAA...................-----------A--G--G..TG................................GT-CA-G-C-GGACT.--GC----TTTT 8950
CPZ.CM.98.CAM5                AGCGCCTGCG-AGTAGAA..CTGCTGACAA....................----------TAA...................-----------A--G--A..-G................................-TCA--GC--GACA.--GC----TTTC 9251
CPZ.GA.88.GAB1                GGCGCATGCG-ACAAGAA..CTGCTGACTCTGC.................----------AA....................----------G-----AG-.T.............................G-T-G--AG-C-GGGTT---GT-----TTT 9589
CPZ.GA.88.GAB2                GGCGCATGCG-TCT..A-AACTGCTGAGCCTGAACTTTTACAA.......----------AA.................................GG.......................GGGGTGG-TTCAA-G-A-GAGTC---GT-----TTT 8926
CPZ.US.85.CPZUS               CGCGCAGGCG-A-TAAAA..CTGCTGACTG.................A--------TAA......................-----------A--GACGT-C.C..................AAGGGGGT-GG-CA-G-C-GAACA.--GC----TTTT 9561
CPZ.CD.90.ANT                 GCTGACAG-G-AGA----TGGTCT..........................----------A--AT...........................................................GT-A--AT-G-C-GGCA-T--G------TTT 8953
CPZ.TZ.00.TAN1                CAGAACTGC-GACAGAG-ACTTTTGGACTC....................----------AA...........................................................TGTGGGTG-TTACT-G-C-GGACA---G-----TTTT 9182
CPZ.TZ.2001.TAN2              CTTCCGGGTGAC--GAA-AAGA-CTGCTGATGA.................--------TG-ACTC.................-----------A.........................TATGGGTG-TTACA-G-C-GGGCA---G-----TTTT 9202
CPZ.TZ.2002.TAN3_1            CTTCCGGGTGAC--GAA-AAGA-CTGCTGATGA.................--------TG-ACTT.................-----------A.........................TGTGGGTG-TTACA-G-C-GGACA---G-----TTTT 9256
GOR.CM.2007.SIVgor2139_287    TTGCTGA--G-AT-AAAAGCTGCTGACACTGCA.................------------T.....................-----AAA-T..........................CCCCGGGGGC-GAACT-G-A-GAGCA---G-----TTTC 9114
GOR.CM.2007.SIVgorCP2135con   CTAAAAAAC-T-CTGACAGC-TG-AAAAGCTGCTGACACTGCA.......------------AT...................--G--AAA-T..........................CCCCGAGGGC-GAACT-G-A-GAGCA---G-----TTTC 9108
GOR.CM.2004.SIVgorCP684con    TTGCTGA--G-AT-AAAAAGCTGCTGACACTGCA................------------T....................---G-AAAGTC..........................CCCGGGGGCG-AACAT-G-A-GAGCA---G-----TTTC 9045
H2A.DE.x.BEN                  GTCAAGG--GGA--TAGCTACT--GAAACAGCTGAGGCTGCA........----------AGAA...............---G--G-AACC.............................AAGGGAGGGACAT----GGA-CT-GT-G--AACG10013
H2A.GW.x.ALI                  GTCAGGG--GGA--TAACTACT.GAAAACAGCTGAGACTGCA........----------AGAAG..............---G--G-AACC.............................AGGGGAAGGACAT----GGA-CT-GT-G--AACG10009
H2A.SN.x.ST                   GTCAGGG--GGA--TAACTA---GAAAACAGCTGAGACTGCA........----------AGAA...............---G--G-TACC.............................AGGGGAGGGACAT----GGA-CC-GT-G--AACG 9452
H2B.CI.x.EHO                  CTAGCTAACA-TGCA-A-AG-AGGAAACTAGCTGATACTGCA........----------AGAA...............---G--G-AACA.............................GTGGGAGGGACAT----GGA-CA-G.-A--AACG 9903
H2B.GH.86.D205                CTAGCTGACA-TGCAC.AAG-AGGAAACTAGCAGACACTGCA........----------AAAA..............-A-G--G-AACA.............................TGGGAGGA-CAA----GGG-TT-GTTG--AACG 9932
H2G.CI.92.ABT96               ..............GCAGA-AG-GGACAAGCTGACACAGCA.........----------AG-A..............T--G--G-AACA--G-AG............................GTACTAT----GGGACT-GTTG--AACG 9362
H2U.FR.96.12034               ...............CAGA-AGGAAACAAGCTGAGACAGCA.........----------A-CA.............GA--GAGG-AACCA-G-AG.............................GA-CAA----GGA-CCA.T-G--AACG 9462
SMM.SL.92.SL92B               ..................................GACAGCA.........----------A-AA.............G----G-G-A--A--G..............................A-GAGTCTG--CGGGG--AA..CG--AACG 9390
SMM.US.x.H9                   ..................................GACAGCA.........----------A-AAA..............---GA-G-TAC.............................GGGGAGG-TCTG----GGA-CT-G.C---AACG 9416
SMM.US.x.PGM53                ..................................GACAGCA.........----------A-AAA..............---GA-G-TA..............................CGGGGAGG-ACTG----GGA-CT-G.C---AACG 9870
SMM.US.x.SIVsmH635F_L3        ..................................GACAGCA.........----------A-AA.............-----------CAAG--GC--TCA.....................TGGGGAGG-ACTG----GGA-CT-G.C---AACG 9968
                                                                                        NF-κ-B-II                                      NF-κ-B-I

MAC.US.x.239                  ..................................AACAGCA.........----------A-AAG..............---.A-G-TA...............................CGGGGAGG-ACTG----GGA-CC-GTCG--AACG 9933

MAC.US.x.251_1A11             ..................................GACAGCA.........----------A-AAG..............---..-G-TA...............................CGGGGAGG-ACTG----GGA-CC-GTCG--AAC- 9928
MAC.US.x.251_BK28             ..................................GATAGCA.........----------A-AAG..............---.A-G-TA.........................................TG----GGA-CC-GTCG--AAC- 9903
MAC.US.x.EMBL_3               ..................................................----------A-AA..................................................GGGGATG-TATG----GGA-CC-GTCG--AAC- 9043
STM.US.x.STM                  ..................................GACAGCA.........----------A-AA...............---G--G-AACA.........................GGGGAGG-ACT.----GGAACT-GT-G--AACG 9587
MNE.US.x.MNE027               ..................................GACAGCA.........----------ATAA...............---GA-G--..........................ATGGGG-GG-ACTG----GGA-CT-GTCG--AAC- 9410
DEB.CM.99.CM40                CATCCTA--G-AGCGCCCG...............................----------A................................GTTGCCTAGGAAATATCATGGCAACCAGGGGA-GGCCT-G-T-GTAC.---G------CA. 9184
DEB.CM.99.CM5                 CATCCTCTTG-AGAGAGC-C..............................--T-------A................................GTAGTCAAGGAGACTCCATGGTTACAGGGGAG.GGACT-G-C-GTAC.---G-----TTTG 9115
DEN.CD.x.CD1                  ..................................................----------A..-..........GTTGCCTA-CA--CC-GGGC---C.......................TCAGGGGA-GGCCT-G-A-GTCT.CT-GGAGT-GCTT 9553
LST.CD.88.447                 C.................................................A-------G-TGATGAGC..............A-----------..........................................--A----CGGG--A-G.T-T-.... 8559
LST.CD.88.485                 C.................................................A-------G-TGATAAGC..............A-----------..........................................--A----CGGG--A-G.T-CA.... 8562
LST.CD.88.524                 C.................................................ACA-GC--G-TGAACAGC..............A-----------..........................................--A----CGGG--A-G.T-T-.... 8565
LST.KE.x.lho7                 GCGC....-AT-GCT-GCTG-...CGAGCA....................----------AGGACGGGCGG.............----GG--TG..................................................................  9640
SUN.GA.98.L14                 AGAGTA............................................----------GAC............................................................AGGG..A--GG-A---TT- 9684
MND-1.GA.x.MNDGB1             CAAAACCACATCCTAC---AGA-CTGTAGTTGCTTGGCAACCTGCTT.....A-C-ACC-GGA---GCGCTTGCGCGCTAGGA..A-----------AA.........................................CAGG-A--GG-A---TCG 9000
OLC.CI.97.97CI12              ..................................................................................................................................................................  8678
WRC.CI.97.97CI14              ....................-GACATG..........TTGCAGTAGGAA-T-G----G--G-TAAGCTACA..................................TCCTGAGACTGTGTACCAACGGACTTA............................  9484
WRC.CI.98.98CI04              ..................................................................................................................................................................  9461
WRC.GM.05.Pbt_05GM_X02        ....................-GACTTG..........TCAGA...GGAA-CTGG---G--G-TAGGCTAGTA.................................TCCGCTAGCTGTGTACCAACGGCCTTAGCAA--AT-CTAATGAACT---ACTCT-GGAG 8446
MND-2.CM.98.CM16              CAGAAGGTTG-C-TGGA-A-GACCTGCAGAAATGCTGACTCTGGA......----------A--AGTGCATGCGCACTAGACT.G--------T-CA-ACT-A..........................................................  9518
MND-2.GA.x.M14                CAGAAGGTTG-T-TGGA-A-G-TCTTTAATTCAGCTGACTCTGGA......----------A--AGTGCATGCGCACTGGACT.G--------T---ACT-A---CGGG.......................................T----G-----TCAG 9474
MND-2.x.x.5440                CAAAGGGTTG-T-TAAGCAACTGCTAACTCTGGA.................----------A---GTGCATGCGCACTGGACT.G-----------.A-ACT-ACGTCGGG......................................T----G-----CAG 9223
DRL.x.x.FAO                   CAGAAGGTTGTC-TGGA-AC-ACCTTCAGAAGTGCTGACTGAGAA......----------A-A-GAGCATGCGCTCTGTGGC.A-----------............................AGGGTGAC-T-G-TTGGGG-A...----TCTG 9304
MON.CM.99.L1                  ....................-ACTCTGCCC.....................----------A...........................-T-ACCATGACT-C-G-GCCCGGTTGCTGCGGCAACC.GGGCGGTCCAAAGGACTTG-T-G-T-GACCC---G------TTG 9398
MON.NG.x.NG1                  ...................................................----------A...........................-TTGCC-TGACT-C-G-GCCCGGTTGCCGAGGCAACCGGGGCGGACTCGTGGGAGGGACTG--AGG-AC---G-----....  8032
GSN.CM.99.CN166               ....................-ACTCTGCCC.....................----------A...........................-T-ACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGAGGGACTG--AGG-.C---G------TCT 9363
GSN.CM.99.CN71                ....................-ACTCTGCTC.....................----------A...........................-TTACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGTGGGACTG--AGGC.C.--G------TCT 9359
TAL.CM.00.266                 ..................GCATCCTCTTGAAGGACTC..............----------A.........................................CGTTGCTAAGGCAACGGGGGACGGACTGGGGCG-GGA-C-G-A-GAGTT.--G------TTG 9290
TAL.CM.01.8023                ............................AAGGACTC...............-----------.........................................CGCTGCTAAGGCAACGGGGGACGGACTGGGGCG-GGA-C-G-A-GAGCT.--G------TA. 8840
MUS-1.CM.01.1085              ....................-AC-TTGCCC.....................----------A...........................-T-ACCATGACT-C-G-GCCCGGTTGCTGAGGCAATCGGGGCGGACTCGTGGGTGGGACTG--CGG-ACT--G-------.... 9366
MUS-1.CM.01.CM1239            ....................-ACTTTGCCC.....................----------A...........................-T-ACCATGACT-C-G-GCCCGGTTGCTGAGGCAACCGGGGCGGACTCGTGGGTGGGACTG--AGG-ACT--G-------.... 9335
MUS-2.CM.01.CM1246            ...........................................................................................-T-ACCATGACT-C-G-GCCCGGTTGCCGCAGCAACCGGGGCGG.ACTCAGGGCGGACT-G-A-GGACCT-AG------TTG 9434
MUS-2.CM.01.CM2500            ................CTACGGG-CTATAAG..........................................................GCCAGGTTGCCTGGGCAACCTGGGCGAACTCGTGGGTAAAACTG--AGGAGC---A------TT- 9380
SYK.KE.x.KE51                 GGGCGGGGGAG-CT-ACC-A--CTGCATAAAAGCCGATGCTCTGTAGCTATCC...........................................................................................................  9068
SYK.KE.x.SYK173               ......................................................................................--AG-C-GGGC---G-AA-GCTCA.....................................................  9336
RCM.GA.x.GAB1                 GGGTTGC-T-G-TAACCAGGCAG-AGAATCTGCTGATGCAAAA........----------A---GTGCATGCGCACTGGGGAA.-----------..........................................GGGAT-AC-TGGGA----G-----TCAG 9086
RCM.NG.x.NG411                GGAAACAGCG-AT--C-A--GAG-AGATGCTGA..CACAGAA.........----------A---GCGCCTGCGCGCTGGTGTAA-----------..........................................AGACT-A--T-GGAG----GT----TTG 9178
GRV.ET.x.GRI_677              ...................................................------------CA.........................-----------CA-T--G.........................TGGATCGGAGGCG-TACA--G-C-GTAC.T--G------TTT 9362
SAB.SN.x.SAB1C                TGGAGTTGTCATG-TGATGACATTAAGAACTGCTGACTGAGA.........----------AG...............................................................GGT-GAGACTG--CGG-ACT--G------TTG 9705
TAN.UG.x.TAN1                 GGCCGTTGCCTAG-AGATGACAT-AAGAAGTGCTGATCAGCA.........----------AG..............................................................GGTG-TGC-A-G-C-GTACTT--G------TT. 9525
VER.DE.x.AGM3                 GGCTTGTGG-TA---ACATC--C-TGGTGATGACATTAAGAACTGCTGAC..----------AGCAA.................---------------....................................C-GG-CATG--CGG-AC---G------TTT 9052
VER.KE.x.9063                 CAGGGAGATGAC-TT-GAAACTGCTGACAG.....................----------AA..................------------......................................GGGCG-G-CAT-G-C-GTAC.---G-----TTTT 9552
VER.KE.x.AGM155               CTAGGAGATGAC-TTAA-AACTGCTGAC.......................----------AGCAC..................-----------A--C--...................................GACAT-G-C-GTAC.---G------TTT 9533
VER.KE.x.TYO1_patent          CATGGAGATGAC-TTAAAAACTGCTGAC.......................----------AGC-AA.................-----------A--C--G.................................ACAT-G-C-GTCC.---G------TTT 9516
COL.CM.x.CGU1                 ..........................GTTGCTGAGACAGCA..........-----------T-C..................-----------A-TT--GAG-..............................-A-CT--GC---GGGAA--GC------.... 8679
```

```
                                                                TAR element start
                           TATA box                    3' LTR U3 end 3' LTR repeat start
H1B.FR.83.HXB2             GCCCTCAG.ATCCT....GCATATAAGCAG..CTGC....TTTTTGC.CTGT..ACTGGG..TCTCTC...TGGTTAGA.CCAGAT.CTGAGCCTGGGAG.CTCTCTGGC.....TAA.........................CTAG................... 9590
H1A1.UG.85.U455            --------.--G--....------------..----....-----C-.----..------..------...-T------.------.C------------.---------.....-G-.........................-G--................... 9049
H1B.US.90.WEAU160          --------.--G--....------------..----....-------.----..------..------...--------.------.-------------.---------.....-G-.........................----................... 9590
H1C.ET.86.ETH2220          A-------.--G--....------------..----....-----C-.-T--..--C---..------...-A-G----.------.-------------.---------.....-T-.........................-G--................... 8990
H1D.CD.84.84ZR085          ................................................................................................................................................................... 8975
H1F1.BE.93.VI850           A-------.--G--....------------..-C--....-----C-.----..------..------...-A------.------.T---------G--.---------.....-G-.........................-G--................... 8888
H1G.SE.93.SE6165           A-------.-AG--....------------..-C--....---C-C-.----..------..------...-T------.------.T------------.---------.....-G-.........................-AG-................... 9034
H1H.CF.90.056              A-------.--G--....------------..----....-----C-.-T--..------..-----T...--------.------.-------------.---------.....-G-.........................---A................... 8908
H1J.SE.93.SE7887           A-------.--G--....------------..----....-----C-.-T--..------..------...-T------.------.-------------.---------.....---.........................----................... 8889
H1K.CM.96.MP535            ................................................................................................................................................................... 8604
H101_AE.TH.90.CM240        A-------.--G--....-----A------..-C--....-----C-.-T--..------..------...-T------.---G--.-----C-------.---------.....-G-.........................-A--................... 9156
H102_AG.NG.x.IBNG          A-------.-C-G-....------------..----....---C-C-.----..------..------...-T-C----.------.-------------.---------.....-G-.........................-GGA................... 9117
H103_AB.RU.97.KAL153_2     .-------.--G--....------------..----....-----C-.----..------..------...--------.------.-GAGC--------.---------.....-G-.........................---A................... 8766
H104_cpx.CY.94.CY032       .-------.--G--....-----A------..-C--....---C-C-.----..------..------...--------.------.-------------.---------.....-G-......................................... 9005
H1N.CM.95.YBF30            .-------.-.G--....-----A------..-C--....---C-C-.-T--..------..------...-T-C-G--.------.TA---T-------.-ATAT---G.......................................................... 9182
H1O.BE.87.ANT70            .-------A-.G--....------------..-C--....----C--.-T--..--C---..----GG...-TAGAG--.----G-.-------C-----.---C-----.....CTC.........................---C................... 9711
H1O.CM.91.MVP5180          .-------.--G--....------------..----....----CC-.-T--..--C---..---TAG...-TAGAG--.----G-.-------C-----.---C-----.....CTC.........................---C................... 9707
H1P.FR.06.RBF168           ................................................................................................................................................................... 8654
CPZ.CM.01.SIVcpzCAM13      --------..AG--....------------..-C--....----C--.-T--..------TC...T..CAC.-G--.------.------C------.------C--.....-G-.........................-GA-................... 9238
CPZ.CM.05.SIVcpzEK505      .-------..AG--....------------..----....-GC-C--.-T--..------TC-----T.........-----.TA-----------.---------.....-G-.........................-A-.................... 9086
CPZ.CM.05.SIVcpzLB7        .-------..AG--....------------..-C--....-C-GC--.-T--..------TC-----T........-T-------.------------.---------.....-G-.........................---A................... 9123
CPZ.CM.05.SIVcpzMB66       --------..AG--....------------..-C--....---GC--.-T--..------TC-----T........-T-------.------------.---------.....-G-.........................---A................... 9234
CPZ.CM.05.SIVcpzMT145      ..--CA-CA-AG--....-----------.CGC--.....-C-C--.-T---..CAA---TC-----T......................................--CAAG.................................................. 9162
CPZ.CM.-.SIVcpzMB897       .G--CTCAG.AG--....------------..-C--....-C-C--.-T--..------TC-----T........................................--T...................................................... 9119
CPZ.CM.-.SIVcpzDP943       .A--CTCAG-AG--....------------..-C--....-C-C---.-T--..------TC-----T....-C------.TA-----C---------------T.....C-GT...................................................... 9199
CPZ.CM.98.CAM3             A-------.-AG--....------------..-C--....-C-C---.-T--..------..------..T--C-----.------.-------------.---------T.....C-G...................G---................... 9045
CPZ.CM.98.CAM5             AGTT-TCAG-......................................................................................................................................................... 9261
CPZ.GA.88.GAB1             --------.-.G--....------------..----....----C--.----..------..-----T..TCAC.-G--.------.T------------.------AC-.....AGT.........................G-A.................... 9682
CPZ.GA.88.GAB2             --------A..G--....------------..----....-C-C---.-T--..-T----..-----T...A--CAGA..------.-------------.---------.....-G-.......................................... 9019
CPZ.US.85.CPZUS            A------A.-GA--....-----A------..-C--....-C-C---.-T--..------..------..TC-C-AGA..------.AA-----C------.---------.....-G.........................-GTA................... 9655
CPZ.CD.90.ANT              C----G-CT..G--....------------..----....---GC--.-T--..TAA---.C-----G...CC.-AG--G------.-A------------.---------.....--TTTTT...................GGCAG.................. 9053
CPZ.TZ.00.TAN1             .G--.-GCTGAG--....------------..----....---GC--T----..AAA--..----TG.CCTAA-CTG.------.-------------.---------.....-G.........................-G-CTGGC................. 9280
CPZ.TZ.2001.TAN2           .G--CGCTG.AG--....------------..----....---GC--TG---..-AA---.C---TG-.....C-A-ACTG-CAGAT--------------------T.....AGTGG.................................................. 9296
CPZ.TZ.2002.TAN3_1         .G--CGCTG.AG--....------------..----....---GC--TG---..-AA---.C---TG-.....C-A-ACTG-CAGAT--------------------T.....AGTGG.................................................. 9350
GOR.CM.2007.SIVgor2139_287 .A--CTC-G.AG--....------------..----....---GC--.----..-GT--TC---G-G.....AC---.------.-A-------------------T.....CTTA................................................... 9204
GOR.CM.2007.SIVgorCP2135con .A--CTC-G.AG--....------------..----....---GC--.----..------TC---G-G.....AC---.------.-A-------------------T.....CTTA................................................... 9198
GOR.CM.2004.SIVgorCP684con .A--CTCAG.AG--....---------C-T..........---GC--.----..------TC-Y-G-G.....AC---.------.-A-----C---AG--------T.....CTTACCAAAACC........................................... 9143
H2A.DE.x.BEN               .------TAC.TTA....CTG----ATGTAC-C--...T-C--GCAT.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.-GAGC----A---GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................10139
H2A.GW.x.ALI               .----.TCAT-.-TC...CTG----ATGTAC-C--...TGC--GCAT.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................10134
H2A.SN.x.ST                .------TAC.TT-....CTG----ATGTAC-C--...TAC-CGCAT.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--..................... 9578
H2B.CI.x.EHO               .---ATCAAG.A--....CTG----ATGTAC-C--...T--G-GCAT.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.-GAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................10029
H2B.GH.86.D205              .---ATTAA-.T--....CTG----ATGTAC-C--...T-C-CGCAT.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................10058
H2G.CI.92.ABT96            .---C.-AA-.T--....CTG----AT-TAG----...A---CGCTC.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--C-----GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--..................... 9487
H2U.FR.96.12034            .----.TCAA--GA-...CTG----ATGTAC-C--...TC---GCAT.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--..................... 9588
SMM.SL.92.SL92B            .---CT-AAC--TG....TAG----AT-TAC-C--...TT-GCCGCTC.---.ATTCA-TC.G---AC.GGA-AGGCTG------.AGAGC--C-A---GT------CAGCACT-.......AGTAGGTGAGCCTGGGTGTTCT--..................... 9518
SMM.US.x.H9                .---A-TTAT.T--....CTG----AT-CAA----...A---CGCTC.---.ATTCA-TC.A---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--..................... 9542
SMM.US.x.PGM53             .---A-TTAT.TT-....CTG----AT-CAA----...A---CGCTC.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--..................... 9996
SMM.US.x.SIVsmH635F_L3     .---A-TTAT.T--....CTG----AT-CAA----...A---CGCTC.---.AGTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................10094
                                                                TAR element start
                           TATA box                    3' LTR U3 end 3' LTR repeat start
MAC.US.x.239               .---A-TTTC.TTG....ATG----AT-TCA----...A---CGCTC.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................10059
MAC.US.x.251_1A11          .---A-TTTC.TTG....ATG----AT-TCA----...A---GCTA.---.AGTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................10054
MAC.US.x.251_BK28          .---A-TTTC.TTG....ATG----AT-TCA----...A---CGCTC.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--.....................10029
MAC.US.x.EMBL_3            .---A-TTTC.TTG....ATG----AT-TCA----...A---CGCTC.---.ATTCA-TC.G---GG..CA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--..................... 9168
STM.US.x.STM               .----TTTAC.TT-....CTG----AT-CAA-C--...T-A-AGCTC.---.ATTCA-TC.A---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--..................... 9713
MNE.US.x.MNE027            .---A-TTTC.TTG....ATG----AT-TCA----...A---CGCTC.---.ATTCA-TC.G---GC.GGA-AGGCTGG-----.TGAGC--------GT-----CCAGCACTA.......GCAGGTAGAGCCTGGGTGTTCC--..................... 9536
DEB.CM.99.CM40             .G--CGCTG-A---....A----A----C-..A---.........T----..-GCTATCCA................................................................................................... 9227
DEB.CM.99.CM5              .----GCTGT-A--....---------T-..A---..........T----..-GCTATCCA................................................................................................... 9158
DEN.CD.x.CD1               .A--CGCTTC-A--....-----A----T-..A---.........T----.A--A-TCGG-T.............................-C---.------CCAGCGGCCTGG.............................................. 9620
LST.CD.88.447              .---A-..TT.G--....------------..-----...-C-........................................................................................................................ 8588
LST.CD.88.485              .---A-..TT.G--....-----A.-----..-----...-C-........................................................................................................................ 8589
LST.CD.88.524              .---A-..TT.G--....------------..-----...-C-........................................................................................................................ 8594
LST.KE.x.lho7              .---A-..TT.G--....------------..-----...-C-CC--T.---TCGGGA-..TC--TG-TCTG--GA...C-AGC-G.TC-CT------T-AGCACACT-GT...CAGCTAAGC.AGTCAGCCCTGGGAGGCTACA.....CTTAGC........... 9766
SUN.GA.98.L14              .---A-..TT.A--....------------..-----...-C--C--C.---..T-G--AG.------.GC-CAGAG-..AT--GCGAC--C----A---.G----ACCT.....AG-GCTGGCTTGTCGCTCCTGAGTGAGCACTTGCTAAG............. 9809
MND-1.GA.x.MNDGB1          .---CA..TG.-TG....CT----------..-----...A---CGCT..----..T-G--AGTCTCTA-TACA-AGGCT-A...............---.TTGTATCTCTGA.GC.............................................. 9080
OLC.CI.97.97CI12           ................................................................................................................................................................... 8678
WRC.CI.97.97CI14           ................................................................................................................................................................... 9484
WRC.CI.98.98CI04           ................................................................................................................................................................... 9461
WRC.GM.05.Pbt_05GM_X02     -AGACTT-TGGTT-TGCCCT---A---T--ATGCT-TGTAGC-A-CTGG-TGCAGTCTTCTGC--GAGGCTA---GCCTGAGT-G-GG-CCCA--T--CCTAG----CTTGGTGCCTGAGTGGTGGTCCCACTTGGCCAAGAGACT-AG-TAGGAG............ 8602
MND-2.CM.98.CM16           ................................................................................................................................................................... 9518
MND-2.GA.x.M14             .---CA..T..-TGAC.T------------.C-C--...T-GCCGCT..---.ATT.-A-TC----C-T.............................................................................................. 9529
MND-2.x.x.5440             .---AA...C.-TGGC.T------------.CT---...T-GCCGC-..---.A-G--A-TC----C-T..GA-AGGCTG------.TGAGC.------.TTGT---CT.GG.TGAGTTCTTG.AAGGAG..TGCCTGCTTGTAGC-CT-GGCGGT............ 9355
DRL.x.x.FAO                .-------CT.G-A....-----A------.CT---...T-AC-GCATG-...A-AC--TC.----C-..TG--AGGCTA------.---------T-.T-------TG....AGTCTGCTGAAGGAGGCTACCTGCTCGAGCC-CG-GTGGTTCGCAGGGCCTTGGC 9451
MON.CM.99.L1               .-----TCGT-T.....CA----A------..AC--....CC..........AG--ATCCT--GGAGTC.............................................................................................. 9448
MON.NG.x.NG1               ..-T-ACCCTCAGATCCT-----A------..AC--....-C..........AGM-ATCCGC-GGAGTC...............................--.-----GTCT.................................CCAC--CGGCCT........... 8109
GSN.CM.99.CN166            A---CGCTAT.--A....-----A-----A..AC--....-C..........AG-GACTCT--GGAGTC.............................................................................................. 9414
GSN.CM.99.CN71             A---CGCTAT.--A....-----A------..AC--....-C..........AG-GACTCT--GGAGTC.............................................................................................. 9410
TAL.CM.00.266              .-------A-.--A....------------..A---....-C..........AG--ATCCT-T..G-TC.............................................................................................. 9338
TAL.CM.01.8023             .----.............................................................................................................................................................. 8845
MUS-1.CM.01.1085           ..-T-ACCCTCAGATTCC-----A------..AC--....-C..........AG--ATCCC--GGAGTC.............................................................................................. 9420
MUS-1.CM.01.CM1239         ..-T-GCCCTCAGATTTC-----A----CA..GC--....-C..........AG--ATCCC--GGAGTCAGTTC---.CT-C-CA.GC-GC.-------C-CTCGGAA-......-G.GTAA.AGCGGCCTGAGAGCCCTTGCCTAGGAGAACTG............. 9465
MUS-2.CM.01.CM1246         .-------ATC-A.....-----A------..AC--....-C..........AG--ATCCAA-GGAGTC.............................................................................................. 9484
MUS-2.CM.01.CM2500         .-------A--T-.....-----A------..----....-C...........G--ATCCAGACGGAGTC............................................................................................. 9430
SYK.KE.x.KE51              ................................................................................................................................................................... 9068
SYK.KE.x.SYK173            .----.....A--.....-----A----C-..A---....-C.........---AGCTATCCGCGGAG-TGCAACTA..........................G------CCAGCGGCCT...GGG.AGCCCTGAGAGACTGGTAGAGCGGCCC.GGGA........ 9438
RCM.GA.x.GAB1              .-----TCCT.G--....------------..-----...-C-GC--T.---AA-AC---..------CCTG--AGGCTA--G---.---------T-.T-------T.....A-GTCTCTAGGAACTCCAGCTTGAGCCTGGGTGTTCGCT................ 9217
RCM.NG.x.NG411             .-----TCCT.G--....------------..-----...A---C--T----.A--CA-..------CCTG--AGGCTA------.---------TTGT-------T.....A-GTCTCTAGGAGACT.AGCTTGAGCCTGGGTGTTCGCT................ 9308
GRV.ET.x.GRI_677           C-----.AG-G-TG....-----A------..A---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-..A---------T-.T--G----T.....-GCCTA...ACCCGGT.TGGCCA.CCGGGGG--A................... 9481
SAB.SN.x.SAB1C             .---CGT-CT.G--....CA----------.-----.......GCGC----AA-ACA-..------GC.T---AGGCTG------.---------T-.T-------T.....A-GTCTG..AACCAGCT.TGAG...CCTGGGTGTTCGCT................... 9828
TAN.UG.x.TAN1              C-------.-.G--....-----A------..A---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-.----------.T--G----T.....G-GCCTA...ACCTGGTCTGGCCATCCAGGGG--A.................... 9646
VER.DE.x.AGM3              A-------.-.G--....-----A------..A---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-.---------T-.T--G----T.....-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A..................... 9171
VER.KE.x.9063              A-------.-.G--....-----A------..A---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-.---------T-.T--G----T.....-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A..................... 9671
VER.KE.x.AGM155            A-------.-.G--....-----A------..A---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-.---------T-.T--G----T.....-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A..................... 9652
VER.KE.x.TYO1_patent       A-------.-.A--....-----A------..A---....-CGC--GCT---..-ACTCAG.-----T.AC-A-GAGA..----C-.---------T-.T--G----T.....-GCCTA...ACCTGGT.TGGCCA.CCAGGGG--A..................... 9635
COL.CM.x.CGU1              .----GCTA-A.--GCTT-T---A---TCAG.AGC-....-CGAGAGCT-CGAGG-........................................................................................................... 8728
```

```
                                                  TAR element end
H1B.FR.83.HXB2                 ...............................GGAACCC.....ACTGCTTAA......................................................................................GCCTC........................... 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                ...............................-------.....---------......................................................................................-----........................... 8604
H101_AE.TH.90.CM240            ...............................-------.....---------A.....................................................................................-----........................... 9178
H102_AG.NG.x.IBNG              ...............................-------.....---------......................................................................................-----........................... 9138
H103_AB.RU.97.KAL153_2         ...............................-------.....---------......................................................................................-----........................... 8787
H104_cpx.CY.94.CY032           ...............................-------.....---------......................................................................................-----........................... 9026
H1N.CM.95.YBF30                .......................................................................................................................................................................... 9182
H1O.BE.87.ANT70                ...............................T------.....G--CG----......................................................................................CG---........................... 9732
H1O.CM.91.MVP5180              ...............................T------.....G--------......................................................................................CG---........................... 9728
H1P.FR.06.RBF168               .......................................................................................................................................................................... 8654
CPZ.CM.01.SIVcpzCAM13          ...............................-------.....--------.....................................................................................AA-----........................... 9259
CPZ.CM.05.SIVcpzEK505          ...............................-------.....---------......................................................................................-----........................... 9107
CPZ.CM.05.SIVcpzLB7            ...............................-------.....---------A.....................................................................................-----........................... 9145
CPZ.CM.05.SIVcpzMB66           ...............................-------.....---------......................................................................................-----........................... 9255
CPZ.CM.05.SIVcpzMT145          ...............................-------.....---------......................................................................................-----........................... 9183
CPZ.CM.-.SIVcpzMB897           .......................................................................................................................AGGAGAACCCACTGCTTAA-----........................... 9143
CPZ.CM.-.SIVcpzDP943           ........................................................................................................................AGAGAACCCACTGCTTAA-----........................... 9222
CPZ.CM.98.CAM3                 ...............................A------.....---------......................................................................................-----........................... 9066
CPZ.CM.98.CAM5                 .......................................................................................................................................................................... 9261
CPZ.GA.88.GAB1                 ...............................-------.....---------......................................................................................-----........................... 9703
CPZ.GA.88.GAB2                 ...............................-------.....---------......................................................................................-----........................... 9039
CPZ.US.85.CPZUS                ...............................-------.....---------......................................................................................-----........................... 9675
CPZ.CD.90.ANT                  ..............................-AG----.......-----........................................................................................-T-.............................. 9068
CPZ.TZ.00.TAN1                 .............................TA-AG---......G--------......................................................................................CG-.TC.......................... 9302
CPZ.TZ.2001.TAN2               .....................................................................................................................TGGCTGGAGACCGCTGCTTAACG---........................... 9322
CPZ.TZ.2002.TAN3_1             .....................................................................................................................TGGCTGGAGACCGCTGCTTAACG---........................... 9376
GOR.CM.2007.SIVgor2139_287     ........................................................................................................................GCAGAACCCACTGCTTAACG---........................... 9227
GOR.CM.2007.SIVgorCP2135con    ........................................................................................................................GCAGAACCCACTGCTTAACG---........................... 9221
GOR.CM.2004.SIVgorCP684con     .......................................................................................................................................................................... 9143
H2A.DE.x.BEN                   ............................GCT---CT-TC.ACC-G-A---GGCCGGTACTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAA..AGA----TTC........................10210
H2A.GW.x.ALI                   ............................GCTA--CT-TC.ACC-G-----GGCCGGCACTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAA..AGA----TT.........................10204
H2A.SN.x.ST                    ............................GCTA--CT-TC.ACC-G-----GGCCGGCACTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAAA.AGA----TT......................... 9649
H2B.CI.x.EHO                   ............................GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAA..AAC----TT.........................10099
H2B.GH.86.D205                 ............................GCTA--CT-TC.ACC-GCA--AGGCCAGTGTTGGGCAGA........................................CGGCTCCACGCT.TGCTTGC...TTAA..AAC----TT.........................10128
H2G.CI.92.ABT96                ............................GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................CGGCTCCACGCT.TGCTTGC...TTACAACGA----TT......................... 9560
H2U.FR.96.12034                ............................GCTA--CT-TC.ACC-AAAG.......................................................................................................................... 9607
SMM.SL.92.SL92B                ............................ACTAA-CT-TCAACC-GCAACAGGCCATTGCTGGGTAGAC.......................................GGCTCCACGCT.TGCTTGC...TTGA.....---TC.......................... 9585
SMM.US.x.H9                    ............................GCT---CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TRGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC........................ 9614
SMM.US.x.PGM53                 ............................GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC........................10068
SMM.US.x.SIVsmH635F_L3         ............................GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC........................10166
                                                                                                                                  TAR element end
MAC.US.x.239                   ............................GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TGACTCCACGCT.TGCTTGC...TTAA.AG.C----TTC........................10131
MAC.US.x.251_1A11              ............................GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGGGCAGAG.......................................TGACTCCACGCT.TGCTTGC...TTAA.AG.C----TTC........................10126
MAC.US.x.251_BK28              ............................GCTA--CT-TC.ACC-GCA---GGCCAGTGCTGGGCAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC........................10101
MAC.US.x.EMBL_3                ............................GCTA--CT-TC.ACC-GCA---GGCCAGTGCTGG.CAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC........................ 9239
STM.US.x.STM                   ............................GCTA--CT-TC.ACC-GCA---GGCCAGTGCTGGGTAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAT.AT.AT---TTC........................ 9785
MNE.US.x.MNE027                ............................GCTA--CT-TC.ACC-GCA---GGCCGGTGCTGTGCAGAG.......................................TGGCTCCACGCT.TGCTTGC...TTAA.AG.A----TTC........................ 9608
DEB.CM.99.CM40                 .......................................................................................................................................................................... 9227
DEB.CM.99.CM5                  .......................................................................................................................................................................... 9158
DEN.CD.x.CD1                   ................................-AGC--TGAG..-GGC----GCAGCGGCCTGGGAGCCCTCTAAGGTATTGCAACAAGTACC............................................................................. 9679
LST.CD.88.447                  .......................................................................................................................................................................... 8588
LST.CD.88.485                  .......................................................................................................................................................................... 8589
LST.CD.88.524                  .......................................................................................................................................................................... 8594
LST.KE.x.lho7                  ...................ACCCACCTGGGCCTG-T-AGCCTAGGGAGC---G.................................GGCTCCTAGCTTGCTACTT..TAAA...GCCTTC.................................................. 9829
SUN.GA.98.L14                  .........................CACCTCACCCGAGCCTGAT-A----GGGGAGCTAGAGG............................................CTCCTTGCTTTGCAGTACAAG..........C-T--........................... 9873
MND-1.GA.x.MNDGB1              .................................................................AGATCCCCTTAGAGCAAGGACCAGAGTCCTGAGTGACTGG..GTCTGAGCACCTCACTCGGGGCTGATCACCTCGAGGTAGTGGAACTCCTTGCTTGCTTGCTAT 9183
OLC.CI.97.97CI12               .......................................................................................................................................................................... 8678
WRC.CI.97.97CI14               .......................................................................................................................................................................... 9484
WRC.CI.98.98CI04               .......................................................................................................................................................................... 9461
WRC.GM.05.Pbt_05GM_X02         ...................CTGACCACGCTT-CTTGTTAGGGTGT-TGCAG-T..................................................................................................................... 8636
MND-2.CM.98.CM16               .......................................................................................................................................................................... 9518
MND-2.GA.x.M14                 .......................................................................................................................................................................... 9529
MND-2.x.x.5440                 ...................TCGCAGGCCCCT--CTTGTGGCTCTGG-TAGCTCG.TCAGGTG..TTCTGGAAAGGCTTACGGGAGGACACTCTTGCTTGGTCTTG..ATAGACCTCTAGCAATCCCTTAGGCCTGGAGAT.-GTGGGATTGACTACCGCTTGCTTGCTAT 9500
DRL.x.x.FAO                    CTTGGGCTCTGGGTAGCTCGGTCAGGAGGCTTCTGGAGGTAGAT-ACTAGGGGCTGCTCTGGTCTTGGTCTGGTAGACCTTCGCTACTTCTAGTGGCCATAGAAGTAGGGGAGTAGGCACCGCTTGC...TTAT..CCTAA-TGCTC....................... 9593
MON.CM.99.L1                   .......................................................................................................................................................................... 9448
MON.NG.x.NG1                   ...................GGGAGCCCTTGGA-CTAGGTAAAGCGGC--GGGAGCCCTTGCCTAGAGAGGTGCCGGCTCTCGGCTCGCTTGCTGAATAGGCTTTACC...............................AG-C-........................... 8202
GSN.CM.99.CN166                .......................................................................................................................................................................... 9414
GSN.CM.99.CN71                 .......................................................................................................................................................................... 9410
TAL.CM.00.266                  .......................................................................................................................................................................... 9338
TAL.CM.01.8023                 .......................................................................................................................................................................... 8845
MUS-1.CM.01.1085               .......................................................................................................................................................................... 9420
MUS-1.CM.01.CM1239             ...................GTTCTCGGCTTGCTTTG-TTCGAAG--TA--G-GCCCACAATA............................................................................................................ 9508
MUS-2.CM.01.CM1246             .......................................................................................................................................................................... 9484
MUS-2.CM.01.CM2500             .......................................................................................................................................................................... 9430
SYK.KE.x.KE51                  .......................................................................................................................................................................... 9068
SYK.KE.x.SYK173                ...................GCCCTTGCCAGTAA-GG-TAGTGAGG-C-GGGTTTGGCCGCCCGGTCTTCGGCAGCTCTCCGTTGCTT.............................GCTTTGCTTACGC.TGATCAGCC-.A............................ 9530
RCM.GA.x.GAB1                  .............................GGT-TCT-TGAACAGGCTTGCTGGGGTGCCTCTCGCTCTTCGGGTAGACCGCCAGTTGAGGCTCGGCCGGCCTCAA..CGGGAGAGATCACCGCTTGC...TTATAGCCTTGAAGCTC....................... 9330
RCM.NG.x.NG411                 ..............................A-TTCT-TGAAGGG--TGC-GGGAGCGCCTTGGCTGGGTCTTGGTAGACCTTCAGCTGGGTACTTGGCCAGTCCCAGTGGGAGTAGGCTCCGCTTGCT..TTCT.GGTTAA-TGCTC....................... 9422
GRV.ET.x.GRI_677               ...............................---CT--.....TTG----C-TAT...................................................................................AG---........................... 9505
SAB.SN.x.SAB1C                 ............................GGG-A-GG-AAGCCT-AGAGCACTCTGGTGAGGTCTTGGTAGACCCTCGCCTGGCGACTGGCCATTGCCAGTAGCAG..AGACTCC.......GCTTGC...TTGCTT.GATTG-CTC........................ 9933
TAN.UG.x.TAN1                  ..............................GA-CCT--......T------GAA...................................................................GGCTATT...........G--............................ 9674
VER.DE.x.AGM3                  ...............................---CT--.....TTG-----GAA....................................................................................AG-.T........................... 9193
VER.KE.x.9063                  ...............................---CT--.....TTG----GGAA....................................................................................AG--............................ 9693
VER.KE.x.AGM155                ...............................---CT--.....TTG-----GAA....................................................................................AG--............................ 9674
VER.KE.x.TYO1_patent           ...............................---CT--.....TTG-----GAA....................................................................................AG--............................ 9657
COL.CM.x.CGU1                  .......................................................................................................................................................................... 8728
```

```
                                                        Extensive secondary structure in this region?
                      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...GGT..AA.CTAGAGA..............TCCCT..CAGA...................CCCTTTTAGTCAGTGTGGAA 9708
H1A1.UG.85.U455           ........-------..------.......-----..GTG---G-..------...----------.--.--..A-T.-------...---..--.-------..............-----..----...................--AC-A---A-T.GTGTA-- 9167
H1B.US.90.WEAU160         ........-------..------.......-----..--.------.------...----------.--.--..---.-------...---..--.-------..............-----..----...................--A----------------- 9708
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           ........................................................................................................................................................ 8604
H101_AE.TH.90.CM240       ........-------..------.......-----..--.---A-........................................................................................................... 9203
H102_AG.NG.x.IBNG         ........-------..------.......-----..--.------.------...----------.--A--..---.-------...---..--.-............................................................ 9201
H103_AB.RU.97.KAL153_2    ........-------..------.......-----..--.-............................................................................................................... 8808
H104_cpx.CY.94.CY032      ........-------..------.......-----..--.----............................................................................................................ 9050
H1N.CM.95.YBF30           ........................................................................................................................................................ 9182
H1O.BE.87.ANT70           ........-------..------.......-----..--.AG.............................................................................................................. 9754
H1O.CM.91.MVP5180         ........-------..------.......-----..--.AGAAGC.AGT---...--CT-A..--TG..--CAACC.CTGG-G-...CTA..G-......--TC.............................................. 9793
H1P.FR.06.RBF168          ........................................................................................................................................................ 8654
CPZ.CM.01.SIVcpzCAM13     ........-------..------.......-----..--CT-CA............................................................................................................ 9284
CPZ.CM.05.SIVcpzEK505     ........-------..------.......-----..--CT-AA............................................................................................................ 9132
CPZ.CM.05.SIVcpzLB7       ........-------..------.......-----..--CTGCA............................................................................................................ 9170
CPZ.CM.05.SIVcpzMB66      ........-------..------.......-----..--A---A-........................................................................................................... 9281
CPZ.CM.05.SIVcpzMT145     ........-------..------.......-----..--CTGCA............................................................................................................ 9208
CPZ.CM.-.SIVcpzMB897      ........-------..------.......-----..-AA---G............................................................................................................ 9167
CPZ.CM.-.SIVcpzDP943      ........-------..------.......-----..--A---G............................................................................................................ 9247
CPZ.CM.98.CAM3            ........-------..------.......-----..--A---G..A-C-----..------TA--CA..--C..A-.-A-----..---..--.-------TCC...............................CTCAGAA.-TA--A--A..GTA-GT--- 9162
CPZ.CM.98.CAM5            ........................................................................................................................................................ 9261
CPZ.GA.88.GAB1            ........-------..------.......-----TGTAT--G..A-C-----..------TA--TA..--C..A-.AC-----..---..--.-------TCC...............................CTCAGAT.TAAA--ATA..GTCAA---- 9800
CPZ.GA.88.GAB2            ........-------..------.......-----..--CTC-A..A-------..-----T-A--TG..--C..A-.-A-----..---..GT.-------TCC...............................CTCAGAC.TAAA.CATAA.G-A------ 9135
CPZ.US.85.CPZUS           ........-------..------.......-----..--CT--A..A-------..------TA--.C..--C..A-.-------..---..--.-------TCC...............................CTCAGAA.-T---AA-A..GTA-GT--- 9770
CPZ.CD.90.ANT             ........................................................................................................................................................ 9068
CPZ.TZ.00.TAN1            ........---------..-C----.......-----..A-.TG-T.-........................................................................................................ 9326
CPZ.TZ.2001.TAN2          ........---------..-C----.......TGA---..--TTA........................................................................................................... 9346
CPZ.TZ.2002.TAN3_1        ........---------..-C----.......TGA---..--TTA........................................................................................................... 9400
GOR.CM.2007.SIVgor2139_287 .......-------..------.......A-----..--AA--.............................................................................................................. 9252
GOR.CM.2007.SIVgorCP2135con ......-------..------.......A-----..--AA--.............................................................................................................. 9246
GOR.CM.2004.SIVgorCP684con ........................................................................................................................................................ 9143
H2A.DE.x.BEN              ........-------..--.---AG.....--AGA..A-CAAGTT-.AGT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-..TCGG-GTTC-TCTGA....GTAACAAGA-CCTGGTCTGTTAGGACCCTTTCTCGCTT.TGGGAATCCA--GCA----10348
H2A.GW.x.ALI              ........-------..--.---AA.....--AGA..A-CAGGTT-.AAG---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-..TCGG-GTTC-CCTGA....GTAACAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CGG--GCA----10342
H2A.SN.x.ST               ........-------..--.---AG.....--AGA..A-CA............................................................................................................... 9672
H2B.CI.x.EHO              ........------A...--.---TA.....--AGA..A-CAAGAC-.AGT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-..TCGG-..GCTCCAC......TA..GAAA-CCTGGTCTGTTAGGACCCTTCTGCTT.TGGGAA-CCA--GCA----10231
H2B.GH.86.D205            ........-------..--.---AA.....--AGA..A-CAAGT-..AGT---TGT-CCCAT-TCT.CCTAGTC-C..GCCTGG-...CAC..TCGG-..GCTCCAC......TGA.TAGA-CCTGGTCTGTTAGGACCCTTCTTGCTT.TGGGAA-CCG-AGCG----10258
H2G.CI.92.ABT96           ........-------..--.---AAA.....--AGA..A-TAAG---GTGT---TGT............................................................................................... 9599
H2U.FR.96.12034           ........................................................................................................................................................ 9607
SMM.SL.92.SL92B           ........-------..G--.--AAC.....--AGA..A-CAAGTC.......................................................................................................... 9613
SMM.US.x.H9               ........-------G..--.---AG.....--AGA..A-TA.............................................................................................................. 9638
SMM.US.x.PGM53            ........-------..--.---GC.....--AGA..A-TAAG---..GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-C.TCGG-..-CTCGAAT...CATAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGC.....................10184
SMM.US.x.SIVsmH635F_L3    ........-------..--.---AT.....--AGA..A-TAAG---..GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CA-C.TCGG-..-CTCGACA..CATAAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CCG--GCA----10304
                      Poly-A signal         3' LTR R repeat end 3' LTR U5 start

MAC.US.x.239              ........-------..--.---AT..T..--AGA..A-TAAG-T-..GT---TGT-CCCAT-TCT.CCTAGCC-CC.GCCTGG-...CAAC.TCGG-..-CTCAATA....ATAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CCG-AGCA----10268
MAC.US.x.251_1A11         ........-------..--.---AT..T..--AGA..A-TAAG-T-..GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CAAC.TCGG-..-CTCGGTA....ATAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGGGAA-CCG-AGCA----10263
MAC.US.x.251_BK28         ........-------..--.---AT..T..--AGA..A-TAAG-C-..GT---TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CAAC.TCGG-..-CTCGGTA....ATAAGAAGA-CCTGGTCTGTTAGGACCCTTTCTGCTT.TGAGAA-CCG-AGCA----10238
MAC.US.x.EMBL_3           ........-------..--.---CA..T.T--AGA..A-TA............................................................................................................... 9264
STM.US.x.STM              ........-------..--.---AA..T..--AGA..A-TAAG---..GT-G-TGT-CCCAT-TCT.CCTAGTC-CC.GCCTGG-...CAAC.TCGG-..-CTCAA.............GA-CCTGGTCTGTTAGGACCCTTTCTGC..................... 9892
MNE.US.x.MNE027           ........-------..--.---AT.....--AGA..A-TAAG-C-..GT---TGC-CCCAT-TCT.CCTAGTC-CC.......................................................................... 9664
DEB.CM.99.CM40            ........................................................................................................................................................ 9227
DEB.CM.99.CM5             ........................................................................................................................................................ 9158
DEN.CD.x.CD1              ........................................................................................................................................................ 9679
LST.CD.88.447             ........................................................................................................................................................ 8588
LST.CD.88.485             ........................................................................................................................................................ 8589
LST.CD.88.524             ........................................................................................................................................................ 8594
LST.KE.x.lho7             ........-------.....AA-TTGCTCATTG-T--AAA-TACAAGCTAGT--CTCA-AGTAGTC--T-TC-CCTTC-CCC.........TGGTTCAG-GATC.................T--CA-TAGAGAGATTGGAGCCTTGTGAT--GGGAC--GGCT-CCA--. 9957
SUN.GA.98.L14             ........-------...-C----------.--AG...-TTTACTGCAAGCAA-TGT-CCCGTGT--TCTC--TCTCC..............TTCTAA-CCCTGTTCGACTGCTCTCAG-TAGGGAGC-A............TTACTGG--GCC-AGTGATCC-G--CT 9995
MND-1.GA.x.MNDGB1        TGTCTTC.-------T.AAC-TAGAA..TTAGA-CA.................................................................................................................... 9215
OLC.CI.97.97CI12          ........................................................................................................................................................ 8678
WRC.CI.97.97CI14          ........................................................................................................................................................ 9484
WRC.CI.98.98CI04          ........................................................................................................................................................ 9461
WRC.GM.05.Pbt_05GM_X02    ........-------...AG-CTTTTGCTTAGCA--AAGCGTG-GGTTAGTT-TTGT---AAGAAGTAGAACCCT-G-AG-CT-.................................................................... 8709
MND-2.CM.98.CM16          ........................................................................................................................................................ 9518
MND-2.GA.x.M14            ........................................................................................................................................................ 9529
MND-2.x.x.5440            TATTGCCC-------....C-TA-CTGATTAGAATAGCAAGTAAG-G.TACTC-TCCAT-TTGTC-TCAATAGAA...ACTCTGG-TACC-GAG-.......................----..T----TCTGTGGCAGAGATTTGATAT-TAG-G--AGTAGAGAA-- 9636
DRL.x.x.FAO               ........-------C..TA..--CA.....--T--ATG-ATGAGT-.AGT--ACTC-TGT-TTT-G........T-CT-AGTGAGACTCTGGTT-CTGGAGA..............TC--TCAGACT-GGTGAGAGCTATCCTGAGCTTTAG-..GAGAGTAAA-A-- 9720
MON.CM.99.L1              ........................................................................................................................................................ 9448
MON.NG.x.NG1              ........------....G-GCTTGAAAGCAAAA-AAAAAAAAAA--AAA-AAAAAAAAAAAAAAAAAAA................................................................................... 8260
GSN.CM.99.CN166           ........................................................................................................................................................ 9414
GSN.CM.99.CN71            ........................................................................................................................................................ 9410
TAL.CM.00.266             ........................................................................................................................................................ 9338
TAL.CM.01.8023            ........................................................................................................................................................ 8845
MUS-1.CM.01.1085          ........................................................................................................................................................ 9420
MUS-1.CM.01.CM1239        ........................................................................................................................................................ 9508
MUS-2.CM.01.CM1246        ........................................................................................................................................................ 9484
MUS-2.CM.01.CM2500        ........................................................................................................................................................ 9430
SYK.KE.x.KE51             ........................................................................................................................................................ 9068
SYK.KE.x.SYK173           ........-------T.GTGGTAAACGCAAG-A.....ACGTC-GCT.C-CA-CTGT-C---TG-GT.CACC-CCCGA.A-GTG..................................................................... 9597
RCM.GA.x.GAB1             ........------C......ATGCCAG..--AGTTTAC.TG-A-GC.A----TGTGCCTGTTT-A....CCTCTCAGCAGT-AACGACTCTGGGGTAG-GA..............TC--TCAG........ATTCTTGTGGCAGAAGAG-C--G..GGCTAAGAA-- 9453
RCM.NG.x.NG411            ........------C....-.--ATGCTCTTAGTT-AACCAGAATC-.AGT---TG..CCCATTT-A......TCTCTCA-TGAAACGACTCTGGGGT-G-GA..............TC--TCAG........ATATTTGTGGCAGAAGT.AC-.G-AGGCTAAGAA-- 9545
GRV.ET.x.GRI_677          ........------C....C---TCGC.....--AGT..C-CTA-ATTG.G---CAAG--CT-ATTG-TGCGCCGA-CCT.CTAGAGG......TG-A-CTCTCTTACTGGGTTCTCCTG-A--CAGGT.............................G-G..-G 9612
SAB.SN.x.SAB1C            ........-------T..------........--AGTTTACAAG-A--GCA----TGTGCCCATTTATTCCTCAG...................................GACAA-CCTGGTTACTAAGGATCCC.........TGA..........-AGAA--10025
TAN.UG.x.TAN1             ........------....G-----TG....C--A--..A-T--ATCTGAG.CAAGTGCC-CATT-CGCGCCC-TCC-CAG-.......................GAACCCTCGTTACTGGG-.............TTTCTTATC-AGGC-G--............ 9773
VER.DE.x.AGM3             ........------T...---CG-TG...CA--AG-.CA-AGACTT..--GAGAGA....CATT-C-TG-TGCATCCAGCACAAG..................TGAGGGAAGACCCTGATGGAAT--ACCATGGAGAAGTCTTGGTCTGGAAG.........-T--ATC 9314
VER.KE.x.9063             ........------AC..T-----TG...CA--AGA..G.CT-ATCTGAGTCAAGTGCCCT-ATTGACG.CC-CAC-C..A-GCAGG.....GG--C-GTTCCTTACTGGGTTCTCTCTCTGACCCAG..................................GCG---G 9804
VER.KE.x.AGM155           ........------....------TG...CA--AGA..G.CT-ATCTGAGTCAAGTGCCCT-ATTAATG.CC-CAC-C..TTGAACG.....GG-GAAGTTCCTTACTGGGTTCTCTCTCAAACCCAG..................................GCGA--G 9783
VER.KE.x.TYO1_patent      ........------....------TG...CA--AGA..G.CT-ATCTGAGTCAAGTG-CCT-ATTGACG.CC-CAC-CTCTTGAACG......GG-ATCTTCCTTACTGGGTTCTCTCTCT........................GACCCAGGCGAG--............ 9767
COL.CM.x.CGU1             ........................................................................................................................................................ 8728
```

```
                                3' LTR U5 end
H1B.FR.83.HXB2                  AATCTCTAGCA..   9719
H1A1.UG.85.U455                 -----------..   9178
H1B.US.90.WEAU160               -----------..   9719
H1C.ET.86.ETH2220               ............   9031
H1D.CD.84.84ZR085               ............   8975
H1F1.BE.93.VI850                ............   8903
H1G.SE.93.SE6165                ............   9074
H1H.CF.90.056                   ............   8953
H1J.SE.93.SE7887                ............   8943
H1K.CM.96.MP535                 ............   8604
H101_AE.TH.90.CM240             ............   9203
H102_AG.NG.x.IBNG               ............   9201
H103_AB.RU.97.KAL153_2          ............   8808
H104_cpx.CY.94.CY032            ............   9050
H1N.CM.95.YBF30                 ............   9182
H1O.BE.87.ANT70                 ............   9754
H1O.CM.91.MVP5180               ............   9793
H1P.FR.06.RBF168                ............   8654
CPZ.CM.01.SIVcpzCAM13           ............   9284
CPZ.CM.05.SIVcpzEK505           ............   9132
CPZ.CM.05.SIVcpzLB7             ............   9170
CPZ.CM.05.SIVcpzMB66            ............   9281
CPZ.CM.05.SIVcpzMT145           ............   9208
CPZ.CM.--.SIVcpzMB897           ............   9167
CPZ.CM.--.SIVcpzDP943           ............   9247
CPZ.CM.98.CAM3                  ------------..   9170
CPZ.CM.98.CAM5                  ............   9261
CPZ.GA.88.GAB1                  -----------..   9811
CPZ.GA.88.GAB2                  -----------..   9146
CPZ.US.85.CPZUS                 -----------..   9781
CPZ.CD.90.ANT                   ............   9068
CPZ.TZ.00.TAN1                  ............   9326
CPZ.TZ.2001.TAN2                ............   9346
CPZ.TZ.2002.TAN3_1              ............   9400
GOR.CM.2007.SIVgor2139_287      ............   9252
GOR.CM.2007.SIVgorCP2135con     ............   9246
GOR.CM.2004.SIVgorCP684con      ............   9143
H2A.DE.x.BEN                    ----C------..  10359
H2A.GW.x.ALI                    ----C------..  10353
H2A.SN.x.ST                     ............   9672
H2B.CI.x.EHO                    ----C------..  10242
H2B.GH.86.D205                  ----C------..  10269
H2G.CI.92.ABT96                 ............   9599
H2U.FR.96.12034                 ............   9607
SMM.SL.92.SL92B                 ............   9613
SMM.US.x.H9                     ............   9638
SMM.US.x.PGM53                  ............  10184
SMM.US.x.SIVsmH635F_L3          ----C------..  10315
                                3' LTR U5 end
MAC.US.x.239                    ----C------..  10279
MAC.US.x.251_1A11               ----C------..  10274
MAC.US.x.251_BK28               ----C------..  10249
MAC.US.x.EMBL_3                 ............   9264
STM.US.x.STM                    ............   9892
MNE.US.x.MNE027                 ............   9664
DEB.CM.99.CM40                  ............   9227
DEB.CM.99.CM5                   ............   9158
DEN.CD.x.CD1                    ............   9679
LST.CD.88.447                   ............   8588
LST.CD.88.485                   ............   8589
LST.CD.88.524                   ............   8594
LST.KE.x.lho7                   ............   9957
SUN.GA.98.L14                   G-CGG------..  10006
MND-1.GA.x.MNDGB1               ............   9215
OLC.CI.97.97CI12                ............   8678
WRC.CI.97.97CI14                ............   9484
WRC.CI.98.98CI04                ............   9461
WRC.GM.05.Pbt_05GM_X02          ............   8709
MND-2.CM.98.CM16                ............   9518
MND-2.GA.x.M14                  ............   9529
MND-2.x.x.5440                  -----C------..  9647
DRL.x.x.FAO                     ------C-----..  9731
MON.CM.99.L1                    ............   9448
MON.NG.x.NG1                    ............   8260
GSN.CM.99.CN166                 ............   9414
GSN.CM.99.CN71                  ............   9410
TAL.CM.00.266                   ............   9338
TAL.CM.01.8023                  ............   8845
MUS-1.CM.01.1085                ............   9420
MUS-1.CM.01.CM1239              ............   9508
MUS-2.CM.01.CM1246              ............   9484
MUS-2.CM.01.CM2500              ............   9430
SYK.KE.x.KE51                   ............   9068
SYK.KE.x.SYK173                 ............   9597
RCM.GA.x.GAB1                   .T--C---C--GT  9465
RCM.NG.x.NG411                  ----C-A-C--G.  9557
GRV.ET.x.GRI_677                --A---C-----..  9623
SAB.SN.x.SAB1C                  ----CT------..  10036
TAN.UG.x.TAN1                   ------C-----..  9784
VER.DE.x.AGM3                   CCATGT------..  9325
VER.KE.x.9063                   --A---C-----..  9815
VER.KE.x.AGM155                 --A---C-----..  9794
VER.KE.x.TYO1_patent            --A---C-----..  9778
COL.CM.x.CGU1                   ............   8728
```