

VII

Primate Lentivirus Proteins

Introduction	659
Table of sequences in PLV protein alignments	660
Gag	663
Pol	668
Vif	677
Vpx	680
Vpr	681
Tat	683
Rev	685
Vpu	686
Env	687
Nef	696

Primate Lentivirus Protein Alignments

As has been noted in the PLV Complete Genome alignment section (IV), the phylogenetic relationships between genes, and thus their corresponding protein sequences, is not completely clear. See the tree and discussion in section IV.

The selection of Primate Lentivirus Protein sequences for the following alignments was based on the sequences in the complete genome alignment as a starting alignment, and complete or nearly complete genes from other isolates were added if they increased the diversity of samples represented. For example, several diverse African green monkey virus isolates have been sequenced only in a region of the *env* gene. When necessary, some of the more common sequences (such as HIV-1 M group) were removed to make room on the Compendium pages for these diverse virus sequences. More complete alignments are available from our web site: http://www.hiv.lanl.gov/content/hiv-db/ALIGN_CURRENT/ALIGN-INDEX.html where space limitations are not an issue.

The annotation is mainly based on knowledge from HIV-1, and should therefore be taken “with a grain of salt.”

Table 1: Table of sequences in the Primate Lentivirus (PLV) complete genome alignments

Name	Accession	Region	Author	Reference
ASC.-.-.Qu	AJ551401	POL	Verschoor, EJ	<i>J Gen Virol</i> 85 (1):21–4 (2004)
BAB.TZ.85.2010E	U10897	ENV	Jin, MJ	<i>J Virol</i> 68 (12):8454–60 (1994)
COL.CM.00.COL243	AF478607	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
COL.CM.01.COL247	AF478606	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
COL.CM.99.COL11	AF478608	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
COL.CM.-.CGU1	AF301156	ENV GAG NEF POL REV TAT VIF VPR	Courgnaud, V	<i>J Virol</i> 75 (2):857–66 (2001)
CPZ.CD.-.ANT	U42720	ENV GAG NEF POL REV TAT VIF VPR VPU	Vanden Haesevelde	<i>Virology</i> 221 (2):346–50 (1996)
CPZ.CM.98.CAM5	AJ271369	VPU	Muller-Trutwin, MC	<i>J Med Primatol</i> 29 (3–4):166–72 (2000)
CPZ.CM.-.CAM3	AF115393	VPU	Corbet, S	<i>J Virol</i> 74 (1):529–34 (2000)
CPZ.GA.88.GAB2	AF382828	VPU	Gao, F	Unpublished
CPZ.GA.-.CPZGAB	X52154	ENV GAG NEF POL REV TAT VIF VPR VPU	Huet, T	<i>Nature</i> 345 (6273):356–9 (1990)
CPZ.TZ.-.TAN1	AF447763	VPU	Santiago, ML	<i>J Virol</i> 77 (3):2233–2242 (2003)
CPZ.US.85.CPZUS	AF103818	VPU	Gao, F	<i>Nature</i> 397 (6718):436–41 (1999)
DEB.CM.01.DEB1083	AF478600	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
DEB.CM.01.DEB1161	AF478604	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
DEB.CM.01.DEBS1014	AF478602	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
DEB.CM.99.CM40	AY523865	ENV GAG NEF POL REV TAT VIF VPR	Bibollet-Ruche, F	<i>J Virol</i> 78 (14):7748–7762 (2004)
DEB.CM.99.CM5	AY523866	ENV GAG NEF POL REV TAT VIF VPR	Bibollet-Ruche, F	<i>J Virol</i> 78 (14):7748–7762 (2004)
DEB.CM.99.DEB CNE1	AF478605	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
DEN.CD.-.CD1	AJ580407	ENV GAG NEF POL REV TAT VIF VPR VPU	Dazza, MC	Unpublished
DRL.-.-.FAO	AY159321	ENV GAG NEF POL REV TAT VIF VPR VPX	Hu, J	<i>J Virol</i> 77 (8):4867–4880 (2003)
GRV.ET.-.GRI_677	M66437	ENV GAG NEF POL REV TAT VIF VPR	Fomsgaard, A	<i>Virology</i> 182 (1):397–402 (1991)
GRV.-.-.GRI2E	U03995	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
GRV.-.-.GRI3E	U03994	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
GSN.CM.99.CN166	AF468659	ENV GAG NEF POL REV TAT VIF VPR VPU	Courgnaud, V	<i>J Virol</i> 76 (16):8298–309 (2002)
GSN.CM.99.CN71	AF468658	ENV GAG NEF POL REV TAT VIF VPR VPU	Courgnaud, V	<i>J Virol</i> 76 (16):8298–309 (2002)
GSN.CM.99.GSN CN7	AF478589	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
H101_AE.TH.90.CM240	U54771	VPU	Laukkanen, T	<i>J Virol</i> 70 (9):5935–43 (1996)
H102_AG.NG.-.IBNG	L39106	VPU	Howard, TM	<i>ARHR</i> 10 (12):1755–7 (1994)
H103_AB.RU.97. KAL153_2	AF193276	VPU	Liitsola, K	<i>AIDS</i> 12 (14):1907–19 (1998)
H104_cpx.CY.94.CY032	AF049337	VPU	Gao, F	<i>J Virol</i> 72 (12):10234–41 (1998)
H1A1.UG.85.U455	M62320	VPU	Oram, JD	<i>ARHR</i> 6 (9):1073–8 (1990)
H1B.FR.83.HXB2	K03455	ENV GAG NEF POL REV TAT VIF VPR VPU	Wong-Staal, F	<i>Nature</i> 313 (6000):277–84 (1985)
H1B.US.90.WEAU160	U21135	VPU	Wei, X	<i>Nature</i> 422 (6929):307–312 (2003)
H1C.ET.86.ETH2220	U46016	VPU	Salminen, MO	<i>ARHR</i> 12 (14):1329–39 (1996)
H1D.CD.84.84ZR085	U88822	VPU	Gao, F	<i>J Virol</i> 72 (7):5680–98 (1998)
H1F1.BE.93.VI850	AF077336	VPU	Laukkanen, T	<i>Virology</i> 269 (1):95–104 (2000)
H1G.SE.93.SE6165	AF061642	VPU	Laukkanen, T	<i>Virology</i> 247 (1):22–31 (1998)
H1H.CF.90.056	AF005496	VPU	Murphy, E	<i>ARHR</i> 9 (10):997–1006 (1993)
H1J.SE.93.SE7887	AF082394	VPU	Laukkanen, T	<i>ARHR</i> 15 (3):293–7 (1999)
H1K.CM.96.MP535	AJ249239	VPU	Peeters, M	<i>ARHR</i> 16 (2):139–51 (2000)
H1N.CM.95.YBF30	AJ006022	VPU	Simon, F	<i>Nat Med</i> 4 (9):1032–7 (1998)

H1O.BE.87.ANT70	L20587	VPU	Vanden Haesevelde	<i>J Virol</i> 68 (3):1586–96 (1994)
H1O.CM.91.MVP5180	L20571	VPU	Gurtler, LG	<i>J Virol</i> 68 (3):1581–5 (1994)
H2A.DE.-.BEN	M30502	ENV GAG NEF POL REV TAT VIF VPR VPX	Kirchhoff, F	<i>Virology</i> 177 (1):305–11 (1990)
H2A.GW.-.ALI	AF082339	VPX	Azevedo-Pereira	Unpublished (1998)
H2A.SN.-.ST	M31113	VPX	Kumar, P	<i>J Virol</i> 64 (2):890–901 (1990)
H2B.CI.-.EHO	U27200	VPX	Rey-Cuille, MA	<i>Virology</i> 202 (1):471–6 (1994)
H2B.GH.86.D205	X16109	VPX	Dietrich, U	<i>Nature</i> 342 (6252):948–50 (1989)
H2G.CI.-.ABT96	AF208027	VPX	Brennan, CA	<i>ARHR</i> 13 (5):401–4 (1997)
H2U.FR.96.12034	AY530889	VPX	Damond, F	<i>ARHR</i> 20 (6):666–672 (2004)
LST.CD.88.447	AF188114	ENV GAG NEF POL REV TAT VIF VPR	Beer, BE	<i>J Virol</i> 74 (8):3892–8 (2000)
LST.CD.88.485	AF188115	ENV GAG NEF POL REV TAT VIF VPR	Beer, BE	<i>J Virol</i> 74 (8):3892–8 (2000)
LST.CD.88.524	AF188116	ENV GAG NEF POL REV TAT VIF VPR	Beer, BE	<i>J Virol</i> 74 (8):3892–8 (2000)
LST.KE.-.lho7	AF075269	ENV GAG NEF POL REV TAT VIF VPR	Hirsch, VM	<i>J Virol</i> 73 (2):1036–45 (1999)
MAC.US.-.239	M33262	ENV GAG NEF POL REV TAT VIF VPR VPX	Kestler, H	<i>Science</i> 248 (4959):1109–12 (1990)
MND-1.GA.-.MNDGB1	M27470	ENV GAG NEF POL REV TAT VIF VPR	Tsujimoto, H	<i>Nature</i> 341 (6242):539–41 (1989)
MND-2.-.-.5440	AY159322	ENV GAG NEF POL REV TAT VIF VPR VPX	Hu, J	<i>J Virol</i> 77 (8):4867–4880 (2003)
MND-2.CM.00. MND2S46	AF478598	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
MND-2.CM.00.MND2S6	AF478599	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
MND-2.CM.01. MND2S109	AF478597	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
MND-2.CM.98.CM16	AF367411	ENV GAG NEF POL REV TAT VIF VPR VPX	Takehisa, J	<i>ARHR</i> 17 (12):1143–54 (2001)
MND-2.CM.99.MND254	AF478596	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
MND-2.CM.-. MND302ENV	AF328294	ENV	Souquiere, S	<i>J Virol</i> 75 (15):7086–96 (2001)
MND-2.GA.-.M14	AF328295	ENV GAG NEF POL REV TAT VIF VPR VPX	Souquiere, S	<i>J Virol</i> 75 (15):7086–96 (2001)
MNE.US.-.MNE027	U79412	ENV GAG NEF POL REV TAT VIF VPR VPX	Kimata, JT	<i>J Virol</i> 72 (1):245–56 (1998)
MON.CM.99.L1	AY340701	ENV GAG NEF POL REV TAT VIF VPR VPU	Courgnaud, V	<i>J Virol</i> 77 (23):12523–12534 (2003)
MON.NG.-.NG1	AJ549283	ENV GAG NEF POL REV TAT VIF VPR VPU	Barlow, KL	<i>J Virol</i> 77 (12):6879–88 (2003)
MUS.CM.01.1085	AY340700	ENV GAG NEF POL REV TAT VIF VPR VPU	Courgnaud, V	<i>J Virol</i> 77 (23):12523–12534 (2003)
MUS.CM.01.MUSS1239	AF478592	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
OLC.CI.-.SIVolc_ 97IC_12	AY138269	POL	Courgnaud, V	<i>J Virol</i> 77 (1):744–748 (2003)
PAT.SN.-.PAMG31NEF	U26299	NEF	Bibollet-Ruche, F	<i>J Gen Virol</i> 77 (4):773–81 (1996)
RCM.GA.-.GAB1	AF382829	ENV GAG NEF POL REV TAT VIF VPR VPX	Gao, F	<i>Science</i> 300 (5626):1713 (2003)
RCM.GA.-. SIVrcmGB1c1	AY336733	VIF	Gaddis, NC	Unpublished
RCM.GA.-. SIVrcmGB1c3	AY336734	VIF	Gaddis, NC	Unpublished
RCM.NG.-.NG411	AF349680	ENV GAG NEF POL REV TAT VIF VPR VPX	Beer, BE	<i>J Virol</i> 75 (24):12014–27 (2001)
SAB.-.-.C5	U59191	ENV	Vidal, N	Unpublished

SAB.-.-SAB3E	U03997	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
SAB.-.-SAB4E	U03998	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
SAB.SN.-.SAB1C	U04005	ENV GAG NEF POL REV TAT VIF VPR	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
SMM.SL.92.SL92B	AF334679	ENV GAG NEF POL REV TAT VIF VPR VPX	Chen, Z	<i>J Virol</i> 70 (6):3617–27 (1996)
SMM.US.-.BPZ_m12	AY603050	VIF	Glenn, AA	<i>Virology</i> 325 (2):297–307 (2004)
SMM.US.-.H9	M80194	ENV GAG NEF POL REV TAT VIF VPR VPX	Fultz, PN	<i>PNAS</i> 83 (14):5286–90 (1986)
SMM.US.-.PGM53	AF077017	ENV GAG NEF POL REV TAT VIF VPR VPX	Novembre, FJ	<i>J Virol</i> 72 (11):8841–51 (1998)
STM.US.-.STM	M83293	ENV GAG NEF POL REV TAT VIF VPR VPX	Novembre, FJ	<i>Virology</i> 186 (2):783–7 (1992)
SUN.GA.98.L14	AF131870	ENV GAG NEF POL REV TAT VIF VPR	Beer, BE	<i>J Virol</i> 73 (9):7734–44 (1999)
SYK.KE.-.KE51	AY523867	ENV GAG NEF POL REV TAT VIF VPR	Bibollet-Ruche, F	<i>J Virol</i> 78 (14):7748–7762 (2004)
SYK.KE.-.SYK173	L06042	ENV GAG NEF POL REV TAT VIF VPR	Hirsch, VM	<i>J Virol</i> 67 (3):1517–28 (1993)
TAL.CM.00.TAL266	AF478595	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
TAL.CM.00.TAL271	AF478594	POL	Peeters, M	<i>Emerg Infect Dis</i> 8 (5):451–7 (2002)
TAN.CF.-.AGMB05	M81070	ENV	Muller, M	<i>J Virol</i> 67 (3):1227–35 (1993)
TAN.CF.-.AGMTB14	M80208	ENV	Nerrienet, E	Unpublished (1992)
TAN.-.-.TAN17E	U04000	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
TAN.-.-.TAN40E	U04001	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
TAN.-.-.TAN49E	U04002	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
TAN.UG.-.TAN1	U58991	ENV GAG NEF POL REV TAT VIF VPR	Stivahtis, GL	<i>Virology</i> 228 (2):394–9 (1997)
VER.DE.-.AGM3	M30931	ENV GAG NEF POL REV TAT VIF VPR	Baier, M	<i>Virology</i> 176 (1):216–21 (1990)
VER.ET.-.VER385E	U10898	ENV	Jin, MJ	<i>J Virol</i> 68 (12):8454–60 (1994)
VER.KE.-.9063	L40990	ENV GAG NEF POL REV TAT VIF VPR	Hirsch, VM	<i>J Virol</i> 69 (2):955–67 (1995)
VER.KE.-.AGM155	M29975	ENV GAG NEF POL REV TAT VIF VPR	Johnson, PR	<i>J Virol</i> 64 (3):1086–92 (1990)
VER.KE.-.TYO1	X07805	ENV GAG NEF POL REV TAT VIF VPR	Fukasawa, M	<i>Nature</i> 333 (6172):457–61 (1988)
VER.KE.-.VER266E	U10896	ENV	Jin, MJ	<i>J Virol</i> 68 (12):8454–60 (1994)
VER.-.-.VER1E	U04003	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
VER.-.-.VER2E	U04004	ENV	Jin, MJ	<i>EMBO J</i> 13 (12):2935–47 (1994)
WRC.CI.97.14	AY138268	POL	Courgnaud, V	<i>J Virol</i> 77 (1):744–748 (2003)

	/ Gag p17 Matrix start	
SAB.SN.-.SAB1C	MGASNSV.LSGRKLDAFESVRLRPNNGKKYKLRHLVWASKELDRFSLSANLLETKEGVVKILSVLLPLVPTGSENIALFNLCCVLACIHAEIKVKDTTEAKAKVKEVPAEMTESATATSSG...QTK	125
H1B.FR.83.HXB2	---RA---GE--RW-KI---G-----K-I---R-E--AVNPG---S--CRQ--GQ-Q-SLQ----E-RS-Y-TVAT-Y-V-QR-EI---K--LD-IE--Q...NK-KKKAQQA...AAD	121
CPZ.CD.-.ANT	---GA---R-E---TW--I---G----MIK-----RS--Q--A--SS---S--CE-AIHQ-S-SIEIR-PEI-S---TI--W-V-KGE-I---Q-VKT--MK...MQTQAE-G-SQTASRGM	125
CPZ.GA.-.CPZGAB	---RA---T-G---RW-K---G-R-R-MMK-----R-E--ACDPG-M-S---CT-L-QQ-E-ALK---G-RS---TLA--W---SD-T-E--QK-LEQL.....RHHGEQQ-KTESNSGS	121
H2A.DE.-.BEN	---R---R-K-A-EL-K---G----R-K-I---AN--K-G-AES---S--CQ--R--D-----KS--TV--IW-L---E-----KLA.QRHLVAE.....TGT	114
COL.CM.-.CGU1	--NEQGL--GKKT-EDLQK---KKGK-GC--IK-VR-MCT-VS-CV-IFE--KSAT--AQ--EKVT--D---V-RS-YG--S-CY-L-RKWNI--Q--EK--E--AYKKQAMIEMA.....	116
DEB.CM.99.CM40	---GR-AP-T-AA--R--K---K---MIK-----G-----G-A-A--N---CQE--TR-Q--EA-----KS-YGIV--IWAC--NLS-D-----KA-YRIKQVEQ-EELEMALIQRQKEK.	129
DEB.CM.99.CM5	---GR-AP-T-AA--R--KI---K---LIK-----R-----G-AGA-----CQE--TR-Q--EAS---G-KS-YGIV--VWAC--NLS-E-----KA-ARRIKQAEQ-EELEMALIQKKAEK.	129
DEN.CD.-.CD1	--SGS--PMT-AA--R--K---GS---MIK--I-SK--ME--G-GEQ---Y--CQT-IET-A--E-Q---S-KS--RT--VW-C-RGLE-A-----LRE-RKRIG-KKKNIVIGNADQGDDQPPQ	130
DRL.-.FAO	---A-G--R-E---K---S-----K---VA-----G-HER--SQ--CQ--G--F-----S---T---VW-V-SKV--T---VQ--QHCHLVDKNENA-SKNENGETA-S	129
GRV.ET.-.GRI_677	--GGH-A-----S--T--KI-----QIK--I--G--ME--G-HEK-----CQ--IE--T--E-----G-K-----IW---Q-----VVT--QHYHLVDKNEKA.....	118
GSN.CM.99.CN166	---RH-AM-T-T---RY-K---K---MIK-----E-YA--DA--NQ--CRR--EIVF--E---A-KS--GIVS--Y---D-Q-E--Q--QQ-RIRCHL-E-KGKEKQNKD.....	123
GSN.CM.99.CN71	---RH-AM-T-T---RY-K-S---R---ATK-----E-YA--GS--S---CQQV-E-TS--E-N---S-KS-YGIVS--Y---D-W-E--Q--KQ-QIRCHLVG-KEKETQNKD.....	123
LST.CD.88.447	--SG---RQIEGE-C-I---DS--T-QK-----TR-----G-G-H---AD-CK--G-CW--YS---K-K--VGTV--C-C-LG--A--Q--MQ-I-I-PTQ-ERQ.....	114
LST.CD.88.485	--SG---RQIEGE-C-I---GS--T-QK--I---T-----G-G-H---AD-CKR--G-CW--YS---K-K--VGTV--C-C-LG--IAN-A--MQ-I-I-PTQ-DRQ.....	114
LST.CD.88.524	--SG---RQIEKD-CN---GS--T-QK--VD---T-----G-G---A-CK--G-RW--Y---K-K--VGTV--IC-C-LG-R-N--Q--IN--I-QNKPV.....	112
LST.KE.-.lho7	--SG---RQIEKD-C---GS--T-QK--VE--T-----G-GSQ---A-CK-----CW--YA---K-K--VGTV--IC-C-LGVRIS--Q--IS---I-PAP-AAG.....	114
MAC.US.-.239	--VR---K-A-EL-KI-----M-K-V---AN---G-AES--N---CQ-----A---KS-Y-TV--IW---E--H---QI-.QRHLVVE.....TGT	114
MND-1.GA.-.MNDGB1	--NG--A--L-TD--K--KI--KRG---C-R-K--C-CKG---G--DK---QQ-CE---CW--YDQ--D--K--VGTV--V---G-EI-S-QD-LK-L-VITRK-EKQEDE.....	117
MND-2.-.5440	---A-G--R-E---EL-KI---S---Q-K-VI-V-----G-HEK--SQ--CE---F-----S-Y-T--CIW-V--KV--T---E---QRYHLVVERENA-SEEEKGATA-P	129
MND-2.CM.98.CM16	---A-G--R-E---EL-KI---S---Q-K-II-V-----G-HEK--S---CE---F-----S-Y-T--CVW-V--KE--T---E---QKLHLVAEKENA-SEKEQRAIV-P	129
MND-2.GA.-.M14	---A-G--R-E---EL-TI---S---Q-K-II-V-----G-HEK--Q--CE---F-----L--Y-T--CIW-V--KV--T---K--EQCFHLAAKGESA-SEKBEKATA-P	129
MON.CM.99.L1	---RH-AM--T---KY-K---R---LIK-I--A---G--DS---QD-CK--E-I--Q---SIKS--GIAS--Y---G-EIE--Q--QQ--IRCHLAGEQGEQKAAAAA.....	125
MON.NG.-.NG1	0
MNE.US.-.MNE027	---R---K-A-EL-KI---G---M-K-V---AN---G-ADS--N---CQ---A-A-----KS-Y-TV--IW---E--H---QI-.QRHLVVE.....TGT	114
MUS.CM.01.1085	---RH-AM-T-N---RY-K---K---LIK-II--R---G--DS---SRD-CQ--IE-I--E-A--S-KS--GIVS--W---VH-E---KQIRARCHL-EKKGEKESQAQAA.....	125
RCM.GA.-.GAB1	---RA-L--K---W---G---M-K---C-K-NK-G--DH---AT-CE--G-----G-KS-----W-V-K-V-----V---CCHLVEKAEN-TEKEK...GAT	125
RCM.NG.-.NG411	---RA-I--K---W-Q---G---M-K---CR-E-G--DT--SA--IQ--IG-I-----G-KS-----W-V-K-V-----V-H--KQCHLVDRDENAGEQEK...GAT	125
SMM.SL.92.SL92B	---RG---K-A-EL-K---G-R--M-K-II--AR---GSAES--S---CQR--A--A--M-----KS--STV--VW-L--M-----KT-.QSHLVVE.....SGT	114
SMM.US.-.H9	---VR---K-A-EL-KI---G---M-K-I---AN---G-AES--N---CQ---A-----KS-Y-TVR--W---E--H---QI-.QRHLVVE.....TGT	114
SMM.US.-.PGM53	---R---K-A-EL-KI---G---M-K-I---AN---G-AES--S---CQ--I--A---A---KS-Y-TV--IW---E--H---QI-.QRHLVVE.....TGT	114
STM.US.-.STM	---RS---K-A-EL-K---G---M-K-V---AN---G-AES--S---CQ--IT--E-----KS--TV--IW---E--H---QV-.KRHLVVE.....TGT	114
SUN.GA.98.L14	---G--T.VDREVVRS--R-A-K-G---T-QA--V--G-----G-NKE--R-V--CQ-----CW--YAS--K-K--VGTV---C-QG-P---Q--LK--RL-PAKNEE.....	113
SYK.KE.-.KE51	---GG-AI-T---KY-KI---K---R-LV---G-----G--DQ--S---CEQ-IKTI--EKH---KS--GITA-VWAV--KD-E--Q--Q---ACNW-DE-TV-SSGQK...ENS	126
SYK.KE.-.SYK173	---AG-AI-T--E--RY-KI---K---R-LV---K-----G--DQ-M-S---CE--T---EAN---KS--GIIS-VWAV--KKE-E--Q--Q---ACNWKDDPP--SGGQS...ENS	126
TAN.UG.-.TAN1	---GH-A-----N--T--KI-----Q-K-I--G--KME--G-HEK-----CQ--IE--S--E-----G-KS-----IW---Q-----VVL--QRCHLVEK-KTA-AP--...GQQ	125
VER.DE.-.AGM3	---AT-A--NR-Q--K--HI---T---QIK--I--G--ME--G-HER--SE-CK--IE--Y--E-----G-KS--V---F-V-KDKE-----V-I-RQCCHLVEK-RNAERNTT...E-S	125
VER.KE.-.9063	---AT-A--K--Q--KI---G---QIK--I--G--ME--G-HDR--ED-CK--IE--F--E-----G-KS--V---Y--R-Q-----V-T-RQCCHLVEK-KRATEP-S...GE	125
VER.KE.-.AGM155	---AT-A--NR-Q--E--HI-----QIK--I--G--KM--G-HEK--E-CK--IE--S--E-----GMKS--Y--V---L-V-Q-K-----L-I-RQCCHLVK-KTAV-PP...GQQ	125
VER.KE.-.TYO1	---AT-A--NR-Q--Q--KI-----QIK--I--G--ME--G-HER--E-CKR--IE--Y--E-----G-KS--V---Y-L-K-Q-----V-T-RQHCHLVEK-KSATET-S...GQ-	125

Gag p17 Matrix \ / p24 Capsid Cyclophilin A binding site |

```

SAB.SN.-.SAB1C .....ELQAKKKNEPTV.....TPS.GGS.RNY.PIVSVN.NQVWHQPLSPRTLNAWVKVIEEKKFAEVVPMFSAALAEIPAIDYNQMLNAVGEHQGALQIVKDVINEEAADWDLRHP.PPQQ 234
H1B.FR.83.HXB2 .....TG.....HSN.QV-.Q---.QNIQG-M---AI-----V---A---P---I-----S---T-Q-L-T---T-G--A-M-ML-ET-----E--RV--.VHAG 221
CPZ.CD.-.ANT .....LLRLLLLNK-.WCQR.....HLS-EG.---.IVDAGGIAR---T-----C---N-NP--I-----S---T-H-L-T---D---M-VL-E-----E--RL---.THAG 237
CPZ.GA.-.CPZGAB .....REGGASQGAS.....S.....AG.I-.G---.L-QNAQG-M---AI-----V---A---P---I-----S---L-Q-V-T---G---M-VL-E-----E--RL---.THAG 229
H2A.DE.-.BEN .....AEKMPNTS.....RPTAP--.KR.G---.VQQAG.GNY-V-----LV---G---G-Q-S-CT-----C-D--A-M--IREI-----SQ--.I-.G 222
COL.CM.-.CGU1 .....SKEEBEKKAKEAEKCLDM--TGP.QGP-----G---CV-.GGIAPSLA--L-YST--A--M-L---ILD-T--FP-VL--E--KK-EEY--L--.V-- 218
DEB.CM.99.CM40 .....EVEQKQQKQEQRQOEPOAAPOTTSPV---LRQG.Q-F-M-----IVKT-INAV-----P-I--L-QV---CT---GL---I-DL---M-I-----E---Q---.QQ-P 246
DEB.CM.99.CM5 .....ELE-K.QKTQAAAADTQQKTNQ..PA-F---LRQG.Q-Y-M-----IVKT-INAV-----P-I--L-QV---CT---GL---I-DL---M-I-----E---Q---.QQ-P 242
DEN.CD.-.CD1 .....G.....AAGGGSSAEQGASAGGSSFRDPN-.Q---.RTAQG-YQ--A--T-V-KT-I-EV-DR---P-I--L-V---L---L---I-----M-VL-----H---A--ANL-- 248
DRL.-.-.FAO .....SGR.....-.-.QV--.Q-A---AI-----L--S-C-S-----I-D-----I-----R---.QVGP 224
GRV.ET.-.GRI_677 .....KKKNETTA.....P-G-.E---.V-NQ--.A-----CV---RWG-----Q-S-CLS--V---VI-D-----L-E-----E--RT-R--.AG 223
GSN.CM.99.CM166 .....PPGAAGGQAVV-.Q---.VIRNAQG-YQ--A-NS-I-KT--SIV-----AP-T-AL-Q-S-C---M-----I-DY--V--I-----Q-E---L--.V-AA 227
GSN.CM.99.CM71 .....PPGAAGGQAVV-.Q---.RNAQG-FQ--A-NS-V-KT--SIV-----AP-T-AL-Q-T-C---M-----DY--V--I-----Q-E---L--.Q-AA 227
LST.CD.88.447 .....PK-.G---.L-RE-.QR---T-----IQT--IV-DRGWKP-T-A-----T-K-L-D-L-V---I-D---M-I--H-V--G-E--RK--.QQ-P 208
LST.CD.88.485 .....QK-.G---.LIRE-.QR---T-----IQT--IV-DRGWKP-T-A-----T-K-L-D-L-V---I-D--V-M-I--H-V--G-E--RT--.QQ-P 208
LST.CD.88.524 .....PE-.K-F-.LIRE-.QR---T-----IQT--IV-DRGWKP-T-A-----T-K-L-E-L-V---I-D---M-I--H-A--G-E--RQ--.QQ-P 206
LST.KE.-.lho7 .....KKQ.QTG.G---.LIRE-.QR---T-----IQT--IV-DRGWKP-T-A-----T-K-L-D-L-V---I-D---M-I--H-V--G-E--RQ--.QQ-P 211
MAC.US.-.239 .....TETMP-TS.....RPTAPS-.RG.G---.VQQIG.GNY-L-----L-----G---G-Q-S-CT-----C-D--A-M--IR-I-----Q--.Q-A. 222
MND-1.GA.-.MNDGB1 .....-K-F-.VQRDAAG-YQYT-I--IIQT--TV---WKP--I-L---T---SH-L-I-----D---M-VL-----Q-E---T---.QQ-P 210
MND-2.-.-.5440 .....AVR.....-K---.QVI-.QTP--GI-----C---P-I---I-S-C---L-G---I-D-----G-----VGP 224
MND-2.CM.98.CM16 .....SGR.....-K---.QII-.QTP--GI-----C---P-I---I-S-CL--L-G---I-D-----N---QVGP 224
MND-2.GA.-.M14 .....SGR.....-K---.QII-.QTP--GI-----C---P-I---I-S-CL--L-G---I-D-----I-----QVGP 224
MON.CM.99.L1 .....P-TG-VPSP--.V-RTQGGGPQ--AVE--L-KT-Q-----AP--AL-Q-S-M-----L--I-D-----I-----N--V--.Q-P- 225
MON.NG.-.NG1 .....0 0
MNE.US.-.MNE027 .....AETMP-TSR--A.....PS-.KR.G---.VQQIG.GNYT-L-----L-----G---G-Q-S-C-----C---A-M--IREI-----Q--.QQA. 222
MUS.CM.01.1085 .....-A---.VIRNAQG-FQ--A-NA-I-KT--SIV-----A---A-Q-----L--I---I-I-----Q-E---LN--.Q-- 218
RCM.GA.-.GAB1 .....AP--.QR.G---.ITI-.Q-PE-N-I-----V-----A---S-C-----I-----E-----A---.VPG 222
RCM.NG.-.NG411 .....VT.....S-.QR.G---.TI-.Q-PE--I-----V-----G-----S-C---V-----I-----E---D-----NPG 222
SMM.SL.92.SL92B .....AEKLPQAS.....RPTAP--.G.G---.VQQ-G-.NY-T-----LV---G---G-Q-S-CT-----C---A-M--IREI-----Q--RQG-P 222
SMM.US.-.H9 .....ADKMPATS.....RPTAP--.RG.G---.VQQ-G.GNYT-L-----LV---G---G-Q-S-CT-X-----C---A-M--IREI-----Q-X-Q-.G 222
SMM.US.-.PGM53 .....ADKMPITS.....RPTAP--.RG.G---.VQQIG.GNYT-L-----LV---G---G-Q-S-CT-----C---A-M--IREI-----Q--Q-.G 222
STM.US.-.STM .....ANKMPATS.....RPTAP--.RG.G---.VQQ-G.GNY-L-----LV---G---G-Q-S-CT-----L-C---A-M--I-EI-----MQ--.Q-PG 223
SUN.GA.98.L14 .....A.QA-.Q-F-.VQREG.QNYI-----VQT--IV--GWKP-T-A-A--MT--E--V---I-----M-MI--H-V--S-E--RM--.QQ-P 208
SYK.KE.-.KE51 .....NDT-TSSGREG.KMQLPAAMPSSG..GSG--.LIRNPQ---I-VGVNT--KT--EAVNS--D-SL--L-QI-T-F---L-D---I-D---M-VI-----GSE--Q--.Q-- 243
SYK.KE.-.SYK173 .....SQNMASETSSGQKVQEQEKQAAT.PPPRG--.LLRNPQ---I-TGVVPV--KT--EAVNS--D-SI--L-Q-T-F---L-G---D---M-VI--I---G-A--E--.Q-- 246
TAN.UG.-.TAN1 .....QNYNTAA.....-RH.G---.V-QQ-.-----T-----TV--R-G-I-----Q-S-CLS-----VI-D---M-I-E--D---Q--IT--.AG 229
VER.DE.-.AGM3 .....SG-K-MDKGV--.....-P---.Q-F-.AQOQG-.A-I-V-----AV---G-I---Q-S-CT-----VL-D-----EI-----Q-IA--.AG 233
VER.KE.-.9063 .....KSNRETTA.....P-G--.Q-F-.AQOQG-.A-V-----AV---G-I---Q-S-CT-----VL-D-----EI-----Q-VT--.AG 229
VER.KE.-.AGM155 .....KNNTGGTA.....-G--.Q-F-.AQOQG-.A--V-----AV---G-I---Q-S-CT-----VL-D-----EI-----Q-VT--.AG 229
VER.KE.-.TYO1 .....KNDKGIAA.....P-G-.C-.Q-F-.AQOQG-.A--V-----AV---G-I---Q-S-CT-----VL-D-----EI-----Q-VT--.L-AG 229

```

SAB.SN.-.SAB1C	PPAQQVLRDPQGSDIAGTTSTIQEQIEWTTRA...QNAVNVGNIYKGWIIILGQKCVKMYNPNVILDIKQGPKEPKDYVDRFYKALRAEQTDPAVKNWMTQSLLIQNANPDCKTVLKGLGMNPTLEEM	360
H1B.FR.83.HXB2	-I-P-QM-E-R-----L---G-M-N...NPPIP-E---R-----N-I-R--S-TS---R-----R-----T---ASQE-----ET--V-----I--A--PAA-----	346
CPZ.CD.-.ANT	-VQA-Q--E-T-----V---MQ-MSTP...QQGG-P--D---R---M--N-V-RX-S--S-E-----R-----TI---ASQP--A---ET-----HI--A--TGAS-----	365
CPZ.GA.-.CPZGAB	-I-P-Q--E-R-----L---G---A...NPPIP-DV-RR-V---N-V-R-C-S---R-----R-----T---ASQE-----DT--V-----QI--A--PGA-----	354
H2A.DE.-.BEN	-LPA-Q---R-----VD---Q-MY-P...P-P---RR--QI---RK--T---QS---S-----T-----L-----	348
COL.CM.-.CGU1	-QQ--A--Q-TA---T-N--SVA--VA-GE...PIA---R---VQS-E-VIQUIAR-SSV---R--S--D--S-----S---PAAGEI-A--ANN---H---RI---QK.-S--D-	337
DEB.CM.99.CM40	.P-Q--E-S-A---N--VE---A-M--PAGQG-GPID-Q--RR-V---R-----T---V-----T---A-Q---TT-M---RII---Q-----	374
DEB.CM.99.CM5	.P-Q--E-S-A---N--VE---A-M--PAGQGAGPID-Q--RR-V---R-----T---V-----T---A-Q---TT-M---RII---Q-----	370
DEN.CD.-.CD1	-LPPAAQ---R-----S-ED--T-M--Q...GPI--D---R-V-M---V---Q--S-----S-----A-QS--L--T--V-----QI-EA--P-----	374
DRL.-.-.FAO	.LP---N-S-----S-E-----...DSI--A--RQ-VV---R-T---V-----H--A-----T-----VI-----S---	349
GRV.ET.-.GRI_677	-LPA-Q---T-----S-----FN...NPRID--AQ-RK-V---V-Q---QKV--R-----Q-----APQD-----T-----LI-----	349
GSN.CM.99.CN166	-QPVAG---S-A---V---PN---I--Q...NQP--SD--RK-----R-----E---F-C---A-Q---V---E---LI--AMPG.AN---	352
GSN.CM.99.CN71	-QPVAG---S-A---V---PN---I--Q...N-P---AD--RK-----R-----G-----E---F-C---A-Q---L--V---E---LI--AMPG.AS---	352
LST.CD.88.447	AQPG-G--T-NAT-V--V---VE--LA--AD...TP-D--K---E-V-QAME-V-RIHQ--SVI-----E-A--F---GGSH-E--E--KEKM-V-----RM-I-A--EGA-----	333
LST.CD.88.485	AQPG-G--T-NAT-V--V---VE--LA--AD...TP-D--K---E-V-QAME-V-RIHQ--SVM-----E-A--F---GGSH-E--E--KEKM-V-----RM-I-A--EGAS--D-	333
LST.CD.88.524	AQPG-G--T-NAT-V--I---VE--QA--AD...TPID--K---E-V-QAME-V-RIHQ--SVM-----A--F---GGSH-E--E--KEKM-V-----M-I-A--EGA-----	331
LST.KE.-.lho7	AQPG-G--T-NAT-V--V---VE--LA--AD...TP-D--K---E-V-QAME-V-RIHQ--SVM-----E-A--F---GGSH-E--E--KEKM-V-----RL-I-A--EGAS---	336
MAC.US.-.239	.Q-Q--E-S-----SVD---Q-MY-Q...PIP---RR--Q---R---T---V-----QS---S-----A-----T-----L-----V-----	347
MND-1.GA.-.MNDGB1	AQPG-G--T-S-----VE--LA-MNMQ...I--T---S---MNRL--SHC-IS-T-VR---A-----NVM---ASGE--M--Q-H--E--E--QI-RS--KGA-----	336
MND-2.-.-.5440	.MP---N-T-N-----S-E-----Q...DQ---G---Q--V-----S-----R---Q--T---T-----AT-----	349
MND-2.CM.98.CM16	.MP---N-T-----S-E-----E...D---G---Q--V-----S-----R---Q--T---T-----SI-----S---	349
MND-2.GA.-.M14	.LP---N-T-----S-E-----Q...EQ---A--Q--V-----S-----R---Q--T---T-----SI-----S---	349
MON.CM.99.L1	-QPNAG-G--T-A---VS--P-Q---I---...N-P-Q-SD--RK-V---R-----C-----Q-----T--V-----LI--AMPG.AS--D-	350
MON.NG.-.NG1LI--AMPG.AS---	19
MNE.US.-.MNE027	.Q-Q--E-S-----VD---Q-MY-Q...PIP---RR--Q---T---R---QS---S-----T-----L-----	347
MUS.CM.01.1085	-QPNAG--N-T-----VS--P--I---...N-PI--AE--K-V-M--R-----F-C---S-Q---ST--V---E--LI--SMPG.A---	343
RCM.GA.-.GAB1	-IPA-Q--E-T-----S-A--A---...N-PI--L-RN--V--W-----S-----A-----M-----S---	348
RCM.NG.-.NG411	-LPA-Q--E-T-----N-A--A---...N-PIA---RN--V-----A-----	348
SMM.SL.92.SL92B	AQPA-G--E-S-----PS---MY-Q...P-P--D--RR--Q---R---T---V-----QS---S-----T-----L-----	348
SMM.US.-.H9	-IPA-Q--E-X-----XD---Q-MX-Q...PIP---XR--Q---R---T---V---S-QS---S-----X---T-----L---XI-----	348
SMM.US.-.PGM53	-IPA-Q---R-----VE--Q-MY-Q...PIP---RR--Q---R---T---V---S-QS---S-----T-----L-----	348
STM.US.-.STM	-LPA-Q--E-S-----SPE---Q-MY-Q...PIP---RR--Q---R---T---V---T-QS---S---A--S---RT-----L-----	349
SUN.GA.98.L14	AQPG-G--T-N-T---V---VE---Q-IYGL..NGATR-A-QD---V-E-MERV-RLHQTTSVVE-R---R--T--F---SGSEE--E--KEKM-V-----L-I-A--ET-S---	336
SYK.KE.-.KE51	QQPVAG---SA-----I--Q...N-PIQ--Q--RQ---QV-----R---T---N---HC---A-----L-----E-RSI--AMVK.---	368
SYK.KE.-.SYK173	-AP-AG---SA-----S-A---I--Q...N-P-Q--E--RR-----QV---S-----I---HC---A--S--G-L--T-----E-RQI--AMVK.---	371
TAN.UG.-.TAN1	-LPA-Q---R-----SVA-----FN...NPR-D--R--R--V-----ISV--R--A-----Q-----PQD-----ET-----L-----IH-----	355
VER.DE.-.AGM3	-LPA-Q---R-----V---L--IYT...NPR-D--A--RR-----SV--R---A-----I---ASGE--Q--E-----VI---H-----	359
VER.KE.-.9063	-LPP-Q---R-----V---L--IYT...NPR-D--A--RR-----SV--R---I---ASGE--Q--E-----VI---H-----	355
VER.KE.-.AGM155	-LPA-Q---G-----V---L--IYT...NPR-D--A--RR-----SV--R---I---ASGE--Q--E-----VI---H-----	355
VER.KE.-.TYO1	-LPA-Q---R-----SV---L--IYT...NPR-D--A--RR-----SV--R---I---ASGE--Q--E-----VI---H-----	355

	Gag p24 Capsid \ / p2	p2 \ / p7 NC	
SAB.SN.-.SAB1C	LTACQIGGAQHKA	RMALMAEAMTAAFQQ	.QTV..GNIFVQQGAR.PRGPLGGRGRPL.....NP...NIKCYNCGKPGHLARFC.KAPRRQ.....GCWKCSPDHQMKDC..Q 453
H1B.FR.83.HXB2	M-----V--PG---VL---SQV	TNS.A.....MM-R-NF.RNQRK...IV--F---E--T--N-.R---KK.....KEG-----T 427
CPZ.CD.-.ANT	-----V--PA---VL---AS-N	NAOQT.....AV-L-R-NG-N-GKR..PL--F---E--T--N-.K.....R--QEG--L-N...P 449
CPZ.GA.-.CPZGAB	M-----V--PS---VL---SMV	QNGRA.....DV-F-K-QG.AGPKR..K--F---E---N-.K.....R--QEG-----T 437
H2A.DE.-.BEN	-----V--PGQ-----LKE	-MGP.SPI..PFAAARK...A-RYW---E--S--Q-.R-----K-G-I-AN...P 426
COL.CM.-.CGU1	-A-----V--PD---KVL---Q	FQE.....RTNMI	EVKTA.....-F--QGI-----M-P-R-IGGAGRGRGRGGFRGAPRRPVR-FT-NQEG-MQR---P 432
DEB.CM.99.CM40	-H-----V--P-Q-----AS	-LKE.....AGSLGM---RRGPPGSRR..Q-R-F---QI--QKD-.R--KT.....K-F---QEG-IA-N... 457
DEB.CM.99.CM5	-H-----V--P-Q-----AS	-LKE.....AGSLGM---RRGQGGGPRR..QLR-F---QI--VQ-D-.K--KV.....K-F---QEG-IA-N... 454
DEN.CD.-.CD1	-Q-----V--PGQ---PT---L	AS---K.QGLCHGRKGA-PPVER.RGLR-F---QI---KD-.RK-K-V.....TQPGA-F---KLG--A-N-.RS 463
DRL.-.-.FAO	-L-----V--PG---V---RE	QQA.IMMQN..PPRGPP-G---PPK...PR-P---QF--TL-Q-.TK--KK.....FR--AL--MLRN...P 436
GRV.ET.-.GRI_677	-I-----V--P---K--V-M	-SNGQ.....NM--V-PQ.KKRG...PL--F---F--MQ-E-.---QI.....K-F---KIG-MA---K 427
GSN.CM.99.CN166	-----V--PT--S--L---AT	-LKG.TSS..YNMK--P-R.....Q GK..TP-----QF---D-.PK-KER.....K-F--RAG-FS-P-R.T 436
GSN.CM.99.CN71	-----V--PT--S--L---AT	-LKG..SS..YNM-P-R.....Q GK..TP-----F--I--D-.PK-KER.....K-F--KAG-LARQ-K.T 435
LST.CD.88.447	MR-----V--PA--G-IL---AT	-M-S..MR..QNM--.VTP	--NAQ-RFV-TGGGG..PRK..PLT-F-----T--M-.RQ--QE.....N--KE-RFAQ...P 428
LST.CD.88.485	MK-----V--PA--GKIL---AT	-M-S..MR..QNM--.VTP	--NEQ-RFV-TGGGG..PRK..PLT-F-----T--M-.RQ--QE.....N--KE-RFAQ...P 428
LST.CD.88.524	MK-----V--PA--RKIL---AT	-M-G..MK..HLM--.QTP	--NAQ-RFV-TGGGG..PRR..PLT-F-----S--M-.RQ--QE.....N--AK-R-AQ...P 426
LST.KE.-.lho7	MK-----V--PA--GKIL---AS	-I---VR..QNM--.VTP	.L-NAQ-RFV-TGGGG..PRK..PLT-F-----T--M-.RQ--QE.....N--KE-RFAQ...P 431
MAC.US.-.239	-----V--PGQ-----LKE	-LAP.VPI..PFAAA-RGPRK..P--W---E--S--Q-.R-----KM--V-AK...P 428
MND-1.GA.-.MNDGB1	-E-----V--P-----M	-RTVVG..S.....QN---Q	RGQP---VRQPTG.....R..KPI-F--N-E--V--F.----K.....N--AM--KAQ...P 423
MND-2.-.-.5440	-L-----V--PKY--QM-----Q	QAA.VMMQNSGGPPRGPP-Q-PR...PR-P---F--VL-D-.R---KR.....F--D-G-L-RN...P 434
MND-2.CM.98.CM16	PL-----V--PKY--QM-----K	ESA.VM..MQNSGGPPRGPP-Q-PR...R-P---F--GL-D-.IS--KK.....F--DLG-I-RN...P 434
MND-2.GA.-.M14	-L-----V--PKY--QM-----Q	KMQSE.VMMQNSGGPPRGPP-Q-PR...PR-P---RY--VL-D-.RL--KK.....F--DTG-M-RN...P 434
MON.CM.99.L1	-Q-----V--PM--S--IL---AG	-IA..NMP..MNM--.ARGPPQRKG.QP--F---F--M-KN-.---Q-R.....K-YN--Q-G-LA---P 432
MON.NG.-.NG1	-Q-----V--PA--S--L---AT	-INS.NMP..MNMGRGGXQPRRQGXQ-R-----F--V-KN-.T--KT.....FR--KEG-XS-N...P 105
MNE.US.-.MNE027	-----V--PGQ--K-----LKE	-LAP.GPL..PFAAA-KGPRK..P--W---E--S--Q-.R-----QMG-V-AK...P 428
MUS.CM.01.1085	-Q-----V--PT--S-----L	AS--K..AGT..LNM-AKGARGRNGSQGNRG-PR-----QY--V--D-.PK--NK.....K-F---REG-LARQ... 433
RCM.GA.-.GAB1	-----V--P-----V	L---QMQSN.....MA--S-N	.RGP-RRSG.....NPN...LR-----IS-Y-.----K.....LL---T 433
RCM.NG.-.NG411	-----A--P-----V	L---QMVQSN.....MA--G	PRKGP-KL-G.....PR..FL-----T--T--Y-.TS-KK.....R--EEG-L...P 434
SMM.SL.92.SL92B	-----V--PGQ-----K	D-LTGSLVAAQFRGAAGK-KP..I-R-F---T--S--Q-.R---K.....EEGRIQAN...P 432
SMM.US.-.H9	-----VX-PGQ--X-----L	KD-LT.GPL.PFXA--KGQXK..I-X-X--X-E--S--QF.R---X.....KAG-V-AK...P 429
SMM.US.-.PGM53	-----V--PGQ-----LKE	-LTP.GQL.PFAA--KGPRK..T--W---E--S--Q-.R-----KTG-V-AK...P 429
STM.US.-.STM	-----V--PGQ-----LKE	V--P.DPL.PFAAA-QGRR..TV--W---A--T-KQ-.G---.....QQG--AK...P 429
SUN.GA.98.L14	MR-----V--PS--GKIL---AS	-R--VG..RQAM--NLP	--NSQ-RFV-IGGGG..PRK..PMT-F-----NQ-.RE-KKG.....PPGS----KMG-KQAQ...P 436
SYK.KE.-.KE51	-Q-----V--P-Y--K--L---V	MTQHS.....LGM1--PQGSNPRR.GPTR-F---QL--QKD-.PR-KKL.....K-FN--GTG-IARQ-RQP 452
SYK.KE.-.SYK173	-Q-----V--PL--K--K---V	M-Q-S.V.....NM--G	PSK.G.....RS..M-----QI--MQKD-.K-LKA.....K-FN--KTG-LARA-RQP 451
TAN.UG.-.TAN1	-----V--PG--K--V---Q	MQGV.....NM--GAP	.GGRGR.....RG...PPR-FK--QI--IQKD-.PRAGPN.....K-L--K-G-LA---R 437
VER.DE.-.AGM3	-----V--PSY--KV---M	-QMQS.....NMM--G	RGR.....R...PV-----F--MQ-Q-.PE--KM.....R-L--K-G-LA---R 439
VER.KE.-.9063	-----V--PSY--V---M	-QMQT.....NM--A	GGVRRQ.....R..PL-----F--MQ-Q-.TE--KI.....K-L--KLG-LA---R 437
VER.KE.-.AGM155	-----V--PSY--KV---M	-QLQS.....NM--G	RGR.....R..PP-----F--MQ-Q-.PE--KI.....K-L--K-G-LA---R 435
VER.KE.-.TYO1	-----V--PSY--KV---M	-QTMQN.....NM--G	P-K-Q.....R..PLR-----F--MQ-Q-.PE--KT.....K-L--KLG-LA---R 434

	p7 \ p1	p1 \ p6	PTAP motif	Vpr binding site	
SAB.SN.-.SAB1CKQVNFLGF.GPW.GRGGK.P.RNFPLTISI.....RPTAPPMERDYSRPEENWYADRPPTRGPGPDDPATALLKQYAVQGRKQKQWQNHSPQQSPYEEAYSSLSRSLFGEDQ\$RPTAPPMERDYSRPEENWYADRPPTRGPGPDDPATALLKQYAVQGRKQKQWQNHSPQQSPYEEAYSSLSRSLFGEDQ\$KELYP.LT-----N-PSSQ\$	554
H1B.FR.83.HXB2ER-A---K.IWPSYK-R-.G--LQS.RP.....E----E-SFR-GV-TTTPPQ....KQEP.....E----E-SFR-GV-TTTPPQ....KQEP.....ID.....KELYP.LT-----N-PSSQ\$	500
CPZ.CD.-.ANT	AT....NTGK----KPT-TWNGCR-.G--VQKKEEVV.....E----I-IYQEEHKR.....TQK.....E----I-IYQEEHKR.....TQK.....LKG.....EELPP.SY--K---K----	522
CPZ.GA.-.CPZGABGR-----K.-WPSRS-R-.G--VQNR.....E----I-SYGYQE.....EEKSQE.....E----I-SYGYQE.....EEKSQE.....KKEGE.....S-L-PP..T--K----S-PSSQ\$	508
H2A.DE.-.BENER-AG--L.-RG-.K-----V-QAQQL.....I----ADPAEELL.....ERYMQQGRKQREQRERPYKEVTEDLLHLE-RETPhREETEDLLHL..N---K----I----ADPAEELL.....ERYMQQGRKQREQRERPYKEVTEDLLHLE-RETPhREETEDLLHL..N---K----N---K----	521
COL.CM.-.CGU1N--A---A.....TRGV.....ELQ--IFPPKMSKDL.....ELQ--IFPPKMSKDL.....P-RREKG-SL-P--K----D----	482
DEB.CM.99.CM40GQTPRA---N.T--AS.RR.-.-LEGIGKY.....I-ESTADQVYSTTRVVPVQGENN-Q-EETTRRGKYPD-MGLKSLFGEDQ\$I-ESTADQVYSTTRVVPVQGENN-Q-EETTRRGKYPD-MGLKSLFGEDQ\$I-ESTADQVYSTTRVVPVQGENN-Q-EETTRRGKYPD-MGLKSLFGEDQ\$	535
DEB.CM.99.CM5GQVPR---N.A--AS.RR.-.-LEGIGKY.....I-ESTADQVYSTTRVVPVQGERNA-QEDTTTRRGKYPD-MGLKSLFGEDQ\$I-ESTADQVYSTTRVVPVQGERNA-QEDTTTRRGKYPD-MGLKSLFGEDQ\$I-ESTADQVYSTTRVVPVQGERNA-QEDTTTRRGKYPD-MGLKSLFGEDQ\$	532
DEN.CD.-.CD1	N....TNDQG-----N.-LGRK.NP.-.-VQYYPQMNQD.....L-WS-VRVQQPQY.....-QNFQE-GNRQSCQED-RQN..KQ..SHQEDEETTRESLYPSL..K---D----L-WS-VRVQQPQY.....-QNFQE-GNRQSCQED-RQN..KQ..SHQEDEETTRESLYPSL..K---D----K---D----	558
DRL.-.-FAO	K.....-----N.T--GS.-.-QA-L.....T-S--LPGYLQED.....-AE-MLQKYMEOG-QQ-RQQ..RQQ.-.-KRGPY..EEAYNSL..S---S--LQ\$T-S--LPGYLQED.....-AE-MLQKYMEOG-QQ-RQQ..RQQ.-.-KRGPY..EEAYNSL..S---S--LQ\$-AE-MLQKYMEOG-QQ-RQQ..RQQ.-.-KRGPY..EEAYNSL..S---S--LQ\$	521
GRV.ET.-.GRI_677	N.....G-A---Y.-H-G-.A---VQYRGDTVGL.....E-----TA-DPA.....E-----TA-DPA.....KK-L-QYAEKQRLREERBQTRK-KEK-VEDV--S---G----	513
GSN.CM.99.CN166GTPI-ELY--DL.-RGR.RA.-.-V-FL.....L---AE-EEPOQL.....NLSTPVGEP-SKGQKEESRNTL-P--S-----L---AE-EEPOQL.....NLSTPVGEP-SKGQKEESRNTL-P--S-----NLSTPVGEP-SKGQKEESRNTL-P--S-----	512
GSN.CM.99.CN71GTPA-EL--DL.-RGR.RA.-.-V-FL.....L-S--AE-VEPQQL.....NLSTPVGEP-SKGQKEESRNTL-P--S-----L-S--AE-VEPQQL.....NLSTPVGEP-SKGQKEESRNTL-P--S-----NLSTPVGEP-SKGQKEESRNTL-P--S-----	510
LST.CD.88.447	..K...PKGK----Y.-.-RS.-P-.G---MGGNAGV.....V-S-----SPTQA-RALETY-NLQQLRR.....V-S-----SPTQA-RALETY-NLQQLRR.....QQ-.IP-KCVDEPCLSFPPFPDDQ\$	512
LST.CD.88.485	..K...PKGK----Y.-.-RN.-P-.G---MGGNAGV.....V-S-----NSTKA-RALETY-NLQQLRK.....V-S-----NSTKA-RALETY-NLQQLRK.....QQ-.VP-KCVDEPCLSFPPFPDDQ\$	512
LST.CD.88.524	..K...PKGK----Y.-.-RN.-PS.G---IMGGAAGV.....I-S-----KVPPTRA-RAIETYRNLGQQLRK.....I-S-----KVPPTRA-RAIETYRNLGQQLRK.....Q-Q-.PQ-CVDEPCLSFPPFPDDQ\$	511
LST.KE.-.lho7	..K...PKGK----Y.-.-SS.KP-.G-Y--LGGAAGR.....I-S-----SAPTKA-RALETY-TLQQLKR.....I-S-----SAPTKA-RALETY-TLQQLKR.....QQ-QVP-KCVDEPCLNSLF.PDDQ\$	515
MAC.US.-.239DR-AG---L.-.-G.-K-----MAQVHQL.....M-----EDPAVDLL.....KNYMLGKQ-REK-RESREKPYKEVTEDLLHL..N---G----M-----EDPAVDLL.....KNYMLGKQ-REK-RESREKPYKEVTEDLLHL..N---G----KNYMLGKQ-REK-RESREKPYKEVTEDLLHL..N---G----	510
MND-1.GA.-.MNDGB1	KPA...QQQR----Y.-.-GP.S--.G-Y-AQEV.....T-----L-EKPLQKTLST-QKLGRL.....T-----L-EKPLQKTLST-QKLGRL.....RQKMKKEEKREDFHSLSTLFQEDQ\$	502
MND-2.-.-5440	K.....M-----N.T--GS.-.-AMPL.....T-S--PGLEDPA.....EKMLLDYMKKGQQQRAA..AGA--EKKKGPY..EAAYNSL..S---T--LQ\$T-S--PGLEDPA.....EKMLLDYMKKGQQQRAA..AGA--EKKKGPY..EAAYNSL..S---T--LQ\$AGA--EKKKGPY..EAAYNSL..S---T--LQ\$	518
MND-2.CM.98.CM16	K.....M-----N.T--GS.-.-AMPL.....T-S--PGMEDPA.....EKMLLDYMKKGQQQ-RAE..S-QE-KE.RGPY..EAAYNSL..S---T--LQ\$T-S--PGMEDPA.....EKMLLDYMKKGQQQ-RAE..S-QE-KE.RGPY..EAAYNSL..S---T--LQ\$S-QE-KE.RGPY..EAAYNSL..S---T--LQ\$	517
MND-2.GA.-.M14	K.....M-----N.A--GS.-.-AMPL.....T-S--PGMEDPA.....EKMLLDYMRKGQQQRAAKE.T-QE.KD.KGPY..EAAYNSL..S---T--LQ\$T-S--PGMEDPA.....EKMLLDYMRKGQQQRAAKE.T-QE.KD.KGPY..EAAYNSL..S---T--LQ\$T-QE.KD.KGPY..EAAYNSL..S---T--LQ\$	518
MON.CM.99.L1	QP.PKQ.N-G-----N.PFGP.KKG-----V.....Q-S--TLPAGTAE-VNLNQEIGS.....Q-S--TLPAGTAE-VNLNQEIGS.....KTG-PKETRRDLYPSLASLFGEDQ\$	510
MON.NG.-.NG1	NGGQNQRNR--N.PFGP-KRG-----V-L.....Q---AEPTXGXKYNFSPVPG--ETPAPTAPKE-RKDLYPSLASLFGDD-SEQ\$Q---AEPTXGXKYNFSPVPG--ETPAPTAPKE-RKDLYPSLASLFGDD-SEQ\$ETPAPTAPKE-RKDLYPSLASLFGDD-SEQ\$	192
MNE.US.-.MNE027DR-AG---L.-.-G.-K-----MAQMHQGL.....T-----EDPAVDLLKNYMLGKQRE.....T-----EDPAVDLLKNYMLGKQRE.....SKRKPYPEVAEDLLHLNSLFGEDQ\$	506
MUS.CM.01.1085	..RSDNT-A-----D.PFGLQKRR-----Q--L.....L---EQEGELVNLSPAVGPGEGKQDRKSLY-SLSS-F-DDQSDQ\$L---EQEGELVNLSPAVGPGEGKQDRKSLY-SLSS-F-DDQSDQ\$EQEGELVNLSPAVGPGEGKQDRKSLY-SLSS-F-DDQSDQ\$	509
RCM.GA.-.GAB1	K.....-I---R.L--GQ-----L.....T-S--G--SN-DPA.....EEMLNKYLRRAGEQ-RQQ..RQEESSKKREGAY..QEALTSL..N---S--LQ\$T-S--G--SN-DPA.....EEMLNKYLRRAGEQ-RQQ..RQEESSKKREGAY..QEALTSL..N---S--LQ\$RQEESSKKREGAY..QEALTSL..N---S--LQ\$	518
RCM.NG.-.NG411E-----N.A--KQ-----L.....-GR-EGWNPQ.....YD-AEEMLRKYLALGRQH--EQR..-EN-EKVGRAY..EDALSSL..N---S-.L.\$-GR-EGWNPQ.....YD-AEEMLRKYLALGRQH--EQR..-EN-EKVGRAY..EDALSSL..N---S-.L.\$YD-AEEMLRKYLALGRQH--EQR..-EN-EKVGRAY..EDALSSL..N---S-.L.\$	521
SMM.SL.92.SL92BNQKAG--L.-.-G.-K-----M.QTTS.L.....T-S--D.PAARIV.....T-S--D.PAARIV.....E-LEKAQ-E-TRRSRYPYKEVTEDLLHL..N-----	507
SMM.US.-.H9ER-AG--L.-.-G.-K-----MAQMPQGL.....T--X-EDPAVDLL.....T--X-EDPAVDLL.....N-MKV-R--RENRRPYKEVTEDLLHL..N-----	507
SMM.US.-.PGM53ER-AG--L.-.-G.-K-----MAQMPQGL.....T-----EDPAVDLL.....T-----EDPAVDLL.....RN-MKM-RK-RENRRPYKEVTEDLLHL..N---G----	507
STM.US.-.STMER--G-----G.-K-.H--MAQ-PQGL.....T-----EDPAADLL.....T-----EDPAADLL.....RS-MQL--K-RESRKTYPYKEVTEDLLHL..N-----	507
SUN.GA.98.L14	Q.....G-----Y.-.-NR.-P-.G---V.....M-S--L-DLTLGNRMTTPPQSKAERALETYRLLGQG-R.....M-S--L-DLTLGNRMTTPPQSKAERALETYRLLGQG-R.....AQ--RKERGECQEPCLNMLFPEDP\$	521
SYK.KE.-.KE51	RKGQGNPPP-A---K.-.-GN.R--PA---VMRE.....T---A-DWPWQQQTWGNVYAS-QQKHIAMAT-LQPQSSPKT-KDLLVPGKPS.....T---A-DWPWQQQTWGNVYAS-QQKHIAMAT-LQPQSSPKT-KDLLVPGKPS.....EKK-EMKEKGP-YPSL-SLFGDDQ\$	557
SYK.KE.-.SYK173	KRNQGPVPA-A---K.-.-GVSRR-PA---V.RS.....E-S--L-DIEDG-WLT-S-QMSQQAQAKAQNPSKPPPTNREVLSP.....E-S--L-DIEDG-WLT-S-QMSQQAQAKAQNPSKPPPTNREVLSP.....KESGKEETKSLYPSL-SLFGEDQ\$	554
TAN.UG.-.TAN1	S.....G-A---RMPT-.T---LEQGA.....V-----PAHGFP.....V-----PAHGFP.....GSPA-GAYDP-RKLLLE-YAKKGDQLRK-KEK-LEDY--S---G----	523
VER.DE.-.AGM3G-----Y.-R-M-.A---AATLGV.....E----PPSP-DPA.....E----PPSP-DPA.....KK-L-QYADKGG-LREQRKKP-AVN-DWTEGY--N-----	521
VER.KE.-.9063G-----Y.-R-T-.T---LAATHGV.....E-S---PPTP-DPA.....E-S---PPTP-DPA.....KK-L-QYAEKGG-LREQGKRT-PTN-DWAEGY--N-----	519
VER.KE.-.AGM155G-----Y.-R-M-.T---AATLGA.....E-S---PPNNTSTPY.....E-S---PPNNTSTPY.....DPAKK-L-QYAEKGG-MRNQNRNP-ANN-DWNEGY--N-----	520
VER.KE.-.TYO1G-----Y.-R-M-.A---AATLGA.....E-S---PPSGTTPY.....E-S---PPSGTTPY.....DPAKK-L-QYAEKGG-LREQQRNP-AMN-DWTEGY--N-----	519

	/ Pol reading frame (-1 from Gag) p15	Gag-Pol TF \ Pol Protease	
SAB.SN.-.SAB1C	FRVWPLG....QRET.QEFPD...LHQ...TNSSPNG.TGLQQ...AGGKLVCRQTSQDQTRARRSSNSPVKAVCCSGETAETAVA.....	KPLATTEPLRGLQLPQVSLWRRPMKTVYIEGQ	105
H1B.FR.83.HXB2	--ED.-AF...LQGK.AREF-S...EQT...RAN...SPT...RRELQ-WGRDN.....NSP.....	SEAGADROGTVSFNF---T--Q--LV-IK-G--	73
CPZ.CD.-.ANT	--ET.DPHVV.GVQ...TRELCA...EGGSS.GAN...SST...HRDL.SGGA.....Q-DS-GSQ.....GGG....	GTTTS-VF-EIP-Q--ME-L-Q--	73
CPZ.GA.-.CPZGAB	--ER.-AFPQ.RE...ARQLCA...EQN...RTN...GPT...DRELW-PGGR.....E-PGEERGR...EQ.....	SI.STN---IT--Q--LIP-KV---	71
H2A.DE.-.BEN	--G-T.....KEAS.-LPRDP...SPS...GADTN.....STSGRSSSG-VGEIYA--EKAEGAEGETIQR-DGGLA-PR.....	AERD-SQRGDR--AA--F--K--VV-A---D-	98
ASC.-.-.Qu			0
COL.CM.-.CGU1		.GN-.....GGGATDShL.PPEDEQRP.....AT..QEHEGGES-S-LE--R--IVI-EV---	53
COL.CM.00.COL243			0
COL.CM.01.COL247			0
COL.CM.99.COL11			0
DEB.CM.01.DEB1083			0
DEB.CM.01.DEB1161			0
DEB.CM.01.DEB11014			0
DEB.CM.99.CM40	--EYS--LQET-ELSRNWEVH...PGE.....HS--GILDYQGTSTGAGEQPP.....EGGNN-REV-PNGAEI-----T--H--R		79
DEB.CM.99.CM5	--ECS--LPET-EFFGRDWEIH...PRE.....HS--GLFNNQGT-AR-KECSP.....GGHNDQ-REVPRFNGAEI-----TA--V--K		79
DEB.CM.99.DEB1CNE1			0
DEN.CD.-.CD1	--E-SWEE...ESSQEL.....CTVL.....PSDESGFAMEPSEGAATTVPTKFK-RGGKQTILSGRQ-AEQ-V.....	PSGGRGDNEGES-SFVE-----IGQIWWQ--	94
DRL.-.-.FAO	--EY.....SLGQRETQEFS...SGL.....ADPISSSIARIPPGGSG-ED-PKVHGAGSS-EETT...ET-T.....	EFG-----F--RL--TTI-EV---	87
GRV.ET.-.GRI_677	-----RS--KK.....CAIQRR.HSWSGTNSPPNGNS...L--KEAPP--RE-TAP-RGE.....	RTDKE--GE-S-GCFLELP-----R-I---T	89
GSN.CM.99.CN166	--P-GA...PPSQELSSNFFAANS PG...G.....-RTPAA-SL...DT-RG...TP.....	EQGTGGEGEHS-S--VI--G---RE-S-G--	76
GSN.CM.99.CN71	--P-GA...TSPQELSSNFFAANS PG...G.....-GTPAA-PV...SAGGG...AP.....	EQG-EGRKEHS-S--VI--G---RE-SM-G--	75
GSN.CM.99.GSNCN7			0
LST.CD.88.447	--L-T-E...EWSS.RKLSF-...GGE...CR--T.....ISSPNGEPT-RESLR.NIQ-PRT--EE.....	ATA.DPTEMC-RAMPELL-S---IENILVN--	85
LST.CD.88.485	--L-T-E...EWSS.RKLSF-...GGE...CR--T.....ISSPNKELNK-RESLR.NIQ-PRT-TEE.....	ATA.GSTEMC-RAMPELL-S---IENILVNR-	85
LST.CD.88.524	--S-E...EWTFRQLSNH...GRG...CR-DS...FSA-HGESA--KSREGRDIQ-PRT-TEE.....	ATT.SSTVC-RALPELL-S---IEK-LVN--	86
LST.KE.-.lho7	--I-S-E...QTS.RQLS-T...GRS...CRENS.....ISSPNGEC-NK-RESTG.DLQDFRT--EK.....	AAA-SPTMC-RALPEPTL.S---IESILVN--	85
MAC.US.-.239	--P-SM-...KEAP.-FPHGS...SAS...GADAN...CSP.....RGPSCGSAKELHAVGQA-ER...KAER-QREALQGGDR-FAA--F-----VV-AH----		85
MND-1.GA.-.MNDGB1	-----SLQ-G-LS.....GTRGDSNSSTIRGETS-ENSEHLSEIRER.....	QAEDGEGEE--FSF-EY--S---IEE-SVD-V	78
MND-2.-.-.5440	--EY.....PLGQW-AQELP...C-A.....IDPISTPDARVGRGRKDAVRLYEERATTEGSSR.....	PERGKEGS---S-----F--N--TTV-E----	86
MND-2.CM.00.MND2S46			0
MND-2.CM.00.MND2S6			0
MND-2.CM.01.MND2S109			0
MND-2.CM.98.CM16	--EY--Q...WQTQELPGDAT.....DPNGPSNARDGRP-RENAVRLHEERATAESGKQ.....	TGKEREGSI--S-----F--N--TTV-E----	85
MND-2.CM.99.MND254	--EC...SLGQW-AQELS...C-A.....TDPISTPDARDGGS-REDVIRLYEKGATTEGSKGD.....	QAGEGQGS--S-----F--N--TTV-E----	86
MND-2.GA.-.M14			0
MON.CM.99.L1	--ESFR...SQ-GAK-L--H...LCAA...QRTD-ASRDCGA-QSQPRDWF.....EDGDPKGDEA-S-S-LSI--G--IRS-L-G--		76
MON.NG.-.NG1	--ESFPW...RQ-GAK-LS-N...LA...A--PS...GTNPXGQGELLTSSGTTDRDPS.....	ADS.PEGDEE-S-S-LSIP--G--VR--E-A--	80
MNE.US.-.MNE027	--L-M-...KEAP.-FPHGP...NAS...GADTN...CSP.....RGPSCGSAKELHAVGQ-ER.....	QKEALQGGDR-FAA--F-----VV-AH----	81
MUS.CM.01.1085	--GSLRP...PEKAE-L-P...LSLA...NCS-RTGG-IGQSLSS.....GSRGGETGQES-SF-VIP-SG--VRP-E----		72
MUS.CM.01.MUSS1239			0
OLC.CI.-.SIVolc_97IC_12			0
RCM.GA.-.GAB1	--ET--S...GEAE.ELSFDF...-DS...LC-RDGEQL...RPCRRDAKE...LSEGR-TKET-EAG...REQBERGIS-S-N--FA-K--TTIAQ----		87
RCM.NG.-.NG411	--EC.SLE...AG--EPESLDF...-ET...YC-RER-RMESP...VRSS-RDAQE...IPSIGETAQTGAER...EQGESG-S--CVV--F--K--VTW-E----		91
SMM.SL.92.SL92B	--P-M-...KEAS.-LPHA...NII...DTV-TPRS...SSP...DSQVRSRES-EGKD...EEQALQGGD-FAA--F--R--VVRAL----		78
SMM.US.-.H9	--A-M-...KEAP.-FPHGP...DAS...GADTN...CSX...RG-SCGSTEELH-GGQK-EG...EQRE-LQGGN-FAA--F-----VV-A---E-		81
SMM.US.-.PGM53	--A-M-...KEAP.-FPHGP...DAS...GADAN...CSP...RG-SCGSTEELH-DGQK-EG...EQRE-LQGGD-FAA--F-----VVAA---E-		81
STM.US.-.STM	--L-M-...KEAP.-LPHGP...NTS...GADAN...CPS...RRPSCGSAEKLHAAGQE-ER...EQED-LQGGDR-FAA--F-----VV-AH----		81
SUN.GA.98.L14	-----E...GAA-R-LS-H...AF...CSA...SRGSDIGE...QNDTTPVKGRESSDLO-VR-GPQ...STABEEGKGGMSGAMLEYA-S---LEK-F-N--		87
SYK.KE.-.KE51	--ERVGE...-ET...PSKFCDBGDTNSPT-GGLAMAATDMGKLCFTPTAHNSGHPA...ATVL-KDSQGLPGWETF...REERD-GEGTS-S-SI-----MEIN--T		105
SYK.KE.-.SYK173	--ERMG...L-EAPS-LPSEERACSPS...G...HRRWAMAHMVSTDEPTST-EGTELSLEETPHQQS-L...EGEQWEGDQES-S-S-L-----IE-DV--D		99
TAL.CM.00.TAL266			0
TAL.CM.00.TAL271			0
TAN.UG.-.TAN1	--ENANL...GNKA.AQ-FRA...RRS...RSDG.SPD.ACARIIP...H-FA-CGS...I-P-QETAG--Q...KG...RSIEKA-GEGT-GLFLEFP-----SKI--Q--		89
VER.DE.-.AGM3	----VD...GS--K-SRR...YSW...GGAN...CAP...STES...I-PCKEAPA-I-RQ--AV-GTKE...TTSSER-DR-IPFEL,P-----I-----V		86
VER.KE.-.9063	----TVD...GDK-EKLSRC...HSW...SGTE...RAP...STDT...I-P-KEAPA--RE-K-V-GTRE...ENTTKSR-G-RIPFEL,P-----I---I--V		86
VER.KE.-.AGM155	----VD...GDK-KK-SRS...HSW...GGTK...CAP...STEQ-YT...L-P-KEAPA--RER--N-KSEQ...PSEQSR-ER-IPFEL,P-----IR-CI-G-T		89
VER.KE.-.TYO1	----TVD...GGK-EK-SRR...YSW...SGTE...CAS...STERHHP...I-P-KEAPA-I-RER--T-G-KE...ESTGNEG-DR-IPFEL,P-----I-----V		89
WRC.CI.97.14			0

Asp25 catalytic site

Protease // Pol p51 Reverse Transcriptase

SAB.SN.-.SAB1C	KVTALLDTGADDSVIQIIEI...GDNWKPRIIGGIGGCINVKAYHNQEVKIEDKTK....ATILVGETPVNIIIGRNVLQGLVTLNLTQREIEPI...KVHLKPGQDGPRIQRWPLSRKEKIEALKAI	223
H1B.FR.83.HXB2	LKE-----T-LEEMS...PGR---KM-----F-K-RQ-DQILIE-CGHKAI...G-V---P-----L-T-I-C---FPISP-TV...P-K---M---KVK---TE---K-VE-	191
CPZ.CD.-.ANT	-CQ-----T-VE-H...LKQ---KT-----F-S-QQ-NKVPQI-G-R-VL...-V-L-PN-----CL-C---FPISKV-TV...P-K---E-M---VK-----E-	191
CPZ.GA.-.CPZGAB	LCE-----T-ER-Q...QGL---KM-----F-K-QFD-VHIE-GRKVV...G-V---P-----I-T---C---VFPISS-TV...P-K---M---KVK---A---K-TE-	189
H2A.DE.-.BEN	P-EV-----IVA-----YT-K-V---F--T-E-K-V-I-VLN-RVR...-MT-D-I-F---I-TA-MS---PVAK-----T---K---LK---T-----E-	216
ASC.-.-.Qu	0
COL.CM.-.CGU1	-CQ-----T-FKEEDVQL..TGA-TAVQ-Q---A-R-Q-RDKIL-MGG-EYR...GD--I-H--I--V---M-KEGKLVVAQLSDR-PVT...-T--E-M---KVK-----G-QK-	173
COL.CM.00.COL243	0
COL.CM.01.COL247	0
COL.CM.99.COL11	0
DEB.CM.01.DEB1083	0
DEB.CM.01.DEB1161	0
DEB.CM.01.DEB1014	0
DEB.CM.99.CM40	--QV-I-----TI-SEKDIDL..EAP-T-KTV--L-F---C-PGI-IAMAN-VAV...SDL--D-I-L--Y--KM-SI-FPVLDTY-V...L..E-A---K-K---R-----D-	197
DEB.CM.99.CM5	R-QV-I-----TI-SEKD-DL..DAP-T-KTV--L-F---C-PGI-I-MA--VAV...SDV--D-I-L--Y--KM-SI-FPVLDTY-V...L..E-A---K---R-----E-	197
DEB.CM.99.DEBCE1	0
DEN.CD.-.CD1	R-EV-I-----TI-AEQDINLGTEG.-I-KTV----Y--Q-PGI-I-LM--VAV...D-V--P-I-L--Y-SK-----MAVQOMR-Q.VTP-K---F---K---TT--LK--QQ-	217
DRL.-.-.FAO	--E-----T---L---TGK--Q-----T--RQ-F-C-I-AG-ITH...-SV--P-----CK-A---FPVSQV-V...-T---M---K-K-----E-	205
GRV.ET.-.GRI_677	P-Q-----TI-EKD-HFPHPKP-RSKVV---G-H-E-QGVQ-QL---IIT...GS--I-S-I---I---A-MK-VMGVLSSQIE.ET--Q--E-K---KLK---R-----TE-	213
GSN.CM.99.CN166	V-SM-----TIV-DSAIEL..DTP-T-KMV---L-Q-REHKHV--IFNE-RI...-V--P-----L--C-SK-----MV-QK---E---A-KE--K-K---R-----TQ-	196
GSN.CM.99.CN71	V-SM-----TIV-ESAIEL..DTP-T-KTV---L-Q-EHRHV--IFNE-RI...-V--P-----L--C-SK-----MV-QK---E---S-KE--K-K---R-----TQ-	195
GSN.CM.99.GSNCN7	0
LST.CD.88.447	P-S-----TIFSENSVRI..EGPYT--VV---Q-R-E-KDVFIE-AG-TV...G-V-L-P---D-V---I-TNV-AR-V-S-LSDKIP.IT--K---C---VK-----G-Q-	207
LST.CD.88.485	P-S---R---TIFSENSVKI..EGPYT--VV---Q-R-E-KDVFIE-AG-TV...G-V-L-P---D-V---I-TSV-AK-V-A-LSDKIP.IT--R---C---VK-----G-Q-	207
LST.CD.88.524	E-----TIFSENSVRI..EGPYT--VV---Q-K-EFRDVFQI-AG-TV...G-V-L-P---D-V---I-NP--AR-V-A-LSEKIP.IT--C---C---VK-----G-Q-	208
LST.KE.-.lho7	P-S-----TIFSENSVRI..EGPYT--VV---Q-R-E-RDVFIE-AG-TV...G-V-L-P---D-V---I-TAM-AK-I-A-LSDKIP.IT--S---C---VK-----G-Q-	207
MAC.US.-.239	P-EV-----IVT-----PHYT-K-V---F--T-E-K-V-IEVLG-RI...G-MT-D-I-F---L-TA-MS---FPIAKV-V...-A---K---KLK-----V-RE-	203
MND-1.GA.-.MNDGB1	TIR-----TIFNERNIKL..KG-Q-K-----NLR-Q-D-VY-E-RG-GTF...G-V-I-P-ID---IMEK-GK-I-A-LSDKIP.IT--K---V---K-----VG-QK-	200
MND-2.-.-.5440	--E-----T--KDL---TG-----A-R-RQ-F-CK-E-AG-ITH...-AV--P-----KK-C---FPISK-TV...-T---T---K-----L-QE-	204
MND-2.CM.00.MND2S46	0
MND-2.CM.00.MND2S6	0
MND-2.CM.01.MND2S109	0
MND-2.CM.98.CM16	--E-----T--KDLD...KG---Q---S---QPF-CK-T-AG--TH...-SV--P---V---KK-C---FPVSKV-TV...-T---T---K-K---L-QE-	203
MND-2.CM.99.MND254	0
MND-2.GA.-.M14	--E-----T--KDVD...TG---Q---A-R-Q-F-CKITVAG-STH...-SV--P---V---KK-C---FPISK-TV...-T---M---K---L-QT-	204
MON.CM.99.L1	--SM-----TI-DQHVNL..E--T-KVV---M-E-Q-KAIQ-IF--REVW...-V--P-I-L--L---KM---MAAGLDLQPPSLE-T-A-KE--KLK---R-----TQ-	199
MON.NG.-.NG1	S-SM---X---TIV-XQDINL..E--T-KMV---L-Q-QPKGVT-SFQG-VVR...X-V-L-X-I-L---SRM-I---MX-GIAQSP.SLT-T---K---KLK---TT-	202
MNE.US.-.MNE027	PAEV-----IVA-----PHYT-K-V---F--T-E-K-V-IEVLG-RI...G-MT-D-I-F---L-TA-MS---FPIAKV-V...-T---K---KL---V-RE-	199
MUS.CM.01.1085	V-SM-----TI-DAHIHL..EGP-N-KSV---N-S-E-RGVQ-CFNG--IR...-V--P-I-I-M---C-SKF-I---MI-GKL---T---E-AK--MVK---TQ---R-ED-	192
MUS.CM.01.MUSS1239	0
OLC.CI.-.SIVolc_97IC_12	0
RCM.GA.-.GAB1	--EV-----T--E---ND-T-K---Y---Q-K-C-IE-AG-RTH...-HV--P-----KK-A---FPISQ-T...-E-S---VK-----TE-	205
RCM.NG.-.NG411	--E-----GT---NLD...E--K---Y---QFNHCKIT-AG--TH...-HV--P-----SK-AS-FPISKA-T...-E---VK-----TE-	209
SMM.SL.92.SL92B	PLEV-----IVA-----PGYT--V---F--T-E-RTVKIEVLG-VIR...E-LMT-D-I-F---I--A---S-YPVAKV-YT...-K--E-M---LK-----Q--TE-	196
SMM.US.-.H9	PIEV-----IVA-----P-YT-K-V---F--T-E-KDKVI-VLG-VI...G-MT-D-I-F---L-TAM-MS---PIAKVX...-T---K---KL---I-RE-	199
SMM.US.-.PGM53	P-EV-----IVA-----P-YT-K-V---F--T-E-KDKVI-VLG-VI...G-MT-D-I-F---L-TAM-MS---PVAKV...-T---K---KL---V-RE-	199
STM.US.-.STM	P-EV-----IVA-----LQYT-KVV---F--T-EFK-VNIEVLG-KI...G-MT-D-I-F---L-TA-MS---FPVAKV-V...-T---K---K-K-----E-	199
SUN.GA.98.L14	P-----TIVSEDSVSI..EGH---VV---Q-R-E-ADV-F-E-R--AT...G-V--P--D---I-SA-GR-V-ATLSEKIP.IT--K--E-AA--KVK---R---G-QK-	209
SYK.KE.-.KE51	--M-----TI-REQDIQL..HQP-T-K-V--L--N-R-RQ-NQIRFQ--RPDGRIKEVEGSL--P---L---I-SK--AK-VMAIQKL--M...-Q-Q---KEL-KLK---TR--L--R--	230
SYK.KE.-.SYK173	L-QM-V-----TI-REEDIQL..HQP-S-K-V--L--N-T-RQ-R-IRFTVVKPSG-RKQVEG-L--P---L---I-TK--K-VMV-TAL-V...-S---DKEL-LK---V--L--	224
TAL.CM.00.TAL266	0
TAL.CM.00.TAL271	0
TAN.UG.-.TAN1	P-QV-----TI-KEEDIHL..EGR-R-KVV---G-S-Q-D-IL-N-DG-ETR...GNV-I-P-I-V---I--P--AR-IMG-LNEKIP.IT--Q---MS--KLK---E---K--TE-	211
VER.DE.-.AGM3	PIR-----TI-KEAD-QL..SGT---K---GL--E-SDR--RL--ILR...G--I-S-I---I--PA-AK-VMG-LSEQIP.ITP-K--E-AR--FLK-----K--QE-	208
VER.KE.-.9063	P-KV-----TI-KETDIQL..TGP-R-KV---GL--E-S-V---ILR...G-V-L-A-I---L--PA-AR-VMG-LSEQIP.ITP-K--D-AR--H-K-----K--TE-	208
VER.KE.-.AGM155	A-K-----TI-KDTE-QL..RGS-R-K-V---GL--E-D-V--QL---ILR...G-V-I-A-I---F--A-AK-VMG-LSQTIP.ITP-R--E-AR--LK-----I--QE-	211
VER.KE.-.TYO1	PIK-----TI-KEND-QL..SGP-R-K---GL--E-NDR---ILR...G--L-A-I---L--PAVPR-VMG-LSEKIP.VTP-K--E-AR--CV-----QE-	211
WRC.CI.97.14	0

PLV proteins

Table with columns for protein identifiers (e.g., SAB.SN., H1B.FR.), alignment characters, drug resistance sites (M41L, A62V, D67N, K70R), catalytic site (Asp110), and sequence numbers (347, 315, 315, 313, 340, 65, 297, 0, 0, 0, 0, 0, 321, 321, 0, 341, 329, 337, 326, 319, 0, 331, 331, 332, 331, 327, 324, 328, 0, 0, 0, 327, 0, 328, 323, 326, 323, 316, 0, 69, 329, 333, 320, 323, 323, 323, 323, 333, 353, 347, 0, 0, 335, 332, 332, 335, 335, 69).

	polymerase motif catalytic site Asp185 Asp186	drug resistance T215Y K219Q	
SAB.SN.-.SAB1C	TTANKILQBFQKPNPDVYIQYMDMLIASDRPKAEHLVMMVQQLRDYLETGWGFKTPPEKKFKQDPYPLWLMGYELYPKKWQLQBITLPEREE..WTVNDIQKLVGKLNWASQIYTGIKTKHLRCRLIRG.ARP		474
H1B.FR.83.HXB2	SSMT--EP--KQ---IV-----LYVG--LEIGQ--RTKIBE--QH-LR--LT--D--H--E--F-----H-D--TV-P-V---KDS..-----P---VRQ--K-L--TKA		442
CPZ.CD.-.ANT	ASMT--AP--D-Y-A-E-----LYVG--MEITA-RE-IEK--QH-QV--LE--D--H--E--FQ-----H-D--TV-K-K---PDD..-----P-----Q-----VKS		442
CPZ.GA.-.CPZGAB	SSMT--EP--E-----IT-----LYVG--LEIDQ--RKK-EE--QH-LK--T--D--H--E--F-----H-D--TV-P-Q---K-V..-----I-----P---I-Q--K---TKK		440
H2A.DE.-.BEN	Y-MRQV-EP--KA---ILI-----I-----TGL--DKV-L--KEL-NGL--S--DE-----FQ---C--W-T--K--KLQ--QKDI..-----V---A---S-----KMT		467
ASC.-.Qu	H--AG--EPL-AQYS-MT-I-----L--G--N--IES-RKA--AI-QT-TSQ-IA--LE-Y-E-F-VT-L---H--S-TISK-E--NQQ...-----L---I-----P-----Q--K-L--KG		192
COL.CM.-.CGU1	G-VAR--EP--KRGQL-Q-A---LYLG--L-LE--RRA--E---A-LYY-LE---Y-A---K-----H--Q-K--KVKI---QDK..-----A---L--L-P--R--EISKCTK-.KKH		423
COL.CM.00.COL243		0
COL.CM.01.COL247		0
COL.CM.99.COL11		0
DEB.CM.01.DEB1083		0
DEB.CM.01.DEB1161		0
DEB.CM.01.DEB1014		0
DEB.CM.99.CM40	ASVGRQ--I--E--ELI-V---L-VG--LR-DQ--KK-A---QF-LER-LR--PE-Y-E---FH-----H--R-MPV--DE---HK---Q-----S-----A---VPG		448
DEB.CM.99.CM5	A-VGRQ--KL-KE-Y--LI-V---L-VG-NLD-VS--EQ-K---KF-LEK-LR--PE-Y-E---H-----H--H-LP-E---E---HK---Q-----S-----KA---VPG		448
DEB.CM.99.DEB1083		0
DEN.CD.-.CD1	A--D-M-SQ--KAH---ILI-----L-V---SLEQ-RK--KE--F-AV---E--DA---ET--IQ-----H--H--I-K-D--LP-DGO---A---V---L-S---E--KS---VKD		470
DRL.-.FAO	A--D---P--E-----I-----LFVG---HL-EQ-IKE--TH-KF--LE--D---DN---E---Q-H---TV---R--DK-I..-----V---S-----E--K---K		456
GRV.ET.-.GRI_677	N--AN--E-IKRHT-GLE-V---LWL---HDETR-NQQ-DIV-KM-LEK-LE--D--V-RE--WE---K-H-N--TINK-E--PL-G-E---K---V--V---P---T-AML---KKN		465
GSN.CM.99.CN166	NIVASL-RQI-DQE---VLI-----L--G---KLG--RQV-HKI-NL-TSYNIQ---A-H-P-Y-VK-L---H--G-R--PVE--DQDT..-----V---HP--R--Q--K---KG		453
GSN.CM.99.CN71	A-VASL-YQI-DQE---V-I-----L--G---KL--RQV-HKI-NL-TSYNIQ---A-H-Q-Y-VR-L---H--G-R--PVE--DQDI..-----V---HP--R--Q--C---KG		446
GSN.CM.99.GSN166		0
LST.CD.88.447	A-VAGL-S---KL---IIL-----LF-G---D-KG-DQA-KE---L-V--NL-----AE--H---V-H-DR-EIEKTK--DIDLQKT--N-----V---A-L-E--R--E--K---VKA		460
LST.CD.88.485	A-VAGL-S---KL---IIL-----L--G---D-KG-DQA-KE--EL-M--NLE-----AE--H---V-H-DR-EIEKVK--DIDLQKT--N-----V---A-L-E--R--E--K---VKA		460
LST.CD.88.524	E-VSGL-A---KL---ILL-----LF-G---E-BG-SQA-KE--EL-I--NLE-----DK--H---V-H-DR-EIERVK--IDTRKT--E-----V---A-L-D--R--E--K---VK		461
LST.CD.88.524	G-VAGL-S---KL---MI-----LF-G---ERKG-DQA-KE--EL-M--NLE-----AE--H---V-H-DR-EIEK-K--MDLTKT--Q-----V---A-L-D--R--E--K---VK		460
LST.CD.88.524		460
MAC.US.-.239	Y-MRHV-EP--KA---TLV-----I-----TDL--DRV-L-SKEL-NSI--S--E-----FQ-----W-T--K--K-E--Q--T..-----V---A---P-----KMT		454
MND-1.GA.-.MNDGB1	G-VASL-EV--KNH-T-QL-----LFVG--YTAE--EKAIVE--AL-M--NLE---Y-E--FH-----H-D--KIEKVQ--LA-Q-P---E-----A-L-P---Q--K---GLN		452
MND-2.-.5440	A--D---KA-KE-H-E-L-----LFVG--LSAS--SK-IVE--EH-QF--LE--D---E--FE---V-H---TV-K-Q---K-S..-----S---E--K---K		455
MND-2.CM.00.MND2S46		0
MND-2.CM.00.MND2S6		0
MND-2.CM.01.MND2S109		0
MND-2.CM.98.CM16	A--D---P--ERH--V-----LFVG--VAP--SR-I-E--H-LF--LE--D---E--FE---I-H---TV-KVQ--K---S..-----S---E--K---K		454
MND-2.CM.99.MND254		0
MND-2.GA.-.M14	A--D---KT-KEEY-E-L-----LFVG--LSAS--NK--NR--EH-RF--LE--D---E--FE---V-H-Q--TV-K-Q---K-T..-----S---E--K---K		455
MON.CM.99.L1	H--AELP-I-K-Y-E-TLI-----L--G--HSLP--RKI-GVI-AT-LRK-IQ--PE--D-Y-VQ-L---H--SH-KIAK-E--DQ-Q..---E---I-----A--S---N--KS---K		450
MON.NG.-.NG1	YS-SELKXII-R-Y-E-TLI-----L--G--HSLP--RKI-TAI-TX-SEK-IQ--IE-Y-DKY-LQ-L---H-TH-TV-K-E---QDQ..---G-----A--P-----XC-X--XKA		453
MNE.US.-.MNE027	H-MRNV-EP--KA---TLV-----I-V---TDL--DRV-L--KEL-NSI--S--E-----FQ-----W-T--K--K-E--QK-T..-----V---A---P-----KMT		450
MUS.CM.01.1085	F--A-L-E-IKKA---TLI-----L--G--N-ELSS-RKV-A-I-IM-LKL-IQ--QE--E-Y-VQ-L---H-RG-KI-PVA--QQ-V..---V-----A-A-P---Q--K---VKS		443
MUS.CM.01.MUSS1239		0
OLC.CI.-.SIVolc_97IC_12	S-VAHLMPEW-KE--EI-L-----LYVG--L-R---EK-K---EL-RSYNLE--P--Y-E---E---T---R-KIAR-K--K---Q-----S-V---A--K--K-.PKG		196
RCM.GA.-.GAB1	AS-----P--EE---I-----LFVG--T-L--DK-IK---H-LF--E--D---DK-----H--S-TV--K--K---K---S-LR--E--K---K		456
RCM.NG.-.NG411	-----EP--KE---V-----LFVG--TAH--SQ-IKR--EH-L-----E--D---DK--FE---V-H-D--TV---K-L-K---K---L---E--K---K		460
SMM.SL.92.SL92B	A-MRQV-EP--KA---LLV-----L--G--N-GLT--DK--T---M-NNL--S--D---N--LQ--L---K--K-E--K--R..-----V---A---P---N--KM---KMT		447
SMM.US.-.H9	H-XRNV-EP--KA---TLI-----I-----TDL--DRV-L--KEL-NSI--S--E-----FQ-----W-T--K--K-E--Q--T..---N-----V---A---P---K---KMT		450
SMM.US.-.PGM53	H-MRNV-GP--KA---TLI-----I-----TDL--DRV-L--KEL-NSI--S--E-----FQ-----W-T--K--K-E--Q--T..-----V---A---P---K---KMT		450
STM.US.-.STM	Y-MRNV-EP--RA---TLI-----I-----TDL--DRV-L--KEL-NSI--S--E-----FQ-----W-T--K--K-E--Q-DV..-----V---A---P---K---KMA		450
SUN.GA.98.L14	G-VASL-EK--RQY-E-QL-----L--G--YE-KK-BEI-K---QL-ME-NLE---Y-GE---K---I-H-DR-EIEK-K--PL--E-P-----I--VI---L-E-LR--E--K---NK		461
SYK.KE.-.KE51	A-V-Q---PI-E--S-II-V-----L-VG-N-SEK--GQI-SEIVKS-LAV-SI-PE-W-DKF-LQ-L--T-H-D--S--KVQ---IT-S-P---EL--II-V-----P---C-A-SKC---KA		481
SYK.KE.-.SYK173	S-I-Q---P--K-Y-S-LTLI-----L--GT--SEKA-QEI--IVTA-LKV---V-KE-W-DQY-MQ-L--T-H-D---K-E--NIDDE.I---QL---I-V-----S---E--KC---TKQ		475
TAL.CM.00.TAL266		0
TAL.CM.00.TAL271		0
TAN.UG.-.TAN1	H--QRV-K-IGK-VQ-LE-I-----W-G--L-EK--DEK-EEV-KL-LLR--E--D--V--R--D---K--SH-S-NK-EI--Q-H..-----V---L-P---A--K-L--QD		462
VER.DE.-.AGM3	N--AS--E-IKKELKPLT-V---LWVG-QEDEYT-DRL-E---MK-SA--LE--D--V--K--E---K-W-H---ISS-E-EDK---R-----A-L-P-LR--N--K---KKN		459
VER.KE.-.9063	N--A---EKIKKELKPLT-V---LWVG-QEDEYT-DKL-E---VK-K--LE--D--V--K--E---K-W-Q---NS-N-E-K-Q-----A-L-P-LR--QI---KKN		459
VER.KE.-.AGM155	N--S---E-IKKELKQLT-V---LWVG-QEEGPK-DQL--T--NR-QE--LE---V-RE--FE---K-W-H--K--S--E-EKK-Q..---L-----A-L-P-LR--NI--K-L--KKN		462
VER.KE.-.TYO1	N--AS--E-IKRNL-ALT-V---LWVG-QENEHT-DKL-E---TK-QA--LE---M--E--E---K-W-H--E-SR-Q-E-KD---L-----A-L-P-L--R-I-K--T-GKKN		462
WRC.CI.97.14	S-VASL-EIY-G---NIQVL-----LYV-G--LD-T---NK-EE--EL-L--NLE---L-LE--CN---I-H-D---I---K---E---EVK..-----V-----A-L-E--Q--Q---NK		195

SAB.SN.-.SAB1C	LTEIVQWTEEALELEENRQILRQKQGGYYDPALPLRAKVLKLDGGQGWYQIYQPE.NKILKVGKYAKIKTAHTNELRMLAGLVQKIGKESIVIWGQI.PIMELPVERELWEQWWSYDQVQVTWIPEWEM	602
H1B.FR.83.HXB2	---VIPL-----A---E---KEPVH-V---SKD-I-EIQ-Q-Q---T-----EP.F-N-T-T-----RMRG---DVQ-Q-TEA---TT-----KT.-KFK--IQK-T-T--TE---A-----F	570
CPZ.CD.-.ANT	--DR--M-R-----K---Q--IE-Y--Q-G---K-TIQ-Q-S---T-----N-.G-L-A---RPTGT---V-Q---V---L-----EV.-KFK--IT-T-DA---A-----F	570
CPZ.GA.-.CPZGAB	--DV-PL-P-----A---E-VSTPVH-V---DKE-I-EIQ-Q-NC-T---F-EP.H-N-T---RQRS---DI-Q-EA---AT-----KT.-KFR--QK-S-A-AE---A-----F	568
H2A.DE.-.BEN	---E-----L-A---KI--S-E-E-Y--QEEKE-E-TIQ-SQGH-T-K-H-E.E-----NT---GV-L-QV-----AL---R.-.KFH-----T-----DN-----DF	594
ASC.-.-.Qu	-LDK-ELS-D-A-R-R-QE--S-EHT-G-RED---EVE-VS---T-----MFK-D.KQV--T-F-V-RSV-V-QYQ-Q-DALTRL-R--L-Y--RLPE.FRI--VQ-Q-NT-AEH--AN---VKA	319
COL.CM.-.CGU1	-L-E--LSA---E--QD-K--IKEEV--V--I-HEDIWVDIFR--K----AVC-K..HGT-RR--HNTG-NQ-F-SMQE--SVI---R-A-T--KV.-K-KV-AK--D-----SA---I-F	550
COL.CM.00.COL243	0
COL.CM.01.COL247	0
COL.CM.99.COL11	0
DEB.CM.01.DEB1083	0
DEB.CM.01.DEB1161	0
DEB.CM.01.DEB1014	0
DEB.CM.99.CM40	--DP-TLS--QA--R--KE--K-EVS-V--KEQE--I-ELS--K-----V-R---KG--T--FSRD-G-Y-DFHQ--KAMY--T---F-R-.QFR--VK-E-DN--HMH--AA---D-A	575
DEB.CM.99.CM5	--DK-SLS--QA--R--KETPK-EVS-V--KEDB--ISEIS-I-R-----V-R-T..KGL--T--PS-D-GT-Y-DFHQ--KAMY--T---Y-KV.-KFR--VK-E-DN--HMH--A---D-A	575
DEB.CM.99.DEBCE1	0
DEN.CD.-.CD1	-G-TISL-G-A-K-AQE--KESIS-S-R-EK--I-EIS--EN---T-K-E..KHL--T--H-D-GS-Y-PYQQ--KVMA---R-AL---RL.-TFR--SKVD-DA---E---N---D---	597
DRL.-.-.FAO	-D-E-E-R-----Y---KL-IKEQMH-V-Q-EK--K-IQ--N---S---E-DD---P-T---V-NS---DM-----A-AL---RL.-FY-----V---QE-----D--F	584
GRV.ET.-.GRI_677	-L-EIV-----A-YKN-QG-VQET-E-T---LKE-I-T-Q-Q-E---T--FT-E..GAV---R--QRET---D--T-H---C-ALT--RL.-RVQ--DKKT-DM-Q---S-----F	592
GSN.CM.99.CN166	--DQ-EL---QI--A---E--K-SSE-G---AEK--VVEITS--EQ---MFS-EG..M--SA-F--TRN--I-SYQQF-DALT-V-R--Q-T--KVPDKFRI--VK-Q-DA--MN---A---TI-A	581
GSN.CM.99.CN71	---Q-EL---QI--A---E--K-SSE-S---AEK--VVEITS--EQ---MFS-DS..M-RS-F--TR--I-SYQQ--DAMT-V-R--Q-T--KVPDKFRI--VK-Q-DT--MN---A---TI-A	574
GSN.CM.99.GSNCN7	0
LST.CD.88.447	-D-V-T--DA-LE-Y-Q-Q---KE-L-A---TKE-VVR-QQSKK-IITF-WR-G..S--RA-R-ORQ-S---P-QK--EA-----HV.-KIQV--T--V-D---H-----L-F	587
LST.CD.88.485	-E-V-T--DA-LE-Y-Q-Q---KE-L-A---TKE-VVR-QQSKK-IITF-WR-G..S--RA-R-ORQ-A---P-QK--EA-----HV.-KIQV--T--V-D---H-----L-F	587
LST.CD.88.524	-EQL-T---LE-Y-T-KT--KE-V-A---SKE-VVR-QQSKK-IITF-WR-G..A--RA-R-ORQ-S--M-P-QQ--EA-----FV.-KIQV--T--T-N--INF-----DL-F	588
LST.KE.-.lho7	E--IN---LE-YGQ-KEV-KE-M-A---EKE-IVR-QQSKK-IITF-WR-G..N--RA-R-ORQ-A---P-QK-VEAI-----FV.-KIQV--T--V-H---H-----DL-F	587
MAC.US.-.239	---E-----M-A-Y---KI--S-E-E-C--QEGK--E-T-I-SQ-N--S-K-H-E..D-----F---NT---GV-L-HVI-----A-----V.-KFH---KDV---T-----DF	581
MND-1.GA.-.MNDGB1	I--K-TM---R--Y-Q-KE--AEE-E-S---NKE-VYRFQ-TTG-DISP-WK-G..V-RA--G-Q---S-D-MK---AT--V-R-----FV.-K-QI-TT--I--D--HE---C-----V-F	579
MND-2.-.5440	-DK--E--R-----Y---KL-VQEEVH-V-Q-EK--M--Q--TQ---S---E-ED.--P-----RT-N-----V-----VA-CL---KL.-KFY--L--V-D--H-----DF	583
MND-2.CM.00.MND2S46	0
MND-2.CM.00.MND2S6	0
MND-2.CM.01.MND2S109	0
MND-2.CM.98.CM16	-D-K-E--R-----Y---KL-VQEEVH-V-Q-EK--M--Q--TQ---S---E-ED.--P-----RT-N-----V-----A-AL---KL.-KFY--I--V-D--PE---A-----F	582
MND-2.CM.99.MND254	0
MND-2.GA.-.M14	-D-V-E--R-----Y---KL-VQEEVH-V-Q-GK--M--Q--TQ---S---E-E.--P-----RT-H-----V-----A-C---EL.-KFY--L--V-D--H---A-----F	583
MON.CM.99.L1	--DT-VLSDL-QA--A---E--KEEST-G---NE--KVELTS-AEN---RFF-E..KSV--T-F---RST-S-TYQQ--DALA-L--A-T--RL.--FR--VK-Q-DA--A-N--IN-V-DI-A	577
MON.NG.-.NG1	-D--ELSDL-QA--A---E--S-QST-G---Q-D-KVEITS-XE---XYF-E..XRV--T-F---N--S-TYQQ--DA-T-L--AL---KL.--FR--NK-Q-DT--A-N-----I-P	580
MNE.US.-.MNE027	---E-----M-A-Y---KI--S-E-E-C--QEGK--E-T-I-NQ-N--S-K-H-E..D-----F---NT---GV-L-HVI-----A-----V.-KFH---KDV---T-----DF	577
MUS.CM.01.1085	---E--L---QA--A--QE--Q-TV--A---EQ--V-EIVS---S---NFT-S..KGV-RT--F--VRSV---SYQQ--DAMA---R-L---KPPDKVRI--TK-Q-D--IAH---S---TI-A	571
MUS.CM.01.MUSS1239	0
OLC.CI.-.SIVo1c_97IC_12	-LDK-E---KE-----E--KEE-E-VM-KEEB--QCTLCY..TKGTVAIRIQWKHGT--T-RFQRKM---KF-QMT-AIE-----KL.--FQV-TT-KD--A-AF-D---H---DL-F	322
RCM.GA.-.GAB1	-D-K-EM-K--I-Y---KM--KE-LH-V---EKK--V-NIQ--EG---S---E-ES-G-P-T---Q-----I-----A-A---RL.-TFR--I--V-D\$-R-Q-----D--F	583
RCM.NG.-.NG411	-DDK-EF-R---Y---KAV-KE-IH-V--N-EK--I--Q--EG---S---E-GD-S-P-T---Q-----I-----C-A---KL.-KF--I--V-----S-----F	588
SMM.SL.92.SL92B	---E-----L-L-A-A---I--N-E-E-R--REDE--E-T--NQ-N--S-K-H-G..DR-----F---NT---GI-L-NV-----L---KT.-FFH-----V-D--T---A-----DF	574
SMM.US.-.H9	---E-----M-A-Y---KI-PS-EPE-C--QEGK--E-P-I-SQ-N--S-K-H-E..D-----F---NT---GVSL-HV-----A-I---V.-RFH---X-I---T-----DF	577
SMM.US.-.PGM53	---E-----M-A-Y---KI--S-E-E-C--QEGK--E-T-I-SQ-N--S-K-H-E..D-----F---NT---GV-L-HV-E---A-----V.-KFH-----I---T-----DF	577
STM.US.-.STM	---E-----M-A-YA--KI--S-E-E---REDK--E-T-V-DQ-N--S-K-H-E..D-----F---NT---GV-L-HVI-----A-----V.-KFH-----V---A-----DF	577
SUN.GA.98.L14	-A-K-KM---RE-YQS-QEV-QESVS--S-E-DKB-ICR-Q-VKQ-ILTF-WL-G..KQV-R--R-Q-RGA--E-PCQQ--AAL---R-----FV.-KIQV-IQ--I-S---A---C---L-F	588
SYK.KE.-.KE51	--D--EL---A-MA---E--KVE-T-S-Q-EK--E-HIS---QQ---I-K-GPQE-P-IT--TG-AYAT-H-DYQA--Q-MN---IQALW--K.-EFH--K--E--K--T-H--A--L--VKK	610
SYK.KE.-.SYK173	---VLTL--A-A-----KEE-A-S---KK--E-HIT--SQ---M-K-EQKGP-IT--T--TFA--S-DYQS-Q-LN---IQ-LWY--KV.-TFH--K--E--K--T---A--V--VKF	604
TAL.CM.00.TAL266	0
TAL.CM.00.TAL271	0
TAN.UG.-.TAN1	--AV---Q--V-FRQ-QE--KEE---A---SKS---TIT--D-----FK-E..G-V--T---T-S-NT---F-V---M--RVC-A-TF--KL----V---V-----L--I-F	589
VER.DE.-.AGM3	-L-T-T-----A-YA--KE--KTE-E-T--K-GR-I--A-Q--EG---S--FK-E..GQV-----T-Q-NT---F-V-----LC---L---EL.-VL---I--V---A---S---D--F	586
VER.KE.-.9063	-L-E-----A-YA--KE--KTE-E-T--A-GK---A-Q--E---S--FK-E..GQV-----Q-NT---T-----C-AL---EL.-F-----V---A---S---F	586
VER.KE.-.AGM155	-LDV-E--P---A-Y---KE--KTE-E-T--A-EK---A-Q-----S--FK-E..G-----F--Q-AT---V---V-----AL---L.-TF-----DT---A---S---DF	589
VER.KE.-.TYO1	-L-L-A--P---A-YA--AE--KTE-E-T--K-GI--I--A-Q--EG---S--FK-E..GQV-----T-Q-NT---T-----C-AL---IL.-VL---I--V---A---S---DF	589
WRC.CI.97.14	---E-E-----LE---I-K--KEE-S-A--K.GKD-I-ELI-T-P---A-T-REIGEK-P-----GRQRHT-S---SM-HAM---T-----RV.-KFRI-I-KDI-DS-----S---DT-H	323

Accession	Sequence	Length
SAB.SN..SAB1C	VSTPQLRLWYKLVKDPPIGAEVYVVDGAANRNSKEGKAGYLTRDGRQKVVALENTTQKAELEAILLALRDSGSKVNIITDSQYAMGIAGEPTESDNNIVQOIIIEELIKKEAVYIAWVPAHKGVGGNE	732
H1B.FR.83.HXB2	-N-P-VK---Q-E-E-V-AETF-----ET-L----V-N-R----T-TD-----T-Q-Y---Q---LE---V-----L---QAQ-DQ-ESEL-N---Q---K-L-----I---	700
CPZ.CD..ANT	-N-P-----N-LA---EAETF-----QL-----V---RSR-KH-QK---Q---Q---M-E---TGP---V-----L-VLQ-T-DQ-ESPL-EE---QK---R-QI-LS-----I---	700
CPZ.GA..CPZGAB	IN--P-VK---S-ETE---TTDT-----ET-T-----V-K-K---IIS-----Q---K-L---DQQ---V-----VL---QSQ-DH-ESEL-N---Q---KI-LS-----I---	698
H2A.DE..BEN	----P-V---TFN-G-----AETF-T---SC-Q-----V---KD--KV-Q---Q---VFRM--A---P-----V---V---V---Q---E-R--N-----M-----V-----I---Q	724
ASC.-.-.Qu	-H--P-L-S-FT-ATA-LD.APTF-T-----E-L----V-NT-KRACIG-----VK---KEGPR-M-LV---VLN-F--A-EV--HPL-ED-QL-LG--S--LT-----I---	448
COL.CM..CGU1	--SSYVVK-VWN--LE-LKEAPT-WT--GCS-KLGA---WINSK-EEE--N-HEGS--Q---TGV---KHGPKRM-LV---L-L---T-Q-YDQLPSSRGE--QAGMA---IHV--C-----I---	680
COL.CM.00.COL243	0
COL.CM.01.COL247	0
COL.CM.99.COL11	0
DEB.CM.01.DEB1083	0
DEB.CM.01.DEB1161	0
DEB.CM.01.DEB1014	0
DEB.CM.99.CM40	IH--H-VK---E-SE---DADT-----E-L----V-EW-K-S-KC-----EEGP-M--V-----L---LEH-S-TEHK--EKV-QA-QG-QI-L-----I---D	705
DEB.CM.99.CM5	-H---V---E-SE---DADT-----E-M---ISEW-K-G-KS-----EEGP-M--V-----L---LLEH--TEHE--EK-QA-QE-Q-LN-----I---D	705
DEB.CM.99.DEB1081	0
DEN.CD..CD1	----P-L-NF-N--SE--DKAPT--T-----QT-Q-R---VATN-K---T-E-----IGV-M-Q-PN---V-----L---T-EHQ--E---QAMQG--EI--T-----I---T	727
DRL.-.-.FAO	I---H--K---N-LS-----E-----SKVT-L-----V-S--KE--I-I-E-----Q---Q---K---P---V-----SSS-EI---P-N---QM-S-----LN-----I---	714
GRV.ET..GRI_677	----L-VK---S--E-K-D-----SKVT-L-----SE--KSRIRE-----Q---T-VKM--E---EN--V-----V-N-LTAC-Q--NSPL-E---QA-M--RQ--LQ-----I---T	722
GSN.CM.99.CN166	IH--H-L-Q--T--SE-LEDAVT-----T--L---V-NT-KHRAIE--E--Q---H-V---KEGPP-M-LV---V---LQSQ-EV-TSAL-E--Q--L---I-LS-----I---T	711
GSN.CM.99.CN71	-H--H-L-Q--T--SE-LENAVT-----T--L---V-NT-KYRAIE--G--Q---Q-V---KEGPPRM-LV---V---LQSQ-EV-TSPL-E--Q--L---LS-----I---T	704
GSN.CM.99.GSNCN7	0
LST.CD.88.447	I---E-QE--S-EAE--E-VDT-----EKVG-T---V-QS-KE--KE-TD---Q---V---K-R--V---V-KVLSQR--TEHPV-K---CK-DQ-LG-----I---Q	717
LST.CD.88.485	I---E-QE--S-EAE--E-VDT-----EKVG-T---V-QS-KE--KE-TD---Q---V---K-R--V---V-KVLSQR--TEHPV-K---CK-DQ-LG-----I---Q	717
LST.CD.88.524	I---E-QE--S-EAE--E-TDT-----E-VG-T---I-QS-KE-IKE-QD---Q---V---Q-K-T--V---V-KVLSQR--TEHP--KN--K--Q--D--L---LG-----I---Q	718
LST.KE.-.lho7	I---E-QE--IWEAE--I-VDT-----EKVG-T---I-QS-KE--KE-TD---Q---V-M-Q-N---V---V-K-LSQR--TEHP--KD---QCKQ-DQ-LG-----I---Q	717
MAC.US.-.239	I---P-V--VFN-----E-ET--T--SC-KQ-----I---KD--KV-Q---Q---F-M-T---P-A--V---V---T-C---ESRL-N---M---SEI-V-----I---Q	711
MND-1.GA..MNDGB1	I---M-E-E--S-SPE-LE-VET-----D--M---I---F-R-EEYL---QT--H-VK--E---Y--V---VV--L-SR---T-HP--KE---LMKG--KI-LS-L-----I---	709
MND-2.-.5440	I---P-----N-L-----D-----Y-S--KS-----E-----T--KM--Q---PR--V---L--LSAA-DQ---P--RE---LM-G--EI-LS-----I---	713
MND-2.CM.00.MND2S46	0
MND-2.CM.00.MND2S6	0
MND-2.CM.01.MND2S109	0
MND-2.CM.98.CM16	---H--G--N-LRE-V---D-----V-A-NKSR-I-----KM--Q---P---V-----LSAA-DQ---P--RE---LM-H---L-----I---	712
MND-2.CM.99.MND254	0
MND-2.GA.-.M14	I---P-----N-L-E---D-----Y-A--KS--I-----L-E---K---PRA--V---L--LSAT-DR---P--RE--NQM-A---L-----I---	713
MON.CM.99.L1	-YS-H-L-Q--T--Q--E-APT-----T-----V-NF-AR-A-T--Q-----V---GPPQM-----VL--L-SC-EI-ESP--EA--Q--L--DQ-FLS-----I---	707
MON.NG.-.NG1	-HS-H-L-Q--T--QX--EHAPT-----A-----V-ST-QR--IP--Q-----V---QNGPASM--V---V---L-SK-EV-ESP--EK--QQ-LG-KQ--LS-----I---	710
MNE.US.-.MNE027	I---P-V--VFN-----E-ET-----SC-KQ-----I---KD--RV-Q---Q---F-M-T---P-A--V---V---T-C---ESRL-N---M---TEI-V-----I---Q	707
MUS.CM.01.1085	-H-TH-LKQ-FT--PE-LKEAPT-----V--L---V-NT-KE--IS--S-----V---KEGPPRM-LV---VL--V-SQ-Q--TSPL-E--Q--LS--I-LS-----I---	701
MUS.CM.01.MUSS1239	0
OLC.CI.-.SIVolc_97IC_12	--NSENQ-TFWE--PA-M.EVETF---Y--K-----VSSG-KH-IIHIQESS---IR--IE---K-E--V---L---VLEK-EV-EEPLI-EL-K--E--RI--G-----L---Q	451
RCM.GA.-.GAB1	---P---G-N-----E-----I---V-N--KE--KE-E-----V---K--P---V---VY--LEAQ-DT--SGL-TE--NQM-G---LS-----I---	713
RCM.NG.-.NG411	---P---N-----E-----L---V---KE--KE-----G-V---K---KR--V---RHVF--L--S-DT-ESP--L--Q---QM-G--E--LS-----I---	718
SMM.SL.92.SL92B	I---P---VFN-----EK-E--I--SC-----V---KE--LP-QA--Q---Q-L---K--P---V---VLN--T-Q-S---SD--A---Q-VQ-----G---I---N	704
SMM.US.-.H9	---P-V--VFN-----E-Q-AETF--RSC-Q-R-----V---RG-TKL-Q---Q---FY--A--P-A--V---V---Q---ESRL-N---KM---I-V-----XI-E-Q	707
SMM.US.-.PGM53	---P-V--VFN-----E-Q-AETF--SC-Q-R-----V---RD-AKP-Q---Q---FY--A--P-A--V---V---Q---ESRL-N---M---I-V-----I---Q	707
STM.US.-.STM	---P-V--VFN-----E-LE-TETF--SC-Q-----I---KN--K--Q---Q---FAM--A--P-A--VV---V---T-Q---ESKL-N---M---I-V-----I---Q	707
SUN.GA.98.L14	---K-EQE--T-ATE-V---DT-----EKLE-R---I-QG-RSR-KK-----Q---KM--E--R-S--V---LRLLSKR---T-SEL-KE-V-LIRQ-DQ--LG-----I---Q	717
SYK.KE.-.KE51	-H--I-V-WY-N--SE-V.EAETF-----Q---I---K-TLKR-----V-M-Q---PE--V---L--LMNC--S-EHPV-E--MQKAME--KI--T-----I---Q	739
SYK.KE.-.SYK173	I---P-V-WY-N--PE---EAETF-----D--T-N---VASD-T-R-QY--Q---Q---GL-M-Q--KD---VV---SY--LMTC--NTEHP--E--Q-A---I-VT-----I---	734
TAL.CM.00.TAL266	0
TAL.CM.00.TAL271	0
TAN.UG.-.TAN1	I---P--K--N-----KAK---C---C-----L---F--NT-ESQ-IE-D---Q---Q-V---KH-K---V---V---LTAQ-Q--SVL-N-----M-N-D-I-LT-----I---Q	719
VER.DE.-.AGM3	---P-VK---T-T-E---K-D-----C---R-----I-QY-K-R-EK-----Q---M--KM--E---PN--V---LTAQ-Q--SPLIE---ALMVQ-HQI-LQ---D---I---	716
VER.KE.-.9063	-V-P-VK---T-T-E---D-----C---Q---I-QY-K-R-EK---Q---Q---V--KM--E---E---V---LTAQ-Q--SPL-E---ALMV--R-I-LQ---I---	716
VER.KE.-.AGM155	-V-P-VK---T-T-E---D-----C---Q---I-QY-K-R-EK---Q---Q---T--KM--E---P---V---LTAQ-Q--SPL-E---AQMVQ---I-LQ---I---	719
VER.KE.-.TYO1	-P-LT---T-T-E---K-D-----C-----ISQY-K-R-ET---Q-K-T--KM--E---PN--V---LTAQ-Q--SPL-E---ALM-Q-QQI-LQ---I---	718
WRC.CI.97.14	---F-A-E---LE-L-EVET--I--S---D-----WSKSKE---SVQ---QT--I-LK--E-G-Q-M-----V---VKDH--M-ESPL--E--L-KG--EI--D-----I---T	452

Pol p15 RNase H // Pol p31 Integrase		
SAB.SN.-.SAB1C	EIDKLVSQG.IR..QVFLFLDRIE.FAQBEHDK.YHANWRSMQOEFGLPAIVAKEIVAACPCKQIKG.ESVHGQVDASPGVWQMDCTHLEGGKIIIVAVHVASGFIEAEVIPQETGKATAHFLKLKLSRWPI	856
H1B.FR.83.HXB2	QV-----A-...K----G-D.K-D-E-...S--A-ASD-N--PV-----S-D--L--.AM-----C--I--L-----V-L-----Y-----A--QE--Y-----G--V	824
CPZ.CD.-.ANT	QV-----...EG-D.K--D--...S--LAD-YN--P-----I-Q-D--HV--.AR-----C--I--V-----V-----S-----MAD--S--Y-----V	824
CPZ.GA.-.CPZGAB	QV-----A-...K----G-D.R-----ER...S--KA-ASD-N--P-----H-D--V--.AM-----C--I--V-----V-----Y-----A--QE--Y-----G--V	822
H2A.DE.-.BEN	-V-H-----...EK--.P-----E-...SIIKELTHK--I-LL--RQ--NS-AQ--Q--.AI--N-EI-----Y-----S-RQ--L-----	848
ASC.-.-.Qu	-V-----S-...G-N.Q-----ER..T--KYLRTD-QI-TL-----INH-----SQ--.PKT--N--L-L-----VVLV-----YVW-RI--N--RI--LE--N-CAT--V	572
COL.CM.-.CGU1	L--QK-...V-...MWI-K--.A-E-D-Q-.F-S-VQYLKEQ---TV---WER-SE--N--.QA---L-Y-Y-L--L---E--V-L---CTL-CW-TILKR---EE-GRA-I---Q-EV	802
COL.CM.00.COL243XWI-K--.A-E-D-Q-.F-S-VQYLKEQ---TV---WER--D--NR-.QA---L-Y-Y-L--L---E--V-L---CTL-CW-PILKR---EE-GRA-I---Q-EV	111
COL.CM.01.COL247XW-NK--...EVD-Q-.F-G-IQY-KEQY---TL---WER--E--NR-.QPI---L-Y-Y-L--L---E--V-L---CTL-CW-TILKR---EE-GRA-I---Q-EV	111
COL.CM.99.COL11XWIEK--.A-E-D-Q-.F-S-VQYLKEQY---TL---WER--E--NR-.QP---L-Y-Y-I--L---E--V-L---CTL-CW-IILRR---EE-GRA-I---Q-EV	111
DEB.CM.01.DEB1083X-E-V-...D-E-...G-KQLRD-Q-TLI-Q-ITQ---T--.PM-----M-I--I---M-N-----Y--TKIL-A---E-L--Q-GA--V	111
DEB.CM.01.DEB1161X-E-V-...D-E-...G-KQLRD-Q-TLI-Q-ITQ---T--.PI-----M-I-----M-N-----Y--TK-L-A---RE--L--Q-GA--V	111
DEB.CM.01.DEB1014X-E-V-...D-E-...G-KQLRD-Q-TLI-Q-I-Q-----T--.PT-----M-I-----M-N-----Y--TK-MAS---E-L--Q-GA--V	111
DEB.CM.99.CM40	QV---K-...RI--E-V-...D-E-...G-KQLRD-Q-TLI-Q-ITQ---TR--.PT-----M-I-----M-N-----Y--TKIL-A---E-L--Q-GA--V	829
DEB.CM.99.CM5	QV---R-...RI--E-V-...D-E-...G-KQLRD-Q-TLI-Q-I-Q-----R--.PM-----M-I-----M-N-----Y--TK-LSA---E-L--Q-GA--V	829
DEB.CM.99.DEB1083X-E-V-...D-E-...G-KQLRD-Q-TLI-Q-I-Q-----R--.PT-----M-I-----M-N-----Y--TK-LSA---E-L--Q-GA--V	111
DEN.CD.-.CD1	-----K-...I--EQ-P.Q-----S--KEIRDKY---TM---INV---TH--.PIR---L-T-----N-V---NA---YL-TKIL-A---E-L--QI-A---	851
DRL.-.-.FAO	-E-----R-...NM--.K-----L-N---LA-----G-----Q-----H--.PI-----T-----H---EK--Y-----G--V	838
GRV.ET.-.GRI_677	-----K-...I-----DD-A--N---V-----N-----R--.PK-----IET-----V-----R---E-----LA--V	846
GSN.CM.99.CM166	-V-----K-...EN--.P-V-D-E-...S--KYLRDQYKI--LL---NK-S---VH--.PK---N-EL-M-----V-L-----YVW-RI---RQ--LK--E--AT--V	835
GSN.CM.99.CM171	-V-----R-...MEN--.P-V-D-E-...S--KYLRDQYKI--LL---NK-S---H--.PK---N-EL-I-----V-L-----YVW-RI---RQ--LK--E--AT--V	828
GSN.CM.99.GSN17X-EN--.P-I-D-E-...S--KYLRDQYKI--LL---NK-S---H--.PK---N-EL-I-----V-L-----YVW-RI---RQ--LK--E--AT--V	111
LST.CD.88.447	-V-H---K-...QK---GK--.P---YS-.F-N-AKDLEEK-HI-PM--Q--NS-N---T--.AIT---T--I--I---M-QV--N---MV---D---T---C---V	843
LST.CD.88.485	-V-H---K-...QI---EK--.P---S-.F-N-AKDLEEK-HI-PM--Q--NS-N---T--.AIT---T--I--I---M-QV--N---C-MV---D---T---C---V	843
LST.CD.88.524	--Q--IK-...TK--M--EK--.P-V--G-.F-N-AKDLEEKY-I-PM--Q--NE-A---Q--.AIT---V-V-I--L---Q-MN--M---MV---D---T---C---V	844
LST.KE.-.lho7	-V-H---K-...QK--M--EK--.P-V--S-.F-N-AKDLEEK-N-PM--Q--ND-AN--K--.AIT---V-V-I--L---QV--N---MV---D---T---C---V	843
MAC.US.-.239	--H---X-...EK--.P-----S-VKELVFK---R--RQ--DT-D--HQ--.AI---ANSDL-T-----RQ--L-----G--V	835
MND-1.GA.-.MNDGB1	Q-----S-...K--QN--.P-----E-...S-EAQLREK-H--L--Q--QS-S--CHH--.PIK--T--L--I---NQ-----MK---TA---K--E---AQ--	833
MND-2.-.5440	Q-----K-...EG-D.K-----N--ALS-D-N--P-----Q-----T--.P---N-D--T-----Y-----A---E-----G--V	837
MND-2.CM.00.MND2S46X-EG-D.K-----S--ALA--SI-P-----Q-----V--.PI-----T-P-----V-----YL---L-S--RE-----G--V	111
MND-2.CM.00.MND2S6X-EG-D.K-----N--ALS-D-NI-P-----Q-----V--.PI---N-D--T-----V-A-----YL---A---E-----G--V	111
MND-2.CM.01.MND2S109X-EG-D.K-----N--LA-D-CI-N-----Q-----T--.PI-----A-T-----M--V--A-----YL---T---RE-----G--V	111
MND-2.CM.98.CM16	QV---R-...V-...EG-D.K-----N--ALA-D-CI-N-----Q-----T--.PI---T---T-----M--V--A-----RYL---T---E-----G--V	836
MND-2.CM.99.MND254X-EG-D.K-----N--ALA-D-CI-N-----Q-----T--.PI---A---T-----M--V--A-----YL---T---E-----G--V	111
MND-2.GA.-.M14	Q-----K-...EG-D.K-----N--LS--SI-P-----Q-----V--.PI---A---T-----M--V--A-----YL---A---E-----G--V	837
MON.CM.99.L1	-V-R--G-...MENL-.P-R-D-E-...S--KYLRTDYHI-TLL---NH-H---NH--.PKS--N-EV-----S--L-----YVW-KIL-R---C-GIA--E--AM--V	831
MON.NG.-.NG1	-X---X-...MEN--.X-R-D-E-...S--KYLRTDYI-TLL---NH-H---HTQ--.PK---XSDI-I-----Y-S-L-----YVW-KIL-R---E-GIA--E--AT--V	834
MNE.US.-.MNE027	--H---X-...EK--.P-----S-VKELVFK---RL--Q--DT-D--HQ--.AI---NSDL-T-----RQ--L-----G--V	831
MUS.CM.01.1085	-V---NS-...MEQ--.P-K-D-E-...S--KYLRDQYKI--LL---NH---TH--.PKT--N-EI-----L-----YMW-KIL-R---RQ--LG--E--L--A	825
MUS.CM.01.MUSS1239X-MEQ--.P-K-D-E-...S--KYLRDQYKI--LL---NL---TH--.PKT--N-EL-----L-----YTW-KIL-R---RQ--LG--E--AL--V	111
OLC.CI.-.SIVolc_97IC_12	-V-Q---K-...TK--M--AD--.P---S-N--.Q-AQ-L---QI-LV--R--DE-EE-KG--.VPR--I-AVGVYR--I-----V-CL-A-TG---LYG-ILRA-D-EN-VK-V-R---Q--V	577
RCM.GA.-.GAB1	-V-----K-...G--K-----E-...N--ALAED-QI-Q-----Q-----V--.AI-----T-----Y-----A---E-----A--V	837
RCM.NG.-.NG411	QV-----...G-G.K-----E-...S--A-AE-QI-Q-----Q-----V--.A-----T-----Y-----A---E-----A--V	842
SMM.SL.92.SL92B	-V-R-----...ES--.P--D---S-VKELA-KYNI-QL--Q--N--N--Q--.AI---TN-EV-T-----V-----R---RQ--L---I-----	828
SMM.US.-.H9	-V-H---E-...EK--.P-----E-...S-VKELVFK-X-RL--Q--DT-D--HL--.XI---NXXL-T-----G---X--RQ--L-----	831
SMM.US.-.PGM53	-V-H---E-...EK--.P-----E-...S-VKELVFK-I-RL--Q--DT-D--HQ--.AI---N-EL-T-----RQ--L-----	831
STM.US.-.STM	-V-H-----...EK--.P-----E-...S-VKELVFK-I-RL--Q--DT-D--HQ--.AI---N-EL-T-----RQ--L-----V	831
SUN.GA.98.L14	--Q-----...KR--M-IEK--.P-V--G-.F-N-AA-L-EM-DI-LV--Q--NE-AQ--Q--.AIT---V-I--I---M-E-V-----YMW--L-N-Q--T--TW---CAM--V	843
SYK.KE.-.KE51	-V-Q---K-...KI---N-P.K-----E-...T-VEYIR--H-RQ--A-A-EM-----R--.PK---TDL-T-----V-C---NT---YT-TKILKR---QE-GL---QI-A---	863
SYK.KE.-.SYK173	AV---K-...KI---E--P.Q---D-ER...S--MEYLR--H-RQ--A-IQQ---NR--.PK---VDIYN-----E--V-C---NT---Y--TKILKR---DE--L--MQI-----	858
TAL.CM.00.TAL266X-VEQ-P-----ER..N---DLKAR-K-T---A-IE-----VQ--.PKT--NN-AV-T-----Q-C-----Y--TKIL-R---RE--L--QV-----	111
TAL.CM.00.TAL271X-MEQ-P-----ER..N---DLKAR-K-T---A-IES-----VH--.PRT--TN-DV-T-----QV-C-----Y--TKIL-R---RE--L--QI-G---	111
TAN.UG.-.TAN1	-----M-...M-----S-----T-QFIRDA--I--L-----R--.PI-----V-----L---AR---E--K---II-----	843
VER.DE.-.AGM3	-----M-.KI---EK---ER..N---NLADTY-Q-----M-----M-----V-----R---E--K---IL-----	840
VER.KE.-.9063	-----R--IG---R--S--KNLADT-Q-----M--NAK--.PI-----V-----V-----R---RE--K---IL-----	840
VER.KE.-.AGM155	-----K-.V-.RI--IG---R--S--NLADT-Q-----M-----V--.PI-----I-----I-----R---E--K---IIG-----	843
VER.KE.-.TYO1	-----K-.R--EK---K-ER..N---KNLADTY-Q-----M-----M-----T-----I-----K-VV-----R---E--K---IL-----	842
WRC.CI.97.14	--R---A-...SK-.MV-EN-T.P-I-S-T-.QG-KAL-E-H-QLI--R-IEE-ST--G--GAPM--M--YEV-T-----LN-----YL---KI-SA-DAYQ--L-T-Q-----	578

SAB.SN.-.SAB1C	TQLHTDNGTNTFTSQVAAICWVGKIEHTFGVYPNPQSQGVVESMKNQLKEIGQIRDDAERLETAVIMAVHIHNFKRKGGIGGYSAAER.LINIIHTELETKTLQOKI..SKIQNFVRYREGRDPVWKG	983
H1B.FR.83.HXB2	KTI-----S---GAT-R-A---AG-KQE--I-----E-K---V--Q--H-K---Q---F-----G--.IVD--A-DIQ--E--KQ-.T-----DS-N-L---	951
CPZ.CD.-.ANT	KTI-----A---AA-K-A---AN-QQE--I-----Q---Q-K---V-----TPGQ-.ILD-LA-DIQ-TQ--NQ-.L---Q---H--DS-----	951
CPZ.GA.-.CPZGAB	KTI-----P---AA-K-A---AD-KQE--I-----L--E--K---V--Q--H-K---Q---F-----T-G--.I-D--A-DIQ-SE--KQ-.L-V-K-----DS--I---	949
H2A.DE.-.BEN	-H-----P---E-KMVA--VG--QS-----A--HH--NQ-SR--EQ-NTI--I-L-----CM----R---DMTP---.---M-T--Q-IQF--R-N..-NFK--Q-----QL---	975
ASC.-.-.Qu	HHI-----P--I--EF--AV--AN-----T-----I-NV-----KT-AKV--EVQ	632
COL.CM.-.CGU1	R-V-----P--V--HFV-AV--LG-A--T-H-----QR--DV-RK-KMK-Q--T--SK-A---YAL-----L--K-PW--QVERA-IE.-D-QN-.T-LQNQ-FK--KA-WK-H.TGE-Q-	929
COL.CM.00.COL243	R-V-----P--V--HFV-AV--LG-A--T-T-----I--QR--MI--SLRKM-K-Q--TF\$	173
COL.CM.01.COL247	K-V-----P--I--HFV-AV--LG---T-A-----I--QR--MI--S-KMK-Q--T--	174
COL.CM.99.COL11	K-V-----P---HFV-AV--LG---T-T-----I--QR--MI--S-RMK-K-Q--T--	174
DEB.CM.01.DEB1083	K-I-----P--I--P-K-A---LGV---T-----NK-RV--DT-K--E-----	174
DEB.CM.01.DEB1161	K-I-----P--I--P-K-A---LGV---T-----K-RV--T-K-----	174
DEB.CM.01.DEB11014	K-I-----P--I--P-K-A---LG---T-----K-RV--T-K--E-----	174
DEB.CM.99.CM40	--I-----P--I--P-K-A---LG---T-----K-RI--DT-K-----A--TY-I---R---DM-PI--.V-M-T--I-----..LLG-K-----A--T---	956
DEB.CM.99.CM5	K-I-----P--I--P-K-A---LG-D--T-----K-RV--T-K--E-----A--TF-I---R---DMA-I--.V-M-T-D--Q-----..LLGYK-----A--T---	956
DEB.CM.99.DEBCE1	--I-----P--I--P-K-A---LG-D--T-----K-RV--DT-K--E-----	174
DEN.CD.-.CD1	-K-----P--I--AMQ-A--AN-K--T-----I--NK-----V-ALV-E-----LA--TL-L---R--L-DMT---.Y--MLN-H--IQHTK.QP...FS--K-----AN-Q-Q-	976
DRL.-.-.FAO	SH-----P---EK--TV---A---T-----N---K-----Q--K---Q---L-----E---.IVD--ASD-L-SK---N-.L-----QL---	965
GRV.ET.-.GRI_677	KH-----P---N---V---N---T-I-----S---R-----C-----Q--T-----I-S---.V-MLT-Q--LN---NQ-.Q--L--K-----	973
GSN.CM.99.CN166	-H-----P--V-KELE-A--AD-Q--T-----N-----T-TK--EEVTY---AQ-CY-----DMCPT--.IV-M-----QH-NTQN..F-K-----Q-AN-L-Q-	962
GSN.CM.99.CN71	-H-----P--I--KELE-A--AN-Q-ST-----N-----T-QK--EVTY---AQ-CY-----DMCPT--.V-M-----QH-NTQS..F-K-----Q-AN-H-Q-	955
GSN.CM.99.GSNCN7	-H-----P--I--KELE-A--ANVQ--T-----N-----T--K--EEVAY--	174
LST.CD.88.447	Q-V-----P--V-KE-Q-VI--LG---T-I-----AK--A--Q--ERV-E--QQ-K--L--T-----QR--L--L-P---.F--M-NAD---QY--KLN..--LK-K---Q---Q---	970
LST.CD.88.485	Q-V-----P--V-KE-Q-VI--LG---ST-I-----AK--V--Q--ERV-E--QQ-K--L-----QR--L--L-P---.F--M-NAD---QY--KLN..--LK-K---Q---Q---	970
LST.CD.88.524	K-I-----P--V-KD-Q-VT--LS---T-I-----AK--V--Q--ERV-E--QQ-K--L-----QR--L--L-P---.F--M-NAD---QY--KLN..--LK-K---Q---Q-R-	971
LST.CD.88.524	K-I-----P--V-KE-Q-VI--IG---T-I-----AK--V--S--ERV-E--QQ-K--L-----QR--L--LTP---.F--M-NA---QY--KLN..--LK-K---Q---Q---	970
LST.CD.88.524	K-I-----P--V-KE-Q-VI--IG---T-I-----AK--V--S--ERV-E--QQ-K--L-----QR--L--LTP---.F--M-NA---QY--KLN..--LK-K---Q---Q---	970
LST.CD.88.524	K-I-----P--V-KE-Q-VI--IG---T-I-----AK--V--S--ERV-E--QQ-K--L-----QR--L--LTP---.F--M-NA---QY--KLN..--LK-K---Q---Q---	970
MAC.US.-.239	-H-----A--A--E-KMVA--AG-----A--HH--NQ-DR--EQ-NSV--I-L-----CM----R---DMTP---.---M-T--Q-IQFQ-S-N..-FK-----QL---	962
MND-1.GA.-.MNDGB1	SK-----P---E-ETM--LG---I-----NK--Y--L-EK--E-CKE-K--A--TF-----QR--L--MT-G--.IV-M-N---YQYQ-NQ-.NL--K--F-----QL---	960
MND-2.-.-.5440	KH-----P--V-EK--TV---AQ---T-----AK-HH--K--E-V--Q-----Q--L-----E--PG--.IVD--T-DIL-TK-R-N..-Q-----QQ---	964
MND-2.CM.00.MND2S46	KH-----P--V-EK--TV---AQ---T-I-----AK-HH--K--E-V--Q-----	174
MND-2.CM.00.MND2S6	KH-----P--V-EK--TV---AQ---T-----AK-HH--K--E-V--Q-----	174
MND-2.CM.01.MND2S109	KH-----P--V-EK--TV---AQ---T-I-----AK-HH--K--E-V--Q-----	174
MND-2.CM.98.CM16	KH-----P--V-EK--TV---AQ---T-I-----I-AK-HH--K--E-V--Q-----Q--L-----E--PG--.IVD--A-D-L-TK--HN-.Q-----QQ---	963
MND-2.CM.99.MND254	KH-----P--V-EK--TV---AQ---T-A-----AK-HH--K--E-V--Q--K--	174
MND-2.GA.-.M14	KH-----P--V-EK--TV---AQ---T-I-----AK-HH--V--E-V--Q-----Q--L-----E--PG--.IVD--T-DIL-TK--N-.Q-----EQL---	964
MON.CM.99.L1	--I-----P---EFV-AA--AN-H--T-----N--R---T-TK--EV-Y-P--LAQ-LF-L-----DMPVD-.F--M---LQ-SNNQ-.T-FSK---F-T-A--Q---	958
MON.NG.-.NG1	HHV-----P---EFV-AA--AN--ST-----N---I--TTL-K--EVQF-S--LAQTLY-L---X---DI-PI--.V-M---D--LQHSNNQ-.T-FSK---F-T-A--S---	961
MNE.US.-.MNE027	-H-----A---E-KMVA--AG-----A--HH--NQ-DR--EQ-NSM--I-L-----CM----R---DMTP---.L-M-T--Q-IQFQ-S-N..-FK-----QL---	958
MUS.CM.01.1085	--I-----A--I-AEFA-A--AS---T-----NK-----A-HK--EEVTY---AQ--F-L--K---V-DL-N--.I--ML---LQH--QK..Y-F-----T-S--S---	952
MUS.CM.01.MUSS1239	--I-----A--I-EEFG-A--AS-D--T-----NK-----T-QK--EEVTY--	174
OLC.CI.-.SIVolc_97IC_12	KII-----VN-KMQ-V---LG---ET-I-W-----II-AK-RV---L-R--EQ-TY---LA--IF-----R-	656
RCM.GA.-.GAB1	RK-----A---AA-Q-V---AQ---A-----I--E-V-EQ--K---Q---LV-----G--.I-D--A-D-A-NK--NQ-.L-----QL-R-	964
RCM.NG.-.NG411	KH-----A---AA-Q-D---AQ-----H--I-----Q--K---Q---L-----G--.I-D--ASD-Q-TK--NQ-.L-----F-----QQ---	969
SMM.SL.92.SL92B	KH-----A---E-KMVA--LGV-QS-----A-DLH--KN-DK--EQ--SV--L-L--A-CM---R---DMTP---.IV-M-T---QY-NSQN..F-----QL---	955
SMM.US.-.H9	-H-----A---E-KMVA--AX--Q--X-----A--HH--TQ-DR--EQ-NSI--I-L--I-CI---R---DMTP---.V-M-T--Q-IQFQ-S-N..-FK-----QL---	958
SMM.US.-.PGM53	-H-----A---E-KMVA--AG--Q-----A--HH--TQ-DR--EQ-NSI--I-L--I-CI---R---DMTP---.V-M-T--Q-IQFQ-S-N..-FK-----QL---	958
STM.US.-.STM	-H-----A---E-KMVA--AG--Q-----A--HH--TQ-DR---Q-NTV--V-L---CM---KR--L-DMTP---.V-M-T--Q-IQFQ-S-N..-FK-----QL---	958
SUN.GA.98.L14	K-I-----P--I-KD-E-V--LG-Q--T-I-----AK--V--Q--SR--E--QE-K--L--L-----QR--L--MT---.F--M-NAD---QY--KIN..--LK-K---Q---Q---	970
SYK.KE.-.KE51	KHI-----P--I-DAF--A---AGV---T-I-----I--NK-R---T-N--EEV---A--TF-L-----TTP-D..Y--MLY---QLQ.NTHT..Q-FS--K---H--TSD-Q-	987
SYK.KE.-.SYK173	K-I-----P--V-DKFK-A--CG---T-I-----I--K-RY--A-S---VTH-Q--A--TF-L-----I-PG--.Y--MLY---QLQNTTTS..P-FS-----Q.GKNE-Q-	983
TAL.CM.00.TAL266	SH-----P--V-AEMQ-MV--L---ST-----A--NK---KT-T---EVQY-S	174
TAL.CM.00.TAL271	SH-----P--V-EM--MV--L---T-----N--NK---KT-Q---EVQY-S	174
TAN.UG.-.TAN1	-K-----P--V--E-QT---QV---T-I-----S---R---T-EK--E-CAF---L--C-----MTP---.---M-T-Q--IQHI-TQQ..Q--S--K-----E---	970
VER.DE.-.AGM3	-----P---E--M---T-----SI-----K---CQYT--L--C-----LTP---.---M-T-Q--LQH--T--.Q--L-----	967
VER.KE.-.9063	V-----P---E---T-----S-----K---CQYT--L--C-----LTP---.---M-T-Q--LQ--T--.Q--L-----	967
VER.KE.-.AGM155	-H-----P---E--M---V---T-----SI-----K---CQYT--L--C-----LT---.---M-T-Q--IN--T--.Q--L-----	970
VER.KE.-.TYO1	-----P---E---T-----SI-----K---CQYT--A--L--CIL-----QTS---.---T-Q--IQH--T--.Q--L-----	969
WRC.CI.97.14	KKI-----K--KMQ-----LG--QE--I-----I--K-RR--D--HR--E--K---LA--LF-----R-	657

	Pol p31 Integrase end \	
SAB.SN.-.SAB1C	PAKLIWKGEHAVVIEQEG.ELKTIPIRRKAKIIKDYGK...ALDSQAPLEGNRGTAGEVD\$	1038
H1B.FR.83.HXB2	---L-----DNS.DI-VV-----R---Q.M-G-DCVASRQDED\$	1002
CPZ.CD.-.ANT	--Q-L-----KD-E.-I-VV-----E-R-K.IEDRDDLAGRQ-ED\$	1002
CPZ.GA.-.CPZGAB	--T-L-----D---VV-----R---Q.M-G-DCVASRQ-ED\$DL-\$	1003
H2A.DE.-.BEN	-GE-L-----IVKVGT.DI-VV-----R---GR.QE---SPH---ARED.--MAC.PCQVPEIQ	1040
COL.CM.-.CGU1	-GE-V-----RNSQGT-FVK---V--TRLQYGEDVGSENLLSNGQKEAETVKGMD\$	989
DEB.CM.99.CM40	--T-L-----CKTEVGDI-VV-----V--E---D..M--KVNTQNSNES\$	1006
DEB.CM.99.CM5	--T-L-----CKTEVGDI-VV-----V-----G..V--KTDTQNPSES\$	1006
DEN.CD.-.CD1	--E-L-----KT-AGDIFPV-----P-NA.....ESRKGEQNKGMSETDI\$	1030
DRL.-.-.FAO	--E-V-----K-GT.D--V-----N.VDSHTNMES\$	1011
GRV.ET.-.GRI_677	--R-----K-GE.DI-VV-----ER.KTM--EGSM--VREANKQMEGSDSLQDQES\$	1039
GSN.CM.99.CN166	--V-L-----V-T-AG-II-V-----P--EA.KTE-VGGATHT-ND\$EEGRLVA\$	1021
GSN.CM.99.CN71	--V-L-----V-T-AG-II-V-----P--EA.K-E-VGSAAHTSNDREEGRMAD\$	1015
LST.CD.88.447	--Q-L-----VK-GD.TIFSV-----LV----E..GPK--EGG-HN-\$	1018
LST.CD.88.485	--QIL-----VK-GD.TIFSV-----LV----E..GPK--EGS-DN-\$	1018
LST.CD.88.524	--Q-L-----VK-GE.NIFSV-----LV----E..GPK--SSMDNQ\$	1019
LST.KE.-.lho7	--Q-L-----VK-GE.NIFSV-----LV----E..GPK--ESS-DN-\$	1018
MAC.US.-.239	-GE-L-----ILKVGT.DI-VV-----GG.KEV--SSHM-DT-EA.R--A--	1018
MND-1.GA.-.MNDGB1	-GI-L-----LKY-E.-I-IV---C-----E..SGKN--VN--SV\$	1008
MND-2.-.-.5440	--E-----K-GV.D--VV-----T.VDSNPHMGR\$	1010
MND-2.CM.98.CM16	--E-----K-GT.D--VV-----R---T.VDS-PNVEA\$	1009
MND-2.GA.-.M14	--E-----K-GT.D--VV-----R---T.VDSNPHMEDRQETA\$	1015
MON.CM.99.L1	--R-L-----KTEEGDIL-V-----P--TT.ENVGGDTNQYNLRKQD-LAN--	1016
MON.NG.-.NG1	--H-L-----L--KTDQG-VI-V-----RP--S..K.E-VGSKPSAHQIREVDGMAD\$	1019
MNE.US.-.MNE027	-GE-L-----ILKVGT.DI-VV-----GG.KEV--SSHM-DT-EA.R--A-.PHKISEI\$	1021
MUS.CM.01.1085	--V-L-R-----KTDSG-VI-V-----P--A..K.E-VGSKSDTGDLRKEDRLDN\$	1010
RCM.GA.-.GAB1	-----ETGD--VV-----E--RK.DVDSEANLAGRQEEN.\$	1017
RCM.NG.-.NG411	-----E-----D.E.D--VV-----C-----RKEVDRETNMGEGRQEE\$	1021
SMM.SL.92.SL92B	--E-L-----KVGT.-I-VV-----GR.QEMG-S-SM-DQ\$KNS-ALA--	1011
SMM.US.-.H9	-GE-L-----ILKVGT.-I-VV-----GG.KE--GSH--DT-EA.R--A--	1014
SMM.US.-.PGM53	-GE-L-----ILKVGT.-I-VV-----GG.KE--GPH--DT-EA.R--A--	1014
STM.US.-.STM	-GE-L-----IVKVGT.DI-VV-----GG.KEV--GSH--DTREA.---A--	1014
SUN.GA.98.L14	---L-----KQGE.NILVV-----LV----G.....ESSSV-MV-\$	1015
SYK.KE.-.KE51	--S-L-----TPDQQVIIV-----SSD-ER...V--GTH--ITSKSN\$	1039
SYK.KE.-.SYK173	--R-L-----V-TEEGDIFAV-----T-H-E...RM--GSHV-NDPK-D\$	1035
TAN.UG.-.TAN1	--T-L-----K-GE.---VV-----E--ER.KTVGDKTNM--\$	1018
VER.DE.-.AGM3	-GQ-----KGGV.---EY-----EPR.KRMGDESN--A-GADN\$	1021
VER.KE.-.9063	--Q-----K-GE.---IV-----EPR.KRMGDEGNM--LRGPDNQMAR.NSQILDD\$	1032
VER.KE.-.AGM155	--R-----LK-GE.---VV-----EPR.KT-GDETH--A-GSDHQMAG.DS\$	1030
VER.KE.-.TYO1	--Q-----LKDGS.D--VV-----EPK.QRVGNEGDV--TRGSDN\$	1023

SAB.SN.-.SAB1C	M.....EKHWIVRP...LWKVTGG..QOERWTSLVKYHMHVS.KQCQVHW.RYTPH.TKIRWNW.YSYQEWWIPLKD.GALIKVT.NYWH.LTPEKGWLETYA.....	85
H1B.FR.83.HXB2-NR-Q-MI...V-Q-DRM..RIRT-K---H-Y--.GKARG-.F-RH-.YESPHP.RISS-VH---G..D-R.L-ITT--GLH-G-RD-H.....	80
CPZ.CD.-.CPZANTI-Q-DRM..RINI-K---H-IWET..-VLKP-.K-RH-..YENDHP.KKGE-VH---PTLDDK.L-VTVF-GLQCG-RP-H.....	74
CPZ.GA.-.CPZGAB-NR-Q-MI...V-Q-DRM..RIKT-N-----IYR-..KARG-.F-RH-..YDHP-P.KVAS-IH--FR-..YSK.LIVTT--ALSPG-RA-H.....	81
H2A.DE.-.BENEEDRN---V...T-R-P-R...M-K-HA---LKYRT..DLEEVRYVPH-..-VG-A-.WTCSRVI-F--Q..-KSHLEIQ-A--N-.....SSH-.....	86
COL.CM.-.COLCGU1-M-A--I.CYPMG-RKPK..PLRE-I...-KQF...PE-QYYLGGKTES..PE-..-TEGKIEFY-G.SAIKL..SII-LGT---DRYHRGPPV.....	80
DEB.CM40-T---LTHKIPMNKLDL..WCSAIKR.E--KT...-ELDVTQ.WIH-F..ELR-SY-TQTKI-F-IT...KEGG-VVELLWC-----P-M-V.....	86
DEB.CM5-E---LTHKIPVKNLKD..WCSAIKR.G--KT...-ELQAIK.WIH-Y..ELRRAP-TQTKI-F-VT...EEG--VVEIFWS-A-----MPSM-V.....	86
DEN.CD.-.CD1-QNLGRENRL...T-E---LTYKIGETKLEHWMSTLKSIIYKTN.....ELKGTVIIIPHR..-T-E-..TMTKIH---R-AYMEVKL---Y-SQV-ISIAYV.....	98
DRL.-.-.FAO-M---I...T-R-ERYM..V--N-C--K-KGE-YLER-HYAPHF..QCS.G..WTHSQ-T-F..DKSK.I-VTVL-N-...D---SK-I.....	86
GRV.ET.-.GRI_677ER-Q---V...V-R-SER..-IS--RGI-T-KIRNK.QLPWEY...RH-WQVQWF-..T-SQFI---S..KDDYIEVNI-HN-...R---SS.....	83
GSN.CM.99.CM166ERQRQ-M-V-Q-...IQIM-ER..KVD..LLRATK-HI...WSGKTPFV-VH-Y..QLQHRFTQNKIRLAMDTRKVGEE.VEATY.IEIT.IL-DT-S.HGPASLSRSTYYQAVI	103
GSN.CM.99.CM71ERQRQ-M-V-Q-...IQVM-ER..KV-..LIRATK-HI...WSGKTPFV-IH-Y..QLQHRFTQNKIRLA-DTRQVGEE.VEATY.IEIT.IL-DT-S.HGPASLSRSTYYQAVT	103
LST.CD.88.SIVlhoest447ERDQR---V...AYIT-RS..LV-KFLDIRRA-K...ET-NDWVSL..HGT.GVG-EF-..T-NNI---E..TGTL..-VRC-A-L.AAGR--ISQW.....	82
LST.CD.88.SIVlhoest485ERDQR---V...A-IT-RN..LI-KFLDIRRA-K...ET-VDWVSL..HGT.GVG-EF-..T-NKI---E..IGTL..-VRC-A-L.AAGR--ISQW.....	82
LST.CD.88.SIVlhoest524ERDQR---V...A-ITNRS..LV-KFLDIKRN-S...ESKNDWVSL..HGT.GVG-EF-..T-NKLL--D..TGTL..-VRI-A-L.AAGR--ISQW.....	82
LST.KE.-.SIVlhoestERDQR---V...A-IT-RS..LI-KFLDI-RLNT...EAKENWVGL..HGT.GVG-EF-..T-HKL---E..SGTL..-VRI-A-L.AAGR--ISQW.....	82
MAC.US.-.MM239EE--R--AV...T-RIPER...L--H--I--LKYKT..DLQKVCYVPHF..-VG-A-.WTCSRVI-F--Q..EGSHLEVQG---S---.....	86
MND-1.GA.-.MNDGB1ERVER--L...T--SSQRI..-K-HW--RRQ-A.....WATANNEEG.CW--YPHFMAYNWYTCISKVVIIN	65
MND-2.-.-.5440-Q--AI...T-EDRQDV.KLRD-N-I---KYKGE-HLEN-QL-PHF..QCS.G..WTHSQKK--F..D-SK.IIITVL-NL-----SQ--F.....	87
MND-2.CM.98.98CM16-Q--I...T--LDRHA..V--N-----KYKGE-HLDK-DLKAHF..QCS.G..WTHSQKY--E.EDEQ.IIITIL-NL-----S--M.....	86
MND-2.GA.-.MND14CG-Q--I...T-RIDRRQ..L-K-N-----KYKGE-HLDK-ELF.H-F.QCS.G..WTHSQKI--F..D-SK.IIITAL-NL-----SQ--I.....	86
MNE.US.-.MNE027EE--R--AV...T-RIPER...L--H--I--LKYKT..DLQKVCYVPHF..-VG-A-.WTCSRVI-F--Q..EESQLEVQG--NL---R--S---.....	86
MON.CM.-.SIVmon_99CMCML1APR-M-V-T...TNIISES..KMD..LIRCTKWHI...LTGSAPFV-VH-Y..QLH-QRF-QNKIKL--DLGITQEGESWATY.LEIT.IY-DL-N.VGPSALSPTVYNKQAYT	102
MON.NG.-.NG1APR-M-V-S...VHIRSER..WM-..LIRATKYNI...YTGKVPFV-VH-Y..QLTHQRF-QCKIKL--EMGITQXXEQWATY.XEIT.IL-DX-N.VGPASLSPSVKWKQAYI	102
MUS.CM.-.SIVmus_01CM1085APR-M-V-S...IQETSER..KID..IIRATKYI...LSGRAAFT-VH-Y..QLYHQRF-QSKIR--TQNTRESGAVESTY.IELT.VL-DV-N.VGPASCSSKSTWQKQSPI	102
RCM.GA.-.GAB1A--M-----I-R-DRRK..I-Q-H-----QYKGE-AAKE-EYVPHF..-VP-G-.W-HS-VH---E.E-SK.LKITT--NL.-V-----GTYG.....	87
RCM.GB1c1A--M-----I-R-DRRK..I-Q-H-----QYKGE-AAKE-EYVPHF..-VP-G-.W-HS-VH---E.E-SK.LKITT--NL.-V-----GTYG.....	0
RCM.GB1c3A--M-----I-R-DRRK..I-Q-H-----QYKGE-AAKE-EYVPHF..-VP-G-.W-HS-VH---E.E-SK.LKITT--NL.-V-----GTYG.....	0
RCM.NG.-.RCMNG411A--R--E-...I--DRRK..V-Q-H-----QYKGE-EAKD-EYVPHF..-VP-G-.W-RS-VH---G.NNTK.IKVT--SL.-I-----GTYG.....	87
SMM.FGBEE--N--V...T--IP-R...L-K-H--I--LKFRT..DLQKACYVPHF..-VG-A-.WTCSRVI-F--R..DESHLEVQG--N.-A-----S---.....	86
SMM.SL.92.SL92BEEG-R-VAV...V-RISRRI..V--H-CI-F-KYKT.RELEKACYVPHF..-VG-A-.TASRVIF--E..EGSHLEVQV--N-.....SS---.....	87
SMM.US.-.SIVSMMH9EE--N--V...T-RIP-R...L-K-H--I--HLKYNT..DLQKACYVPHF..-VG-A-.WTCSRVI-F--R..DEAHLEVQG--N-.....X---S---.....	86
SMM.US.-.SMMPGMEE--N--V...T--IP-R...L-K-H--I--LKFRT..DLQKACYVPHF..-VG-A-.WTCSRVI-F--R..DESHLEVQG--N.-A-----S---.....	86
STM.US.-.STMEE--R--V...T-RIP-R...L--H--I--HLKYNT..ELSKACYVPHF..-VG-A-.WTCSRVI-F--Q..-EAHLEVQG--N-.....SE---.....	86
SUN.GA.98.SIVSUN	-EENPPQWRWDEREW-DRRQYKI.VRIV-LID..RIAV-KFLD-RRM-R...ETKDDWVSM..YGT.GTG-E-..T-NKII-VT..YGTV..-VRI-G-L--A--LINQWG.....	99
SYK.KE.-.SYK173-E--V-TWRMTPRQIDR..L-HIIKT.H--KS...-ELEKAT..-KH-Y.Q-E-Q-..T-CQ-T-VG...DGTIWIWF-HNLA.--R--HMQGI.....	85
SYK-KE51-E---VLTWKLHPNQIDK..M-HILKV.H--KS...-LEKAT..-KH-Y.Q-T-E-..WTRAQ-E-VG...QGVLIIFK-HNLT-...L-R-EGV.....	85
TAN.UG.-.TAN1	-E...R--L-VT-L...T-R-S-E..HIDK-KGI-----R.NRLQD-TYLMHY..QCG-A-..TCSRFL--G-..EGKI-VDC---Q---S---.....	85
VER.DE.-.AGM3	-N.....Q--E-VM-V...T--PEE..LITK-QGI-R-W-..R.TRKLD-KYRMHY..Q-T-A-..TMSRYE--G..QHGSIHVDL---S---.....	85
VER.KE.-.AGM155	-S.....Q---VM-L...T---QEE..VITK-QGI-R-W-..N.-RNLK-EYKMHY..Q-T-A-..TMSRY---P-..SGEIHVDI---A-KQ---S---.....	85
VER.KE.-.AGM_VER_9063	-N.....Q--E-VM-V...T--S-D..LITK-QGI--W-..T.-RNLK-QYMHY..Q-N-A-..TMCRYI--IG..LEGEIQIDC---R--S---.....	85
VER.KE.-.TYO1	-N.....PN-E-VM-V...T--P-D..LITK-QGI-R-W-..R.QRNLK-NY-MHY..Q-T-A-..TMSRY---IG..KHGEICVDL---Q---S---.....	85

	phosphorylation site Thr96	 Cys114	involved in inhibition of APOBEC 3G Cys133		
SAB.SN.-.SAB1CTGIGYSKGEWFTELDPWTADHIIHWSYFPFCFTDRAVQQAIRGEKYLWCKHQVGHQPTGQVPSLQYLALRVYTNGLRR...VAPTSRRGS 171
H1B.FR.83.HXB2LGQGVSEWRKRYSLQV--EL--QL--LY--D--S--S--IRKLLHIVSPREYANK-G-----AALITP 156
CPZ.CD.-.CPZANTLGHGVS-EWRC-KYLQV--E--QM--QY----S-Q-R--L--RI--TY-HYKK--S--GT----FKILEFRGYP 154
CPZ.GA.-.CPZGABLGHGVS-QWRL-SYVQV--F--RL--SQ--D--AET-IRR--L-QLVAPREYKE--R--G--F--KALISE 157
H2A.DE.-.BENV.....RLTWYTEKFWDVT-DC--IL--ST--S--AGE-RR-----L-SC-NYPQA-KA-----V-VQQND-PQRKGTARKQWR 173
COL.CM.-.COLCGU1SYELQIGKWE-D-SWQQVGYM-IIRGK-VRIEEEAR--LPWNP-DF---L-LK--E-C-LAL.....REKER 153
DEB.CM40AIAW.ETKTKW--T-DL--L--LR-----SQQ-R-----IL-QQ-QQY-RG...PEGKGP-I---AVKHGKVPRLTPT-PHLKN 175
DEB.CM5GISW.EQT-WQ--T-DL--L--LR-----SQQ-T--L-K--L-QQ-QQY-RG...PEGKGP-TAVKYGKVTSTLPT-PHIKN 175
DEN.CD.-.CD1HLSTGWA--V--V--E--KDH-S--RQ-R--L--VV-ST-ARKKAYHPRTGPGREP-IV-QAAIRYKGVTTITPTSISFKN 189
DRL.-.-.FAOT-EHIQEKFC-LI--T--R--LE-----Q-IRGVLL-YRIQA-YFPR--KG-G-----QAHIKY-KHGRKTVSGESSG 172
GRV.ET.-.GRI_677HGV-LS-YHQKGYK-V--G--RM--LY--N-----I-----TTFKE--KG-Q--L--VA.....YTNGIRK-SKR 165
GSN.CM.99.CM166	IEWVYNRSK.....LGAR.EGD-IWYS-N-T-GV-MQ--GK----QNEQIRR----QI-GN-E-PEA-L.....QV-TPQT-EQ--FFA-VKHGKNTQSSYALAQP 203
GSN.CM.99.CM71	IEWVYNRSK.....LGAR.E-D-TWYS-N-T-DV-MQ--TR----QNEQRR----QL-GY-D-PEA-P.....QV-TPHT-ER--FFA-VRYGKNTPKS-STLDQP 203
LST.CD.88.SIVlhoest447V.....S-EWIY-NYQ-I--IL--NQ--CK--N--ST-DIRK-LW--IMAF-AMPT--KGC-L---F-C--QLQHVQNKAAKESRKTS--LC-L---F-C--QLQHVQNKAAKESRKTS--L 170
LST.CD.88.SIVlhoest485V.....S-EWIY-NYQ-I--IL--QM--CK--N--ST-DIRK-LW--IMAI-AMPT--KGC-L---F-C--QLQHVQSKAAKESRKTS--LC-L---F-C--QLQHVQSKAAKESRKTS--L 170
LST.CD.88.SIVlhoest524V.....S-EWIY-NYQ-V--IL--QML-VK--D-WVH-DIVR--I--HS--MPT--KGC-L---C--QLQHVQTKTAAQPRKAP--LC-L---C--QLQHVQTKTAAQPRKAP--L 170
LST.KE.-.SIVlhoestV.....S-EWVYRDYQ-I--IL--QM--YK--N--IQ-DIRR-LL--RITMC-QPL--KGC-L---FIC--QLQHVQAEAAKKSRETT--L 169
MAC.US.-.MM239V.....R-TWYSKNFW-DVT-NY--ILL-ST-----AGE-RR-----QL-SC-RFPRA-KY-----K-VSDVRSQGEN--WKQWR 172
MND-1.GA.-.MNDGB1	RDIRLIVRSYWHLQIEVGLCLSTYAVSIEAVVRPPPFK-C--IT-EV--L--LHFYD--M-S-MK----EV-KV-RFPA--KAQG-L---F-C---IYGPEE-\$ASIRRSRTSQR 183
MND-2.-.-.5440TVE-Q-ENY--YI--V--R--GE-----Q-IRK-LF--RLVA-YFPW--RG--GT--F--QA-LR-RKHGREGTRGA--SR 173
MND-2.CM.98.98CM16T-E-KPTNY-HI--Q--RM--K-L--EQ-IR--LL-KRLTV-YFHW--KGK-G-----LS--AYCNRGRGRDRP--SR 172
MND-2.GA.-.MND14CGTAE-N--DYY-HI--V--RM--E-----AA-RKVLY--RIVA-YSPW--KG--GT--L---A-IKFC-NGRKSRTGTTG-R 172
MNE.US.-.MNE027V.....R-TWYSRNFV-DVT-DY--ILL-ST-----AGE-RR-----QL-SC-RFPRA-KN-----VSVYRSQRENPTWKQW-RD 174	
MON.CM.-.SIVmon_99CMCML1	IEWVYWMQK...VERVPLDPVWNPNP.RDREVRTWYA-H-T-DL-NQ-V-TH--A--QQLDRR-----L-GQ-Q-LAT-YKV-TPLT-EK--FLALVRYGATPTKSSAPLDLPD 215
MON.NG.-.NG1	IEWVYWRQPRVPQQEGVDRVWFPSRRPDREIDRVWYA-Y-T-DV--Q--TR--S--QASDRR-----AX-GN-H-IQT-YNKV-TP-T-EK--FLACPKYKCTTEPSS-LDLPR 221
MUS.CM.-.SIVmus_01CM1085	LEWVYLRRH..PQDR-Y-S-TL-MR--TQ--S--YSQIDPRVIRGQHTLGDCEHPDAHTKVPP---V--LALKKHGLKALQPSSALALQ-202
RCM.GA.-.GAB1V.....R-TWYSRKFV-DVT-DY--TLL-GT-----SEGE-RR-----L-SC--FPKA-KN-----T-VSHV-SQRED--WKQWR 172
RCM.GB1c1V.....R-TWYSEKFW-DVT-DV--QLA-ST-----AAH-R-----QV-SY-GYA-A-HSS-Q--L--K-VLQND-PKGKN--RKQWR 174
RCM.GB1c3V.....R-TWYSRNFV-DAT-DY--TLL-GT-----SEGE-RR-----L-SC--FPKA-KN-----T-VSHVXSQGEX--WKQWR 172
RCM.NG.-.RCMNG411V.....R-TWYSRKFV-DVT-DY--TLL-GT-----SEGE-RR-----L-SC--FPKA-KN-----T-VSHV-SQRED--WKQWR 172
SMM.FGBV.....R-TWYSRKFV-DVT-DY--TLL-GT-----SEGE-RR-----L-SC--FPKA-KN-----T-VSHV-SQRED--WKQWR 172
SMM.SL.92.SL92BV.....R-TWYSEKFW-DVT-DV--QLA-ST-----AAH-R-----QV-SY-GYA-A-HSS-Q--L--K-VLQND-PKGKN--RKQWR 174
SMM.US.-.SIVSMMH9V.....R-TWYSRNFV-DAT-DY--TLL-GT-----SEGE-RR-----L-SC--FPKA-KN-----T-VSHVXSQGEX--WKQWR 172
SMM.US.-.SMMPGMV.....R-TWYSRKFV-DVT-DY--TLL-GT-----SEGE-RR-----L-SC--FPKA-KN-----T-VSHV-SQRED--WKQWR 172
STM.US.-.STMV.....R-TWYTRNFVSDVT-DC--QLL-GT-----AGE-RR-----L-SC-RFTKA-KN-----K-VEHV-SQRENTARKQWR 172
SUN.GA.98.SIVSUNC.....SMEWIYNSYQ-I--LV--QM--FK--D-W-S-CIRR-ML--I-HE-RN--A-KGL-L---F-C--LH-QQERASRTA-R--A186
SYK.KE.-.SYK173RIQYQWNQWN-D-T-AV--RL--NF-----A-N--V--LLTSH-WTP-TD-----Q-LKDDGGFLOSL-ACA-NT 170
SYK-KE51GLIWAHPS-W-VT-Q--ALV-SQ-----S-L--L--LTSW-WNF-KEL-----LQSKDGTGFLQSLPAAARG 171
TAN.UG.-.TAN1VAI.SFENWQWTYK-VT-DV--M--CH-----ISF--TYKE--VAENHW--R--F--T--DF--NGRRKRFQGGKTR 179
VER.DE.-.AGM3EGI.QYLSNRDPWYR-A--SL--TH--T--E-IRK-LL-QRFTF-QFPE--KK-----LA...HONGLRQRSQRSKTG 173
VER.KE.-.AGM155VGI.QYVSLVNDKYR-N--SM--CH--T-----IS--L--NRFIF-QFPG--KL-----LA...HONGLRKRSQRGETR 173
VER.KE.-.AGM_VER_9063IGI.QYL-AVG-NFR-A--GM--LH--S--E-I-K--L--RFVFP-QFPE--KT-----K-----LA...HONGLRERSQRKQAR 172
VER.KE.-.TYO1VGI.QYVSNLESKYR-A--S--GH--N--KE-I--L--HRFVFP-QFPE--KS-----LA...HONGLRERSKRGKTR 173

	Vif end \	phosphorylated
SAB.SN.-.SAB1C	SQGSQPESQRDRTRMARNMGFAQRAVRRMAPRHVTGPQFRGVPVLPKESFPFSLVEYCGRTS.....H\$	234
H1B.FR.83.HXB2	KKIK.....P---SVTKLTEDRWKPKQKTKGHRGSHTMNG--	192
CPZ.CD.-.CPZANT	KGPR.....RQF-SL-ILTEDRWKPK-RMRGHRENQ.....	185
CPZ.GA.-.CPZGAB	RRHR.....P---SVAKLTEDRWKPKQRTKVHQENLTRNG--	193
H2A.DE.-.BEN	RDHWRGLRVA-EDHRSL.....KQGGSEFSA-RAHFPPGVAKVL-IL.....A-	215
COL.CM.-.COLCGU1	THQKMRVLKG.....TS...HGCRTR-	171
DEB.CM40	NTHDKARVAAGHA-SK-G.....SPKTFHQRRTIWSLESLESLCRLTRPRLDGRISM-CIG.....---	230
DEB.CM5	-TNDKASTTSHAGSK-G.....SFKTLHKRRVMGSLLESLCR-SCPTRLDRRTSM-GV-.....N-	230
DEN.CD.-.CD1	GRNHSKAATGSSMGYH-GGKKTQQR-DHWNMGSLVI..TSA-TLDRRTSHGSRRRRFHKNHAE...RNMEAL-	258
DRL.-.-.FAO	AN--RRGRV--MATEES.....LKNQPGGSITLP-RVPLPSLEH-HGAL.....P-	217
GRV.ET.-.GRI_677	TFTRMAGNLGSRQAMGR-ATRHAQGSKRRSQKALWNE..HAN-SMELLCRGGKE.....T-	219
GSN.CM.99.CM166	KGAHNHGYSRKRVSNGNR.....GG-ETLLKGRAPWNL.....A-	237
GSN.CM.99.CM71	KGAHNHG-CSKRHVGNKR.....RG-KTLLKGRAPWNL.....A-	237
LST.CD.88.SIVlhoest447	WATKGAMGSMASRRHHGG.....KRGGQTPFPRNNFPGPSLRI-CQOYRVRGESLHEIHYH.SK-	227
LST.CD.88.SIVlhoest485	WTTKRAMGSMVSRHHGG.....K-GGQTPFPRSHFPGPSLGI-CQOYRVGGSLHEIHYH.SK-	227
LST.CD.88.SIVlhoest524	WSSKRAMGSMASRRHHGG.....Q-RGKTPFPRSNFPGPSLGI-CQOYVRRGGLHEIHHF.S.-	226
LST.KE.-.SIVlhoest	WTTKGAMGSVSRHHGG.....KRGGKTPFPRSNVPGPSLAI-CRQYVRRGGLHEIHHF.SQ-	226
MAC.US.-.MM239	RDNRRLRMAKQNSRGD.....KQGGKP-TKGANFPLAKVLGIL.....A-	214
MND-1.GA.-.MNDGB1	AIQ-VASRYGNGKGRSK.....KALPSHYPKCSIRI-CAKHR.....E-	222
MND-2.-.-.5440	RGRTGAMAGNVTGENQPG.....GPVTLF-RVFPFSLHEMCRITL.....A-	213
MND-2.CM.98.98CM16	RG-KYLVGKISGENQSG.....GKITLP-RVFPFSLERMHRTL.....A-	212
MND-2.GA.-.MND14CG	RS-TRTVAGKIIGTSEQR.....GSITLP-RVFPFSLHEH-CRTL.....A-	212
MNE.US.-.MNE027	NRR-LRMAKQNSRGDKQR.....GSKP-TKGADFPGLAKVLGIL.....A-	214
MON.CM.-.SIVmon_99CMCML1	GSHRKTGPTG-HLGAK-R.....SPKTLKGRAPWHLGTGNGAA.....C-	255
MON.NG.-.NG1	GSHRRARR-AG-HMGPK-G.....GPXALLKGRAPWHLGSG.....---	257
MUS.CM.-.SIVmus_01CM1085	GNHHAPGT-SHVGSK-R.....G-KTLFQRRAPWNL.....E-	235
RCM.GA.-.GAB1	HW-WRSRVPALA-GH--KGO.....LGSQVTLSSRVHFPSPVAH-CGTL.....A-	216
RCM.GB1c1	HW-WRSRVPALA-GH--K.....GQPGSQVTLSSRVHFPSPVAH-CGTL.....A-	111
RCM.GB1c3	HW-WRSRVPALA-GH--K.....GQPGSQVTLSSRVHFPSPVAH-CGTL.....A-	111
RCM.NG.-.RCMNG411	HW-WRCRVCPLASQN-NR.....SQLRSQVALSSRVHFPSPVAY-CGTL.....A-	222
SMM.FGB	RNNRRGIRMA-KNSRRD.....KQGSSES-AEGANFPLAKVLGIL.....A-	214
SMM.SL.92.SL92B	RNNRRGLRMVTQYSGGP.....ESNSSTP-TKRVNFPGLEKVLGIL.....GR-	217
SMM.US.-.SIVSMMH9	RNXRRGXRLATQNXRRN.....KQGSSEFAXGANFPLAKXLGIL.....A-	214
SMM.US.-.SMMPGM	RNNRRGIRMAKQNSRRN.....KQGSSESLAEGANFPLAKVLGIL.....A-	214
STM.US.-.STM	RGNRGSIRVATQNGRGH.....KPRGSKPSTEGTDFPGLAKVLGIL.....A-	214
SUN.GA.98.SIVSUN	.TER.AICTMAGRYHG--.....Q-RSGKAFPPSSNSCPSLAI-CGLYRVKGRVYESNHF.AQ-	241
SYK.KE.-.SYK173	MVLHSSKCRVDPK-DQ.....CHCKGRTGSDRSIQAFYS-RNIWSL-SILKRRGRD.....-	221
SYK-KE51	TTILHSSKCKQLGRGRSNK.....CHCQSRGTGSPRSMQAFYSGRNIWSL-SML-GGSRSRRNTHHGLD-	233
TAN.UG.-.TAN1	MVRNLGSQ-GAVG--IKRH.....SRTQSGS-T-F.WERT---SMELLSGRRGK.EWGTNDR...KGL...-	238
VER.DE.-.AGM3	GTRNMGFE-GAVG---KRHA.....RYQSGSQDAF.WARA-V-SMELLSGGGRK.ESH-HAR...KGL...-	232
VER.KE.-.AGM155	RTRNLGSQ-GAVG---QRY.....RNQQRSQTA.F.WPRT-I-SMELLSGGRGE.TGKTHSG...KGI...-	232
VER.KE.-.AGM_VER_9063	RSRDLGSK--TMG---KGHV.....RSQSGSQITF.WQRA---SLELLSGRRRE.TGSPNDG...EGL...-	231
VER.KE.-.TYO1	RSRNLGSK-GAVGQ--KRYV.....TRSQPGGEAAF.WERT-V-SMELLSGGRRK.TWY-HDG...KGLQIL-	235

```

                                vif CDS \
H2A.DE.-.BEN      MT..DPRERVPPGNSGEETIGEAFEWLERTIEALNREAVNHLPRELIFQVWQRSWRYWHDEQGMASASYTKYRYLCLMQKAIPTHFKRGCTCWGEDMGREGL..EDQGPPPPPPPLV$ 113
H2A.GW.-.ALI      -A..N--T-----D-V-----Q-----MY--MK---L-GGH-PG-W...RS----- 112
H2A.SN.-.ST       -A..G--TI-----D-V-I-----I-----M-I-S---L-GGH-PG-W...RS----- 112
H2B.CI.-.EHO      -D...-----V-----T-L-H--V-----K--A--RE---I-----M-I--AK--G-LR-GH-PG-W...RS-----A- 111
H2B.GH.86.D205    -D...-----D--V---A---TE--V-----A--RE---I-----L---M-V-YTK--R-LQ-GH-PG-W...RS-----A- 111
H2G.CI.-.ABT96    -D...--X--E--D-----XTEI--I-----R---E-----L---M-V-C-K---LQKGH-PG-W...R-----A- 111
H2U.FR.96.12034   -G...--I---E---V---S---VRDI--A--Q-----Q--R-----I---L-M-C-K--Q-L--GH-PG-W...RS-----I- 112
MAC.US.-.239      -S...--I-----N--V-EI-----E-----P--V-----I---L-M-C-K--R-L--GH-AG-W...RP-----A- 112
MNE.US.-.MNE027   -S...--I-----N--V-EI-----E-----P--V-----I---L-M-C-K--R-L--GH-AG-W...RP-----A- 112
SMM.SL.92.SL92B   --...--I-----HN-V---QT--Q-----R-C-E--V---Y-P-A---VQ---M-Q--RK---R--GHSQG-W...RT-----A- 112
SMM.US.-.H9       -X..X--XI-----XX--X-D--H--X-EI--A---X-----R-X-E---M--V---X---I---L-M-C-K--R-L-GEH-AG-W...RP-----A- 112
SMM.US.-.PGM53    -S...--I-----N--V-EI--A---R---E--R--M--E-----I---L-V-C---R-L--EH-AG-W...RS-----A- 112
STM.US.-.STM      -S...--I-----A-E---H--V-DI-----R---E-----PG--V-----I---L-M-S---R-L--GH-AG-W...RP-----A- 112
MND-2.-.-.5440    -A.....-A-E.APEGAGEVGLAQ---TSL-RI---RL-FHP-FL-RL-NTCVEH---RHQR-LD-A---L--H--MY--MQQ--P-RNG.....RPRG----.-MA- 99
MND-2.CM.98.CM16  -A.....-GA-E.IPEGAGEVDLNT---SL-KI-Q--RL-FHP-FL-RL-NACIEH---RHQR-L-A---L--N--M--MQQE-P-RSG.....HPRG----.-M- 99
MND-2.GA.-.M14    -A.....-A-E.APQGAGEVGLAQ---SL-QV---QL-FHP-FL-RL-NTCVEH---RLRRTLN-A---L---M-V-MQQ--P-RSG.....HSRG----.-A- 99
DRL.-.-.FAO      -AERQSV--A-A.EPMGAGEV-LE---Q-SLLRI-Q--RL-FHP-FL-RL-NTCMEHY--ALQL-FT-S---L-L---M-M--QQ--S-LQGRHPPPLRPAG-RL-----S- 115
RCM.GA.-.GAB1    -A.....-A-E.VPTGAGEA-FQP--RDML-KV-L--RL-FHP-F--RL-RTCVEH---VHQR-LE-AA---L---L-I-CQT--SQRHGP.....NPRAV-ERITIL--M$ 104
RCM.NG.-.NG411   -A..EG----E.APTGAGDV-FAP--H-MLTEV-L--RL-FHP-F--RL-RTCVEH---RL-R-LE-AG---L---L-I-CQS--SQRHGQ.....QAREA-ERIQIL--M$ 108

```

Accession	Sequence	Position
SAB.SN..SAB1C	MASGGWLPVGGDPPKDPKPNREEIPGW..LETWDLPREPFDEWLRDLMLQDLNSEAQCHF.PRNLLFRLWVNIVEEPAI...DHGQTRLEGW.YKYCRILQKALFVH.MKGRCK...PKT...HPAYG	116
H1B.FR.83.HXB2	-EQA.....P-DQGPQ---HN--TLEL-EE-KN--VR---IW-HG-QQH-Y-TYGD.....TWA-V.EAII---QL--I-XFRIG-RHSR.IGV....TR.	86
CPZ.CD.-.ANT	-EQA.....PQDEGPQ---MN---LET-EEIKN--VR---QPT-QH-GNWWYANYGD.....SWR-V.QVLIIL-----T-.YRHG-AHSR.IGH....GP.	85
CPZ.GA.-.CPZGAB	-EQA.....P-DQGP---YQ--ALET-EE-KN--VR---PW-HQ-QGF-YDTYGD.....TWV-V.EAII---HL--I-.FRLG-QHSR.IGI....LP.	85
H2A.DE.-.BENTEAPTEFPPEP-TP.....R-DLGS-D-VIET-REIKE--LR--D--L-.IA-GYY-HNRHGD.....T---A.RELIKT--R-----FRAG-NRSR.IG\$.....TR	89
COL.CM.-.CGU1	-R.....VLKGT\$X.....-Q-QVENITR--RR--PLTHISVIM..GC...TRR..HW-RPW.TAG.LA-V-MEC--YL-----V.	65
DEB.CM.99.CM40	-ER.....Y-PSHPHF\$TR\$TIPMTRLG-QQAM---E-LK--T-EE--GV-NHC-DL--QPDWTGE-AWAASV.ID-IK-V--M-WL-.LREA-FHRE.REATRGY-NIR	104
DEB.CM.99.CM5	-ER.....Y-PSHPHF\$TR\$TIPMTRLA-QQAM---E-LK--T-EE--GV-NHC-DL--QPDWTGE-AWAASV.ID-IK-V--M-WL-.LREA-FHRE.REATRRY-NIR	104
DEN.CD.-.CD1	-ERL.....PPSHPPA...LVS\$RMA-TTQRQ-QEAVW-ITE--RK--SKEEITGI-D.HCWSL\$PALPHWTEGQVMAAAVIDFI--M--EIWK-.YRVG-FHREABRV...RH-P	101
DRL.-.-.FAOE-PPDEGP.....Y---V-T-VEIQE--LK--DR-L-.HNVGSWVY-RYGD.....S---V.HQLIIL--R---L-.FRYG-QESR.LG.....PA	84
GRV.ET.-.GRI_677	---R.....D---PL---I---D---W---Q---R---E--RR--GM-M-I-V-N.YCV-EGRR..HNT.PWN-IG.---Y--V--SM---FRCG-RRRG.-FS....P-	99
GSN.CM.99.CN166	-ER.....IPPSHPM...WLSRRVPTTMAIAQNA-WEI-E-EK--S-DE-RGI-H.D-T-LPADPDWTV\$DQAAIACAID-I-.V-TL--R-.F-DG-FHRVN.LS...GR-P	99
GSN.CM.99.CN71	-ER.....TPPSHPLP...WISRRVPTTMAVAQ\$AMWEI-E-EK--S-EE-RGI-H.D-T-LPADPDWTV\$DQAAIACAID-V-RV-TL--R-.FRDG-FHRYNRIV...RR-P	101
LST.CD.88.447	-SRTR.....QQ-SPEKPP...DFGP---WGQ--A-TMEEIKE--RR--ETI-AQVSEYC-SSTGS.....BEKAC.M-FIT-VNR---I-.LP.T-PSYA.AAR..AGT-HP	97
LST.CD.88.485	-SRAR.....QQ-SPERPP...DFGP---WGQ--A-TMEEIKE--RR--EAI-AQVSEYC-SSTGS.....EE-AC.MRFIT-VNR---I-.LP.T-PSYA.ATR..AGT-HP	97
LST.CD.88.524	-SRQR.....QQ-SPERPP...DFGP---WAQ--A-TMEEIKE--RR--EAI-AQVSEYC-SSTGS.....EE-AC.MRFITLVNR--YL-.LP.Q-PSYA.AAR..AGT-HP	97
LST.KE.-.lho7	-SRQR.....QQ-SPERPP...DFGP---WGQ--A-TMEEIKE--RR--EVM-AQVSQYC-DSTGS.....EE-AC.M-FITLVNR--L-.LP.T-PSYA.AAR..AGT-HP	97
MAC.US.-.239EE-PPNEGP.....Q---W---VVEV-EE-KE--LK--D--L-.TA-GNH-YNRHGD.....T---A.GELI---R---M-.FR-G-IHSR.IGQ.....P-	86
MND-1.GA.-.MNDGB1GQKRDEQVS...DQGP---YQ--A-TMEEIKE--RK--LII-NAVSEYC-QNTGS.....EE-AC.E-FITLMNR-IW--LAQG-DGTF.RER....RP.	90
MND-2.-.-.5440EQPPEDEAP.....Q-G-YNK--VGT-AEIQE--LK--DR-L-.HAVGSWVY-QQGD.....T---V.QQLIS---R---L-.FRYG-RGSR.IGQ.....E-	85
MND-2.CM.98.CM16EQPPEDEAP.....Q---YN---I-T-AEQE--LK--DR-L-.HAVGSW-YKTYGD.....T---V.RQLIS-----SM-.YRYG-RESR.IGQ.....E-	85
MND-2.GA.-.M14EQPPEDEAP.....Q---YN---I-T-AEIQE--LK--DR-L-.HAVGSW-Y-TYGD.....T---V.Q-LITL--R---L-.FRHG-RESR.IGQ.....E-	85
MNE.US.-.MNE027EE-PPDEGP.....Q---W---VVEV-EE-KE--LK--D--L-.TA-GNY-YDRHGD.....T---A.GELIK---R---M-.FR-G-NHSR.IGQ.....S-	86
MON.CM.99.L1	-V.....PIERQA-QAAIWE--E-LK--S-EE-RG.I-EQ-T-LPADPAW\$NADQAWAACAID-T-WV-TI-YR-.YREG-YHRYAEQIRRYVPLRP	90
MON.NG.-.NG1	-H.....XPS...HPPHWT\$SRVVPIERAA-QAAIW--E-LX--SKEE-RG.I-DQ-T-LPADPGW\$NADQAWAAXAID-T-RV-TXI-R--RDG-YHR.HARV...XRR-RP	99
MUS.CM.01.1085	-R.....QAQEAMWEV-E-EK--S-EE-RG.I-E-T-LPADPDWTV\$DQAAIACAID-V-RV-TLV-R-.FRDG-YHRYNRVRRYPTIRP	85
RCM.GA.-.GAB1ELPPEDEGP.....Q---Y---M-T-IEQE--KK--TYAL..TQIGDYVY-OHGD.....SI--V.QAMI-L-R---L-.FRNG-AGSR.IG.....S.	84
RCM.NG.-.NG411EMLPEDEEP.....Q---Y---M-T-IEIQE--KK--T-EL..TQEGNY-Y-OHGD.....S---V.KAMITL-N---L-.FRHG-EGSR.IGG.....R.	85
SMM.SL.92.SL92BEHAPEDTN.....W---I--V-EE-KE--LK--D--L-.TA-GNYVDTYGD.....TI--A.GEIIK-----L-.FRHG-THSR.IGQ.....PR	85
SMM.US.-.H9TE-PPDEAP.....Q---W---XVEV-EEXKE--LN\$X-D--L-.XA-GNY-YDRHGD.....T---A.GELI---G---I-.FR-G-RHSR.IGQ.....S-	86
SMM.US.-.PGM53TE-PPDEAP.....Q---W---VVKV-EVKE--LK--D--L-.TA-GNY-YDRHGD.....T---A.GELI---R---I-.FR-G-AHSR.IGQ.....SR	86
STM.US.-.STMTH-PPDEGP.....Q---W---VVEV-EEIKQ--LR--D--L-.SA-GNY-YDRHGD.....T---A.GELIK---R---M-.FR-G-RHSR.IGQ.....P-	86
SUN.GA.98.L14	---RREPVEQPP...DEGP---AQ--A-TMEEIRD--ER--AAI\$VAVLQYCADSTGS.....EK--C.MRAITL-NR---I-.LP.M-PSYIRIRSGSGNVNPR	98
SYK.KE.-.KE51	--QAF.....FRVYQRQ-G-PLFYIP...RNVNWD-AGVISATAKAEQVAQ--CK--TAEEI-GV-NQCL-VE-GPEETPTMAWIRCM.LDMH-A-NFM--E-.FAAG-PQRT.KYA....RH.	108
SYK.KE.-.SYK173	--EAF.....FNPSQHVG\$TPWF\$FIP...RNVELTP.NVINVTVKAE\$LVVT--SK--TPQBIY.GV--QSLNEBAGTDSPTMAWERTM.LDMV-A-NLM--E-.FAAG-PQRT.RYA....RH.	106
TAN.UG.-.TAN1	--E-R.....DS--RR--LEIWDLS...W--\$---VAE--Q---R-.G-E--QV-N.FCQ-EGER...N-APMI-RA.-R-Y-LV-----FRCG-RRRT.-FE....P-	98
VER.KE.-.AGM155	---R.....D---R--G.-I---S--W-----VEEI-N-KL-.G-E--YQV-N.YCQ-EGER...Q-RPIA-RA.--Y-LV-----FRCG-RRRQ.-FE....P-	99
VER.KE.-.9063	---R.....G---NR--E.V-I---N--W-----E--Q--KL-.G-E--QV-N.YCQ-EGER...R-APMT-RA.-R-Y-LV-----FRCG-RRRQ.-FE....P-	99
VER.DE.-.AGM3	---R.....D---AR--E.-I---S--W-----E-I-Q--KM-.G-E--QV-N.YCQ-EGER...NRTPM-RA.--YKLV-----FRCG-RRRQ.-FE....P-	99
VER.KE.-.TYO1	---R.....D---AR--E.V-I---S--W-K-----Q--RL-.G-E--QV-N.YCQ-EGER...TPMM-RA.--Y-LV-----FRCG-RRRQ.-FE....P-	99

SAB.SN.-.SAB1C	PGAGGPPPLGGASGGAASAAPGL.....\$	140
H1B.FR.83.HXB2QRRARN---.....RS-	97
CPZ.CD.-.ANTR.....RR-	88
CPZ.GA.-.CPZGABQRRRSN-SN.....RS-	96
H2A.DE.-.BEN	RR.TPC-AAPTPRAMH.....-	104
COL.CM.-.CGU1H.A-QCRFPSTQNP.....-	78
DEB.CM.99.CM40	-LT-RNREV.....RDGE-	117
DEB.CM.99.CM5	-LT-RNREV.....RDGE-	117
DEN.CD.-.CD1	NIRPLR-RREEP.....-	113
DRL.-.-.FAO	GRWNPFRTPFRANQQL.....-	100
GRV.ET.-.GRI_677EERRN-QGG--PPPP--A.....-	118
GSN.CM.99.CN166	AIRPSRGTAPPDSNSVSHADPEQPRRPSRYRMDE-	133
GSN.CM.99.CN71	VIRPLRGTAPPDSNSVPHADPEQPRRPSRYRMDE-	135
LST.CD.88.447	TE-APR-TKRDIQR-RY.....-	114
LST.CD.88.485	TE-APK-TERDIQR-RY.....-	114
LST.CD.88.524	TE-APR-T-RDIQR-HY.....-	114
LST.KE.-.lh07	TE-APR-NEGDIQR-SD.....-	114
MAC.US.-.239	G-.NPLSAIPPSR-ML.....-	101
MND-1.GA.-.MNDGB1QL-PS-FRPR-DRL.....-	104
MND-2.-.-.5440	GRYNPLRSFPSPNNPL.....-	101
MND-2.CM.98.CM16	G-CYPLRSFPRSDNPL.....-	101
MND-2.GA.-.M14	GKYNPLRSFPRSPNNPL.....-	101
MNE.US.-.MNE027	G-.NPLSTIPPSR-ML.....-	101
MON.CM.99.L1	MRGTA-G-TSSVPQADPDN...PRRPSRYRMDE-	120
MON.NG.-.NG1	XLRPLRGTAP-PT-SVPHADPERPLHPSRYHHDE-	133
MUS.CM.01.1085	LRGTQ--SNSTPNADPVP...PLSPSRYMDE-	115
RCM.GA.-.GAB1	..R-SN-LRSIPQTRNI.....	99
RCM.NG.-.NG411	...-N-LRSILH-RNI.....	99
SMM.SL.92.SL92B	GP.NPLGSIIPSARDVL.....-	100
SMM.US.-.H9	GX.NPLSTIPPXRGVL.....-	101
SMM.US.-.PGM53	G-.NPLSTIPPSRGVL.....-	101
STM.US.-.STM	G-.NPLATIPPTRGVL.....-	101
SUN.GA.98.L14	-PTRPR-GQDIRR-LQD.....-	116
SYK.KE.-.KE51RGYPHPS.....-	115
SYK.KE.-.SYK173RGYPHPS.....-	113
TAN.UG.-.TAN1EERRN-VGG-RDG.....	111
VER.KE.-.AGM155EERRN-QGG-RPGRVP--D.....-	119
VER.KE.-.9063EERRD-QGG-R-NR-P--D.....-	119
VER.DE.-.AGM3EERRD-QGG-R-GRVP--D.....-	119
VER.KE.-.TYO1EERRD-QGG-R-NRVP--E.....-	119

		Cys-rich domain	Lys acetylated			87		
		* * * * *	* * * * *	exon 1	\ / exon 2			
SAB.SN.-.SAB1C	MDQEQAEPQVWEELQEELHRPLQACDNTCFCKVCCFHCILCFHKKALGIRIYVPRPRRASKKIS	HNQVSLHKSISTWT....RDSQTKKK	87	
H1B.FR.83.HXB2	-EPVDPRL-PWKHPGSQ-KT--	TN-Y--K----	QV--IT-----	S-GRKKRR..Q-RAHQ	...NSQT-QASLSKQPT-Q..PRGDP....GPKE\$	86	
CPZ.CD.-.ANT	--PVDAETPPWLHPPAT-ATP-	-N-Y--C--Y--P--T--G--S-GRK-RAR..NRRITAE	..SSENNQDPVSK\$-LPK	..TSRIGS....SQQK\$		87	
CPZ.GA.-.CPZGAB	--PIDPDL-PWKHPGSQ-RTV-	-N-Y--A--Y--Y--T--G--S-GRKKRTT	..RTAPA	..GSKNNQDSIPKQPL-Q..SRGNKE	...GSEKST-E	91	
H2A.DE.-.BEN	-ETPLKAPESLKPYNPSSCTSERDVTQAQLAKQG--	LAQ-----EP-T-K-Y--R-S--Q--S-G--S-ERKGRR	..TTPRTKTPSPSAPD----	RTGD	.SQPTKEQKKTSEAT	121	
COL.CM.-.CGU1SLSLYAK--E--	NDK-W--A----	Q--LS-G-----	KKNK	..SATQTQ-VAKTGSQNSQ-TQKKYTESK-TA	73	
DEB.CM.99.CM40	-ESIDPFEERANTPHPA-M--A-IL--Q--	IQ-G--S-GPR-RKRKC-REYTPAAESNPQDFVPEQPLPQ	RRRK	.CFDSQKKEAEVEET	92	
DEB.CM.99.CM5	-EEIDPFKETANTPHPN-M--A-TM--Q--	MQ-G--Y-GSR-RKR--G-SFAANDKDHQEFIPE	..HCVEDIPLCRPLPQRWKCFNSEE-		91	
DEN.CD.-.CD1NADSIDPFAGNKTP-	---Y--H-AY--Q--	LQ-G--V-GRAPRK	..PRARQS	...GTPDQEAVSQ-L-DPRRQISQQKE-E-	83	
DRL.-.-.FAOARKVDLQDQAGT.HF-PFQGRTN-	---W--Y-AY--Q--	LQ-G-----N-S--R	..K-TATV	..VSRVSETGDTPVARTHTAQRRTT	TDSQTEKKEKS-TE	102	
GRV.ET.-.GRI_677DKEEHPHL--D	-----P-T-K-Y--E--E--	LQ-G--V-H.VSRK	..KTSTQDNQDPIRQQ----	VQ.RNGQ-TEEGKTEVE-AA	91	
GSN.CM.99.CN166PILSSLDGPDIAWMNRDPPECLLPNW-QPGAA-ATP-S.A-Y--K-AL--Q--	LR-G--LSLHGRSRKRK-PKATEPD.PSVSSNQDSA\$VQ-L-HKQWNAVRQT-P--A				110	
GSN.CM.99.CN71LILSSPEDLQDTAWMNRDPAECLLPKW-QPGAA-NTP-S.A-Y--R-A--Q--	LR-G--LSLHGRSRKRK-SEASEPA.PSVPTNQDPIQIQ-LPHSKWNAIRQTEQE-T				111	
LST.CD.88.447QQEQEQQLTRQKQHODPLK-TYK-AVTK--S--N-K-C--K-----	QQ-G-----H-RR	..K-VAQVFEDK-DPVDGKER-KNSQTKKER\$			95	
LST.CD.88.485QQEQEQQLTRQKQHONPLK-IYK-AVTK--SV--K-C--R-----	QQ-G-----H-RR	..VTQVFEDK-DPVDGKER-QNSQTKKER\$			95	
LST.CD.88.524QQEQEQHT-QKQHLD-LE-IYK-AITD--S--O-K-H-----	QQ-G-----H-TR	..VTYILETE-DPVDRTERAKNSQTKKERQKTEKSTST			106	
LST.KE.-.lho7QQEQEQPI-QKQHQDLMK-TYN-AVTKA--Q--C--K-----	QQ-G--LH--H-KR	..VTSILETQ-DPVDRKRK-KNSQTEKKRKEIQIRLSTST			106	
MAC.US.-.239	-ETPLREQENSLSSNERSSCISEADASTPESANLG--	ILSQ-Y--E--Y--Y--K--Y--QF--L--G--C-EQS-KR	..TP--AKANTSSASN-P--N	RTRH	.CQPEKAKKETVE-A	121	
MND-1.GA.-.MNDGB1	-EPSGKEDHNCPP-DSGQEEIDYKQ-L--YYQ----	E-K-W--K----	M--Q-G-----	H-Y-KR	..VPGTN-K	...IPGSGEAIRRA-DLSFHRTASRTYTANGQT-E--	111
MND-2.-.-.5440VGEVASDKKEEDITHFDPFRARTTP-	-K-----Y--Q--	LQ-----	VH-H-Y-QR	..PRRGLL	..EKVPSNS-AA-GRLSPAPQPOTARQRRR-ARLR-T	102	
MND-2.CM.98.CM16AGKAVSDKKEGDVTPYDPPFRDRTP-	D--Y--R--Y--Q--	LQ-----	VH-HAY-NR	..SRQRL	..EKISEDSAAPVGRSLSPSPQPTARQRR-ARLR-T	102	
MND-2.GA.-.M14AGKAELEKKEANITPLDFPQGRITP-	-K-Y--K--Y--QC--LQ--	---H-H-Y-TR	..TRKRL	..GKISQDSEAAVGRSLSP-SQPOTARRRKR-ACLRR	T	102	
MNE.US.-.MNE027	-ETPLREQENSLKSSNGRSSCTSEAAATTPESANLE--	ILSQ-Y--E--Y--Y--K--Y--QF--L--G--C-EQS-RR	..TP--VKANTSSASN-P--D	RTKH	.CQPKKEQKETVETA	121	
MON.CM.99.L1	-EPVDPDLPE-HPPAT-RTP-T.N--R--	---QI--LT-G--S-GRK-KR	..---TS	..PVPGLSSSKNPARKQ-L-HPTRDSQRPTQQAQAVA-AAT		97	
MON.NG.-.NG1NSAXIENPALPPE-HPSAT--TP--	-N--I-AL-XQV--LQ-G--S-GRG-KR	..GPRH	..PTIGDRSYQDNPGV--LPKSLGNQERQKKQEKAVES-AG			101	
MUS.CM.01.1085PSVEELPKE-RPGAA-ATP-T.S--K-A--V--Q--	---S-GRRFR	...KR-RA	..DSSLPS-SDPVQI--L-DRRRDQKRQKQEKETVEAEAE			95	
RCM.GA.-.GAB1VQGVGLEH-BEVI	-YDPPFRKRETS	---Y--K--Y--Q--	LQ-G--N-.AS-AR	..---R--E	...ENKADKFPVPNQ--	..STTRGNRKLQEKKEKTEVE--V	98
RCM.NG.-.NG411VKGAGLEEREEI	-LDPFFTRET	-S-Y--R-VY--Q--	LQ-----	N-.AS-RR	..GRPGNK.KKNQINPPTDSEQ-L-AIGRNSQSKKEKKEAVEKAVD	102	
SMM.SL.92.SL92B	--APIQGLDSQGDQIPWD-Y--ET-Y-S--K--Q--	LR-G--T-AK--RR	..VKKKAANPDTSNQ--S	GTRN	.CQPKKEQKEKETE	95
SMM.US.-.H9	-ETPLKEQESSLESSREHSSSISEVDADTPESASLEX--	ILSQ-Y-XVE--Y-KXY--K--Y--QH--X--G--C-EQO-RR	..TPKKTANTFSASD--LXX	RARN	.CQPKKEKETEVE	119	
SMM.US.-.PGM53	-ETPLKEQENSLGSSNERSSYTSBEDVLTPESAK-E--	ILCQ-Y--E--S--Y--K--Y--Q--	L--G--C-EQO-RR	..TP--AKANTFSAPNE-L--	RARN	.SQPKKEKETEVE	121
STM.US.-.STM	-ETPLKEQESSLRSSSEPSSCTSEAVAAATPGLANQE--	ILWQ-Y--EE-C-K--K--Y--Q--	VT-G--T-ERS-RR	..VKKKATYPISASN-L--	RARN	.SQPKKEQKEKETEVE	119
SUN.GA.98.L14STQGH-QDQD-GKGT-E-AYKTN-E--	K-W-RR----	Q--LQ-G--H--Y-GR	..QHQ-ISKDK-EAVRRKPAANPQTKKKKKRQR	.KKESTST	98	
SYK.KE.-.KE51	-STEDKVR-TQGI-TSFL-GTFLSNG-QTP	..SR-Y--K--Y--YR--LQ-G--T-VRE-KR	...RHARTTAEDFAA-SSTT	..R-R-PLPSTTRPQQGQKKE\$		98	
SYK.KE.-.SYK173	-SSTDQICQTRV-PSFL-GTFLEKG-PTP-	-K--N--Y--Q--	LQ-G--T-AR--KR	...A-RSISEDDSAPTGTLPRAGRTQANPQTKKAVE-E-D	100	
TAN.UG.-.TAN1	-ESEGDMAS-LQD----	TP-T-K--R--Y--QV--LQ-G--T-H-S-I	...PK-NHN-QNLVSOQ--AGG	..NSQTTQEE-	87
VER.DE.-.AGM3KGEDEQGAYHQD-I-Q-KA--KR-T-K-Y--C--Y--Q--	LQ-G--VT-HA--I	...KKIAPLDRFPEQKQ--R	GRDSQ-TQK	.GOEKVETS	95	
VER.KE.-.9063KGE-E-TVLHQD-IRQYK--TT-R-K--K--Y--Q--	LQ-G--VT-HA--T	...KKSVPQNRLSQDQ--R	GRDQATQ	.E.SQKVERE	95	
VER.KE.-.AGM155KGE-DQDVSHQD-IRQYRK--ET-T-K--K--Y--QF--L--G--T-HAF-T	...KKIASADRIPVPQ--IR	GRDSQ-TQ	.E.SQKVEEQ		95	
VER.KE.-.TYO1KGEABQIVSHQD-S-DYQK--T-K-K--K--Y--Q--	LQ-G--VT-HA--T	...KKIRSLNLAPLQ-Q--	KWGRDQ	-TPT	.SQEKVETT	96

SAB.SN.-.SAB1C	SKAAVG\$	93
CPZ.GA.-.CPZGAB	VASKTEADQ\$	100
H2A.DE.-.BEN	VVTTC-LGQ\$	130
COL.CM.-.CGU1	AADFGD--	79
DEB.CM.99.CM40	AGPGGQHCREDDSSVSSGRISNNC\$	115
DEB.CM.99.CM5	ETKVEKASRAGGRHRPES\$	110
DEN.CD.-.CD1	VE..TATGSSGGNIRED\$	99
DRL.-.-.FAO	ASGTSDQNPPSQDPPKLGAKKQ\$	124
GRV.ET.-.GRI_677	AAN\$	94
GSN.CM.99.CN166	LARKARSGP\$	119
GSN.CM.99.CN71	VARKATSGPSDISQDSRVSCGTT\$	134
LST.CD.88.524	\$	106
LST.KE.-.lho7	\$	106
MAC.US.-.239	VAT-P-LGR\$	130
MND-1.GA.-.MNDGB1	KAT-\$	115
MND-2.-.-.5440	EHQVREL...QTRIS	113
MND-2.CM.98.CM16	EYQIQQL...QSRWEALEQRANQRLAEDFKGLRVAEESTSSN\$	142
MND-2.GA.-.M14	EHQVREL...QERIWQTLERANKELVEGVNNLHLAEBESAGSN\$	142
MNE.US.-.MNE027	VATVP-LGR\$	130
MON.CM.99.L1	PDRQHI--	103
MON.NG.-.NG1	TGXQSMREDF\$	111
MUS.CM.01.1085	TDP\$	98
RCM.GA.-.GAB1	ATSTTIG\$	105
RCM.NG.-.NG411	TNITT-REDIFIS\$	115
SMM.SL.92.SL92B	VDTDC-LGR\$	104
SMM.US.-.H9	VATDL-LXR\$	128
SMM.US.-.PGM53	VATDL-LGR\$	130
STM.US.-.STM	VESTP-LGK\$	128
SUN.GA.98.L14	\$	98
SYK.KE.-.SYK173	-TSLPSAENL\$	110
TAN.UG.-.TAN1	T-IPAAAETSRRPQ\$	101
VER.DE.-.AGM3	ART-PSLGRKNLAQSSGRATGASD\$	119
VER.KE.-.9063	TTT-QILGRKDLERDKREAVGANAS\$	119
VER.KE.-.AGM155	A--NLRISRKNLGDETRGPVGAGN\$	119
VER.KE.-.TYO1	AGSN\$	100

	exon 1 \ / exon 2	
SAB.SN.-.SAB1C	M.....SLG...QEELLRR...F.RIIKFLYT.TN.PYPP..GQGTARQ.RRRARQRWAKQRQVVIHLAERILETPVSIQIDHLAQEFQDQVLVDNLQPPSLPPGHPTENQTANSSS\$	99
H1B.FR.83.HXB2AGRSG..DSDEELI...RTV-L-L-L-Q.S.-P-N..PE---A--NR-R--RERQR-IHSIS---G-YLGRSAEPVPLQLPPLERLTLDCNEDCGTSG.....TQ.GV.....	97
CPZ.CD.-.CPZANTAGREEL.EGTDEQLL..KAVK---I--Q.S.---K..PA-S.-AA--NK-R--K-RQD--EG--A-V-R-L-GGP-EHNPVDLPDLQNLSE-LGS---TPLQNYFQTVQP-GE.....	107
CPZ.GA.-.CPZGABAGRSEPO.DDARLL...QAVK---I--Q.S.---S..PE---KA--NR-R--RARQK-ISEISG-V-A-YLGRPPKGDLELPELQDLKLSL-CVETTDQVQ-S-TSQPQTATGETVPAG....	112
H2A.DE.-.BENERAD..E-G-QGK...L-LRLL-HQ.--Q.GPGT.-S.-NR-R-RRR-WLRLVA--NKLCVAV-DPPT-SPLDRAI-HLQRLTI-E.LPD-PTDLPESNS-QGLAET\$	103
COL.CM.-.COLCGU1	-TNAGVRPVAFTVSYAFFQKGWESAI-R-RV-HK.H.-W-RPG.V-----NT--AR--Q--ILAIDG-VV-QL-ASLQQWDPYCPASTSA-NLRQEEELADSSVSIQTQG-KPSDS\$	118
DEB.CM40AHAGGRGSA-ENT-QLLVISL--I--Q.S.---K.G-GS.-ST.--R--K-RRRQ--DSI-----QSHLGGSPPT-DVALPDLSQLHLAD\$	90
DEB.CM5AHAGGR.GDADQALLRMIRIISKLYQ...S.---KGG--S-ST.--R--K-RRRQE--DDI-Q---AHLGRHLTASDVALNLEELRLAD\$	87
DEN.CD.-.CD1AGLQE..RDQE-VK...AV-L-K-Q-Q.S.---TPE-SKS-----R-RK-RR-QA--VAIS-----YHLGRSLTA-DVELPDLEKLSVSD\$	84
DRL.-.-FAOA-PE..R-PPW..FQEYL-LVTR..WQGR..I--R.DDQLPQTA-Q-K-R-ARQR-VEHQIRTLQARILQSLERRNSRDLVEGIERIH-AEERESS\$	92
GRV.ET.-.GRI_677KE..EKQA.....K---T-G.S.---QFS.....R--Q--IDKI-G-V-N-FED-QLVAQLQEL--ENKD-VLQHLPD-P-IHQDSSGIPAVWAPATPRGSNRA	110
GSN.CM.99.CM166ADHARGNDQK-QNLI..LAC-L-L-T-HR.S.---TSS-TRS---NR-R--RARQ--REISN---SL-GRPEEPGDLDLDPDLGQLSLS-DREVPVGTAAES-PE-NPASETAQSSS	118
GSN.CM.99.CM71ADPANGRDQK-QNLL..LAC-L-L-T-K.S.---TAS-TRS---NKKR--RARQR--RQISH---SL-GRPEESGNLDPDLGHLSDS-EDRQVPGVAAAEGVPT-ETAGGSGSSS	118
LST.CD.88.SIVlhoest447T-GE..K-SPRY.....LK-SKILW.-GLQVGEKREPR-----D-E-QR-HLH-LRTVQ--FHST-ERGLER-FTRLTVCDSPQCVAVPFVVARPFSDFLPTWAT	114
LST.CD.88.SIVlhoest485T-GE..E-LPRY.....LK-SKILW.-GLQVGERREPR-----D-E-QRRHLH-LRTVQ--FHST-ERGLER-FTRLTVCDSPQCVAVPFVVARPFSDFLPTWAT	114
LST.CD.88.SIVlhoest524T-PE..G-LPTY.....LKLRSILW.-GLQVGERREPR-----D-E-QR-HLH-LRAVQ--FQAT-ERGLER-FTRLTVCDSPQCVAVPFVVARPFSDFLPTWAT	114
LST.KE.-.SIVlhoestT-NG..D-LPRY.....LRLSRILW.-GLQVGERREPR-----E-S-YRDLH-LRAVQ--FQAT-ERGLER-FTRLTVCDSPQCVAVPFVVARPFSDFLPTWAT	114
MAC.US.-.MM239NHER..E--RK.....L-L-HL-HQ.--T.GPGT.-N--QRKR--RRRW--LLA--D--YSF-DPPT-TPLDLAI-QLQNLAIES.IPD-PTN-PEALCDPTEDRSRPOD\$	107
MND-1.GA.-.MNDGB1T-NV..YQ--I-----YLVVVVKKLY.EE.-I-Q-----RK-QLRTR-A-LRE-EG--KQILDRGPDQLCQGVNLLALAEKSES-N\$	83
MND-2.-.5440TDRG..DLDEGFW...RKYQA-VKQLW.EG.RLS-APQPQ-----R-A-LR-TEH--RE-QT--\$	61
MND-2.CM.98.98CM16PTETG..DLDRDFW...KXYQ-VQRLW.EG.RLS-SPQPQ-----K-R-A-LR-TEY-IRQ-QS--W-ALEQRANQLA-DFKGLRVAEESTS-N\$	90
MND-2.GA.-.MND14CGTEPG..ELGRDFW...EKYHR-VKQLW.EG.RLS-TSQPQ-----R-K-R-AQLRRTHE--RE-Q--WQ-LEARANKELV-GVNNLHLAEESAG-N\$	90
MNE.US.-.MNE027SHAE..E--R-----L-L-HL-HQ.--T.GPST.-N--NR-R--RQRW--FLA--D-VYSF-DPPT-TPLDLAI-QLQNLAIES.IPD-PTNPIEILHDPTESPRSPQD\$	107
MON.CM.-.SIVmon_99CMCML1AGNGR..D--RRL..LSLALAAVRI-QE.S.---IPR...-NA.--D-R-RR-QL-IASIS---FYHYLGRSQEPCPLDIPDLERLSISDLPD--ESVP-AA-PPAHTPAPTGVKPS\$	112
MON.NG.-.NG1AGAER..GAAVQDI..LRLAIGA-RIIRE.S.---SPS...-RSA.--NR-R--RARQE--XS-C--FNHYLDRLPRPNSLDLPAIEGLRLDLDLPE--ESAQPPDPSSRPVPTAGSQ\$	113
MUS.CM.-.SIVmus_01CM1085AGVSE...NDREL..IRAC-A-QI--K.S.---TAD...-RSA.--NR-R--KRQR-IREIS---LSLLGRPPQSDLELDPDLNKLKSLH-LVATSESSPPDTEGVNKIVVHAQGTD\$	110
RCM.GA.-.GAB1LP-QD..E-D-RKK...I-L-N--L.I.---OHG.GT.-NS.--KK-R--RRRW---QD-----LDSTDPPEQDLDAAIAD-QK--LNNLPE-PVDF\$	93
RCM.NG.-.RCMNG411L--EE..E-ADQEIR..RRI-L-HLHL.S.---QSG.GT.-N--K-RQ-RRRWT-ILQ-----FLY-DTPP-RDPLQEALESN-QQ-TLSDLPE-PVNPFFSNPSSFAVDRS\$	107
SMM.SL.92.SL92BQNP...E--R-----L-L-HLIHQ.--A.GPGT.-S.-NR-R--KRWI-IVA--D--YSF-DPPA-TDIDLAI-QFDSLIS-S.LPE-PTTVPETLRDQ-AD\$	99
SMM.US.-.SIVSMH9SN...E--R-----L-L-H-HQ.--XX.GPGT.-N--R-R--XQRW--ILAX-D--YSF-DPPA-TPLDLAI-QLQGLAIEX.LPN-PASAPEPLKDAESP\$	100
SMM.US.-.SMMPGMSSAG..E--R-----L-L-H-HQ.--Q.GPGT.-N--R-R--RQRW--ILA--D--YSF-DPPA-SPLDLAV-QLQ-LSLEN.IPD-PTSVPREALTENPRSRQD\$	104
STM.US.-.STMDQEE...-RK.....L-L-Q--HQ.--Q.GPGT.-N--NR-R--RQRW--LLA--N--YSF-DPPASTPLDLAV-QLQGLSI-D.LPD-PPNLPKDLQDTAEN\$	99
SUN.GA.98.SIVSUNT-DD..S.INQY.....LR-SKRLY.EGLAPG..NLPQ-H---R-D-ER-NLH-LRAVQ--FA-TLDSRLGR-F-RLSVSDSSQVAESLGNPSSTKHLPP-KFLVAPTYDFLPSWAT	111
SYK.KE.-.SYK173PDQGS...Q-AVF.....LRMIAHL.QE---GPE..P.--T--R-RQ-RQR-T-RLY-QQ--F-AIFGSRATALEDLQ-QLQISD\$	77
SYK-KE51GRER...-DTQEQLLRLTLR-AQQ...LA.---LPQ..PSRA.---N-N-YRQLQA-RLYVQQ--F--IARRSAETLEQSFQGLQITD\$	81
TAN.UG.-.TAN1S-D-----IIQ--RI--H.S.-Q-----E-----R--FQ-QR--AA-S--PIAQRRDPSSGGESLAAAFDQLVDLNDQ--VIETLPDPPEPHD-SSTA\$	103
VER.DE.-.AGM3P-----S-----LL-L-A--NK.N.---V..E-----R--RQAE-LRA---WHSR-EEQLVQ-IDQLV-DQQH-AIQQLPD-PSS\$	89
VER.KE.-.AGM155P-----P-----LL-L-A--R.S.---SV..E-----R--KNRQK-IYA---WG-RQEDQLVQ-IDQLV-DTQH-VTQQLPD-PSQA\$	89
VER.KE.-.AGM_VER_9063P-----P-----K.....VFSL-AY-NR.---V..E-----R--RERQR-LRF---WNETKERQLEQTLQDLA-DVQH-AIQHLPD-PNSS\$	89
VER.KE.-.PYO1P-----P-----V-L-WL--S.---SG.E-----R--RQ-QD-IRV-V--LQ-QYVAVRDLADEAQHLAIQ.....LPD-P-SA\$	84

PLV protein alignment: REV

685

H1B.FR.83.HXB2	..GSPQIL...VESPTVLESQTKES\$	116
CPZ.CD.-.CPZANTQSQHPLD.NNQNS\$	118
CPZ.GA.-.CPZGABGNYSILGKGAKN\$	124
GRV.ET.-.GRI_677	CSSSGBGCEBSLGTGTCYCPIRLSGSHQSKKSAARP\$	147
GSN.CM.99.CM166	QSDSLH\$	124
GSN.CM.99.CM71	LY\$	120
LST.CD.88.SIVlhoest447	SSESLLQQTGGKRSDECELELEQDQKEQKRHISSCKKPTGTKEEMEIQIRKKNVWSQ\$	173
LST.CD.88.SIVlhoest485	SLESQQRMDGGKRSDECELELEQDQKEQKRHISSCERHQTGTKEKMEIQIREKKNVWSQ\$	173
LST.CD.88.SIVlhoest524	SLANPQQIPGGECSENCELDQEQEQRRAIKNS\$	148
LST.KE.-.SIVlhoest	PLADQQQMDGGKRSDESLELAQEQMEQKQRTVIEH\$	148
SUN.GA.98.SIVSUN	PLADPQRLAGFAPYSGEYDQERVNQQGESIIVSEGGK\$	150

	- transmembrane domain - -	cytoplasmic domain / Env CDS start	-
		phos phos	
		- alpha helix - -alpha helix	
H1A1.UG.85.U455	M....TP.LEIWAITGLIVALILA.IVVWTIVGIE.XK..KKLL...KQKKI..DRLLNRIRE.....RAEDSGNESDGDTEEL.SLLVEMGNY..DLGVDNN..L\$		82
H1B.FR.83.HXB2	T....Q-.IP-V--VA-V--I-I-.----S--I---.Y..R-I....R-R-.....ID-LI-.....E-EISA-VEMG---HHA.PWD--D....		82
H1B.US.90.WEAU160QS.-Q-L--VA-V--G-I-.----S--L---.Y..R-I....R-R-.....ID--D.....E--Q---A---HHA.PWDI-D....		81
H1C.ET.86.ETH2220	-VDLL.AKVYRIV-VAF-----I-.-----AY--.Y..R-.....R-RR-.....IK-T-.....-TM-D---LR.L-D-ND....		86
H1D.CD.84.84ZR085QS.-Q-L--VA-VL---I-.-----F--.Y..RRIK...R-R-...-W-ID-.....E-----E--K---T---HHA.PWD--DD..		82
H1F1.BE.93.VI805SY.-LAIG-AA-----I-.-----Y--.Y..---V...R-R--..NK-YK-----E--A---AA-G---PFI.PGDIN-....		81
H1G.SE.93.SB6165QS.-V-L--V-----F-A-.....F--.Y..REIR.....-R..GK--D-----VT-----DFD.PWVG-....		81
H1H.CF.90.056YI.-GLGIGALVVTFI-A.VI----Y--.Y..---V...R-.....-IE--G-.....-K-M---HLN.LGY-AD-....		80
H1J.SE.93.SE7887I--Q-A--VAF--IF--GM---Y--.Y..---R-R-.....-K-ID-.....AD---R-PHD.LWN-ND....		81
H1K.CM.96.MP535VS.-A-SIVALVVALILA..I---Y--.Y..R--V.....-R..NW-ID-.....E--A---AD.IG-L-HLI.LGNI-....		80
H1N.CM.95.YBF30L.SLGFIA-GA-VSI..VI--ALLYR-.Y..--IK...L-E--KHIRO.....E-----A-W-DGDEEWLVTLSSSKL-QGNWV-		83
H1O.BE.87.ANT70HH.RDLL--IIISAL-FIN.VIL-GFILRK..Y..LEQK.EQDRKEREILE--R-.....IRD--DY--N-EE-QEVMD--LSHGFDNPMFE....P-		85
H1O.CM.91.MVP5180HQ.ENLL-LIA-SALCLIN.VLI-LFNLRI..Y.LVQRKQDRRE-EIL..E--R..-K.....IRD--DY--N-EE-QEVMD--LSHGFDNPMFE....P-		85
H101.AE.TH.90.CM240S--V-----S--A--V..I...R-R-.....VK-----E--D--AK.....DFD.PWVG-....		80
H102.AG.NG.-.IBNGQ.-T-T--V--V--F-A-.....S--Y--.Y..R-IR.....D-----ST..M---YEY.I-DN-....		81
H103.AB.RU.97.KAL153_2QSLAIAAIVALVVVGI.-AIVVGSIVF..IEYR-I...R-R-.....ID-----E--Q-A...M---HLV.PWDA-D....		78
H104.cpx.CY.94.CY032LF.W-----V--V---IV-----L-F--.Y..---R...R-RR...-S-Y-----A---T--G---FD.PWVG-....		81
CPZ.CD.-.ANT-N.IF.....BYAF-AFS.--L-I-CIPIL.Y..-LYKIY...-QQ-D.NKRNO--I-.VLSRRLSI.--AI-E-BEADTYLGSFGFANPVYREGDE.....		83
CPZ.CM.-.CAM3LT.W-QIGLI-IGIEI-I--A-G-AFK-.W..R-GKEENRA.QKI.RT-IE--IS.....Q-----EDQRQLDREIHVYGFDPMPFD....W-		85
CPZ.CM.98.CAM5LI.W-QIGLIA-GIEI-IV-----G-AYKK..W..REIKEERNRT.QRI.YN-AE--I.....Q-----N-EE-AQLDN-LHANGFDNPMFDW....P-		86
CPZ.GA.-.CPZGABL.-VGLVLILVGL.IAWN.-CI-GYIIKW.GY..RRYK..RHRLETEI.E--NLIL-.....N-EE--RLEQ-IHNY-HNNHFANPMFD--		88
CPZ.GA.88.GAB2LSMVAIGLI-IGTL-VIN-----G---SV.Y..-RWK..RH-EEQRI.ID-IIK-T-.....EDK-TLAT-LHNNGFDPMPFE-R...I-		88
CPZ.TZ.-.TAN1	-IKIVVGSVSTNVIGILC-LLILIG.GGLLIGI--RR.E.LERERQ..H-R.VL.E--A.-RLS.....IDSGVEEDEE...FNWNNFDPHNYNPRDW....I-		83
CPZ.US.85.CPZUSLN.WFEIGLIA-GIEG--V.VII-GL-ARL..W..RQIKIKENT-QE-Q.NL-E--I.....E-----EE--TLAK-LSSLELDNPRI.....V-		84
DEN.CD.-.CD1LKIKLGPYVCL-FAVVTWIAAIG--YLA.Y..RAYK...SYREELRYR-RLWSI.....DSGYESSQEDP.....		64
GSN.CM.99.CN166H-AAVW-WGAAI-TFIY-CVALLALYLAWDK.W..V-GK.....PKPTQVAV--L.....IEDEEDSGIYD-ASSETGFNGFA-PGFEV.....		78
GSN.CM.99.CN71SAAALW-WGAAV-TFIYFCLAIFALYLAWDK.W..I-GK.....PKIP.VAV--L.....VED-EESGIFE-ASSE.PNAYGFA-PGFEV.....		76
MON.CM.99.L1N.YWWSLVAITYSLILI..LPVAAWAW..WRYY-ITKRFKRIDQEIQRLIQIHE-R.....HDSGVDTSESESEQHEETH..GFV-PVFNDDFGEW..V-		87
MON.NG.-.NG1N.TYWI-AISVWCVIV-L..APC-LYLL..YQSY-EHKRVSRFE.DFQRLVQXYQEXDSGYEDEDEHNSFDNPLFDGDDPDQW-XFRKQIMGHSILWSAG--		98
MUS.CM.01.1085N.YWYL-AVIVTGIYFVIA-FAFVLAQY...RWC-PKKVEVSVI.....L.....LE-GD-DSGIFEDA-DDMAES-HHAFANPAFEQ.....		76

	signal peptide \/ gp120	***	*	*	***	
SAB.SN.-.SAB1C	M...KLLTVL.....LWLSGCWSLWLV.....QY.VTVFYGIPVWKNS...VQAFCK.....TPNTNLWASTNCIP.DDEPEGT.IAEVPIPNITEKFDANKNRNPLVQQAESNIHL					95
H1B.FR.83.HXB2	-RVKE-YQHLW.....R-GWRWGTMLLGLMLM..ICSATEKLV.---Y-V---EAT...TTL--ASDAKAYD-EVH-V--THA-V-.T-PNPQEVVLVNVTE-FNMW.....K-DM-E-MHED-IS					110
CPZ.CD.-.ANT	...R-PIHII..WG...-A-LI.QFIEKGT.....NED-.....V---R-AT...PTL--ATNASMTS-EVH-V--T-S-V-.I-PDPI.VVRLN.TS...VW-..NAYK-YM-ESMTEDMXQ					100
CPZ.GA.-.CPZGAB	...KVMEKKKRDWNSLSIITITITII-LTPCLT.....SELW.---Y-V---HDAD...PVL--ASDAKAHS-EAH-I--TQA-V-.T-PSQEVFLPNV-ESFNMW.....K-NM-D-MHED-IS					109
H2A.DE.-.BEN	...EPGRNQ.....FVVILLTSAC--Y...CS.....-A--A...IPL--A.....-K-RDT-GTIQ-L-.NDDYQE.-ILNVTEAFDAWN.....TVTE--VEDVWH					92
BAB.TZ.85.2010E					0
COL.CM.-.CGU1	...LRY-FIN.....-LVLGIVL.....SN..KW.---YQ-V-A-EBADVNDQ-F--F.....SSSPEIQVVLG-L-PPPGKPVQNMMPNVTEAFDLF.....K-SFS-EV...WI					88
DEB.CM.99.CM40RLIQL.LLLILGLIVIGTQKQKQKQ...-L-Y-V---VDAK...-DL--T.....ANSSESG--V-A-L-.HALVREEVPMNPVQ-FNAFD...-P...IEEQWLQDWT					95
DEB.CM.99.CM5LSLLIIIA-LTLVGAAPT.....-QW---Y-V---H-AT...-GL--T.....ANSSESG--V-A-L-.HAMVREQVPMNPVQYFDSFS...-V...ME-QLWDDMT					91
DEB.CD.-.CD1SQVRRTLKV-LLILWTIGIICTTEK--.....A-R--A...APL--T.....AMDRGV--AAS-V-.E-PYQREVWM-NATDYFDIY.....H--I-D-VVQ-MKD					96
DRL.-.-.FAOIRHI.....IIGL-ILGFLGISI...GK...HW.-----T-K-RPA...THLI-A.....-D-HSF-VT-S---.SLLHYEETEINNIEE-F-VPM.....TENEVIKQAWG					89
GRV.-.-.GRI2E					0
GRV.-.-.GRI3E					0
GRV.ET.-.GRI_677	...GR--TKI.....IIAI.GISI.GI.....GNLY.....T.....M.....M--T-----HDNTE.VPLNITEAFEAW.-N-----K-----					90
GSN.CM.99.CM166	...R-TVESM.....MM-AVSLALMALLI..PGLKC.ENW.T--Y-V---RDAT...PPL--ASDPDIASNEPG-I-I-A-L-.S-PSPAEVPL.NITEKFNIY.....K-YM-DEVRDDMVS					105
GSN.CM.99.CM71	...MRKVG.....KM-VVSLT.LMALLI..PGSRSGNW.T---V---RDRAK...PPL--ASDADITSNEPG-I-I-A-L-.S-PSPEVPL.NITEEFNIY.....K-YM-DEVRDDMVS					105
LST.CD.88.447	...NCPGLI.....LLII-RVQG.....EKTKK...-V-K-EEAV...-PLI-A.....SA-NS-VT-S---.LQSYAQ.-PIHN-SLNFT-E.....IKDN-IIQQAWS					90
LST.CD.88.485	...DC-KLL...FL-F-GAQQ.....EAKK...-V-K-EDAV...-PLI-A.....SA-NS-VT-S-L-.LQSYAQ.-PIHN-SLNFTQE.....IKDN-IIQQAWS					89
LST.CD.88.524	...ACPGNI.....-L-LCLGLTAG.....EKEK...-V-K-EDAV...-PLI-A.....SA-NS-VT-S-L-.LQSYAQ.-PIHN-TLNFTQE.....IKDN-IIQQAWG					91
LST.KE.-.lho7	...ACPLG.....IL-LLLGII..W...GK...-V-N-DDNVS...-PLI-A.....SA--S-VT-S-L-.LQSYAE.VPIYN-SENFITIP.....VKDN-VIQQAWS					88
MAC.US.-.239	...GC-GNQ.....LIAIALLS-YGIY...CT...L-...-V-A-R-AT...IPL--A.....-K-RDT-GT-Q-L-.NGDYSE.V-LNVTEAFDAWN.....TVTE--IEDVWQ					92
MND-1.GA.-.MNDGB1	...KCPGLF.....ILCLAXLVGIG.....K.....-V---EAK...THLI-A.....-D-SS-VT---SLPDYE.VEIPD-KENFTGL.....IREN-IVYQAWH					89
MND-2.-.-.5440	...LRY.....LRYIVLGIIVS-IV...GE...-W---Y--T-K-HKAR...THL--A.....-D-NSF-VT-S-V-.SLLHYEQHIPNITE-F-GPI.....EENEIVTQAWG					88
MND-2.CM.-.MND302ENV	...LRY.....IILGIIIVGLG...GN...-W---Y--T-K-HKAE...THL--A.....-E-DS-VT-S-V-.SLLHYEQIPINITW-F-GPM.....EENEVVMQAWG					85
MND-2.CM.98.CM16	...LRH.....LITAIIVIIGIGI...GK...-W---Y--T-K-HPAR...THL--A.....-D-NS-VT-S-V-.SLLHYEQIIPNITE-F-VPI.....QENEVVRQAWG					86
MND-2.GA.-.M14	...TQG--QNLNQSMMRKLMLSTLCLM...MTLVNGSNW.T--Y-V---PAT...PPL--ASDPNYSKEAG-N-LASS-L-.T-PTPQSLYL.NITEEFNAY.....QNYMVEEMVEDMK					108
MON.CM.99.L1	...IQGMRMRMKTIIIVLTILC-MMGTL.TSGSXSGNKS...-V--Y-PAT...PPL--ASDPXYGSKEAG-N-LASS-L-.S-PVPRALPL.NITESFNAY.....NYMVEEMVEDIK					109
MON.NG.-.NG1	...GC-GNQ.....LIAIALLSAYGIY...CI.....-V-A-R-AT...IPL--V.....-R-RDT-GT-Q-L-.NDDYSE.L-LNITESFDAWE.....TVTE--IEDVWH					92
MNE.US.-.MNE027	...PRMTWLRASHTTL-LLLSSSIIWGR-W...-V---DAT...PPL--ASDANIANREPG-V-IT-A-L-.T-PAPQEVQL.NISGEKFDIY.....KNYMVDQMTDDVR					103
MUS.CM.01.1085	...D-.....K-VIVLIV-IGIIL.VQGSQ.P--I---V---R-T...-PM-V.....-D-OS-GTL---EGGISPE.VSINVSEKFDAW...N-S-YE--KD-VWN					93
RCM.GA.-.GAB1	...A-.....ISI-IIIGLLGII...VGEK...-V---T...-PM-A.....-D-AQS-GTL---EGGISPE.VPVNVSEKFDAW...E-S-YE--K--VWH					91
RCM.NG.-.NG411					0
SAB.-.-.C5					0
SAB.-.-.SAB3E					0
SAB.-.-.SAB4E					0
SMM.SL.92.SL92B	...ACPLGH.....LIDILFLS-LGTW...CA...-I---A-R-AT...IPL--A.....-Q-RDT-GTVQ-L-.NGDYSE.L-LNVTEAFDAWD...TVTE--IEDVWN					92
SMM.US.-.H9	...GC-GNQ.....LIALSLXSASGIY...CV.....-V-A-R-AT...PL--A.....-R-RDT-GT-Q-L-.NGDYSE.L-INVTEAFDAWD...TVTE--IEDVWN					92
SMM.US.-.PGM53	...GC-GNQ.....LIAIALLSAYGIY...CT.....-V-A--AT...IPL--A.....-K-RDT-GT-Q-L-.NDDYSE.L-INVTEAFDAWN...TVTE--IEDVWN					92
STM.US.-.STM	...ACPGNQ.....LIAIALLSAC-TY...CT.....-V-A-R-AT...IPL--A.....-K-RDT-GT-Q-L-.NGDYSE.L-INVTEAFDAWD...TVTE--IEDVWN					92
SUN.GA.98.L14	...RCPEI.....IGFSLLLG-.IAI.....-T-K-EPAV...-PLI-A.....SA-NS-VT-S-L-.LQTYAE.VPITGLEENFTGEG...ISNN-IVQAWQ					87
SYK.KE.-.KE51	...KLFNIIISVVLGIVLSLIGVARLE...-V-N-EDAQ...-PM-A.....-H-GG-TK-V-.SS...QIEVRVNISGEYFTAW...SSHGIR-QLLQDMS					94
SYK.KE.-.SYK173	...AAFR-YIVCLFSLISLGF.....MEK..Q...-H-EDAY...APL--T...-SHKGG-TK-V-.SA...DQIEVRVNITGEYFPW...SSHMIR-QILEDMS					93
TAN.-.-.TAN17E					0
TAN.-.-.TAN40E					0
TAN.-.-.TAN49E					0
TAN.CF.-.AGMB05					0
TAN.CF.-.AGMTB14	...GP-RGKG.....VL-VILGLSLIGL...LYGT-.I-----M.....-T-----HNYTE.VQLNVSEKFEAWK-R...-A-----					55
TAN.UG.-.TAN1	...GP-RGKG.....VL-VILGLSLIGL...LYGT-.I-----M.....-T-----HDYTE.VQLNVSEKFEAWK-R...-A-----					96
VER.-.-.VER1E					0
VER.-.-.VER2E					0
VER.DE.-.AGM3	...KLT-LIG.....IL-I-IGVVLNTR...Q...-W---V-----M.....-T-R---T-S---HDYTE.VPLNITEPFEAWA-R...-A-G---					93
VER.ET.-.VER385E	...KP-GIV.....IAII.GISS..L...SER..W...-M.....-T-R---T-----HDYTE.VPLNITEPFEAWA-R...-A-G---					93
VER.KE.-.9063	...T-F-GIF.....IV-GI.GIGI.GI...STKQ-W.I---V-----M.....-T-R---T-----HDYTE.VPLNITEPFEAWA-R...-A-G---					94
VER.KE.-.AGM155	...RYTIIT.....GIIVI.GIGI.VL...SK..-W.I-----M.....-T-S---T-----HDYTE.VPLNITEPFEAWG-R...-IA-A---					92
VER.KE.-.TYO1					0
VER.KE.-.VER266E					0

Accession	Protein Name	Sequence	Score
SAB.SN..SAB1C	..AT...HTV.....	GCYMIHCNTSVIKEADCKTYWDTFRLRYCAPAGYALLRCADTDYSGHK..ACRNVTVSACTRLINTTVSTGIGINGSYVA.NRTEIW.....QKNGNSNDSVVIIRLNR	299
H1B.FR.83.HXB2T.....	SYKLTSS---TQ--P-VSFEPIPIH-----F-I-K-NKTFN-TGP..-T--STVQ--HG-RPV---QLLL---LAE.EEVR-R.....SNVFTDNAKT-IVQL	288
CPZ.CD.-.ANTCY-LW---TT-TQS-E-STFEPIPIH-----I---E-E-FT-VGM..-K--S-VH--HG-SPM-A-WLLL--T-QTNTSVMM.....GRKN..ESVLVRFGKE	288	
CPZ.GA.-.CPZGABTYRI-N---TA-TQ--P--SFEPPIPIH-----F-I-K-N-K-F--KGK..-T--STVH--HG-KPV-T-QLL---LAE.GNITVR.....VENK-KNTDVWLVQL	286	
H2A.DE.-.BEN	DNT-....AG.....	T--R-----I---S---H---AM-F-----P-F-----N--N---FEP.K-TK-VAAS---MME-QT--WF-F--TRAE.---Y-Y.....WHGRDNRT--S--K	297
BAB.TZ.85.2010ELLL---SE.-K-Q-----HRVN---L-L--K	34
COL.CM.-.CGU1NSI.....	RAT-KD-ANYVTQV--M-IV-PV-TGF--AP--M-----D-KKWD-TGA..-N---AVS--HEF-I--MSHVLV-A-KEL.SDWAKD.....REGVWKNDSGT-EYYWF	288
DEB.CM.99.CM40	ADNK.....	DGSGNRTG--K--D--TQS-ETSRFKP-KI-----G---D-KNFN-TGL..-N---AV--N--H-MA--WVQF---DEE.RAE-LHIIRK.EV-GEVQ-G-IT--VPA	300
DEB.CM.99.CM5	ETEN.....	QTNTNKSIRHG--N--N---TQS-ETSRFKP-KI-----G---D-RNFN-TGL..-T---AV--N--H-MA--WMQF---DEE.RAED-HIIRK.....G.LE-KTITV-IPP	286
DEN.CD.-.CD1NQT.....	YY--RY--Q-T-TQ--E-VVFEP-KI-M-----F--K-N-HPW--YGL..-K-MSARV--DE-H-MA--WLIM-STVEL.LD.NHTQ...IIRNTSKD-QTIA-LFRK	296
DRL.-.-.FAO	KCGNETESPESNCTNKKR.T-VTS--ST--TQD-N-AST-MMKF-L---P--I---NEKLNISS...-K-I-AVQ---HMPA-I-SHF-F--TKHG.QDEL-ET...RG-PDFL-HKYVF-V-K	313
GRV.-.-.GRI2EL-----F-----L-KN-N-FAP.S-K---TH-----L--RAE.-----EK---NT--K-K-K	79
GRV.-.-.GRI3EL-V-----K-N-K--N-FAP.K-K--S-VH-----IT---L---R-E.-----EK---T--K-K-K	79
GRV.ET.-.GRI_677GSK.....	D-----D-----L-V-----K-N-K--R-FAP.K-K--S-VH-----IT---L---R-E.-----G--D--T--K-K-K	292
GSN.CM.99.CN166NQS.....	Y-IRN---Y-TQK-V-SSFQVPPIH---P--M-K-N-VNFT-VGT..-Y--SAVT--HG-QPL-A-WLHL--T-QP.GN.NTR.....IMM--MK-E-IV-GFGE	300
GSN.CM.99.CN171NDT.....	DY-I-N---YVTQK-V-SFQVPPIH---P-F-M-K-N-ANFT-VGK..-S--SAVT--HG-QPL-A-WLHL--T-QP.GN.NTR.....VMM--KK-E-IV-GFGE	302
LST.CD.88.447	DKNG.....	T--T--D-I-QQD-N-GVISNAYF-L-----M--K-NEKLNFSAN...-T-I-ATP--DYMISS--SFF-F--TNHT.DDEL-PLTP.KKM.GNLLLEAKFVYKVG	310
LST.CD.88.485	DKNG.....	T--T--D-I-QQD-N-GVISNAYF-L-----M--K-NEKLNFSAN...-T-I-ATP--DYMISS--SFF-F--TNHT.EDKL-PLT...PKKMGDNLGAKFVYKVG	309
LST.CD.88.524	NENG.....	T--T--D-I-LQD-N-GVMSNAYF-L-----M--K-NEKLNFSAN...-T-I-ATP--DYMISS--SFF-F--TNHT.EDEL-PLT...PKKMGDNLGAKFVYKVG	312
LST.KE.-.lho7	SGERENR.....	T--T--D-I-TQD-N-GVMQNAIF-L-----M---NEQLNFSK...-E-I-ATP--GYMLSS--SFF-F--TNHT.RDEL-PLT...PNKMDLNGAKFVYKVG	305
MAC.US.-.239	EQGN.NTGNES.....	R--N--H--Q-S---H--AI-F-----P-----N--N---FMP.K-SK-V--S---MME-QT--WF-F--TRAE.---Y-Y.....WHGRDNRT--S--K	303
MND-1.GA.-.MNDGB1	ENN.....	T--N--E--NT-D-Q-GL.LRCLIG-V-P-VM--YNEKLNK...-S-ISAVQ--QHLVA--SFF-F--TMHK.EGEL-PIDDKYRGPEEFHRQRKVFVYKVG	295
MND-2.-.-.5440	KLKNTGSATNSTEPY.E--TS--AT--TQDRN-AST-RMTF-L---P-FV--K-NEKLNKTKL...-G--SAVQ--APLPA-I--MF-F--TKHD.YDEL-QTNP.RKG-DEFHDKHYVY-VDK	317
MND-2.CM.-.MND302ENV	0
MND-2.CM.98.CM16	NGTNN-Y.....S---TQ-----TQD-N-AST-EIKF-L---P--V---LERLNVSK...-T-I-AVQ--QPLPA-T--MC-SS-TKHD.YNEL-QTNT.KKG-EEFPHDKHYVY-VDE	296
MND-2.GA.-.M14	KKEND-E.....TE.E---T-----TQD-N-AST-KMTF-L---P--I---REKLNKTKL...-A--SAVQ--DPMPA-I--MF-F--TKHN.YDEL-L...VNPQKFEFPHDKHYVY-V-K	293
MON.CM.99.L1SNGSHY-ILN---Y-TQ--E-SNYEVPV-H---P-----D-PAFT-QGS...-S--SAVT--HA-QPI-A-WFQL-STGN..PN.TTV.....MM-KQK-E-IVV--AK	298
MON.NG.-.NG1NTDDYY-LLN---XYXTQ--E-SNYEPIPIH---P---I---D-PXFT-QNN..-X--SAVT--HG-MPI---WFQL-ATSAT..-X.TTV.....MM--KR-E-IV---XK	300
MNE.US.-.MNE027	EQGN.STEDES.....	R--N--Q-S---H--AI-F-----P-----N--K---FMP.N-SK-V--S---MME-QT--WF-F--TRAE.---Y-Y.....WHSKDNRT--S--K	304
MUS.CM.01.1085	DSY-ITN---Y-TQ--E-SSFQVPPIQ-----S--K-N-ANFT-QGE..-N---ARH--HG-LPL-A-WLQL--T-L.E.GN.NTAV.....MM--DK-E-IG-KFGE	301
RCM.GA.-.GAB1	EKE-NSTNS-KK.....	K-FVR---TS-QQF-EPK--EP-----P-F--V-K-KN-T-FDT..-V---ATS--HM---AS-F-L--INV.-E-W-Y.....QRQRNRTV-G--S	301
RCM.NG.-.NG411	KAG..NNNTR.....	Q-FIR---TS-QQF-EPK--EP-----P--I-K-L-AN-T-FDT..-Q---ATS--HM---SF-L--ISE.---W-Y.....QRKQSNRTV-G--S	295
SAB.-.-.C5	0
SAB.-.-.SAB3EK-K-----	78
SAB.-.-.SAB4ER-S--R---N-Q--S-----E---NTA-----	78
SMM.SL.92.SL92B	EQGG...NESS.....	R--N--Q-S---H--AI-Y---P-----N--N---FAP.N-SK-V--S---MME-QT--WF-F--TRAE.---Y-Y.....WHGGSNRT--S--K	306
SMM.US.-.H9	EQNSNGNETDS.....	K--N---Q-S---H--AI-X---P-----N--N---FAP.N-TK-V--S---MME-QT--WF-F--TRAE.---Y-Y.....WHXRSNRT--S--K	306
SMM.US.-.PGM53	EQS..TSRNES.....	K--N---Q-S---H--AI-F---P-----N--N---FAP.N-SK-V--S---MME-QT--WF-F--TRAE.---Y-Y.....WHSRSNRT--S--K	310
STM.US.-.STM	EQN..VTGEES.....	R--R---Q-S---H--AL-F---P-----N--N-T-FAP.N-SK-V--S---MME-QT--WF-F--TRAE.---Y-Y.....WHGRDNRT--S--K	302
SUN.GA.98.L14	NGND.....	T--T--D-I-TQD-H-GIMQNAIF-L-----M---DEKLNATK...-K-I-ATP--NYMYS--SFF-F--TRHK.EDEL-PIN...NKVQQAEGEYVWVAA	308
SYK.KE.-.KE51	EEGN.....	ESYY-LQ---MRN--E-QTFRP-PIQ---P--S--K-N--EFE-DAE..-Q---AVS--HPF--LA--WFQL--T-K-.KDKVRFIR...-DKNES---LVPE	290
SYK.KE.-.SYK173	EEGN.....	SSYY-LL---SA--E-QTFQP-PIQ---P--S--K-N--NFE-DDV..-T---AVS--QEF--LA--WFQL--T-K-.KDKVRFIK...-DKNES---LVPE	283
TAN.-.-.TAN17EF-I-K-K-N-T-FG..I---S-VS--G-M---SAF-L---QAE.---V-----H-V-----K-K-K	78
TAN.-.-.TAN40EL-----F--K-R--N-A-FG..I-N--S-VS--G-M---I-SAF-L---QAE.---H-----K-K-K	78
TAN.-.-.TAN49EL-----F-I-K-K-N-T-YG..I---S-VS--G-M---SAF-L---QAE.---HRIN-----K-K-K	78
TAN.CF.-.AGMB05D-----E-----L-----AV-I-K-K-N-T-FGT..-S-VS--G-M---SAF-L---QAE.---H-V-----K-K-K	15
TAN.CF.-.AGMTB14NSR.....	D-----E-----L-----AV-I-K-K-N-T-FGT..-S-VS--G-M---SAF-L---QAE.---H-V-----K-K-K	257
TAN.UG.-.TAN1G-R.....	D-----E-----L-----F-I-K-K-N-T-FGV..-S-VS--G-M---SAF-L---QAE.---H-V--N-----K-K-K	300
VER.-.-.VER1EEL-----K-N-RN-T-FAP.R---S-VH--G-M---LLL---TE.-----HKV--N-----I-K	79
VER.-.-.VER2EEL-----K-N-H--N-Y-Q.N-S--S-VH--G-M---T--LLL---SE.---Q-----R-H-V-----LI-K	79
VER.DB.-.AGM3SNR.....	E-----D-----EL-----F--K-N-Y--A-F-T.N-S--S-VH--N---T--LLL---SE.---Q-----HRV-----LVLF-K	309
VER.ET.-.VER385EEL-----F--K-N-Y--D-F-Q.N-S--S-VH--G-M---T--LLL---SE.---Q-----HRV-----L-QI-K	79
VER.KE.-.9063S-K.....	E-----D-----E-----EL-----F--K-H--N-Y-Q.N-S--S-VH--G-M---LLL---SE.---Q-----HRVN-----VLF-K	309
VER.KE.-.AGM155G-K.....	E-----D-----E-----EL-----K-N-W--A-F-P.E-S--S-VH--T-M---T--LLL---SE.---Q-----H-V-----L-L-K	304
VER.KE.-.TYO1TSK.....	E-----D-----QL-----F--K-N-E--N-Y-Q.N-S--S-VH--G-M---T--LLL---SE.---Q-----HRVN--NT-L-LF-K	303
VER.KE.-.VER266EEL-----F--K-N-A--D-F-T.N-S--S-VH--N-M---T--LLL---SE.---Q-----HRT-----L-L-K	79

	gp120 \/ gp41	*
SAB.SN.-.SAB1C	WANELGDYKLVKPIGFAPTKVRRYTG...PERQKRVPF.VLG..FLGFLGAAGAAMGAAATALTQVSQQLLAGILQQQKLL.AAVEQQQMLKLTIWGVKNLNARVTALEKYLEQDARLNWGCASF	615
H1B.FR.83.HXB2	-RS--YK--V-K-E-L-V---AK-RVVQ.....E--A.VG.I-AL-----ST---SMT---AR---S-V---N---.R-I-A--HL-Q-V--I-Q-Q--IL-V-R--K--QL-G---SG	600
CPZ.CD.-.ANT	-RA--SR-QV---X-LSV---TXK-PEIKQHS.....GI.G.I-LF---L-S---ST---SI---A-TRN-XH--V---A---.Q-I-T--HL-Q-SV---Q-Q--ML-V---R--QL-SL---D	593
CPZ.GA.-.CPZGAB	-RS--YK--V-R-E-LSV---A-H-VARQKD.....AA-G.-AL-----ST---VT---AR---S-V---N---.K-I-A--HL-Q-S---Q-Q--LL-V-R--Q--QI-GL---SG	590
H2A.DE.-.BEN	YRL-----I--T-----DQ---SS.T..-V-N-GV-.....-AT--S---RSLT-SA--RT-----V--QQ---DV-KR--E--R-V--T---Q---I---KH--Q--S---S	598
BAB.TZ.85.2010E	--A---R-----T-----E-----G...Q-----.....-A-----T-----D-	302
COL.CM.-.CGU1	STL---R---KMRITTYV--DIK-SVNVNHHG---GI-AFSI...-AL-SG-----S-SV---I-A-S-NGRASASSNRM-LKL--T-SAL-Q-V-----QV--ATI-G---E--K-ASI---N	599
DEB.CM.99.CM40	-RAD-FK--II-V-----SA---E-P..ESV-H--AAGIAP-.LVA-ST-----S-----RS-S-V---QE--.K---AHG-L-T--A---R---T-L-I---K---K--E-----	606
DEB.CM.99.CM5	-RAD-YK--II-V---M---SI---E-P..ESA---AAGIAP-.LVA-ST-----S-----RS-S-V---QE--.K---AHGHL-S-A--R---T-L-I---K--SK--E-----	586
DEN.CD.-.CD1	-RA--ARW---LT---L---S---E-PESAESVSR-RREVS-VLGLI---S---TV---VG-G-A---RD---V---QE--.R---GHSAL-Q-SV-I-----L--I---K---L--E---W	607
DRL.-.-.FAO	FTAVGAH---KVE---D-H-NLPG...H---GA.V---.I--L-SL-S--CSV--M---A--T-MVE--Q--.RL---EL-----Q--L---E-IG---M-SL---S-	649
GRV.-.-.GRI2E	--Y-----T-----D-----P...TK-----.....T-----RH-----R-----L-----	367
GRV.-.-.GRI3E	--Y-----T-----D-----P...S-E-----.....T-----RH-----V-----L-----	371
GRV.ET.-.GRI_677	--Y-----T-----D-----PT...-E-----.....T-----T-----RH-----L-----S---W	610
GSN.CM.99.CN166	-NY--SR--V---D-LSM---PAK-KEHP...AVE--A.LS---IS--T--S---TT-----E--RS-----M--E---.R---A--SL-QPSV--I-Q-QT-LSS-----R--TI-QA---N	612
GSN.CM.99.CN71	-KY--SR--V---D-LSL---SAQ-RPHPGV..HVK--A.IS---IS--T--S---T-----RS-----V--E---.R---A--SL-Q-SV--I-Q-Q--LSS-----R--TI-QA---N	621
LST.CD.88.447	FVATGAH---IRLR-----D-H--AP...SSK---AAPLA-.A--L-S---T---LVS-I---A-VVIQ-----Q--.VL-K-EL-R-----Q--L--I-E-K--TL-AS---QW	639
LST.CD.88.485	FVATGAH---IR-R-----AEHQ-AP...SSK---AAPLA-.A--L-S---T---LVS-I---A-VVIQ-----Q--.VL-K-EL-R-----Q--L--I-E-K--L-AS---QW	638
LST.CD.88.524	FVATGAH---IRLR-----SEH--AP...RG--T-AAPLA-.A--L-S---T---LVS-I---A-VVIQ-----Q--.VL-K-EL-R-----Q--L--I-E-K--L-AS---QW	641
LST.KE.-.lho7	FVA-GSK--IKLN-----DEH--AP...RG--T-AAPLA-.A--L-S---T---LVS-I---A-AV-Q-----Q--.VL-K-EL-R-----Q--L--E-VKH--L-AS---QW	636
MAC.US.-.239	YRL-----T---L---D-K---TGG..TS-N--GV-.....-AT--S---SLT-A--RT---V--QQ---DV-KR--EL-R-V--T---QT---I---K---Q--A---S	613
MND-1.GA.-.MNDGB1	FVAYGAH---K-M---I---D-K-H-LPE..HHKE--GA.VI---.I--L-SL-S--SVSV-----S-VT-VE---Q--.KLI---SEL-----QT-L-S--N-IK---L-SQ---SW	632
MND-2.-.-.5440	YTAVGAH---KVD-----DIH-HHLPN...T---GA.VL-.M--L--L--S---SV-V-----T-N-VE--V--.SLID-HSEL-----QV-L---E-VA--S--SV---S	645
MND-2.CM.-.MND302ENVXVSV-----A-N-VE--I--.SLID-HSEL-----Q-L---D-VA---SM--S-	69
MND-2.CM.98.CM16	FTAVGAH-Q--KVD-----E-A--HLPE...A---GA.VL-.MF-L-L--ST--SV-V-----A-N-VE--V--.SLID-HSEL-----Q-L---D-VA---SM--S-	622
MND-2.GA.-.M14	FTAVGAH---KVD-----D-H--HLPD...AK---GA.VL-.M--L--L--S---SV-V-----A-N-VE--I--.SLID-HSEL-----Q-L---D-VA--S-AV--S-	623
MON.CM.99.L1	-RS--SR--V---D-LSM---TAQ-R--V...H-E--AI..T-MA---ST--GT-----RS--V--E---.R--TA--SL-Q-V---Q-Q--L-V-FIK--TL-A---N	616
MON.NG.-.NG1	-RS--SK--V---D-LSV---AAQ-RH-L...M-ER-AI..T-MA---ST--GT-----RS--V--E---.R--TA--SL-Q-SV--I-Q-Q--L-V-FIK--TL-S---N	621
MNE.US.-.MNE027	YRL-----T---L---N-K---TGG..TP-N--GV-.....-AT--S---SLT-A--RT---V--QQ---DV-KR--EL-R-V--T---QT---I---K---Q--A---S	616
MUS.CM.01.1085	RMEPKKVQ-V---D-LSM---E-Q-RPVP...HVHARAI.SLGIT...-S---GT-----RS--V--A---.R--A--HL-Q-SV--I-Q-Q--L---FIK--L-L---N	624
RCM.GA.-.GAB1	--A---R--AI--R-----EIK--QT.....K---.PL.VLG.---S--T-----RH-----DI-KR--NL---V--T---Q---I---A--SL--TF--W	604
RCM.NG.-.NG411	--A---R--V---M--Y---N---ET.....K---.PL.VLG.---S--T-----RH-----K---.DI---EL---V--T---Q---I---A--SL--TF--W	598
SAB.-.-.C5E-----P-----	88
SAB.-.-.SAB3EE-----P-----	368
SAB.-.-.SAB4EN-----P-----	368
SMM.SL.92.SL92B	YRL-----T---M---H-K---T.S..TSKN--GV-.....-AT--S---SLT-A--RT---V--QQ---DV-KR--EL-R-V--T---QT---I---K---Q--S---S	620
SMM.US.-.H9	YRL-----T---S-K---TTG..AS-N--GV-.....-AT--S---SLT-SA--RT---V--QQ---DV-KR--EL-R-V--T---QT---I---K---XX-S---S	620
SMM.US.-.PGM53	YRL-----T---L---S---T.G..AS-N--GV-.....-AT--S---SVTRSA--RT---V--QQ---DV-KR--EL-R-V--T---QT---I---RK--Q--S---S	624
STM.US.-.STM	YRL-----T---L---N-K---T.S..TS-T--GV-.....-AT--S---SLT-A--RT--T-V--QQ---DV-KR--EL-R-V--T---QT---I---K---Q--S---S	616
SUN.GA.98.L14	FVSTGAH---R-R-----DEH--AP...RKE--AAPVA-.A-AL-S---T---LVS-I---A-AV-Q-----Q--.VL-K-EL-R-----Q--L---E-VQ--SL-AS---QW	642
SYK.KE.-.KE51	-RAD-FN--IR-----SQ---ELP...TK---AAPLA-.--L-S---T---S---L--T--V--QK--.E--A--HL-G-V-----L---T--R---IMSN---	595
SYK.KE.-.SYK173	-RAD-FN--IQ-----DQ---ELP...NT-E--AAPLA-.--L-S---T---G---L--T--V--QK--.E--A--HL-G-V-----L---T--R---I-SN---	589
TAN.-.-.TAN17E	--Y---R--II-----D---V-P...T-E-----.....S---R---T-E-----	371
TAN.-.-.TAN40E	Y-Y---R--I-----E---V-P...T-E-----.....R---T-E-----	373
TAN.-.-.TAN49E	--Y---R--I-----E---V-P...T-E-----.....R---TE-E-----	379
TAN.CF.-.AGMB05	--Y---R---IK--	234
TAN.CF.-.AGMTB14	--Y---R---I-----E--	487
TAN.UG.-.TAN1	--Y---R---I--E---EI---V-PT...-E-----.....T-----T---L---	621
VER.-.-.VER1E	--S---R--T---P-----P...S-E-----.....T-----H-----A-----	372
VER.-.-.VER2E	--A---R--T---D-----G...H-----Q...-T-----T-----H-----A-----	375
VER.DE.-.AGM3	--A---R--T---E-----G...HD-T-----.....T-----H-----A-----A---W	631
VER.ET.-.VER385E	--A---R--T---E-----G...H---I-----.....T-----H-----A-----	375
VER.KE.-.9063	--A---R--T---N-----G...QD-S-----.....T-----H-----M-----A-----V---W	631
VER.KE.-.AGM155	--A---R--T---E-----G...Q-----.....T-----H-----GA-----A-----A---W	626
VER.KE.-.TYO1	--A---R--T---E-----G...H-----.....T-----SS--RH-----A-----S---W	624
VER.KE.-.VER266E	--A---DR-----T---E-----G...H-----.....T-----H-----A-----	379

SAB.SN.-.SAB1C	RQVCHTTVLWKYNN.....TPDWENMTWQEWERQIEKYEANISRILEQAHEQBQKNLDSYQKLVSWSD.FWSWFDLTKWFGWMKIAIMVIAGIIVARVLLVIIGILRKRFRKGYAPLSSLP	729
H1B.FR.83.HXB2	KLI-T-A-P-NASWSNKS.....LEQI-NHT--M--D-E-NN-TSL-HSLI-ESQN-QE--EQELLE-DK-AS.L-N--NI-N-LWYI-LF--IVG-LVGL-IVFAVLV-VNRV-Q--S-.L-FQ	718
CPZ.CD.-.ANT	KVT----P-NNSWVNFQTC.AKNSSDIQCI-----D-LVONSTGQ-YN-QI---QER-KKELYE-DK--S.L-N---I-Q-LWYI--F--IVGA-VGL-I---LVSC--V-Q--H-.L-FQ	720
CPZ.GA.-.CPZGAB	KA--Y--P-NNSWPG.S.....NSTDDI-G-L--Q-DKLVSN-TGK-FGL--E-QS-QE--ERDLLE-DQ-AS.L-N---I--LWYI--FL-AVG--GL-IIMTVFSVV-RV-Q--S-.L-Q	709
H2A.DE.-.BEN	-----P-VNDS.....LS--K-----K-VRYL-----QS--E-QI-QE--MYEL--N--DI.LGN-----S-VKYIQYGVHIVV--AL-IAIYVVLQ-SR-----R-VF-S	713
COL.CM.-.CGU1	M-I-R-I-P-NKTW.....GEEDP-Q---KQ-HERVRN-TDI-EAD-VE-YDLQEE-EKKLAE-GD-TN.WF-G-G-FNI-KYVLY-AY-VG-L-GL-IIM-V-ACI-GAFRVKGF.QQIG	714
DEB.CM.99.CM40	K-I---P-NNSLE.....D--D-----MKVAN-TDEWEGA-QR-Q--QER-VHAL-S-QD-DS.L-N----SR--W-IRLVVYI--AL-LL-IAMFGVN-GS-LCR--S--QT-T	721
DEB.CM.99.CM5	K-I---P-NHTWG.....E--N-----KVAN-TDEWEGA-QR-Q--QER-VHAL-S-TD-DS.L-N----SR--W-IRLVVYI--AL-LL-IAMFGVN-GT-LCR--S--QT-T	701
DEN.CD.-.CD1	K-I---S-E-SNQSIPGL.....ALYQ--FQ---Q--TD-AALFG-V-DA-KK-QA-QER-VHEV-S-YD-DN.L-N----S---W-IRLVVY---LVIL-IIMF--NV-SRLCR--S--LQI-	728
DRL.-.-.FAO	A-----N-V-PNES.....V--N-TSE--M--QKRVDISISN--TLD-QK-Y-----IFEL--GDLTS.WAN---F-W-SKYI--GFFIVMA--GL-I-AALWSTVGR--Q--R--PY.I	763
GRV.ET.-.GRI_677	K-----P-----K-D---L-----NAL-G--TQL--E-QN--S---L---DD--G-----S-ST-L-YV--GFL--VI-LGL-FAW-LW-CI-NI-Q--N-.L-PQ	722
GSN.CM.99.CN166	-PI---I-P-NTSWANGS.....L-----K-SMLV-NDTYT-QQL---NQ-QAS--NELM--SK-DS.L-----ISD-QRYI--FVI-V-AL-AL-IVMF-LNM--RI-Q--S-.L-PQ	729
GSN.CM.99.CN71	QPI---I-P-NDSWAKNS.....-----H-----SKL--NDTYT-QQL--N-NH-QS--MNDLL--SK-DS.L-----ISN-LWYI--F---V-ALVAL-IIMFVLLNM--RV-Q--S-.L-PQ	738
LST.CD.88.447	K-----N-E-N-----I--N-TRD--I--D--VGVL-----TL-QE-YTT-LE-RNAFK--QEPN..--N-L-ILS--QYI-Y-VLI-I---L--VSF-VQIV-MCR--RV.LAPS	751
LST.CD.88.485	K-I---N-E-N-----I--N-TRD--I--D--RVGVL-----TL-QE-YTT-LE-RNAFK--QEPN..--L-ILS--QYI-Y-VLI-I---L-IVSF-VQIV-MCR--RV.LAPS	750
LST.CD.88.524	K-----N-P-N-----V--N-TRD--I---VGSL-----TTL-QE-YTT-LE-RNFK--QEPN..--M--T--QYI-Y-VLI-I---IL-I-SF--QSVV-MCR--RV.LAPS	753
LST.KE.-.lho7	K-----N-E-T-----I--N-TKD--R--SKVAI-DK--TSL-QE-YTT-LE-QNKFK--QEPN..--L-ISH--TVV-Y-VLI-LV--GL---SF--QNVV-MCR--RV.L-PS	748
MAC.US.-.239	-----P-PNAS.....L--K-N-E-----KVDFL-E--TAL--E-QI-QE--MYEL--N--DV.-GN---AS-IRYIQYGVYIVV-V-LL-IVIY-VQM-A-L-Q--R-VF-S-	728
MND-1.GA.-.MNDGB1	A-----S-E-TNTS.....I--N-TSE--K--TRTDYQQ--TEM-K-YDR--R-TYEL--GDLTS.WA---F-W-VQYL-WGVFLVL--GL-I--ALWNTISR--Q--R-VF--	745
MND-2.-.-.5440	S-----S-K-PN-S.....IV-N-TSE--L--D-RVNSIVT-MTID-QR-Y-L--R-IFEL--GDLNFHGLTG---W-LKYV--GLL-VVV--GL-M-ACLWSV-G---Q--R--PY..	759
MND-2.CM.-.MND302ENV	S-----H-P-PN-S.....IQ-N-TSE--L--KRVDLTK-MTVD-QR-Y-L---MYNL--GDL.TSWA-----W-LKYV--GLLIVVV--GL-M-ACLWSV-GR--Q--	177
MND-2.CM.98.CM16	A-----H-P-PNDS.....I--N-TSE--L--DKRVTALTD-MTVN-QK-Y-L---IYELE--GD-TS.WA---F-W-LKYV--GLLIVIV--L-I-ACLWSV-G---Q--R--PY..	735
MND-2.GA.-.M14	S-----N-P-PNES.....I--N-TSE--L--D-RVTAITN-MTID-QR-Y-L---MYEL--GDLTS.WA-----W-LKYV--G-LI-MVV-GL-I-ACLWSTIGR--Q--R--PY..	736
MON.CM.99.L1	KA-----P-NNSWAKGH.....F-E-D-----Q-SELVDNDTMT-QQL--A-Q--QG--QHELM-PGQ-DF.L-N---IS--LWYI--F-I-V-AL-GL-I-MF-L-VISRLGQ--SL.L-SQ	733
MON.NG.-.NG1	-AX---Q--NNTWAKGH.....F-E-D-----Q-SMLVDNDT-L-QXL--E-Q--QG--AHELM-GQ-DW.L-N---IS--LWYI--F-I-V-ALVGL--MF-L--I-RVQG--SL.L-PQ	738
MNE.US.-.MNE027	-----P-PNAS.....L--N-N-E-----KVDFL-E--TAL--E-QI-QE--MYEL--N--DV.-GN---AS-IRYIQYGVYIVV-V-LL-IVIY-VQM-A-L-Q--R-VF-S-	731
MUS.CM.01.1085	--I--R-P-NDSWANHT.....Q-G-----Q-S-LVDNDTTT-QEL--L-QR-QEE-QHKL--LE-DS.L-E---IS--LWYI--FC--V--LVLF-LVMFVL---RV-Q--S-.L-PQ	741
RCM.GA.-.GAB1	-----V-P-TF.....NK--E-QKES-LQ--N-SYL--TIA-QE-QD-HE--VHELE--SN-G.AF--LN-DW-MQYI--GFFIVI--GL--AWLLWNC-SNL-Q--R--PPS	718
RCM.NG.-.NG411	-----E-I.....SQ--E-NRQ--L--N-SRL-G---VA-QD-Q--HER-VHDE--N-G.ML--LNMDW-LKYIR-G-PI-L---GL-I-FLLWSCVS-L-A--T--L-P-	712
SAB.-.-.C5	-----P-F.....-----KS-ANL-G---KT-A-----Q---A-----S-----V-----V---FAV-SLV---KA---.L-S.	200
SMM.SL.92.SL92B	-----P-PNDS.....LV--N-----KKV-FL---TQM--E-RL-QE--MYEL--N--DV.-GN-----S-VRYIQYGVFLVI--VLL-IVIYVVM-SRL-Q--R-VF-S-	735
SMM.US.-.H9	-----X-P-PNDT.....L--X-N-X-----K-VNPL---TXX--E-QI-QEX-MYEL--NXXDX.XGN-X---X-IRYIQYGVFLVI-V-GL-IVIYVVM-ARL-Q--R-VF-S-	735
SMM.US.-.PGM53	-----P-PNAS.....LV-N-N-----VDDL---TQA--E-QI-QE--MYEL--N--DI.-GN-----S-IRYIQYGVFLVI-VVGL-IVIYVVM-ARL-Q--R-VF-S-	739
STM.US.-.STM	-----P-PNDS.....LV--N-----KVDFL---TQL--E-QV-QE--MYEL--N--DV.-GN-----S-VRYIQYGVYIVV-LVML--AIY-MQL-ARL-----R-VF-S-	731
SUN.GA.98.L14	K-----N-P-N-----I--N-TKD--M--D--VKM-DD--TAL-QE-YVT-LE-QNKFKQ-QEPN..--L--SQ--LYI-Y-VLI-GI--A--I-SF--QQIYRMCQ--RV.L-PS	754
SYK.KE.-.KE51	K-I---A-T-QQACGNNSR.....CP-Q-----HT---VDNLTDH-DNL-RE-Q--QE--VHDLT--QE-DS.L-----S---QYL--GFFA--A-VIL---SFAW--V-NMLG--S--LQK-	716
SYK.KE.-.SYK173	K-I---A-T-EKACGNNSN.....FCPK-Q-K---HR--QEVNDLTDH-DGL-RE-Q--QER-VHDLT--QE-DS.L-----S---FYL--GFY--GALVLL-LVSPSV--IKNLLG--V-ILQN-	711
TAN.UG.-.TAN1	K-----P-TF.....-----D-----S--TAL-G---TT-VK-Y-----M-T---GD-TS.W-NI--VSS--W-I-WGFYIVI-L-LF-MAWL-W-CIARV-Q--F-.L-PQ	734
VER.DE.-.AGM3	K-----P-QW.....R---N---L-----SYL-G--TQ--E-RA--E---A---S-----FS--LNIL--GFLDVL--GL-L-YTVYSIARV-Q--S-.L-PQ	745
VER.KE.-.9063	K-----P-QWQ.....M--N-Q---L-----GEL-G--TQ--VK-R--E---A--R-T--N-----FS--LNIL--GFL-VV--GL-L-YT-YSIGRV-Q--S-.L-PQ	745
VER.KE.-.AGM155	K-----P-TW.....E--N---L---K---GL-G--TKQ---R--E---A---S-----FS--LNIL--GFLAVI-V-GL-L-YTLYTCIARV-Q--S-.L-PQ	739
VER.KE.-.TYO1	K-----E-PWT.....R---Q---L-----ADL-S--TQ--VK-R--E---A---T-----PS--LNIL-MGFL--V--GL-L-YTVY-CIVRV-Q--V-.L-PQ	738

SAB.SN.-.SAB1C	SSH\$...QIHIHLKDGQPKDEEQQSGGLNSDSRSTYTWQREFLRHLCHRLITWLRNLTWSFSTIFSNLHR.....CLQD.IQ.....QTRTQL...TAHLEY	813
H1B.FR.83.HXB2	THLPT..PRGPDR....-EGI--E-GERDRDR-IRLVNGSLA-I.....WDDLRLSCL--YHRLRDLLLIVTRIVELLG.....RR.GW.....EALKYW...WN-LQ	801
CPZ.CD.-.ANT	..IPTQN-QDPEQ....-EEIR-E-GRKDRIRW-ALQHGFFA-L.....WVDLTSIQWIYQICRTCLLNLWAVLQHL.....TF-LCNHLENN-ST	804
CPZ.GA.-.CPZGAB	TLIPV..-REQGR...LGEID-G-GEQDR-R-VRLVEGCLP-I.....WDDLRLNLGIW-YQSLTSLACNVWRQKLTG.....HLILH.SL.....RLL-E-LCLL..GGIIQ	799
H2A.DE.-.BEN	PGYLQQIH--K.DRGO.-AN-GTEEDV-GD-GYDLWP-PINYVQF-I-L-TRL-IG-YNICRDLN-KNSP.....TRRLISQS-TA.-R.....DWL-LK....-AQLQ	805
COL.CM.-.CGU1	RTNVSSQIRQDREWE---NA--IHREQE--S--FWRLME-WWS.....SLLRHCSNGIL.TV.....LPQHPP.TTSDKRRS	785
DEB.CM.99.CM40	P.....K--EVLRLAEEGG--GGDG-ARWTA-PRGFFSLIWED...LQQLLTW-YQTFRNCIWLIRRGFELLQEGIYRLSPALW-R.LR.....LLGE-F-V-G...GYLQ	815
DEB.CM.99.CM5	P.....K-VEVLQQLGEGGDK-GEVKSRTWTSRPFRLRLIWD...I-QLLTW-YQTLRNCAWLIRRGFELLQETIYRLSPAFW-R.LR.....LLGEHL-VVW...GYCQ	795
DEN.CD.-.CD1	IR.PQKAAN.....QPBEQ-G-S-D..G-RLKWWQYPRGFLSII.....WEDLSQLLTW-YLTLRNL.VFLIRT.....ACSRWD.....FVH-HSLILLGYFQ	810
DRL.-.-.FAO	FKDAYR.PETTNY.HR--R-KGEEQDRGEW.NIRSEPS...KPGSSKAWSEETVGTWLKE-RGYIW-KNVKAV.....EYGWS.....ELQEAGRGC.....	847
GRV.ET.-.GRI_677	IHIHSS.AERPDNGG--DRGG-SSS-KLIRLQEE-S-PS.RINNWNLNFKSCS--IR-WCYNICLTL-IF.....I-TA...VGY-Q-	800
GSN.CM.99.CM166	ILIPQAVERP.....-AT--GVGAQGVKVR-VR-LTGFSS-L.....WDDLRLNLVIWIYQILASLAWTLRGI...G.....KF-WEQL-KAIQKAI..QOAR.QLRE...VA-RVIA	823
GSN.CM.99.CM71	ILTPQQVERDP.....-GTRKD-GAQQ-VR-VR-LTGFSS-L.....WDDLRLSLVIWIYQILATLWVTLQRI...G.....KF-WEQL...QKAF..QOARQLGEV...A-RAAA	828
LST.CD.88.447	AYVEQDCKWEKRENQE----G-IEKDKE-I.YINLEQYKKESSTPPNVNDWSEPLQD-L-V-LLKWKAFGILLLSLVNQPLSWLGHVILVFPQHGHQHL.RK.....ACS-RV...VENAQ	863
LST.CD.88.485	AYVEQDYKWEKGENPE----G-IEKDKEDI.YINLEQCKKESIPPNVNDWSEPLQD-L-V-LLKWKAFGILLLSLVNQPLSWLGHVILVFPQHGHQHL.WK.....ASSGWM...VENAQ	862
LST.CD.88.524	AYVEQDFKWEKGENSQE----G-TEKDRE-I.YINLEQCKKESFRPPNVNDWSEPLQD-L-V-LLKWKAFGILLLSLVNQPLSWLGHVILVFPQHGHQHL.W.....TSS-WM...VENAQ	865
LST.KE.-.lho7	VYIEQDYKWEKKEENQE--R--K-ADTETI.YINLEQCKKESRPLNVNDWSEPLQD-LLV-LLKWK-KEGGILLLSLVNQPLSWLGHVILVFPQHGHQHL.W.....TSS-WM...VENAQ	860
MAC.US.-.239	P-YFQQTH-QQ.DPAL.-TR-GKERD--EGGNS-WP--I-YIHF-IRQ--RL-TW-F-NCR-LL-RVYQ.....ILQPILQRLSAT--R.-R.....EVL-TE...LTYLQ	824
MND-1.GA.-.MNDGB1	...QDC-QN.-YRKR--NG--ESNSLELG.EHNSENL...KEESLN-SLIEDLTSFARE\$P-WLW-KNLRAA.....IEY.GF.....LAVQEA...IRSLG	824
MND-2.-.-.5440	...VFKGDYLRPHNLKR--R-GGEEPDLKQ.NIKSESS...RQESRKPWKPEQVRSWLKR--LYIW-KNLQAV.....IEYGW.....ELKAAGAALY.....	842
MND-2.CM.98.CM16	...VFKGDYLRPHNLK--G-R-EEPDSKQ.-TKSDSS...KVEFGKPKSKEQIRDWLKT-RGYVW-KNLQAV.....IEYGW.....ELKTAGGKIF.....	818
MND-2.GA.-.M14	...VLKGDYLRPRNLR--E-R-GELNLEEQ.NIKSESS...KKEFGRPWRPEQIRNWLKE--YIW-KNLQAV.....IEYGW.....ELKAAGA-IC.....	819
MON.CM.99.L1	I..PIPSHAGQ...PT--GTGAG-GD-S--R-PA-LKGFFTII.....WEDLRNLVLWTYQILKDSVLVIYRI.....LQRVSRQ...LPPLLHIRLLQLWESLRRLLAYC	826
MON.NG.-.NG1	I..PTQVPREP...-A-BETG-G-GEQGRNRXTV-ARGFLTII.....WTDCLRDILVWTFQLLRDSGLTIYRS.....LQRVSRQ...LIPLLRDLC-QLREASSRLLAYL	831
MNE.US.-.MNE027	P-YFQQTH-RQ.DQAL.-T--GTE-D--DSGNS-WP--I-YIHF-IRQ--RL-TW-F-NCR-LL-RAYQ.....ILQPIFORFSTT--R.VR.....EVL-TE...LTYLQ	827
MUS.CM.01.1085	I..LIRP-TGP...EA-BGT-GD-GS-SRDR-VR-LSGFFS-F.....WDDLRLNLVIW-YQTLISSASTLW.....WLLQRAVHLIQKGSIRLWFTLRALIEPLREALQR-LACL	841
RCM.GA.-.GAB1	YV...Q---.HNT-E-QTPG-KRED-GEEGNGK-NNWLREYC.....WIQLIHLPLSRWITQLSQICRSCSSIIPOS.....LRWILAKIQYGW.....	798
RCM.NG.-.NG411	PCYH..Q--P..RNREEQPIK-GEKG-SGEGGHK-YNWQRGYF.....YI-ILRPIELCRRLYQICNSLSVIYQSLQSIPIRILLHLRWIGAKLQDGW.....	805
SAB.-.-.C5	..LPSYN\$-----D---K-----C---P--T--.LRNLTNWFVTIC-SLH.....H--VIL--V-TW...K-Q---	286
SMM.SL.92.SL92B	P-YHQQIH-QR.DOEL.-A--GTE-E--NGGGY-LWP--I-YIHF-IRQ--RI-TW-YNNLTRLA-RAYQ.....NLQQLCQRLSEIS-P.-R.....ELV-RE...-GYIR	831
SMM.US.-.H9	PAXVXQIP-XT.GOEL.-T--G-E-D--XRGGN--WPX-I-YIHF-IRQ--RL-TW-F-SCRDWLLRXCQ.....ILQPVQLQSLRT--R.AR.....EVI-VE...ITYLQ	831
SMM.US.-.PGM53	PVYVQQIP--K.GQEQ.-T--G-E-D--DRGGN--WP--I-YIHF-IRQ--RL-TW-FNPCR-YLLRIYQ.....ILQPVFQRL---R.-R.....EVI-TE...ISYLQ	832
STM.US.-.STM	P-CRQQIP--K.GQEQ.-T--GTEEG--DRGGIN-WP--I-YTHF-IRQ-VRL-TW-YNN-RACL-RIYQ.....TLHPTFORISR--R.-R.....EVI-LG...-AYLQ	827
SUN.GA.98.L14	AYVEQDWLQE...TCPK-TDK--EETEKERIYINLEQSKKESLPPWTVDWDEPLRD-LLV-LLKWK-KAWGIVLAQNIYHLLSFLWHLTTSPFHGHQRL.W.....TLRGW...GSHLI	864
SYK.KE.-.KE51	QP.....TLYHKAPA-PEEGIETDNDYKPSVSTFSREPLRQSL.A...EAQKLW-RVSESCRSRLIRGLVIAW.....GFIS	785
SYK.KE.-.SYK173	TQ.....GRKDPGKPADEEEG-GDREGLNVSTFSRESLRQSL.E...AGQLWRVTCSSFRSLIRQLTITW.....AGQLWRVTCSSFRSLIRQLTITW.....GFIS	777
TAN.UG.-.TAN1	INIRL.....GRE...--NAGGEDKSSS-RDK-PPSVK-S-LPNRGGIQABE-AWRQHLTNWCLTISSWLLRLRYQILRR.....S-TT.LL.....-LL--E...CQYIQ-	825
VER.DE.-.AGM3	IHIHPW.KGQPDNAE.-GEGGDKRKNSEPPWQKESGTAEWKSNWCKRRLTNWCSISSIWLNSCLTL-VH.....L-SA...FQYIQ-	823
VER.KE.-.9063	IHIHPW.KGRPGNAGE.-EEGG-RDNDSSD-WQKSGTRQKRGWSKRLTNWLSTFS-WLYN-FLTL-IH.....L-GA...WQYIQ-	823
VER.KE.-.AGM155	IHIHPW.KGQPDNAGE.-EEGGRT-KSKSTH\$QKEFGGRDKRTSWCRQLTNWF-TLSIWLHNSCLTL-LK.....L-SA...WQY-Q-	816
VER.KE.-.TYO1	IHIHQVGKGRPDNAE.-GEGGDNSRIK-E-W\$KDSKSRMCQ-TAWLT--N.....-WLYNSCLTL-IQ.....L-KA...FQY-Q-	810

	*	
SAB.SN.-.SAB1C	GW.QELKAAAASLTAVVQAATSASDSVRHACR.....SIVRAVIAHPRRMRQELERWFN\$	867
H1B.FR.83.HXB2	Y-S---NS-V--LN-TAI-VAEGT-R-IEVQ.....GAC--IRHI---I--G--ILL-	856
CPZ.CD.-.ANT	L-.TIIRTEIKNIDRLAIWVGEKT--ILL-LQ.....T---IIREV---I--G--IAL--	858
CPZ.GA.-.CPZGAB	Y-GK---IS-I--LD-TAI-VAEGT-RIIE-FQ.....VTL-IIRNI---I--G--ALL-	854
H2A.DE.-.BEN	YGCWEIQE-FQAFARTTRETLAG-WGWLWE-A.....R-G-GIL-V---I--GA-LALL-	860
COL.CM.-.CGU1	A--T-RYL...-VPRGSLQILETLQRWLSAA...RGW-RAPEYL-GWIYDRPQGP-	837
DEB.CM.99.CM40	YGWK-F-DTITWAGGELW-WLQGIEMALQGLV.....RAGG-ILRV-A-I--GA-LIL--	870
DEB.CM.99.CM5	YGWK-F-DTIIWAGQEIWEWLQGIQVAIQGLV.....WGG-NLL-I-A-I--GA-LLL--	850
DEN.CD.-.CD1	-LQELQTLG..RDLGTSAIQQGR-TAE-VL-ALT...RAA-E-V-I---I--G--IVL--	864
DRL.-.-.FAOTLLEVAQRIWREGRQLGL-SAR-L.....ALAQE-A-I---I--GA-VL---	895
GRV.ET.-.GRI_677	-LQ-LQE--T.G-AQ-LAR--RE-WGRLGAIV.....AY----NS---V--G--KVLG-	854
GSN.CM.99.CM166	YISYGIQELQ-AA-GILDSL-IFTWN..WTEAVLHACRRVW-EPL-I---I--GA-IL---	881
GSN.CM.99.CM71	YISYGIQELQ-AVSGILDSL-VFTWN..WTEPLVQTVGRVW-EPL-I---I--GA-ILL--	886
LST.CD.88.447	KIASWSWNKIRRNRRDISA...HARNL-LGQK....KRW-FRFRGRSGFPS-TTETAL-	915
LST.CD.88.485	KIARWSWNKIRRNRRDISA...HARDI-LGQK....KRW-FRFRGRSGFPS-TTETAL-	914
LST.CD.88.524	KTASWIRNKSRNRNGELSKT...NGTNI-LGQNR....RRW-FRFR...SGLPS-TTETTL-	916
LST.KE.-.lh07	KIQSW-REKCRNRNGQLSST...DRKNIQLGKK....KRW-LRFGGRSGISS-ATETAL-	912
MAC.US.-.239	YGWSYPHE-VQAVWRSATETLAG-WGDLWETL.....RGG-WIL-I---I--G--LTL-	879
MND-1.GA.-.MNDGB1	SLLVSPA-ISWRAGYTGWLQ-NRIPTGC-E-II....AAGTCIWHIL--I--SA--PL--	879
MND-2.-.-.5440QVLQGFQRLWSRGYQLGL-GVGRGAA....AFG-GIWNI---I--GA-ALL--	890
MND-2.CM.98.CM16KVLQGYAQLWSRGHQWGL-TAACF....A-A-GI-NI---I--GA-VLL--	866
MND-2.GA.-.M14QVLQ-FAQRLWSGGHQLGL--IRGAA....A-G--IGNI---I--GA-VLL--	867
MON.CM.99.L1	QYGIQELQ--VTSLLDALARF-IVWTDALLHAGG...RLW--IV-I---I--GA-IFL--	882
MON.NG.-.NG1	QYGLQEXQX-CTGAI DALARF-VIWTDAVLR LGG...RLW-GLV-V---I--GA-ILL--	887
MNE.US.-.MNE027	YGWSYPQE-VQVAWRSATETLAG-WGDLWETLG...RVG-WIL-I---I----LTL-	882
MUS.CM.01.1085	QHGIQELQT-STAVLDQLATF-WNWTETVLQAGR...RVW-EPL-I---V--GA-ILL--	897
RCM.GA.-.GAB1	...-F-EFSSWFAEMAL-N-YTWRGLCAVA.....DFAGWPAMVC--I--G--LC--	850
RCM.NG.-.NG411	...-F-GFSSW-AEMARTN-YTWRGLCAV.....DPAGWPATLC--I--GF--FL--	857
SAB.-.-.C5	-----GA-F-L-----G-I-----IL-----F-----	340
SMM.SL.92.SL92B	YGWNYFIE-CQEAWSAQE-IVG-WGLIWETLG...RVG-GIA-I---I--G--LML--	886
SMM.US.-.H9	YGWRY-QE--XXWKFRETLE--WRDLWETLG...RVG-GIL-IXX-I--G--LTL-	886
SMM.US.-.PGM53	YGCRWPFQE-VQAAR-RGETLA--WRDSWETLG...RVG-W-L-I---I--G--LALL-	887
STM.US.-.STM	YGCIIWQE--QAAWR-AGETLA--GRDLWETLG...RVG-RIG-I---I--G--LTL-	882
SUN.GA.98.L14	QATSRIRN-CRTSRER-SSSQKAR-RTFSLGRK...WRPKWNRTRGS-IPS-TTETTL-	919
SYK.KE.-.KE51	YGVV---E-I--GRE-AHQMVAIWQALLAYA.....RVAEN-A-L---L--G--IYL--	840
SYK.KE.-.SYK173	YGFN---I-----GREIRDWVAIIWQAIYA-T.....RV-E-A-L---L--G--IYL--	832
TAN.UG.-.TAN1	---QF-EG--RSFE-LAS--Q---RTLWN-----AY--ILE-----	879
VER.DE.-.AGM3	-L.G-----QEAV--LARL-QN-GYQIWL-----AY--I-NS---V--G--GIL--	877
VER.KE.-.9063	-L.G---T--QEARL--ARF-QN-GHQIWL-----TY-HI-SS---V--G--EIL--	877
VER.KE.-.AGM155	-L.G-----QEAVNHLASF-RN-AHQIWL-----AY--I-NS---V--G--EVL--	870
VER.KE.-.TYO1	-L.A---TG-QEILQTLAGV-QN-CHQIWL-----AY-NIVNS---V--G--EIL--	864

SAB.SN.-.SAB1C	MGGKSSKQQQRHSLWLWSKLRQAPV.....IQYDMLADPLLGQSSHIQEBCAKSLRDGLIRQGDSSRTEEGVKMKHQGR.....QPSWYDEDEEE.VGF.PVRPCLPLRAMTYKLAIDFG	108
H1B.FR.83.HXB2	---W--SSVIGWPTVRRER-R.....A...EPAADRVAAS--LEKHGAIT.SSNTAATNAA.....CAWLEAQE---.---T-QV---P---A-V-LS	88
CPZ.CD.-.ANT	--SAW--IKWVGARQAIR-I.....H...ETNPDDIGPCGNELAS-GALT.SSTI.GTE-D.....VITYSEDHT--GT--.---QV-M-P---V-LS	86
CPZ.GA.-.CPZGAB	--T-W--SSLVGVPEVRRRI.....REA...PTAA-GVGEVSK-LEKHGAIT.SRNTPETNQ.T.....LAWLEEM-N-.E---.---QV-T-P---A-F-LS	88
H2A.DE.-.BEN	--ASG--KLSK--RG-RER-LR-RGDYQKQDASGGEYS-FQEESGREQNSP-CEGQQYQQ.....GEYMN-PWRNPATERQKDLRYQQNMDDVSD-DDL.I-V.-T-RV-R-E-----MS	120
COL.CM.-.CGU1	--SLL-CLNI-QQPQTRGGVLGRFLGIYWPYQVAFRFLRHFRGGY-LRRVAVGERLNI-ED.....GYMTVPKVLNCPCLCYRHPDVVYHN-EPE-VEDV-.AGCY..MSR--V-QA-I--LV-IS	122
DEB.CM.99.CM40	---N--IPS-GQVGSFGSGRGLLRWRYSRDLSEQEEQFSECLLES-REQSS--TEFW.....GSP.....RREIKCTNKQQDLDQT-A.---.---QVSI-DP---M--YS	103
DEB.CM.99.CM5	---N--TPSPGQKKSQSGSKGLLRWRYSKDLSEAEATFSPFLQES-KEQSCY-TE.....EPLYA.....DPHREIKCTQRAQECYD--HGI--A---RV-I-EP---M--YS	105
DEN.CD.-.CD1	--S-N--ASGIWGLAP-NKEGRQRKRWYRLLPEQQEKWSPDGS-RA-KL--TRLL-HDE.....DEETFVDAEBEIQSSPALDPQHPRH-YVVDHD---.---.---RR--H-P--DM--MS	122
DRL.-.-.FAO	---QN--KREEAYARFYKR--RGYGAKGNLDYHQLPEPSEPLLRKQW-FP-GLDKA-RSSST.....SSLDTTHMQAAAAEKD...VTLPTEEEQPS-E---.---C-QR---EP---DLV--S	118
GRV.ET.-.GRI_677	--SSN--R--QGL-K--RG--GK-GADWVLLSDPLIGQSSVTEQECCKA-KK-WGKGKMPD.....GR-LQEGDTFDEW-D---.---.---Q-RV---Q-----V--S	104
GSN.CM.99.CN166	--S-N---QE-STALLSS-GTQ.....RPFYTLVDEY-ENFWLSPDASD-GRRYYLTEEPKPKRGS-L-Y.EPSC.---RV---DP---VMV-LS	92
GSN.CM.99.CN71	--S-N---SQE-STALLSSHGTQ.....SPSFR-L-DEY-EN-WLSPDASDRGRYYLTEES--RQNCID.Y.EPSC.---QV---DP---MV-LS	91
LST.CD.88.447	--NIFG-PAADGWKTLRR--AGAGTRSEGTEETYQQLMQETSD.....WDKRRDG-S-S-E---.---QR--CSP---TL--LS	81
LST.CD.88.485	--NIFG-PAADGWKTLRR--AGAGTRSEGTEETYQQLMRETS.....WDKRRDG-S-S-GE---.---QR--CSP---TL--LS	81
LST.CD.88.524	--NIFG-PSADSWWRVLR--AGSGTRAEGTEESYQKLM-E-TLG.....WDKTEE-G-S-SE---.---QR--CSP---TL--LS	80
LST.KE.-.lho7	--NAPGRPAADGWKTLRRF-AGSRNAEGTEDSYRALIEKTSN.....WAKRRDG-S-S-E---.---K-QR--CSP---TL--LS	81
MAC.US.-.239	---AT-MRRS-P-GD-RQR-LR-RGETYGRLLGEVEDGYS-SPGGL-KG-SSL-CEGQYQY.....G-FMNTPWNPAGEREKLAYRQNMDDIDE\$-DDL..-V.S---KV---T-S-----MS	119
MND-1.GA.-.MNDGB1	--SSQ--KRSEAWVRYS-A--LVGGPVTPDGYKQIESSQGAEK.....QSLLRG-AYGTYSEGLDKVQNDPLTKDEKLDLTQQ-PE---E---.---CROVS--VPS--DL---S	109
MND-2.-.-.5440	-----REQ-YIKYYKA--RGYGAEGTNLDYQVLEEQQPLAGASGTS-AA-DKVLRPYST.....ERDLRHQQDVTLPSEKEQPSD--EE---.---Y-TQ-VPEA---DL--MS	110
MND-2.CM.98.CM16	-----EKKY-KYYKAM-RGYGAEGTNGDYQQLHASEPLLGALSTSQEEDRE-K.....S-ST-E---.---T---Y-QC-V-EP---DLV-MP	90
MND-2.GA.-.M14	-----REQYAKYYKA--RGYGAEGTNLDYQVLEEQRLLLEGPLGTS--A-DKELKS-ST.....KVDDEQK---R---.---Y-TQ-V-EA---NL--MS	94
MON.CM.99.L1	--S-N---SPL-STPLLSQSSGRMRMYFMLEDDYGEQSWLSPDAS-RERKY-LTEGRNGKQ.....RRQPLDDDDD-DG.....-C---RV---DP-W--MM-LS	98
MON.NG.-.NG1	--S-X-X-PAQEP-MPSHGSQSSGQTRSFAWEDDYGEDSWLSPDAS-RGRRF-LTEGRNHR.S.....NRRPTTVVDDADEQDL.P-C---.---MV---DP-W-IMM-LS	102
MNE.US.-.MNE027	--AT--KRKSPGD-RQR-LR-RGETYGRLEWGEDGYS-SLGG-S-KS-SLL-CEGQYQY.....G-FMNTPWKNPAGEREKLAYRQNMDDIDE--DGL..-V---RV-----V-MS	120
MUS.CM.01.1085	--S-----PAQQ--ISSPPSPGTGRKQYFKLVVEYGENSWLSPDASGRGRRY-LTEGASKR.....PVI--HERTC.---RV-I-DP-----MV-MS	91
PAT.SN.-.PAMG31NEF	-----Q-----QIRYDMLADPLLGQSSNIQEBCAKS-RDG-IK.....QGDSSRTEEGIKYKQGRQPSWYDED--E---.---RV---T--F---V--	107
RCM.GA.-.GAB1	-----NSAAGL-RWRFLM-TT-GEYVVRWHTLLDQGPWCAEGSGRASRDFVIRGGIT-E.....TQASIDDIDWYE-T-DTL.---.---K-QV---P-S-----MS	103
RCM.NG.-.NG411	-----ASAAGLRWRGLMLTT-GEDYARFAETLQDQGPCRCAEGSGRASRDFLTRGGPTIE.....TQQSVDAIDWYE-T-DTL.---.---QV-I-P-S-----LS	103
SMM.SL.92.SL92B	---IT-SKPA-KHGD-RK--LS-RGDSYGRLLWGLEEGSQ-SQGS-RA-N.-CLTRDGY.....YNVGEFMNTPWRNPATERQKQYRQCCMDDIDED-D-L.A-V-V-W-KV-----S-----MS	121
SMM.US.-.H9	---VT--K-RXAGGN-XER-L--RGETYGRLEWGLEGEYS-SXDASGKG-SSL-CEPQKYCE.....G-FMNTPWNPATERGAKLYRQNMDDVDN--L.....-C--S-RV-V-I-----MS	118
SMM.US.-.PGM53	--AAG--K-S-QRGG-GE--L--RGETHGKLEWGEDGYS-SRGLGRDWNHL-FEGQYSE.....G-FMNTPWNPAREKREKLYRQNMDDVD-D-D-L.I-V.S-H-KV---S-----MS	120
STM.US.-.STM	--ASG--K-RKQHG-RER-LR-RGETYGRLEWGLEEGSGPSQAS-KG-NSH-CEPQRYNE.....G-FMNTPWKNPAEASAKLEYRQNMDDVDE--DNL..-V.A-H-RV---E-----LS	120
SUN.GA.98.L14	--NAPGRPSEVGVWRTLFR--AGSGTRAEPAGREYHRLRR-EV.....EPLV.....SAENGGPNGI-Q-E-E-.---QR--CKP---QL--LS	84
SYK.KE.-.KE51	--SQ--K-PSRWDEK-RTRWVFPKPYSPMPDELLRMSQP-HEDF-RGWRSTLT.....EPI.....DPKRDFFS...-GKKWNAGDICH--GD-D.L---.---L-QV---PL-----LS	109
SYK.KE.-.SYK173	--ST--S--LR-EGKYAIGWRLFGKQYTPLELSRPLQPCRGGF-KAWRSTLT.....EPI.....DPHGPRDWG-S-GQKFSPGDIVQ--GDTG.L---.---C-QT---TL-----LS	112
TAN.UG.-.TAN1	---SN--RE-QGL-R--RA--K---.---.---VR-G-----I---N---D--NWNG-ST-R-K.S-.PEGRKLAADD....TWDDWEPE---.---.---RV---Q-----V--S	105
VER.DE.-.AGM3	--LGN--P-HKKQ-S--HA-HKTRA.....TR-GL-----I---TL-----D-A-KES---KRNGKM-.PEGRKLQE-D....KWDEWSD-E-D...---RV---Q-----V--S	107
VER.KE.-.9063	--LGN--P-HKKHVS--HA-HKTO.....TK-GL-----I-T--AV---YD-G--KS---KRNGNM-.PDGRRLOE-D....EWDEWSD-D...---K-RV---Q-----V--S	107
VER.KE.-.AGM155	--LG--P-HKKQ-TI-RA-HATRH.....TR-GL-----I---TL-----D-G--KS---KRNGNM-.PEGRRLOE-D....WDEWSD-E-D...---RV---QI-----V--S	107
VER.KE.-.TYO1	--SQN--PAHKKYSK--QA-HKTH.....TR-GL-----I-T--TV---D-A--KS---KQNGNM-.EBERRLOE-D....TWEWSD---.---.---RV---Q-----V--S	107

SAB.SN.-.SAB1C	HFLKEKGGLEGIYYSERRKKILDLYALNEWGIVDQWQNYTDGPG.....TRYPKCFWGCFLKVPVLDSEEAKNSE..NHC..LLHPAQVAYEDD.....AW.KETLVWKFDPPLAVDY.VAWRLHP	218
H1B.FR.83.HXB2	-----LIH-Q--QD---WIYHTQ-YFPD\$---P-----V---LT---Y----EPDKIEEANKGE-TS.----VSLHGM-----PE.R-V-E-R--SR--FHH.--RE---	199
CPZ.CD.-.ANT	X-----DXLIXLK-AA---MMPF-TQ--FPD-----EE-----V---L-R-----ESPDP.....DERNI.....CTHGDG.....PH.--I-R-E--AS-MRRH.I-RER--	193
CPZ.GA.-.CPZGAB	-----LV--R--QE---WVYHTQ-FFPD-----T-----F-L-----LTEEQVEQAN-GD-N-----ICQHGME.....ED.--V---R--SR--LRH.I-REQ--	200
H2A.DE.-.BEN	--I-----Q-MF--R--HR--I-LEK-E--IPD-----H-----V---MY---LW---S-E-Q--EED-A..N---V---TSRH-----EH.G-----Q--SM--YN-.K-FT-Y-	230
COL.CM.-.CGU1	C-----IRTPE-DDLIEQ--YI---CLK--LE-E-EL-EDGALKED-K-LVA--LW---YIEQLG-YAY-Y.....D.LSL-SVTSRKKKKP.....QQVAI	219
DEB.CM.99.CM40	-----DVF--A--HA--E-H-Q-----IP--LQ--K-----V---Y--FL-----EIADPDYENDE..RN.I---D-HQGMM-----PH.R-R-----ST--YC-.K-GHAEQ	214
DEB.CM.99.CM5	-Y-----D-F--A--HA--EIH-HH---LIP--LK--E-----P--TY--FL-----EIADPDYENDER...I---D-HQGQAE.....DPH--R-----SS--YC-KAGHEA-K	217
DEN.CD.-.CD1	-----QD-F--P--HA-----Q--H-FIT--T-----I---LE--FLW---TIE--YD-K-E..N---DRYEGQQA.....P..R-----R---E--WC-.K-GHKKL	232
DRL.-.-.FAO	-----K--WW-L--ET-----Q---FIK--S--K-----I---M--FL---EVD-DLV-QQ...NR--NSS-MGPI-----PE.G-R-M-Q---G--YTF.Q-CIH--	228
GRV.ET.-.GRI_677	-----S--D-----E--N-----I-D--A-SP-----I---RV--F-----H--R-C--R...M---MGEDP-GI...DH.G-V-----K---E-.RPDMFKD	216
GSN.CM.99.CN166	-----MF-C-D-HQK-EQ--YL---L-P--LSF-P-----TIP-F-TC-R--ATT-DSEPGDDE.YL..T---YQGRSE.....QH.--F--FS-CSK--IKSGIQLDQLQ	204
GSN.CM.99.CN71	-----AMF-C-D-HQK-ES-CYY---P--LQW-P-----I---TMP-F-WC-R--AMT-DSEPGDDQ.YL..N---YQGOQE.....HH.R-I--FS-CSR--LKSGWQMNQLQ	203
LST.CD.88.447	--I-----L-W-Q--QD--C--CE---LIGDFMT-----F-LA---LWQ-E--ACD-YTDP-DYR.Q---SS-LGVQE.....P..G-R-I-H-N-M-----L-KQ-	192
LST.CD.88.485	--I-----L-W-Q--QD--C--CE---LIGDFMT-----LT---LWQ-E--ACD-YTDP-DAS.Q---SS-LGVQE.....P..G-R-I-H-S-M---F--L-KQ-	192
LST.CD.88.524	--I-----L-W-Q--ND-T--CE---LIGDFM-----LA---LWQ-E--ACD-FKDP-DEB.Q---SS-LGVLE.....P..G-R-I-H-N-M-----L-KQ-	191
LST.KE.-.lho7	--I-----L-W-Q--QD--I--CE---LIGDFM-----LT---LWQ-E--ACD-YKDP-DET.Q---SS-LGVLE.....P..G-R-I-H-N-M---F.I-LKKQ-	192
MAC.US.-.239	--I-----A--HR--I-LEK-E--IPD--D--S-----I---T---LW---NV-D--QED-E..Y..M---TSQW-----P..G-V-A-----T--YT-.E-YVRY-	229
MND-1.GA.-.MNDGB1	--I-----G---R--EE-----E---FEP--Q--T-----T--FL---E--SRAIGDEYAANNH...--SS-LCPQE.....PE.G--M-SGTLI-PMTL.QH\$	214
MND-2.-.-.5440	-----WW-K--EE-----Q---LIPN--K-----I---T--FL---AV-PPLEED...NR--NSS-MGIEQ.....PE.G-R-M---SG--YTF.Y-PIIR-	220
MND-2.CM.98.CM16	-----WH-K--EE-----Q---FIPT--S-----I---T-RFL--C--AVPPDQE-N---NK--QSS-LGIEQEE.....P..G-R-----SG--YTF.Y-PIKR-	201
MND-2.GA.-.M14	-----WF-R--EE--N--Q---FIPD--E--S-----I---R--FL--I--QVPPDQE-Q...NR--NSS-LGIEQEE.....P..G-R-M---E--YTF.Y-PIKC-	205
MON.CM.99.L1	-Y-----GEMF-C-D-HRKIEQ--YL---LIP--LQ--E-----V---TMP-FLWC-R--ATT-DSEEGDED.FL..T---YQGRME.....PH.RQF--FS-CSK--KSGRQLAQLQ	210
MON.NG.-.NG1	-----DKLF-CXDXHQK-EQ--YVM--L-P--LQF-X-----V---TIP-FLWC-R--AIQ-DSEGDGXE.FL..T---YQGRDE.....PH..QF--XS-CSK-G-KSGQLDELQ	214
MNE.US.-.MNE027	--I-----H---I-LEK-E--PD--D--S-----P---T---LW---NV-D--QEG-E..NY.....TSQW-----P..G-V-----T--YT-.E-YIRY-	230
MUS.CM.01.1085	-----MFV-TE-LQK-ETW-YV---IP--LM-----IP-FLWC-RE-AMT-GSEEGDDQ.YL..D---FQGR-E.....HH..QF--FS-CSR--LKSGQLDQMQ	203
RCM.GA.-.GAB1	-----W-I--QR--M-LE--H--IPD---P-----I---TL--LWQ---V-D--RED-E..S---ETSGME.....P..G-V-A--N-M-----IGY---	213
RCM.NG.-.NG411	-----F---RR-----ME--H--IPD---P-----I---TM--LWQ---V-D--TEN-R.....ETSGM-----P..G-V-A--K-----AGY---	213
SMM.SL.92.SL92B	--I-----D-----HR--I-LEK-E--IPD---S-----I---F---LWQ---V-D--N-D-T...V---TYQHS.....P..GKV-A--N-H--YT-.E-FVR--	231
SMM.US.-.H9	--I-----V---I--H---X-LEX-E--IPD---A-----I---MF--LW--X--NV-D--QED-T...M---TSQW-----P..G-V-A---X--YX-.K-FVE--	228
SMM.US.-.PGM53	--I-----HR--I-LEK-E--TPD---S-----I---MF--LW---SA-D--QED-T..Y..V---ISQW-----P..G-V-A---SQ--YR-.E-FIR--	230
STM.US.-.STM	--I-S-----HR--M-LEK-E--PD---A-----I---Q---LW---M-N--QEDDG..TH.Y-V---THQW-----P..G-V-----HT-.E-FVR--	231
SUN.GA.98.L14	--I-----LW--RT-EE-----E---FIT--D--K-----V---A--LW--A--TID-DRDPN...HP-QA--SS-QGVNE.....P..G-R-I-T---T--Y-F.R-IQK--	195
SYK.KE.-.KE51	--I-R---Q-MFFCQK-HE--Q--LK--H-VI.DDIT--S-----LI---LWE-A-NEIEGLYVDE-DT...LMM---AGVGASE.....DPHR-N-M-N-N-H--YTP.GWEMARQ	219
SYK.KE.-.SYK173	--I-N---Q-MN-C-K-DE--H--LQ--H--I.DRI--S-----LI---LWE--NEIEGCLYEY-HT...L---SQGSSS...MGEPHV-LQPPPGYTPGW.EMARQLERQ	224
TAN.UG.-.TAN1	-----W-PK-EQ--N-----I-D--A-SP-----K-R--F--E---V-Q--QDE..R...I---EWS.....P..-----N---Q-.NPDSFKD	214
VER.DE.-.AGM3	-----D--N-Q--N-----I-D-NAWSE---I---R--F-----H---ETC--R...V---REDP-GI...NH.G-V-----M---Q-.DPN-KYL	219
VER.KE.-.9063	-----D-----NR--N-----I-D-NAWST-----V--R--F-----EMH---ETC--R.Y..A---LKEDP-GI...SH.G-----M---Q-.DPN-QYF	219
VER.KE.-.AGM155	-----D--D--N--N-----I-D-NAWSK---I---R--F-----A-H---ETC--R...V---LHEDP-GI...NH.G-I-A---M---Q-.DPS-EYF	219
VER.KE.-.TYO1	-----D--D--N--N-----I-D-NAWSK---I-F-----F-----H---QTC--R...V---MGEDP-GI...SH.G-I-----M--IQ-.DPN-EYF	219

SAB.SN.-.SAB1C	EQVPSA.QG\$	226
H1B.FR.83.HXB2	-YFKNC.\$	205
CPZ.CD.-.ANT	-YFRD\$	198
CPZ.GA.-.CPZGAB	-YKDS\$	205
H2A.DE.-.BEN	-EFGHKSLPEKEWKAKLKARGIPYSE\$	257
COL.CM.-.CGU1	-M-D\$	223
DEB.CM.99.CM40	REHTRRCMFPKRK\$	227
DEB.CM.99.CM5	PEHNRRRCMFPKRK\$	230
DEN.CD.-.CD1	QTSQHN.NYCQRAKPRK\$	248
DRL.-.-.FAO	-EFGHVKSIKSVVEEPG.....CCKRKKWQFKPTTEGCHGDNLQKC\$	268
GRV.ET.-.GRI_677	MHEHAKR\$	223
GSN.CM.99.CN166	QEERKMRLTANRFL\$	218
GSN.CM.99.CN71	QEERKKRLTANRFL\$	217
LST.CD.88.447	KSIQASAI-L.....QRK\$	205
LST.CD.88.485	KSIQASSLSL.....NCKRK\$	207
LST.CD.88.524	AS-QATAMRFNCERK\$	206
LST.KE.-.lhc7	AKIQNTAFAPDCRKR\$	207
MAC.US.-.239	-EFG-KSGLSEEEVRRRLTARGLLNLMADKKETR\$	262
MND-2.-.-.5440	-EYKCV.TSLSYEAYKKEKPDCCCKRKKWQF\$	250
MND-2.CM.98.CM16	GEFKHV.-SISHIAYAKEHKPECCCKRKKWQF\$	231
MND-2.GA.-.M14	KEYRHV.TSLSYTAYQE..KSDCCCKRKKWQF\$	233
MON.CM.99.L1	QEERKKRLAANRIL\$	224
MON.NG.-.NG1	QEERKKRLTANRIL\$	228
MNE.US.-.MNE027	-EFG-KSGLSEEEVRRRLTARGLLMADKKETS\$	263
MUS.CM.01.1085	QEERKKRLATNHIL\$	217
RCM.GA.-.GAB1	-FFGERKNKTQ\$	224
RCM.NG.-.NG411	-FFGERKNKTK\$	224
SMM.SL.92.SL92B	-EFGWKSLTEEEVERRLANKPKPKQKMKADKKETS\$	266
SMM.US.-.H9	-EFG-QSGLSKEEVQRRLTARGLLMADKKETS\$	261
SMM.US.-.PGM53	-EFG-KSGLSEEEVKRRLTARGLLMADKKETS\$	263
STM.US.-.STM	-EFG-KSGLPKEEVERRL TARGLLK MADKKETS\$	264
SUN.GA.98.L14	-EFKHVTSLQWEV\$	208
SYK.KE.-.KE51	...QLERQTGKR\$	228
SYK.KE.-.SYK173	TGK-QEL-SALSKNIS\$	240
TAN.UG.-.TAN1	MHGLVK.RK--	222
VER.DE.-.AGM3	TDMHDL.GKRK\$	229
VER.KE.-.9063	-DMHAI.VKRK\$	229
VER.KE.-.AGM155	TDLY-T.V-TGN\$	230
VER.KE.-.TYO1	TDMHGL.VKRK\$	229