

VII

Primate Lentivirus Proteins

Introduction	609
Table of sequences in PLV protein alignments	610
Gag	613
Pol	618
Vif	627
Vpx	630
Vpr	631
Tat	633
Rev	635
Vpu	636
Env	637
Nef	646

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CPZ.GA.88.GAB2	AF382828	VPU	Gao, F	Unpublished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.DEB1014	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.DEB1083	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.GSN1083	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.-. SIVremGB1c1	AY336733	VIF	Gaddis, NC	Unpublished
RCM.GA.-. SIVremGB1c3	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.LSGRKLDAFESVRLRPNPKKKYKLRHLVWASKELDRFSLSANLLETKEGVVKILSVLLPLVPTGSENLIALFNLCCVLACIHAEIKVKDTEAKAKVKEEVPAMETESATATSSG...QTK	125
H1B.FR.83.HXB2	---RA---G---E---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-...MQTOAE-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-RKWNIE--Q--EK--E-AYKKQAMIEMA...116	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-YRRIKQVEQ-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	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-KGKEKONKD...123	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-KEKETONKD...123	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	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	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	112
LST.KE.-.lho7	--SG---RQIEKD-C---GS--T-QK--VE--T---G-GSQ---A-CK---CW--YA--K--K--VGTV--IC-C-LGVVIS--Q--IS--I-PAP-AAG...114	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	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	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	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-V--AN---G-AES---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.VDREVRS--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	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--Q--KI---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-RQCCHLVDK-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

PLV proteins

	Gag p17 Matrix \ / p24 Capsid	Cyclophilin A binding site	
SAB.SN.-.SAB1CELQAKKKNEPTV.....TPS.GGS.RNY.PIVSVN.NQVWHQPLSPRTLNAWVKVIEKKKFAEVVPMFSAALAEIPAIDINQMLNAVGEHQGALQIVKDVINEEADWDLRHP.PPQQ		234
H1B.FR.83.HXB2TG.....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-----CV---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.-.BENAEKMPNTS.....RPTAP--.-KR.G-.-VQQAG.GNY-V-----LV---G---G-Q-S-CT-----C--D--A-M--IREI-----SQ--I-.G		222
COL.CM.-.CGU1SKEEBKAKKEAEKLDLDM--TGP.QGP-----G---CV-..GGIAPSLA--L-YST--A--M-L--ILDLT--FP-VL--E--KK-EEY--L-..V.--		218
DEB.CM.99.CM40EVEQKQKQQEQRQEQEPOAAPQTTTPSV--..LRQG.Q-F-M-----IVKT-INA-V-----P-I--L-QV-----CT-----GL---I-DL---M--I-----E--Q--QQ-P		246
DEB.CM.99.CM5ELE-K.QKTQAAAADTQKKTNQ..PA-F..LRQG.Q-Y-M-----IVKT-INA-V-----P-I--L-QV-----CT-----GL---I-DL---M--I-----E--Q--QQ-P		242
DEN.CD.-.CD1	G.....AAGGSSSAEQGASAGGSSFRDPN-.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_677KKNNETTA.....P-G.-E-.-.-V-NQ--.A-----CV---RWG-----Q--S--CLS--V-----VI-D-----L-E-----E--RT-R--AG		223
GSN.CM.99.CN166PPGAAGGQAVV-.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.CN71PPGAAGGQAVV-.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.88.447PK-.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.88.485QK-.G-.-LIRE-.QR--T-----IQT--IV-DRGWKP-T-A-----T-K-L-D-L-V---I-D---M--I--H-V--G-E--RT--QQ-P		208
LST.CD.88.524PE-.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.-.lho7KKQ.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.-.239TETMP-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.-.MNDGB1K-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-----I-----N---VGP		224
MND-2.CM.98.CM16	SGR.....K-.-QII-.QTP--GI-----C-----P-I-----I--S--CL--L-G---I-D-----I-----N---QVGP		224
MND-2.GA.-.M14	SGR.....K-.-QII-.QTP--GI-----C-----P-I-----I--S--CL--L-G---I-----I-----N---QVGP		224
MON.CM.99.L1P-TG-VPSG--.-V-RTQGGGPQ--AVE--L-KT-Q-----AP--AL-Q-S-M-----L---I-D-----I-----N---V--Q-P-		225
MON.NG.-.NG1		0
MNE.US.-.MNE027AETMP-TSR--A.....PS-.KR.G-.-VQQIG.GNYT-L-----L-----G---G-Q-S-C-----C---A-M--IREI-----Q--QQA.		222
MUS.CM.01.1085A-.-VIRNAQG-FQ--A-NA-I-KT--SIV---A---A-Q-----L---I-----I-----I-----Q--E--LN--Q--		218
RCM.GA.-.GAB1AP--.QR.G-.-ITI-.Q-PE-N-I-----V-----A---S--C-----I-----E-----A---VPG		222
RCM.NG.-.NG411VT.....S-.-QR.G-.-TI-.Q-PE--I-----V-----G-----S--C---V-----I-----E-----D-----NPG		222
SMM.SL.92.SL92BAEKLPAS.....RPTAP--.-G.G-.-VQQ-G-.NY-T-----LV---G---G-Q-S-CT-----C---A-M--IREI-----Q--RQ-P		222
SMM.US.-.H9ADKMPATS.....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.-.PGM53ADKMPITS.....RPTAP--.-RG.G-.-VQQIG.GNYT-L-----LV---G---G-Q-S-CT-----C---A-M--IREI-----Q--Q-.G		222
STM.US.-.STMANKMPATS.....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.L14A.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.-.KE51NDT-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.-.SYK173SQNMASETSSGKVVQEKQKAAT.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.-.TAN1QNYNTAA.....RH.G-.-V-QQ-.---T-----TV---R-G-I---Q--S--CLS-----VI-D---M--I-E--D---Q--IT--AG		229
VER.DE.-.AGM3SG-K-MDKGV--.-P---.Q-F-.AQOQG-.A-I-V-----AV---G--I---Q-S-CT-----VL-D-----EI-----Q-IA--AG		233
VER.KE.-.9063KSNRETTA.....P-G--.-.Q-F-.AQOQG-.A--V-----AV---G--I---Q-S-CT-----VL-D-----EI-----Q-VT--AG		229
VER.KE.-.AGM155KNNTGGTA.....-G--.-.Q-F-.AQOQG-.A--V-----AV---G--I---Q-S-CT-----VL-D-----EI-----Q-VT--AG		229
VER.KE.-.TYO1KNDKGIAA.....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	PPAQQVLRDPQGSADIAGTTSTIQEQIEWTTRA...QNAVNVGNIYKGWII LGLQKCVKMYNPNVILDIKQGPKEPKDYVDRFYKALRAEQTDPAVKNNWMTQSLLIQNANPDCKTVLKG LGMNPTLEEM	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..QQNGG-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-VIQIAR-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--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--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--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--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-MNQ...-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---R---T---V---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---QS-----S-----T-----L-----	348
STM.US.-.STM	-LPA-Q--E-S-----SPE---Q-MY-Q...-PIP---RR--Q---R---T---V---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-----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

PLV protein alignment: GAG

615

PLV proteins

	Gag p24 Capsid \ / p2	p2 \ / p7 NC	
SAB.SN.-.SAB1C	LTACQIGGAQHKKARLMAEAMTAAFQQ	.QTV..GNIFVQQGAR.PRGPLGGRGRPL.....NP...NIKCYNCGKPGHLARFC.KAPRRQ.....GCWKCSPDHQMKDC..Q	453
H1B.FR.83.HXB2	M-----V--PG---VL---SQVTNS.A.....MM-R-NF.RNQ.....RK...IV--F---E--T--N-.R---KK.....KEG-----T		427
CPZ.CD.-.ANT	-----V--PA---VL---AS-NNAQGT.....AV-L-R-NG.N-G.....KR..PL-F---E--T--N-.....K.....R--QEG--L-N-..P		449
CPZ.GA.-.CPZGAB	M-----V--PS---VL---SMVQN-GR.....DV-F-K-QG.AGP.....KR..K--F---E---N-.....K.....R--QEG-----T		437
H2A.DE.-.BEN	-----V--PGQ-----LKE-MGP.SPI..PFAAA.....RK...A-RYW---E--S--Q-.R-----K-G-I-AN-..P		426
COL.CM.-.CGU1	-A----V--PD---KVL---QQFQ-E.....RNTMIEVKTA.....-F--QGI-----M-P-R-IGGAGRGRGRGGFRGAPRRPVPR-FT-NQEG-MQR-..P		432
DEB.CM.99.CM40	-H----V--P-Q-----AS-LKE.....AGSLGM---RRG.....PPGSRR..Q-R-F---QI---QKD-.R--KT.....K-F---QEG-IA-N-..P		457
DEB.CM.99.CM5	-H----V--P-Q-----AS-LKE.....AGSLGM---RRG.....QGGGPRR..QLR-F---QI---VQ-D-.K--KV.....K-F---KEG-IA-N-..P		454
DEN.CD.-.CD1	-Q----V--PGQ---PT---LAS---K.QGLCHGRKGA-PPVE.....R.RGLR-F---QI---KD-.RK-K-V.....TQPGA-F---KLG--A-N-.RS		463
DRL.-.-.FAO	-L----V--PG---V---REQQAA..IMMQQN..PPRGPP-G---PP.....K...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.KK-----RG..PL-F---F--MQ-E-....-QI.....K-F---KIG-MA-...K		427
GSN.CM.99.CN166	-----V--PT--S--L---AT-LKG..TSS...YNM-----K--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--.VTP.--NAQ-RFV-TGGGG..PRR...PLT-F-----S--KM-.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--RGP.....RK...P--W---E--S--Q-.R-----KM--V-AK-..P		428
MND-1.GA.-.MNDGB1	-E----V--P-----M-RTVVG-.S...QN---QRGPQ--VRQPTG.....R...KPI-F--N-E-V--F-....-K.....N--AM--KAQ-..P		423
MND-2.-.-.5440	-L----V--PKY--QM----QQ-QAA..VMMQNSGGPPRGPP-Q-P.....R...PR-P---F--VL-D-.R--KR.....F--D-G-L-RN-..P		434
MND-2.CM.98.CM16	PL----V--PKY--QM----KE-QSA..VM...MQNSGGPPRGPP-Q-P.....R...R-P---F--GL-D-.IS--KK.....F--DLG-I-RN-..P		434
MND-2.GA.-.M14	-L----V--PKY--QM----QKMQSE..VMMQNSGGPPRGPP-Q-P.....R...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--.ARGP.....PQRKG.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...MNM-----GRGGXQPRRQGXQ-R-----F--V-KN-.T--KT.....FR--KEG-XS-N-..P		105
MNE.US.-.MNE027	-----V--PGQ--K---LKE-LAP..GPL..PFAAA--KGP.....RK...P--W---E--S--Q-.R-----QM-G-V-AK-..P		428
MUS.CM.01.1085	-Q----V--PT--S-----LAS--K..AGT...LNM.....-AKGARGRNGSQGNRG-PR-----QY--V-D-.PK--NK.....K-F--REG-LARQ-..P		433
RCM.GA.-.GAB1	-----V--P---VL---QMQSN.....-MA--S-N.RGP-RRSG.....NPN...LR-----IS-Y-....-K.....LL-...T		433
RCM.NG.-.NG411	-----A--P---VL---QMVQSN.....-MA--GPRKGP-KL-G.....PR..FL-----T--T--Y-.TS--KK.....R--EEG-L-...P		434
SMM.SL.92.SL92B	-----V--PGQ-----KD-LTGSLVAAQFRGAAGK--...-KP..I-R-F---T--S--Q-.R--K.....EEGRIQAN-..P		432
SMM.US.-.H9	-----VX-PGQ-X-----LKD-LT-.GPL..PFXA--KGO.....XK..I-X-X-X-E-S-QF.R---X.....KAG-V-AK-..P		429
SMM.US.-.PGM53	-----V--PGQ-----LKE-LTP..GQL..PFAA--KGP.....RK...T--W---E--S--Q-.R-----KTG-V-AK-..P		429
STM.US.-.STM	-----V--PGQ-----LKEV--P..DPL..PFAAA--QG.....RR...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---VMTQHS.....LGM1-.P.....QGSNPRR..GPTR-F---QL--QKD-.PR-KKL.....K-FN--GTG-IARQ-RQP		452
SYK.KE.-.SYK173	-Q----V--PL--K---VM-Q-S..V.....NM--GPSK.G.....RS..M-----QI--MQKD-.K-LKA.....K-FN--KTG-LARA-RQP		451
TAN.UG.-.TAN1	-----V--PG--K--V---QQMQGV.....NM--GAP-.GGRGR.....RG...PPR-FK--QI--IQKD-.PRAGPN.....K-L--K-G-LA-...R		437
VER.DE.-.AGM3	-----V--PSY--KV--M-QMQS.....NM--GG.RGR.....R...PV-----F--MQ-Q-.PE--KM.....R-L--K-G-LA-...R		439
VER.KE.-.9063	-----V--PSY--V--M-QMQT.....NM--GG.VRRQ.....R...PL-----F--MQ-Q-.TE--KI.....K-L--KLG-LA-...R		437
VER.KE.-.AGM155	-----V--PSY--KV--M-QNLS.....NM--GG.RGR.....R...PP-----F--MQ-Q-.PE--KI.....K-L--K-G-LA-...R		435
VER.KE.-.TYO1	-----V--PSY--KV--M-QTMQN.....NM--GP.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.GRGKP.RNFPLTSI.....RPTAPPMERDYSRPEENWYADRPPTRGPGPDDPATALLKQYAVQKQRKQWQNHSPQSSPYEEAYSSLSRSLFGEDQ\$RPTAPPMERDYSRPEENWYADRPPTRGPGPDDPATALLKQYAVQKQRKQWQNHSPQSSPYEEAYSSLSRSLFGEDQ\$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-TWWGCR-.G--VQKEEVV.....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.....EKSQE.....E----I-SYGYQE.....EKSQE.....KKEGE.....S-L-PP..T--K---S-PSSQ\$	508
H2A.DE.-.BENER-AG--L.-RG-.K---V-QAPOGL.....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-ESTADQVYSTTRVPVQGENN-Q-EETTRRGKYPD-MGLKSLFGEDQ\$I-ESTADQVYSTTRVPVQGENN-Q-EETTRRGKYPD-MGLKSLFGEDQ\$	535
DEB.CM.99.CM5GQVPR---N.A--AS.RR---LEGIGKY.....I-ESTADQVYSTTRVPVQGERNA-QEDTTTRRGKYPD-MGLKSLFGEDQ\$I-ESTADQVYSTTRVPVQGERNA-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\$S---S--LQ\$	521
GRV.ET.-.GRI_677	N.....G-A---Y.-H-G-.A---VQYRGDTVGL.....E-----TA-DPA.....E-----TA-DPA.....KK-L-QYAEKGQRLREERBQTRK-KEK-VEDV--S---G----	513
GSN.CM.99.CN166GTFPI-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-L.....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-NLGQQLRR.....V-S-----SPTQA-RALETY-NLGQQLRR.....QQ-.IP-KCVDEPCLSFPPFPDDQ\$	512
LST.CD.88.485	..K...PKGK----Y---RN.-P-.G---MGGNAGV.....V-S-----NSTKA-RALETY-NLGQQLRK.....V-S-----NSTKA-RALETY-NLGQQLRK.....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-TLGQQLKR.....I-S-----SAPTKA-RALETY-TLGQQLKR.....QQ-QVP-KCVDEPCLNSLF.PDDQ\$	515
MAC.US.-.239DR-AG--L.-G-.K---MAQVHQGL.....M-----EDPAVDLL.....KNYMLGKQ-REK-RESREKPYKEVTEDLLHL..N---G----M-----EDPAVDLL.....KNYMLGKQ-REK-RESREKPYKEVTEDLLHL..N---G----N---G----	510
MND-1.GA.-.MNDGB1	KPA...QQQR----Y---GP.S--.G-Y-AQEV.....T----L-EKPLQKTLST-QKLGRL.....RQMKKEEKREEDFHSLSLTFQEDQ\$T----L-EKPLQKTLST-QKLGRL.....RQMKKEEKREEDFHSLSLTFQEDQ\$RQMKKEEKREEDFHSLSLTFQEDQ\$	502
MND-2.-.-.5440	K.....M-----N.T--GS-----AMPL.....T-S---PGLEDPA.....EKMLLDYMKKQQQRAA...AGA--EKKKGPY...EAAYNSL...S---T--LQ\$T-S---PGLEDPA.....EKMLLDYMKKQQQRAA...AGA--EKKKGPY...EAAYNSL...S---T--LQ\$S---T--LQ\$	518
MND-2.CM.98.CM16	K.....M-----N.T--GS-----AMPL.....T-S---PGMEDPA.....E-MLLDYMKKQQQ-AE...S-QE-KE.RGPY...EAAYNSL...S---T--LQ\$T-S---PGMEDPA.....E-MLLDYMKKQQQ-AE...S-QE-KE.RGPY...EAAYNSL...S---T--LQ\$S---T--LQ\$	517
MND-2.GA.-.M14	K.....M-----N.A--GS-----AMPL.....T-S---PGMEDPA.....EKMLLDYMKKQQQRAAKE.T-QE.KD.KGPY...EAAYNSL...S---T--LQ\$T-S---PGMEDPA.....EKMLLDYMKKQQQRAAKE.T-QE.KD.KGPY...EAAYNSL...S---T--LQ\$S---T--LQ\$	518
MON.CM.99.L1	QP..PKQ.N-G-----N.PFGP.KKG-----V.....Q-S---TLPAGTAE-VNLNQEIGS.....KTG-PKETRRDLYPSLASLFGEDQ\$Q-S---TLPAGTAE-VNLNQEIGS.....KTG-PKETRRDLYPSLASLFGEDQ\$KTG-PKETRRDLYPSLASLFGEDQ\$	510
MON.NG.-.NG1	NGGQNPQRNR----N.PFGP-KRG-----V-L.....Q---AEPTXGYKVNFSPPVG--ETPAPTAPKE-RKDLYPSLASLFGDD-SEQ\$Q---AEPTXGYKVNFSPPVG--ETPAPTAPKE-RKDLYPSLASLFGDD-SEQ\$RKDLYPSLASLFGDD-SEQ\$	192
MNE.US.-.MNE027DR-AG-----GK.....MAQMHQGL.....T----EDPAVDLLKNYMLGKQRE.....SKRKPYPEVAEDLLHLSLFGEDQ\$T----EDPAVDLLKNYMLGKQRE.....SKRKPYPEVAEDLLHLSLFGEDQ\$SKRKPYPEVAEDLLHLSLFGEDQ\$	506
MUS.CM.01.1085	...RSDNT-A-----D.PFGLQKRR-----Q-L.....L---EQEGELVNLSPAVGPGEGKQDRKSLY-SLSS-F-DDQSDQ\$L---EQEGELVNLSPAVGPGEGKQDRKSLY-SLSS-F-DDQSDQ\$F-DDQSDQ\$	509
RCM.GA.-.GAB1	K.....I---R.L--GQ-----L.....T-S--G--SN-DPA.....EEMLNKYLRRAGEQ-RQQ..RQEESSKREGAY...QEALTSL...N---S--LQ\$T-S--G--SN-DPA.....EEMLNKYLRRAGEQ-RQQ..RQEESSKREGAY...QEALTSL...N---S--LQ\$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.\$N---S..L.\$	521
SMM.SL.92.SL92BNOKAG--L---G-.K---M.QTTS.L.....T-S---D.PAARIV.....E-LEKAQ-E-TRRSRYPYKEVTEDLLHL..N-----T-S---D.PAARIV.....E-LEKAQ-E-TRRSRYPYKEVTEDLLHL..N-----N-----	507
SMM.US.-.H9ER-AG--L---G-.K---MAQMPQGL.....T---X-EDPAVDLL.....N-MKV-R--RENRRPYKEVTEDLLHL..N-----T---X-EDPAVDLL.....N-MKV-R--RENRRPYKEVTEDLLHL..N-----N-----	507
SMM.US.-.PGM53ER-AG--L---G-.K---MAQMPQGL.....T---EDPAVDLL.....RN-MKM-RK-RENRRPYKEVTEDLLHL..N---G----T---EDPAVDLL.....RN-MKM-RK-RENRRPYKEVTEDLLHL..N---G----N---G----	507
STM.US.-.STMER--G-----G-.K--H--MAQ-PQGL.....T---EDPAADLL.....RS-MQL--K-RESRKTYPYKEVTEDLLHL..N-----T---EDPAADLL.....RS-MQL--K-RESRKTYPYKEVTEDLLHL..N-----N-----	507
SUN.GA.98.L14	Q.....G-----Y---NR.-P-.G--V.....M-S---L-DLTLGNRMTTPPQSKAERALETYRLLGQG-R.....AQ--RKERGECQEPCLNMLFPEDP\$M-S---L-DLTLGNRMTTPPQSKAERALETYRLLGQG-R.....AQ--RKERGECQEPCLNMLFPEDP\$AQ--RKERGECQEPCLNMLFPEDP\$	521
SYK.KE.-.KE51	RKGQGNPPP-A---K.-GN.R--PA---VMRE.....T---A-DWPWQQQTWGNVYAS-QQKHIAMAT-LQPQSSPKT-KDLLVPGKPS.....EKK-EMKEKGP-YPSL-SLFGDDQ\$T---A-DWPWQQQTWGNVYAS-QQKHIAMAT-LQPQSSPKT-KDLLVPGKPS.....EKK-EMKEKGP-YPSL-SLFGDDQ\$SLFGDDQ\$	557
SYK.KE.-.SYK173	KRNQGGPPVA-A---K.-GVSRR-PA---V.RS.....E-S---L-DIEDG-WLT-S-QMSQQAQAKAQNPSKPPPTNREVLSP.....KESSGKEETKSLYPSL-SLFGEDQ\$E-S---L-DIEDG-WLT-S-QMSQQAQAKAQNPSKPPPTNREVLSP.....KESSGKEETKSLYPSL-SLFGEDQ\$SLFGEDQ\$	554
TAN.UG.-.TAN1	S.....G-A---RMPT-.T---LEQGGG.....V-----PAHGFP.....GSPA-GAYDP-RKLLLE-YAKKGDQLRK-KEK-LEDY--S---G----V-----PAHGFP.....GSPA-GAYDP-RKLLLE-YAKKGDQLRK-KEK-LEDY--S---G----S---G----	523
VER.DE.-.AGM3G-----Y.-R-M-.A---AATLGV.....E-----PPSP-DPA.....KK-L-QYADKKG-LREQRKPP-AVN-DWTEGY--N-----E-----PPSP-DPA.....KK-L-QYADKKG-LREQRKPP-AVN-DWTEGY--N-----N-----	521
VER.KE.-.9063G-----Y.-R-T-.T---LAATHGV.....E-S---PPTP-DPA.....KK-L-QYAEKKG-LREQGKRT-PTN-DWAEgy--N-----E-S---PPTP-DPA.....KK-L-QYAEKKG-LREQGKRT-PTN-DWAEgy--N-----N-----	519
VER.KE.-.AGM155G-----Y.-R-M-.T---AATLGA.....E-S---PPNNSTPY.....DPAKK-L-QYAEKKG-MRNQNRNP-ANN-DWNEGY--N-----E-S---PPNNSTPY.....DPAKK-L-QYAEKKG-MRNQNRNP-ANN-DWNEGY--N-----N-----	520
VER.KE.-.TY01G-----Y.-R-M-.A---AATLGA.....E-S---PPSGTTPY.....DPAKK-L-QYAEKKG-LREQKRNPN-AMN-DWTEGY--N-----E-S---PPSGTTPY.....DPAKK-L-QYAEKKG-LREQKRNPN-AMN-DWTEGY--N-----N-----	519

PLV protein alignment: GAG

PLV proteins

Accession	Protein Name	Residues	Length
SAB.SN.-.SAB1C	/ Pol reading frame (-1 from Gag) p15	FRVWPLG...QRET.QEFPSPD...LHQ...TNSSPNG.TGLQQ...AGGKLVCRQTSQDQTRARRSSNSPVKAVCCSGETAETAVA...KPLATTEPLRGGQLQPVSLWRRPMKTVYIEGQ	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...GTTS-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...QEgEGGES-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...EGGNNE-REVP-FNGAEI-----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...PSDESGFAMEPSEGAATTVPTKF-RGGKQTIILSGRQ-AEQ-V...PSGGRGDNEGES-SFVE-----IGQIWWQ--		94
DRL.-.-.FAO	--EY...SLGQRETQEFSS...SGL...ADPISSSIARIPPGGSG-ED-PKVHGAGSS-EETT...ET-T...-EEG-----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...PPSQELSSNFPAANS PG...G...-RTPAA-SL...DT-RG...TP...EQGTEGGEQHS-S-VI--G---E-S-G--		76
GSN.CM.99.CN71	--P-GA...TSPQELSSNFPAAIS PG...G...-GTPAA-PV...SAGGG...AP...EQG-EGRKEEHS-S-VI--G---RE-SMG--		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-RAMPEFLS-S---IENILVN--		85
LST.CD.88.485	--L-T-E...EWSS.RKLSF...GGE...CR--T...ISSPNKELNK-RESLR.NIQ-PRT-TEE...ATA.GSTEMC-RAMPEFLS-S---IENILVNR--		85
LST.CD.88.524	--S-E...EWTf.RQLSNH...GRG...CR-DS...FSA-HGESA--KSREGNDIQ-PRT-TEE...ATT.SSTVC-RALPEFLS-S---IEK-LVN--		86
LST.KE.-.lho7	--I-S-E...QTS.RQLS-T...GRS...CRENS...ISSPNGEC-NK-RESTG.DLQDFRT--EK...AAA-SPTMC-RALPEFTL.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...QAEDEGGEE--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...TGKEREESI--S-----F--N--TTV-E----		85
MND-2.CM.99.MND254			0
MND-2.GA.-.M14	--EC...SLGQW-AQELS...C-A...TDPISTPDARDGGS-REDVIRLYEKGATTEGSKGD...QAGEGQGS---S-----F--N--TTV-E----		86
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	--ESFP...RQ-GAK-LS-N...LA...A--PS...GTNPXGXQGELLTSSGTTDRDPS...ADS.PEGDEE-S-S-LSIP--G--VR--E-A--		80
MNE.US.-.MNE027	--L-M...KEAP.-FPHGP...NAS...GADTN...CSP...RGPSCGSAKELHAVGQ-ER...QKEALQGGGR-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...LSEEGR-TKET-EAG...REQBERGSIS-S-N---FA--K--TTIAQ----		87
RCM.NG.-.NG411	--EC.SLE...AG--EEPSLDF...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...DSQGVSRRES-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...RRPSCGSAKELHAAGQE-ER...EQED-LQGGDR-FAA--F-----VV-AH----		81
SUN.GA.98.L14	-----E...GAA-R-LS-H...AF...CSA...SRGSDIGE...QNDTTPVKGRESSDGLQ-VR-GPQ...STABEEGKGGMSGAMLEYA-S---LEK-F-N--		87
SYK.KE.-.KE51	--ERVGE...-ET...PSKFSCBBDTNSPT-GGLAMAATDMGKLCFTPTAHNSHGHPA...ATVL-KDSQGLPGWETF...REERD-GEGTS-S-SI-----MEINV--T		105
SYK.KE.-.SYK173	--ERMG...L-EAPS-LPSEERAICSPS...G...HRRWAMAHMVSTDEPTST-EGTELSLEETPHQQGS-L...EGEQWEGGDQES-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-RIFEL.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...GDK-EK-SRR...YSW...SGTE...CAS...STERHHP...I-P-KEAPA-I-RER--T-G-KE...ESTGNESG-DR-IPFEL.P-----I-----V		89
WRC.CI.97.14			0



Asp25 catalytic site

	Protease	Pol p51 Reverse Transcriptase		
SAB.SN.-.SAB1C	KVTALLDTGADDSVIQGIEL...	GDNWKPRIIGGIGGCINVKAYHNQEVKIEDKTKC...	ATILVGETPVNIIGRNVLAQLGVTLLNTQREIEPI...KVHLKPGQDGPRIHQWPLSKKEIEALKAI	
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-	
CPZ.CD.-.ANT	-CQ-----T-VE--H...	QG---KT-----F-S-QQ-NKVIQ-G-R-VL...	--V-L-PN-----CL--C---FPISKV-TV...P-K--E-M---VK-----E-	
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-	
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-	
ASC.-.-.Qu	
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-	
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-FPVLDT-V...L...E-A---K-K---R-----D-	
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-FPVLDT-V...L...E-A---K---R-----E-	
DEB.CM.99.DEB1014	0	
DEN.CD.-.CD1	R-EV-I-----TI-AEQDINLGTEG..	I-KTV---Y--Q--PGI-I-LM--VAV...	D-V--P--I--L--Y-SK-----MAVQMR-Q.VTP-K---F---K---TT--LK--QQ-	
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-		
GRV.ET.-.GRI_677	P-Q-----TI-EKD-HPPPHKP-RSKVV---	G-H--E-QGVQ-QL---IIT...	GS--I-S--I-----I---A-MK-VMGVLSSQIE.ET--Q--E-K---KLK-----R-----TE-	
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-	
GSN.CM.99.CN71	GSN.CM.99.CN71	TIV-BSAIEL..	DTP-T-KTV---L-Q-EHRHV--IFNE-RI...	--V--P---L---C-SK-----MV-QK---V...E---S-KE--K-K---TQ-
GSN.CM.99.GSNCN7	0	
LST.CD.88.447	P-S-----TIFSENSVRI..	EGPYT--VV---Q-R--E-KDVFIIE-AG--TV...	G-V-L-P---D-V---I-TNV-AR-V-S-LSDKIP.IT--K---C---VK-----G-Q-	
LST.CD.88.485	P-S---R---TIFSENSVRI..	EGPYT--VV---Q-R--E-KDVFIIE-AG--TV...	G-V-L-P---D-V---I-TSV-AK-V-A-LSDKIP.IT--R---C---VK-----G-Q-	
LST.CD.88.524	E-----TIFSENSVRI..	EGPYT--VV---Q-K-EFRDVIQ-AG--TV...	G-V-L-P---D-V---I-NP--AR-V-A-LSEKIP.IT--C---C---VK-----G-Q-	
LST.CD.88.524	P-S-----TIFSENSVRI..	EGPYT--VV---Q-R--E-RDVFIIE-AG--TV...	G-V-L-P---D-V---I-TAM-AK-I-A-LSDKIP.IT--S---C---VK-----G-Q-	
LST.CD.88.524	P-S-----TIFSENSVRI..	EGPYT--VV---Q-R--E-RDVFIIE-AG--TV...	G-V-L-P---D-V---I-TAM-AK-I-A-LSDKIP.IT--S---C---VK-----G-Q-	
LST.CD.88.524	P-S-----TIFSENSVRI..	EGPYT--VV---Q-R--E-RDVFIIE-AG--TV...	G-V-L-P---D-V---I-TAM-AK-I-A-LSDKIP.IT--S---C---VK-----G-Q-	
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-	
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-	
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-	
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--KDL---KG---Q-----S---QFF-CK-T-AG--TH...	-SV--P---V---KK-C---FPVSKV-TV...-T---T---K-K-----L--QE-		
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-		
MON.CM.99.L1	--SM-----TI--DQHVNL..	E--T-KVV---M-E--Q-KAIQ-IF--REVW...	--V--P--I--L---KM---MAAGGLQPPSLE-T-A-KE--KLK---R-----TQ-	
MON.NG.-.NG1	S-SM---X---TIV-XQDINL..	E--T-KMV---L-Q-QPKGVT-SFQG--VR...	X-V-L-X--I--L---SRM-I---MX-GIAQSP.SLT-T---K---KLK-----TT-	
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-	
MUS.CM.01.1085	V-SM-----TI--DAHIHL..	EGP-N-KSV---N-S--E-RGVQ-CFNG--IR...	--V--P--I--M---C-SKF-I---MI-GKL---T---E-AK--MVK---TQ---R--ED-	
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-		
RCM.NG.-.NG411	--E-----GT--NLD---E---K-----Y---QFNHCKIT-AG--TH...	-HV--P-----SK--AS--FPISKA-T...-E---VK-----TE-		
SMM.SL.92.SL92B	PLEV-----IVA-----	PGYT--V---F--T-E-RTVKIEVLG-VIR...	E-LMT-D--I--F---I--A---S--YPAKV-YT...-K--E-M---LK-----Q--TE-	
SMM.US.-.H9	PIEV-----IVA-----	P-YT-K-V---F--T-E-KVKI-VLG-VI...	G-MT-D--I--F---L-TAM-MS---PIAKV---T---K---KL-----I--RE-	
SMM.US.-.PGM53	P-EV-----IVA-----	P-YT-K-V---F--T-E-KVKI-VLG-VI...	G-MT-D--I--F---L-TAM-MS---PVAKV---T---K---KL-----V--RE-	
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-	
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-	
SYK.KE.-.KE51	--M-----TI-REQDIQL..	HQP-T-K-V--L--N-R-RQ-NQIRFQ--RPDGRIVEGSSL--P----	I-SK--AK-VMIAQKL--M...-Q---KEL-KLK---TR--L--R--	
SYK.KE.-.SYK173	L-QM-V-----TI-REEDIQL..	HQP-S-K-V--L--N-T-RQ-R-IRFVVKPSG-RKQVEG-L--P----	I-TK--K-VMV-TAL--V...-S---DKEL--LK---V--L--	
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-	
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-	
VER.KE.-.9063	VER.KE.-.9063	VER.KE.-.9063	VER.KE.-.9063	
VER.KE.-.AGM155	VER.KE.-.AGM155	VER.KE.-.AGM155	VER.KE.-.AGM155	
VER.KE.-.TYO1	VER.KE.-.TYO1	VER.KE.-.TYO1	VER.KE.-.TYO1	
WRC.CI.97.14	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-	

PLV protein alignment: POL

619

PLV proteins

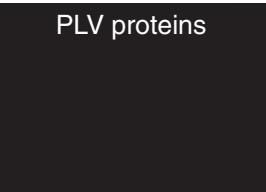
	- drug resistance sites -	catalytic site	
	M41L A62V D67N K70R	Asp110	
SAB.SN.-.SAB1C	CEDLEKQGHLEIRIGPENPYNTVPFAIRKKDKTQWRILMDFRQLNKSTQDFQEVQLGIPHPAGLQOREQIT.....VLDIGDAYFSCPLDPDFQKYTAFTIPSVNNREPGIRYQYKVLPLQGWKGSPTIFQ		347
H1B.FR.83.HXB2	-TEM--E-KISK-----K--S-K--K-V--E--R---W-----K-K-K-S-V-.....-V-----V---E--R-----I--ET-----N-----A---		315
CPZ.CD.-.ANT	-DK--AENKIS---D-----I--K--TSK--K-V--E--R---W-I-----K-K-K-S-V-.....-V-----I---Q--R-----ET---C-N-----A---		315
CPZ.GA.-.CPZGAB	-QEM--E-KISK-----I--K--S-K--K-V--E--R---W-----K-K-K-S-V-.....-V-----K--R-----I--ET--V---N-----S---		313
H2A.DE.-.BEN	--KM--E-Q--EAP-T-----T--K---NK--M-I--E--V---T-I-----AKKKR-S.....I--V-----I--HE--RQ-----L-A--M--K--I-----A---		340
ASC.-.-.QuXQVLRQMEHVSIV-MK--FY-V---A-AP---V---AS--E--FR-----		65
COL.CM.-.CGU1	--R--AE-KI-KAELG-----I-C-----NE--K-I--E--R---M-----G--IEKTH-.....-I---Y-----LA--T--V-----		297
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	LKPMIEA-QIVPAA-T---S--V-K---SK--M-I-L-A--A-EVW---T---SA-P-M-----LA--Y-IA--T-A---V---IQ--E--EFR---SA--A---		321
DEB.CM.99.CM5	IRPMVEA-QIVPTE-T---SS--V-K---NK--M-I-L-A--A--VW---T---S--P-M-----LA--Y-I--Q--AR---TQ--E--EFR---SA--A---		321
DEB.CM.99.DEB1083		0
DEN.CD.-.CD1	VTQM--E-R--KAP-T-----V-P---Y-M-I--K--EA---W-----K-H-.....-LK--Y-V---AQ---V--I--EG--Q--IF-----V---		341
DRL.-.-.FAO	--QM--E-QIS-----C-K--G-K--K-V---E--A--F-----G--KAC-----V-----QG--R-----N-----A---		329
GRV.ET.-.GRI_677	-KQM-EE-K-S--G-----K-K---M-V--E--A--F-----K-K-.....I---Y-I--CKE-R-----TG-----FNC-----		337
GSN.CM.99.CN166	TQEMLR-L-Q--K-----S--K---S--M-I--K--EA---LA-----IPHPAGLEQKEHVTII-MK--Y-I--YE--R---SV---QA-AE--F---A---		326
GSN.CM.99.CN71	TQEMLR-L-Q--K-----S--K---S--M-I--K--EA---LA-----LEQKEHVTII-MK--Y-V--YEE-R---SV---QT-AE--F---A---		319
GSN.CM.99.GSNCN7		0
LST.CD.88.447	-DR--E-KISKADAG---I--K---NE--K-I---AL---H---IKKCK-V-.....-V-----I---YRQ---V--I--Q--R---N---C---		331
LST.CD.88.485	-DR--E-KISKADAG---I--K---NE--K-I---AL---LH---IKKCK-.....-V-----I---YRQ---V--I--Q--R---N---C---		331
LST.CD.88.524	-DR--E-KISPV-D-G---I--K---NE--K-I---K--EL---H-L---IKKCK-.....-V-----I---YRQ---L--Q--K--N---C---		332
LST.KE.-.lho7	-DR--E-KISPV-D-G---I--K---NE--K-I---K--EL---H-L---IKKCK-.....-V-----I---NYR---L--Q--K--N---C---		331
MAC.US.-.239	--KM--D-Q--EAP-T-----T--K---NK--M-I--E--RV---T---AK-KR-.....I---EE-RQ---L---A--K--I-----A---		327
MND-1.GA.-.MNDGB1	-DR--EE-KIS-VD-G-N---I--K---NE--K-I---E--L---H-L---IKKCKR-.....I---YRP---V---QA--K--M-N---C---		324
MND-2.-.-.5440	-NQM--E-KIS-V--G-N---C-K--G-S--K-V---E--V---F---T-G--K-C-----V--R---V---QG--N---A---		328
MND-2.CM.00.MND2S46		0
MND-2.CM.00.MND2S6		0
MND-2.CM.01.MND2S109		0
MND-2.CM.98.CM16	-SQM--E-QISK-----C-K--G-S--K-V---V---F---G--K-C-----E--R---QG--N---A---		327
MND-2.CM.99.MND254		0
MND-2.GA.-.M14	-DEM--E-KIS-----C-K--G-S--K-V---RV---F---G--K-C-----I---ES-R---QG--N---A---		328
MON.CM.99.L1	TQEMLR-L-Q--PTE-N---S-C--K---SK--M-I--E--A--F---EKMDHV-.....IV-MK--FY-I--WEP-R---S---L--A--AK--F---V---		323
MON.NG.-.NG1	XREMIEL-Q--EAE-T---S-C--K---K-M-I--E--A--F---NMN-.....IV-MK--FY-I--EN-R---S---A--AV--FR--X---		326
MNE.US.-.MNE027	--KM--D-Q--EAP-T-----T--K---NK--M-I--E--V---T---AK-KR-.....-V-----I---EE-RQ---L---A--K--I-----A---		323
MUS.CM.01.1085	TQQML-V-QI-E-----S-C--K---SK--M-I--EA--ELA-----IEH-HV-.....IV--K--FY-V---R---SL-Q--LG-AR--F---		316
MUS.CM.01.MUSS1239		0
OLC.CI.-.SIVolc_97IC_12KEG-----I---EYRE-----DR--R--I--N---C---		69
RCM.GA.-.GAB1	-NAM--E-KISK-----I-C-K--S-K--K-V---E--R---F-----G--K-C-R-.....L--YEP-R---A---QG--V---N---A---		329
RCM.NG.-.NG411	-NAM--E-KIS-----I-C-K--S-K--K-V---E--R---F-----G--K-CR-.....YEP-R---LAI--QG--N---A---		333
SMM.SL.92.SL92B	--KM--D-Q--XAP-T-----T--K--N-DK--M-I--E--M--E-T---KEM-R-.....-V-----V---E-RQ---Q--K--I-----A---		320
SMM.US.-.H9	--KM--D-Q--XAP-T-----T--KXX-GNK--M-I--E--V---T---AK-RR-.....-V-----I---EE-RQ---L---A--K--I-----A--X		323
SMM.US.-.PGM53	--KM--D-Q--EAP-T-----T--K---NK--M-I--E--RV---T---AK-RR-.....-V-----I---EE-RQ---L---A--K--I-----A---		323
STM.US.-.STM	--KM--D-Q--EAP-T-----T--K---NK--M-I--E--V---T-I---AK-RR-.....-V-----I---EG-RQ---L---A--K--I-----A---		323
SUN.GA.98.L14	-DR--AE-KIS-AD-G---I--K---NE--K-I---V--GM---H-L---KKCK-.....-V-V-----I---YR-.....QA--K--V-N-----C---		333
SYK.KE.-.KE51	VS-MLEK-Q--KAS-T---V-Q---LK--MP---K--EA---W---G--STKR.M-.....-LK--YTV--EE-RQ---V--I--SS--E--F---		353
SYK.KE.-.SYK173	V--ML-A-Q--KAS-T---V---KK--M-I--K--EA---F---G--KRQK.L-.....II-LK--Y-V--KE-RP---V--I--AS--E--FT---		347
TAL.CM.00.TAL266		0
TAL.CM.00.TAL271		0
TAN.UG.-.TAN1	TQA--E-KISKV-G-----C-K-----M-I--A--A--F---KRKK-.....-Y-I---E-RP---Q---VG--FNC-----		335
VER.DE.-.AGM3	-DQ--E-KISK--G--A---C-K---S--M-V--E--A--F---S-FEKMT-.....-Y-I---E-R---QG--T--FNC-----		332
VER.KE.-.9063	--Q--E-KISK--G--A---I-C---M-V--E--V---F---KMK-.....-Y-I---E-R---QG--V--FNC-----		332
VER.KE.-.AGM155	-KT--EE-K-S-V-GD-A---C---S--M-V--E--A--F---KMK-.....II-V---Y-I---E-R---T--EG--FNC-----		335
VER.KE.-.TYO1	-SQ--QE-KIS-V-G--A---I-C-K---S--M-V--E--A--F---RKMR-.....-V---Y-I---N-R---T--QG--FNC-----		335
WRC.CI.97.14RKKN-S.....-V-----I---EYRQ---V---ES--K--V-N---C---		69



	polymerase motif catalytic site Asp185 Asp186	drug resistance T215Y K219Q	
SAB.SN. . . SAB1C	TTANKILQBFQKNDVDIYQYMDMLIASDRPKAEHLVMVQQLRDYLETWGFKTPPEKKFKQKPPYLWVMGYELYPKKWQLQBITLPEREE. . . WTVNDIQKLVGKLNWASQIYTGIKTKHLCLRLIRG	ARP	474
H1B.FR.83.HXB2	SSMT---EP--KQ---IV-----LYVG--LEIGQ--RTKIEE--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-V-TL---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.DEB1014			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.GSN107			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-EIEKTK--DIDLQKT---N-----V---A-L-E--R--E--K---VKA		460
LST.CD.88.524	E-VSGL-A---KL---ILL-----LF-G---E-EG-SQA-KE--EL-I--NLE-----DK--H---V-H-DR-EIERVK--IDTRKT---E-----V---A-L-D--R--E--K---VK-		461
LST.KE. . . lho7	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
MAC.US. . . 239	Y-MRNV-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--KIEKQV--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--NHSPL--RKI-GVI-AT-LRK-IQ--PE--D-Y-VQ-L---H--SH-KIAK-E--DQ-Q. . .	E---I-----A--S---N--KS---S	450
MON.NG. . . NG1	YS-SELKXXI-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-NM-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---. . .	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---. . .	E--D---DK--S-----E--K---VKA	460
SMM.SL.92.SL92B	A-MRQV-EP--KA---LLV---L--G-N-GLT--DK--T---M-NNL--S---D---N--LQ--L---L---K--K-E---K-R. . .	V---A---P---N--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-----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-----KMT	450
STM.US. . . STM	Y-MRN--EP--RA---TLI-----I-----TDL--DRV-L--KEL-NNL--S---E-----FQ-----W-T--K--K-E--Q-DV. . .	V---A---P-----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--QOI-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-YS-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---EKIKELKPLT-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. . .	A-L-P-L--R-I-K--T-GKKN	462
WRC.CI.97.14	S-VASL-EIY-G---NIQVL-----LY-G--LD-T---NK-EE--EL-L--NLE---L-LE--CN---I-H-D---I---K---EVK. . .	V-----A-L-E--Q--Q---NK-	195

PLV protein alignment: POL

621



PLV proteins

PLV proteins

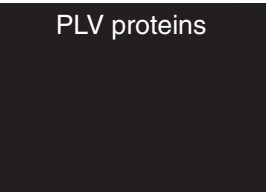
SAB.SN.-.SAB1C	LTEIVQWTEEALELEENRQILRQKQQGYYDPALPLRAKVLKLGDGQWGYQIYQPE.NKILKVGKYAKIKTAHTNELRMLAGLVQKIGKESIVIWGQI.PIMELPVERELWEQWSDYQVWTWIPEWEM	602
H1B.FR.83.HXB2	---VIPL-----A---E--KEPVH-V---SKD-I-EIQ-Q-Q---T-----EP.F-N-T---RMRG---DVKQ-TEA---TT-----KT.-KFK--IQK-T--T--TE---A-----F	570
CPZ.CD.-.ANT	--DR--M-R-----K---Q--IE-Y--Q-G---K-TIQ-Q-S---T-----N..G-L-A---RPTGT---V-Q---V---L-----EV.-KFO--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.-.KPH-----T-----DN-----DF	594
ASC.-.-.Qu	-LDK-ELS-D--A--R--QE--S-EHT-G--RED---EVE-VS---T-----MFK-D..KQV--T--F--VRSV-V-QYGO--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--FS-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.-KPH---KDV---T-----DF	581
MND-1.GA.-.MNDGB1	I--K-TM-----R--Y-Q-KE--AEE-E-S---NKE-YVRFO-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.-KPY--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.-KPY--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.-KPY--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.-KPH---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..TKGTVAYRIWQKHGT--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--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.-KPH---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.-KPH---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.-.KB51	--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..G-----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



p51 Reverse Transcriptase \ Pol p15 RNase H

SAB.SN.-.SAB1C	VSTPQLRLRLWYKLVKDPPIGGEAVYVVDGAANRNSKEGKAGYLTRDGRQKVVVALENTTQKAELEAILLALRDSGSKVNIITDSQYAMGIAGEPTESDNNIVQQIIIEELIKKEAVYIAWVPAHKGVGGNE	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---Q--DQQ---V-----VL---QSQ-DH-ESEL-N---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-BEE--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.DEB1083	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--T-EHQ--E---QAMQG--EI--T-----I---T	727
DRL.-.-.FAO	I---H--K---N-LS-----E-----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--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-BEYL---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-----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-ESPL-----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---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---V-Q---ESRL-N---KM--FY--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---V-Q---ESRL-N---M---I-V-----I---Q	707
STM.US.-.STM	---P-V--VFN---E-LE-TETF--SC--Q-R---I---KN--K--Q---Q---FAM--A--P-A--V---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---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---V--KM--E--E--V---V---LTAQ--Q--SPL-E--ALMV--R-I-LQ---I---	716
VER.KE.-.AGM155	-V-P-VT---T-T-E---D-----C--Q-----I-QQ-K-R-QQ---Q---T--KM--E--P---V---V---LTAQ--Q--SPL-E--AQMVQ---I-LQ---I---	719
VER.KE.-.TYO1	P-LK---T-T-E---K-D-----C-----ISQY-K-R-ET---Q--K-T--KM--E--PN--V---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

PLV protein alignment: POL



PLV proteins

PLV proteins

Table with 3 columns: Gene/Protein Name, Pol p15 RNase H \ Pol p31 Integrase alignment, and numerical values. Rows include SAB.SN, H1B.FR, CPZ.CD, CPZ.GA, H2A.DE, ASC., COL.CM, DEB.CM, DEN.CD, DRL., GRV.ET, GSN.CM, LST.CD, MAC.US, MND-1.GA, MND-2.CM, MON.CM, MNE.US, MUS.CM, OLC.CI, RCM.GA, SMM.SL, STM.US, SUN.GA, SYK.KE, TAL.CM, TAN.UG, VER.DE, VER.KE, WRC.CI.



SAB.SN. .SAB1C	TQLHTDNGTNFTSQVAAICWGWKIEHTFGVYPNPQSQGVVESMNMKQLKEIIGQIRDDAERLETAVIMAVHIHNFKRKGGIGGYSAAER.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.DEB1014	K-I-----P--I--P-K-A--LGV---T-----K-RV--T-K--E-----	174
DEB.CM.99.CM40	--I-----P--I--P-K-A--LGV---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--LGV---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--LGV---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--EEVTV---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. .L-K-K---Q---Q---	970
LST.CD.88.485	Q-V-----P--V--KE-Q-VI--LG---T-I-----AK--V--Q--ERV-E--QQ-K--L-----QR--L--L-P---. .F--M-NAD---QY--KLN. .L-K-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---. .Y--M-NAD---QYI-KLN. .L-K-K---Q---Q-R-	971
LST.CD.88.524	K-I-----P--V--KE-Q-VT--IG---T-I-----AK--V--S--ERV-E--QQ-K--L-----QR--L--LTP---. .F--M-NA---QY--KLN. .L-K-K---Q---Q---	970
LST.CD.88.524	K-I-----P--V--KE-Q-VT--IG---T-I-----AK--V--S--ERV-E--QQ-K--L-----QR--L--LTP---. .F--M-NA---QY--KLN. .L-K-K---Q---Q---	970
LST.KE. .lho7	K-I-----P--V--KE-Q-VT--IG---T-I-----AK--V--S--ERV-E--QQ-K--L-----QR--L--LTP---. .F--M-NA---QY--KLN. .L-K-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-----Q--L-----E--PG-. .IVD--T-DIL-TK-R-N-. .Q-----QQ---	174
MND-2.CM.00.MND2S6	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---	174
MND-2.CM.01.MND2S109	KH-----P--V--EK--TV---AQ---T-I-----AK-HH--K--E--V--Q-----Q--L-----E--PG-. .IVD--T-DIL-TK-R-N-. .Q-----QQ---	174
MND-2.CM.98.CM16	KH-----P--V--EK--TV---AQ---T-I-----AK-HH--K--E--V--Q-----Q--L-----E--PG-. .IVD--T-DIL-TK-R-N-. .Q-----QQ---	963
MND-2.CM.99.MND254	KH-----P--V--EK--TV---AQ---T-I-----AK-HH--K--E--V--Q-----Q--L-----E--PG-. .IVD--T-DIL-TK-R-N-. .Q-----QQ---	174
MND-2.GA. .M14	KH-----P--V--EK--TV---AQ---T-I-----AK-HH--K--E--V--Q-----Q--L-----E--PG-. .IVD--T-DIL-TK-R-N-. .Q-----QQ---	964
MON.CM.99.L1	--I-----P--V--EFE-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---EF-XAA--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--AEFE-A--AS---T-----NK-----A-HK--EEVTV---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--EEVTV--	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-. .-----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-. .-----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-----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--EQ-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. .L-K-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---QLQNTTS. .P-FS-----Q.GKNE---	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---L--C-----QTS---. .M-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



PLV proteins

Accession	Protein Name	Residue Range
SAB.SN..SAB1C	PAKLIWKEGEAVVIEQEQG.ELKTIIPRRKAKIIKDYGK...ALDSQAPLEGNGRTAGEVDS	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--KVNTQNSNE\$	1006
DEB.CM.99.CM5	--T-L-----CKTEVGDI-VV-----V---G...V--KTDTONPSE\$	1006
DEN.CD..CD1	--E-L-----KT-AGDIFPV-----P-NA.....ESRKGEQNKGMSETDI\$	1030
DRL...FAO	--E-V-----K-GT.D-V-----N.VDSHTNME\$	1011
GRV.ET..GRI_677	--R-----K-GE.DI-VV-----ER.KTM--EGSM--VREANKQMEGDSLDQDE\$	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-VGSAHTSNDREEGRMAD\$	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-VGSKSDTGLRKEDRLDN\$	1010
RCM.GA..GAB1	-----ETGD--VV-----E--RK.DVDSEANLAGRQEEN.\$	1017
RCM.NG..NG411	-----E-----D-E.D-VV-----C-----RKEVDRETNMEGRQEE\$	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---PTLDKK.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--Q..KSHLEIQ--N.-----SSH.....	86
COL.CM.-.COLCGU1-M-A--I.CYPMG-RKPK..PLRE-I...KQF...PE-QYYLGTKTES..PE-...TEGKIEFY-G.SAIKL..SII-LGT--DRYHRGPPV.....	80
DEB.CM40T---LTHKIPMNKLD..WCSAIKR.E--KT...ELDVTQ.WIH-F..ELR-SY-TQTKI-F-IT...KEGG-VVELLWC-----P-M-V.....	86
DEB.CM5E---LTHKIPVNKLDK..WCSAIKR.G--KT...ELQAIK.WIH-Y..ELRRAP-TQTKI-F-VT...EEG--VVEIFWS-A-----MPSM-V.....	86
DEN.CD.-.CD1	..QNLGRENR.....T-E---LTYKIGETKLEHWMSTLKSIIYKTN.....ELKGTVIIPHR..-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..QLQHRFTQNKIRLAMDRKVGEE.VEATY.IEIT.IL-DT-S.HGPASLSRSTYYQAVI	103
GSN.CM.99.CM71ERQRQ-M-V-Q...IQVM-ER..KV-.LIRATK-HI...WSGKTPFV-IH-Y..QLQHRFTQNKIRLA-DTRQVGE.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..DLQKVCYVPH..-VG-A..WTCSRVI--Q..EGSHLEVQG---S-----S.....	86
MND-1.GA.-.MNDGB1ERVER--L...T--SSQRI..-K-HW--RRQ-A.....WATANNEEG.CW--YPHFMAWNEWYTCSKVVIIN	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-----KYKEE-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..DLQKVCYVPH..-VG-A..WTCSRVI--Q..EESQLEVQG--NL.---R---S---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.VGPASCSKSTWWKQSPFI	102
RCM.GA.-.GAB1A--M-----I-R-DRRK..I-Q-H-----QYKKG-AAKE-EYVPHF..-VP-G-.W-HS-VH---E.E-SK.LKITT--NL.-V-----GTYG.....	87
RCM.GB1c1A--R--E---I---DRRK..V-Q-H-----QYKKG-EAKD-EYVPHF..-VP-G-.W-RS-VH---G.NNTK.IKVTT--SL.-I-----GTYG.....	0
RCM.GB1c3A--R--E---I---DRRK..V-Q-H-----QYKKG-EAKD-EYVPHF..-VP-G-.W-RS-VH---G.NNTK.IKVTT--SL.-I-----GTYG.....	0
RCM.NG.-.RCMNG411A--R--E---I---DRRK..V-Q-H-----QYKKG-EAKD-EYVPHF..-VP-G-.W-RS-VH---G.NNTK.IKVTT--SL.-I-----GTYG.....	87
SMM.FGBEE--N--V...T--IP-R..L-K-H--I--LKFRT..DLQKACYVPH..-VG-A..WTCSRVI--R..DESHLEVQG--N.-A-----S---S.....	86
SMM.SL.92.SL92BEEG-R-VAV...V-RISRRI..V--H-CI-F-KYKT.RELEKACYVPH..-VG-A..TASRVIF--E..EGSHLEVQV--N.-----SS---S.....	87
SMM.US.-.SIVSMMH9EE--N--V...T-RIP-R..L-K-H--I--HLKYNT..DLQKACYVPH..-VG-A..WTCSRVI--R..DEAHLEVQG--N.-A-----X---S---S.....	86
SMM.US.-.SMMPGMEE--N--V...T--IP-R..L-K-H--I--LKFRT..DLQKACYVPH..-VG-A..WTCSRVI--R..DESHLEVQG--N.-A-----S---S.....	86
STM.US.-.STMEE--R--V...T-RIP-R..L--H--I--HLKYNT..ELSKACYVPH..-VG-A..WTCSRVI--Q..EAHLEVQG--N.-----SE---S.....	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.-.SYK173E--V-TWRMTPRQIDR..L-HIIKT.H--KS...-ELEKAT..KH-Y.Q-E-Q-.T-CQ-T-VG...DGTIWIIF-HNLA.--R--HMQGI.....	85
SYK-KE51E--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---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---S---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---S.....	85
VER.KE.-.AGM_VER_9063	-N.....Q--E-VM-V...T---S-D..LITK-QGI---W..T.-RNLK-QYMMHY..Q-N-A--TMCRYI--IG..LEGEIQIDC---R---S---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---S.....	85

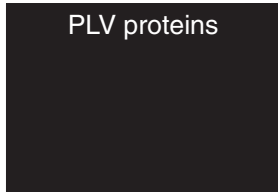
PLV protein alignment: VIF

	phosphorylation site Thr96	 Cys114	involved in inhibition of APOBEC 3G Cys133																							
SAB.SN.-.SAB1CTGIGYSKGEWFTELDPWTADHIIHWSYFPCFTDRAVQQAIRGEKYLWCKHQVGHQPTGQVPSLQYLALRVYTNGLLRR...VAPTSRRGS	171																			
H1B.FR.83.HXB2LGQGVSEWRKRYSLQVELQLLYDSSSIRKLLHIVSPREYAN.....KGDAALI.....TP	156																		
CPZ.CD.-.CPZANTLGHGVSEWRCKYIQVEQMQYSQRLRITYHYKKSGTFCKILEFRGYP	154								
CPZ.GA.-.CPZGABLGHGVSQWRLSYVQVFRLSQDAETIRRLQLVAPREYKERGFKALISE	157							
H2A.DE.-.BENVRLTWYTEKFWDVTDCILSTSAGERRLSCNYPQAKAVVQQNDPQRKGTARKQWR	173								
COL.CM.-.COLCGU1SYELQIGKWEDSWQQVGYMIIRGKVRIEEEARLPWNPDFLLKECLALREKER	153											
DEB.CM40AIAWETKTKWTDLLLRSQQRILQQQQYRGPEGKGPIAVKHGKVPTLTPPHLKN	175									
DEB.CM5GISWEQTWQTDLLLRSQQTLKLQQQQYRGPEGKGPTAVKYGKVTSTLTPPHIKN	175								
DEN.CD.-.CD1HLSTGWAVVEKDHSRQRLVVSTARKKAYHPRTGPGREPIVQAIRYGVVTTITPTSISFKN	189									
DRL.-.-.FAOTEHIQEKFCLITRLEQIRGVLLYRIQAYFPRRGGQAHIKYKHGRKTVSGESSG	172											
GRV.ET.-.GRI_677HGVLSYHQKGYKVGRMLYNIQLVAYTNIRKSKR	165											
GSN.CM.99.CM166LGAREGDIWYSNTGVMQGKQNEDIRRQIGNEPEALQVTPQTEQFFAVKHYGNPTQSSYALAOQ	203			
GSN.CM.99.CM71LGAREDTWYSNTDVMQTRQNEDRRQLGYDPEAPQVTPHTERFFAVRYGKNPTKSSTLDQP	203	
LST.CD.88.SIVlhoest447VSEWIYNYQIILNQCKNSTDIRKLWIMAFAMPTRGCLFCQLQHVQNKAAKESRKTSL	170			
LST.CD.88.SIVlhoest485VSEWIYNYQIILQMCKNSTDIRKLWIMAIAMPTRGCLFCQLQHVQSKAAKESRKTSL	170			
LST.CD.88.SIVlhoest524VSEWIYNYQVILQMLVKDWVHDIVRIHSMPTRGCLFCQLQHVQTKTAAQPKAPL	170			
LST.KE.-.SIVlhoestVSEWVYRDYQIILQMYKNIQDIRRLLRITMCQPLKGCLFICQLQHVQAEAAKKSRETTL	169				
MAC.US.-.MM239VRTWYSKRFWDVTNYILLSTAGERRQLSCRFPPRAKYKVSDVRSQGENWKQWR	172								
MND-1.GA.-.MNDGB1RDIRLIVRSYWHLQIEVGLCLSTYAVSIEAVVRPPPFEKCITEVLLHFDYMSMKEVKVRFPAKAQGLFCIYGPEESASIRRSRTSQR	183		
MND-2.-.-.5440TVEQENYYIVRGEQIRKLFRLVAYFPWRGGTFQALRRKHGREGTRGASR	173						
MND-2.CM.98.98CM16TEKPTNYHIQRMKLEQIRLLKRLTVYFHWRGKGLSAYCNRGRRGPRDPSR	172						
MND-2.GA.-.MND14CGTAENDYYHIVRMEAARKVLYRIVAYSPWKGGTLAIKFCNGRKSTRGTTGR	172						
MNE.US.-.MNE027VRTWYSRNFDVTDYILLSTAGERRQLSCRFPPRAKNVSYVRSQRENPTWKQWRD	174										
MON.CM.-.SIVmon_99CMCML1IEWVYWMQKVERVPLDPVWNPNPRDREVRTWYAHTDLNQVTHAQQLDRRLGQQLATYKVTPLTEKFLALVRYGATPTKSSAPLDDLDP	215			
MON.NG.-.NG1IEWVYWRQPRVPQQEGVDRVWFPSRRPDREIDRVWYAYTDVQTRSQASDRRAXGNHIQTYNKVTPTEKFLACPKYGKCTTEPSSLDLPR	221				
MUS.CM.-.SIVmus_01CM1085LEWVYLRRHPQDREQDIYWYSYSTLMRTQSYSQDIPRVIHQHTLGDHEHPDAHTKVPPVLALKKHGKLAQPPPSALALQ	202										
RCM.GA.-.GAB1VRTWYSRKFDVTDYTLLGTSEGERRLSCFPKAKNGTFIQAWAKSQLDRYGRKPP	172						
RCM.GB1c1VRTWYSRKFDVTDVQLASTAAHRQVSYGYAAHSSQLKVLQNDPKGNRKQWR	67			
RCM.GB1c3VRTWYSRNFDATDYTLLGTSEGERRLSCFPKAKNGTFIQAWAKSQLDRYGRKPP	67						
RCM.NG.-.RCMNG411VRTWYSRKFDVTDYTLLGTSEGERRLSCFPKAKNGIKAWARQQLKDGGRGQRKAP	178								
SMM.FGBVRTWYSRKFDVTDYTLLGTSEGERRLSCFPKAKNTVSHVSQREDWKQWR	172								
SMM.SL.92.SL92BVRTWYSRKFDVTDVQLASTAAHRQVSYGYAAHSSQLKVLQNDPKGNRKQWR	174			
SMM.US.-.SIVSMMH9VRTWYSRNFDATDYTLLGTSEGERRLSCFPKAKNTVSHVXSQEXWKQWR	172								
SMM.US.-.SMMPGMVRTWYSRKFDVTDYTLLGTSEGERRLSCFPKAKNTVSHVSQREDWKQWR	172								
STM.US.-.STMVRTWYTRNFSDVTDCQLLGTAGERRLSCRFTKAKNKVEHVSQRENTARKQWR	172									
SUN.GA.98.SIVSUNCSMEWIYNSYQILVQMFKDWSCIRRMLIHERNAKGLLFCLHQGERASRTARA	186
SYK.KE.-.SYK173RIQYQWQWDTAVRLNFANVLLTSHWTPTDQLKDGGGFLOSLACANT	170									
SYK-KE51GLIWAHPSWVTQALVSQSLLLTSHWNFKELLLLLLLLLLLL	171
TAN.UG.-.TAN1VAISFENWQTYKVTDVMCHISFTYKEVAENHWRFTDFNGRRKRFGQKTR	179								
VER.DE.-.AGM3EGIQYLSNRDPWYRASLTHTEIRKLLQRFTFQFPEKKLAHQNGLRQRSQRSGTGR	173											
VER.KE.-.AGM155VGIQYLSLVNDKYRNSMCHTILNRFIFQFPGKLLAHQNGLRQRSQRSGTGR	173												
VER.KE.-.AGM_VER_9063IGIQYLAVGNFRAGMLHSEIKRFVPQFPEKTKLAHQNGLRQRSQRSGTGR	172								
VER.KE.-.TYO1VGIQYVSNLESKYRASGHNKEILHRFVPQFPEKSLAHQNGLRQRSQRSGTGR	173											



		Vif end \	phosphorylated
SAB.SN.-.SAB1C	SQGSPPQESQRRDTRMARNMGFAQRAVRRRMAPRHVTGPQFRGPVPLPKESFPFSLVEYCGRTS.....H\$	Thr188	234
H1B.FR.83.HXB2	KKIK.....P---SVTKLTEDRWNKPKQTKGHRGSHTMNG--		192
CPZ.CD.-.CPZANT	KGPR.....RQF-SL-ILTEDRWNKP-RMRGHRNQ.....		185
CPZ.GA.-.CPZGAB	RRHR.....P---SVAKLTEDRWNKHQRTKVHQENLTRNG--		193
H2A.DE.-.BEN	RDHWRGLRVA-EDHRSL.....KQGSSEPSA-RAHFPPGAVKVL-IL.....A-		215
COL.CM.-.COLCGU1	THQKMRVLKG.....TS...HGCRTR-		171
DEB.CM40	NTHDKARVAAGHA-SK-G.....SPKTFHQRRTIWSLESCLRLTRPRLDGRISM-CIG.....---		230
DEB.CM5	-TNDKASTTSHAGSK-G.....SFKTLHKRRVMGSLESCLR-SCPTRLDRRTSM-GV.....N-		230
DEN.CD.-.CD1	GRNHSKAATGSSMGYH-GGKTLQQR-DHWNMGSLLEVI..TSA-TLDRRTSHGSRRRRHFKNHAE...RNMEAL-		258
DRL.-.-.FAO	AN--RRGRV--MATEES.....LKNQPGGSITLP-RVPLPSLEH-HGAL.....P-		217
GRV.ET.-.GRI_677	TFTRMAGNLGSRQGAMGR-ATRHAQGSKRRSQKALWNE..HAN-SMELLRCRGKE.....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.....KRGQTPFPRSNFPGPSLRI-CQOYRVRGESLHEIHYH.SK-		227
LST.CD.88.SIVlhoest485	WTTKRAMGSMVSRHHGG.....K-GGQTPFPRSHFPGPSLGI-CQOYRVGGSLHEIHYH.SK-		227
LST.CD.88.SIVlhoest524	WSSKRAMGSMASRRHHGG.....Q-RGKTPFPRSNFPGPSLGI-CQOYRVRRGGLHEIHHF.S.-		226
LST.KE.-.SIVlhoest	WTTKGAMGSVSRHHGG.....KRGKTPFPRSNVPGPSLAI-CRQYRVRRGGLHEIHHF.SQ-		226
MAC.US.-.MM239	RDNRRGLRMAKQNSRGRD.....KQRGGKP-TKGANFPLAKVLGIL.....A-		214
MND-1.GA.-.MNDGB1	AIQ-VASRYGGNKGRSK.....KALPSHYPKCSIRI-CAKHR.....E-		222
MND-2.-.-.5440	RGRTGAMAGNVTGENQPG.....GPVFLP-RVFPFSLHMCRTL.....A-		213
MND-2.CM.98.98CM16	RG-KYLVGKISGENQSG.....KGITLP-RVFPFSLERMRTL.....A-		212
MND-2.GA.-.MND14CG	RS-TRTVAGKIIIGTSEQR.....GSITLP-RVFPFSLHMCRTL.....A-		212
MNE.US.-.MNE027	NRR-LRMAKQNSRGRDQQR.....GSKP-TKGADFPGLAKVLGIL.....A-		214
MON.CM.-.SIVmon_99CMCML1	GSHRKTGPTG-HLGAK-R.....SPKTLKGRAPWHLGTGNGAA.....C-		255
MON.NG.-.NG1	GSHRARR-AG-HMGPK-G.....GPXALLKGRAPWHLGSG.....---		257
MUS.CM.-.SIVmus_01CM1085	GPNNHAPGT-SHVGSK-R.....G-KTLFQRRAPWNL.....E-		235
RCM.GA.-.GAB1	HW-WRSRVPALA-GH--KGO.....LGSQVTLSSRVHFPVSAH-CGTL.....A-		216
RCM.GB1c1	HW-WRSRVPALA-GH--K.....GQPGSQVTLSSRVHFPVSAH-CGTL.....A-		111
RCM.GB1c3	HW-WRSRVPALA-GH--K.....GQPGSQVTLSSRVHFPVSAH-CGTL.....A-		111
RCM.NG.-.RCMNG411	HW-WRCRVPLASQN-NR.....SQLRSQVALSSRVHFPVSAH-CGTL.....A-		222
SMM.FGB	RNNRRGIRMA-KNSRRD.....KQGSSES-AEGANFPLAKVLGIL.....A-		214
SMM.SL.92.SL92B	RNNRRGLRMVTQYSGGP.....ESNSSTP-TKRVNFPGLEKVLGIL.....GR-		217
SMM.US.-.SIVSMMH9	RNXRRGRLATQNXRRN.....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-CGLYRVKGRVYESNHFAQ-		241
SYK.KE.-.SYK173	MVLHSHKCRVDPK-DQ.....CHCKGRTGSDRSIQAFYS-RNIWSL-SILKRRGRD.....-		221
SYK-KE51	TTILHSHKCKQLGRGRSNK.....CHCQSRGTGSPRSMQAFYSGRNIWSL-SML-GGSRSRRNTHHGLD-		233
TAN.UG.-.TAN1	MVRNLGSQ-GAVG--IKRH.....SRTQSGS-T-F.WERT---SMELLSGRRGK.EWGTNDR...KGL...-		238
VER.DE.-.AGM3	GTRNMGFE-GAVG---KRHA.....RYQSGSODAF.WARA-V-SMELLSGGGRK.ESH-HAR...KGL...-		232
VER.KE.-.AGM155	RTNRLGSQ-GAVG---QRY.....RNQQRSTAF.WPRT-I-SMELLSGGRGE.TGKTHSG...KGI...-		232
VER.KE.-.AGM_VER_9063	RSRDLGSK--TMG---KGHV.....RSQSGSQITF.WQRA--SLELLSGRRE.TGSPNDG...EGL...-		231
VER.KE.-.TYO1	RSRNLGSK-GAVGQ--KRYV.....TRSQPGGEAAF.WERT-V-SMELLSGGRRK.TWY-HDG...KGLQIL-		235

PLV protein alignment: VIF



PLV proteins



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-----E-----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.APQAGEVGLAQ---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

630

PLV protein alignment: VPX



		frameshift in HXB2	
SAB.SN.-.SAB1C	MASGGWLPPVGGDPPKDPKPNREEIPGW..LETWDLPREPFDEWLRDMLQDLNSEAQCHF.PRNLLFRLWVNIVEEPAI...DHGQTRLEGV.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-DQGPQ---YQ--ALET-EE-KN--VR--..-PW-HQ-GQF-YDQTYGD.....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-FHREAERV...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.....IPPSHPMP...WLSRRVPTTMAIAQNA-WEI-E-EK--.S-DE-RGI-H.D-T-LPADPDWTV\$DQAAIACAI\$D-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\$DQAAIACAI\$D-V-RV-TL--R-.FRDG-FHRYNRIV...RR-P		101
LST.CD.88.447	-SRTR.....QQ-SPEKPP...DFGP---WGQ--A-TMEEIKE--RR--..-ETI-AQVSEYC-SSTGS.....E\$EKAC.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.AAR..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-PPENEGP.....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---Y\$NQ--A-TMEEIKE--RR--..-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---Y\$N---I-T-AEQE--LK--DR-L-.HAVGSW-YKTYGD.....T---V.RQLIS-----SM-.YRYG-RESR.IGQ.....E-		85
MND-2.GA.-.M14EQPPEDEAP.....Q---Y\$N---I-T-AEQE--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...HPPHWTSRVVPIERAA-QAAIW--E-LX--.SKEE-RG.I-DQ-T-LPADPGW\$NADQAWAAXAID-T-RV-TXI-R-.RDG-YHR.HARV...XRR-P		99
MUS.CM.01.1085	-R.....QAQ\$AMWEV-E-EK--.S-EE-RG.I-E-T-LPADPDWTV\$DQAAIACAI\$D-V-RV-TLV-R-.FRDG-YHRYNRV\$MRRYPTIRP		85
RCM.GA.-.GAB1ELPPEDEGP.....Q---Y---M-T-IE-QE--KK--TYAL..TQIGDYVY-QHGD.....SI--V.QAMI-L-R---L-.FRNG-AGSR.IG.....S.		84
RCM.NG.-.NG411EMLPEDEEP.....Q---Y---M-T-IEIQE--KK--T-EL..TQEGNY-Y-QHGD.....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	---R.....RREPVEQPP...DEGP---AQ--A-TMEEIRD--ER--..AAIVAQVLQYCADSTGS.....EK--C.MRAITL-NR---I-.LP.M-PSYIRIRSGSGNVNPR		98
SYK.KE.-.KE51	--QAF.....FRVYQRQ-G-PLFYIP...RNVNWD-AGVISATAKAEQVAQ--CK--.TAE\$EI-GV-NQCL-VE-GPEETPTMAWIRCM.LDMH-A-NFM--E-.FAAG-PQRT.KYA.....RH.		108
SYK.KE.-.SYK173	--EAF.....FNPSQHVG\$TPWFFIP...RNVELTP.NVINVTVKAE\$LVVT--SK--.TPQEIY.GV--Q\$LNBEAGTDSPTMAWERTM.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-----V\$EEI-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...NRTM-RA.---YKLV-----FRCG-RRRQ.-FE.....P-		99
VER.KE.-.TY01	---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	VIRPLRGTAAPPDSNSVPHADPEQPRRPSRYRMDE-	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	GKYNPLRSFPRPNNPL.....-	101
MNE.US.-.MNE027	G-.NPLSTIPPSR-ML.....-	101
MON.CM.99.L1	MRGTA-G-TSSVPQADPDN...PRRPSRYRMDE-	120
MON.NG.-.NG1	XLRPLRGTA-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.NPLGSI PSARDVL.....-	100
SMM.US.-.H9	GX.NPLSTIPPXRGVL.....-	101
SMM.US.-.PGM53	G-.NPLSTIPPSRGVL.....-	101
STM.US.-.STM	G-.NPLATIPPTRGVL.....-	101
SUN.GA.98.L14	-PTRPR-GQGDIRR-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	exon 1 \ / exon 2	
SAB.SN.-.SAB1C	MD.....	QEQEARPQVWEELQEBELHRPLQACDNTFCCKVCCFHCILCFHKKALGIRYVPRP...	RRASKKIS.....	HNQVSLHKSIST.....	WT....RDSQTKKK 87
H1B.FR.83.HXB2	-EPVDP.....	RL-PWKHPGSQ-KT--	TN-Y--K----	QV--IT-----S-GRKKRR..	Q-RAHQ...NSQT-QASLSKQPT-Q..PRGDP....GPKE\$ 86
CPZ.CD.-.ANT	--PVDA.....	ETPPWLHPPAT-ATP-	N-Y--C--Y--P--T--G---S-GRK-RAR..	NRRRTAE..SSENNQDPVSK\$-LPK..	TSRIQS....SOKK\$ 87
CPZ.GA.-.CPZGAB	--PIDP.....	DL-PWKHPGSQ-RTV-	N-Y--A--Y--Y--T--G---S-GRKKRTT..	RTAPA...GSKNNQDSIPKQPL-Q..	SRGNKE....GSEKST-E 91
H2A.DE.-.BEN	-ETPLKAPESSLKPYNEPSSCTSERDVTQAQLAKQG--	LAQ-----	EP-T-K-Y--R-S--Q---S-G---S-ERKGRR..	TPR.....TKTPSPSAPD----	RTGD..SQPTKEQKKTSEAT 121
COL.CM.-.CGU1	SLSLYAK--E--	NDK-W--A----	Q---LS-G-----	KKNK...SATQ.....TQ-VAKTGSQNSQ-TQKKYTESK-TA 73
DEB.CM.99.CM40	-E.....	SIDPFEEERANTPH....	PA-M--A-IL--Q---IQ-G---S-GPR-RKRKC-REYTPAAESNQPDPDFVPEQPLPQ..	RRRK.CFDSQKKEAEVEET	92
DEB.CM.99.CM5	-E.....	EIDPFKETANTPH....	PN-M--A-TM--Q---MQ-G---Y-GSR-RKR..	-G-SFAANDKDHQEFIPE..	HCVEDIPLCRPLPQRKWKCFNSEE- 91
DEN.CD.-.CD1	NADSIDPFAGNKTP-...	Y--H-AY--Q---LQ-G---V-GRAPRK..	PRARQS...GTPDQEAQVSKQ-L-DPRRQOI....	SQKKE-E- 83
DRL.-.-.FAO	ARKVDLDQQDAGT.HF-PFQGRNTN-...	W--Y-AY--Q---LQ-G---V-N-S--R..	K-TATV..VSRVSETGDTFVARHTAQRRTTTDSQTEKKEKS-TE	102
GRV.ET.-.GRI_677	DKEEHPHL--D-----	P-T-K-Y--K--Y--E--LQ-G---V-H.VSRK..	KTSTQDNQD.....	PIRQQ---VQ.RNGQ-TEEGKTEVE-AA 91
GSN.CM.99.CN166	PILSSLDGPDQDIAMNRDPPECLLPNW-QPGAA-ATP-S.A-Y--K-AL--Q---LR-G--LSLHGRSRKRK-PKATEPD.PSVSSNQDSA\$VQ-L-HKQWNAVR....	QT-P--A	110	
GSN.CM.99.CN71	LILSSPEDLQDTAMNRDPPECLLPKW-QPGAA-NTP-S.A-Y--R-A--Q---LR-G--LSLHGRSRKRK-SEASEPA.PSVPTNQDPIQIQ-LPHSKWNAIR....	QTEQE-T	111	
LST.CD.88.447	QQEQEQQLTRQKHQDPLK-TYK-AVTK--S--N-K-C--K-----	QQ-G-----H-RR..	K-VAQVF....	EDK-DPVDGKER-KNSQTKKER\$ 95
LST.CD.88.485	QQEQEQQLTRQKHQDPLK-TYK-AVTK--S--N-K-C--K-----	QQ-G-----H-RR..	VTQVF....	EDK-DPVDGKER-KNSQTKKER\$ 95
LST.CD.88.524	QQEQEQHT-QKHLD-LE-IYK-AITD--S--Q-K-H-----	QQ-G-----H-TR..	VTYIL....	ETE-DPVDTERAKNSQTKKERQKTEKSTST 106
LST.KE.-.lho7	QQEQEQPI-QKHQDLMK-TYN-AVTKA--Q---K--K-----	QQ-G--LH--H-KR..	VTSIL....	ETQ-DPVDKRR-KNSQTEKRRKEQIQLRSTST 106
MAC.US.-.239	-ETPLREQENSLSSNERSSCTSEADASTPESANLG--	ILSQ-Y--E--Y--Y--K--Y--QF--L--G---C-EQS-KR..	TP--...	AKANTSSASN-P--N.RTRH..	CQPEKAKKETVE-A 121
MND-1.GA.-.MNDG1	-E.....	PSGKEDHNCPP-DSGQEEIDYKQ-L--YYQ----	E-K-W--K----M--Q-G-----	H-Y-KR..	VPGTN-K...IPGSGEAIRRA-DLSFHRTASRTYTANGQT-E-- 111
MND-2.-.-.5440	VGEVASDKKEEDITHFDPRFRARTTP--	K-----Y--Q---LQ---VH-H-Y-QR..	PRRGLL..	EKVPSNS-AA-GRLSPAPQPOTARQRRR-ARLR-T. 102
MND-2.CM.98.CM16	AGKAVSDKKEGDVTPYDPPFRDRTPP--	D--Y--R--Y--Q---LQ---VH-HAY-NR..	SRQRL..	EKISEDAAVGRSLSPSPQPOTARQRRR-ARLR-T. 102
MND-2.GA.-.M14	AGKAELDKKEANITPLDPPQGRITP--	K-Y--K--Y--QC--LQ---H-H-Y-TR..	TRKRL..	GKISQDSEAAVGRSLSP-SQPOTARRRRR-AQLRRT. 102
MNE.US.-.MNE027	-ETPLREQENSLKSSNGRSSCTSEAAAATLESANLE--	ILSQ-Y--E--Y--Y--K--Y--QF--L--G---C-EQS-RR..	TP--...	VKANTSSASN-P--D.RTKH..	CQPKKEQKQKQKTEVETA 121
MON.CM.99.L1	-E.....	PVDPDLPE-HPPAT-RTP-T.N--R-----	QI--LT-G---S-GRK-KR..	TS...PVPGLSSSKNPARKQ-L-HPTRDSQRPTQQAQAVA-AAT	97
MON.NG.-.NG1	NSAXIVNPAALPE-HPSAT--TP--	N---I-AL-XQV--LQ-G---S-GRG-KR..	GPRH..PTIGDRSYQDNPGV--	LPKSLGNQERQKQKQKAVES-AG 101
MUS.CM.01.1085	PSVEELPKE-RPGAA-ATP-T.S---K-A--V--Q-----	S-GRRFR...KR-RA..	DSSLPS-SDPVQI--L-DRRRDQKRQKQKQKTEVETA	95
RCM.GA.-.GAB1	VQGVGLEH-BEVI--YDPPFRKRETS--	Y--K--Y--Q---LQ-G---N-.AS-AR..	R--E...ENKADKFPVPNQ--	STTRGNRKLQEKKEKTEVE--V 98
RCM.NG.-.NG411	VKGAGLEEREEI--LDPPFTRETS--	S-Y--R-VY--Q---LQ---N-.AS-RR..	GRPGNK..KKNQINPFDTSEQ-L-AIGRNSQSKKEKKEAVEKA	102
SMM.SL.92.SL92B	--API.....	QGLDSQGDQIPWD-Y--ET-Y-S---K---Q---LR-G--T-AK--RR..	VKKKA.....	ANPDTSNQ---S.GTRN..	CQPKKEQKQKQKTEVETE 95
SMM.US.-.H9	-ETPLKEQESSLESSREHSSSISEVDADTPESASLEX--	ILSQ-Y-XVE--Y-KXY--K--Y--QH--X--G---C-EQO-RR..	TPKKT.....	ANTFSASD--LXX..	RARN..CQPKKEKQKQKTEVETE 119
SMM.US.-.PGM53	-ETPLKEQENSLGSSNERSSYTSEEDVLTPEAK-E--	ILCQ-Y--E--S---Y--K--Y--Q---L--G---C-EQO-RR..	TP--...	AKANTFSAPNE-L--	RARN..SQPKKEKQKQKTEVETE 121
STM.US.-.STM	-ETPLKEQESSLSSEPSSCTSEAVAATPGLANQE--	ILWQ-Y--EE-C-K---K--Y--Q---VT-G---T-ERS-RR..	VKKKA.....	TYPISASN-L--	RARN..SQPKKEQKQKTEVETE 119
SUN.GA.98.L14	STQGH-QDQD-GKGT-E-AYKTN-E--	K-W-RR----Q---LQ-G---H--Y-GR..	QHQ-IS.....	KDK-EAVRRKPAPNPQTKKKKRQR.KKESTST 98
SYK.KE.-.KE51	-S.....	TEDKVR-TQGI-TSFL-GTFLSNG-QTP--	SR-Y--K--Y--YR--LQ-G---T-VRE-KR..	RHARTTAEDFAA-SSTT..	R-R-PLPSTTRPQGGQKKE\$ 98
SYK.KE.-.SYK173	-S.....	STDQICQTORV-PSFL-GTFLEKG-PTP--	K---N--Y--Q---LQ-G---T-AR--KR..	A-RSI.....	SEDDSAPTGTLPFRAGRTQANPQTKKAVE-E-D 100
TAN.UG.-.TAN1	-E.....	SEGDMAS-LQD----	TP-T-K---R--Y--QV--LQ-G---T-H-S-I...PK-NH.....	N-QNLVSSQ---A.....	GG..NSQTTQEE- 87
VER.DE.-.AGM3	KGEDEQCAHYQD-I-Q-KA--KR-T-K-Y--C--Y--Q---LQ-G--VT-HA--I...KKIAPL....	DRFPEQKQ---R.GRDSQ-TQK.GQEKVETS	95	
VER.KE.-.9063	KGE-E-TVLHQD-IRQYK--TT-R-K---K--Y--Q---LQ-G--VT-HA--T...KKSVPQ....	NRLSQDQ---R.GRDGQATQE.SQKQVERE	95	
VER.KE.-.AGM155	KGE-DQDVSHQD-IKQYRK--ET-T-K---K--Y--QF--LR-G--T-HAF-T...KKIASA....	DRIPVPQQ---IR.GRDSQ-TQE.SQKQVEEQ	95	
VER.KE.-.TYO1	KGEABQIVSHQD-S-DYQK--T-K-K---K--Y--Q---LQ-G--VT-HA--T...KKIRSL....	NLAPLQ-Q---KWGRDQ-TPT.SQEKVETT	96	

PLV protein alignment: TAT



PLV proteins

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..TATGSSGGNIREDS\$	99
DRL.-.-.FAO	ASGTSDQNPPSQDPPKLGAKKQ\$	124
GRV.ET.-.GRI_677	AAN\$	94
GSN.CM.99.CN166	LARKARSGP\$	119
GSN.CM.99.CN71	VARKATSGPDISQDSRVSCGTT\$	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...QSRIWEALEQRANQRLAEDFKGLRVABEESTSSN\$	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-PSLGRKNLAQOSGRATGASD\$	119
VER.KE.-.9063	TTT-QILGRKDLERDKREAVGANAS\$	119
VER.KE.-.AGM155	A--NLRISRKNLGDETRGPVAGAGN\$	119
VER.KE.-.TYO1	AGSN\$	100

634

PLV protein alignment: TAT





	exon 1 \ / exon 2	
SAB.SN.-.SAB1C	M....SLG...QEELLRR...F.RIIKFLYT.TN.PYPP..GQGTARQ.RRRARQRWAKQRQVVIHLAERILETPVVSQIDHLAQEFQDQLVLDNLQPPSLPPGHPTENQNTANSSS\$	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.-.CPZGABAGRSEBQ.DDARLL...QAVK---I--Q.S-.-S...PE---KA--NR-R--RARQK-ISEISG-V-A-YLGRPPKPGDLELPELDKLSL-CVETTQDVG-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-NLRQEEELADSSVIGTQG-KPSDS\$	118
DEB.CM40AHAGGRGSA-ENT-QLLKVISL--I--Q.S-.-K.G-GS.-ST.--R--K-RRRQ--DSI-----QSHLGGSPPT-DVALPDLSQLHLAD\$	90
DEB.CM5AHAGGR.GDADQALLRMIRIISKLYQ...S-.-KGG-S.-ST.--R--K-RRRQ--DDI-Q---AHLGRHLTASDVALPEELRLAD\$	87
DEN.CD.-.CD1AGLQE..RDQE-VK...AV-L-K-K-Q.S-.-TPE-SKS-.-R-RK-RR-QA--VAIS-----YHLGRSLTA-DVELPDLEKLSVSD\$	84
DRL.-.-FAOA-PE..R-PPW..FQBYL-LVTR-.WQGR..I--.RDDQLPQTA-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-WDREVPVGTAAES-PE-NPASETAQSSS	118
GSN.CM.99.CM71ADPANGRDQK-QNLL..LAC-L-L-T-K.S-.-TAS-TRS-.-NKKR--RARQR--RQISH---SL-GRPEESGNLDPDLGHLSDS-EDRQVPGVAAAEGVPT-ETAGGSGQSSS	118
LST.CD.88.SIVlhoest447T-GE..K-SPRY.....LK-SKILW.-GLQVGEKREPR-----D-E-QR-HLH-LRTVQ--FHST-ERGLER-FTRLTVCDSPVEVDQGLWNSSPQCVAVPFVVARPFSDFLPTWAT	114
LST.CD.88.SIVlhoest485T-GE..E-LPRY.....LK-SKILW.-GLQVGERREPR-----D-E-QRRHLH-LRTVQ--FHST-ERGLER-FTRLTVCDSPVEVDQGIWNSSPQCVAVPFVVARPLSDPFLPTWAT	114
LST.CD.88.SIVlhoest524T-PE..G-LPTY.....LRLSRILW.-GLQVGERREPR-----D-E-QR-HLH-LRAVQ--FQAT-ERGLER-FTRLTVCDSPVEAVGNTAPHPRVAIPHLAGTYSDFLPPWAT	114
LST.KE.-.SIVlhoestT-NG..D-LPRY.....LRLSRILW.-GLQVGERREPR-----E-S-YRDYLH-LRAVQ--FQAT-ERGLER-FTRLAVSDSPEVAQGRGNTPEI-SVAEPQLAVAFVDFLPPKWAT	114
MAC.US.-.MM239NHER..E--RK.....L-L-HL-HQ.--T.GPGT.-N-.-QRKR--RRRW--LLA--D--YSF-DPPT-TPDLAI-QLQNLAIES.IPD-PTN-PEALCDPTEDSRSPQD\$	107
MND-1.GA.-.MNDGB1T-NV..YQ--I-----YLVVVVKKLY.EE.-I-Q-----RK-QLRTR-A-LRE-EG--KQILDGRPDQLCQGVNLLALAEKSES-N\$	83
MND-2.-.5440TDRG..DLDEGFW...RKYQA-VKQLW.EG.RLS-APQPQ-----R-A-LR-TEH--RE-QT--\$	61
MND-2.CM.98.98CM16PTETG..DLDRDFW...KQYQK-VQRLW.EG.RLS-SPQPQ-----K-R-A-LR-TEY-IRQ-QS--W-ALEQRANQRLA-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-TPDLAI-QLQNLAIES.IPD-PTNPIEILHDPTESPRSPQD\$	107
MON.CM.-.SIVmon_99CMCML1AGNGR..D--RRL..LSLALAAVRI-QE.S-.-IPR-.-NA.--NR-R--RR-QL-IASIS--FYHYLGRSQEPCPLDIPDLERLSISDLPD--ESVP-AA-PPAHTPAPTGVKPS\$	112
MON.NG.-.NG1AGAER..GAAVQDI.LRLAIGA-RIIRE.S-.-SPS-.-RSA.--NR-R--RARQE--XS-C--FNHYLDRLPRPNSLDLPAIEGLRLDLDLPE--ESAQPPDPSSPRPVPTASGQ\$	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--QQ-----LDSTDPPEQDLDAAIAD-QK--LNNLPE-PVDF\$	93
RCM.NG.-.RCMNG411L--EE..E-ADQEIR..RRI-L-HLHL.S-.-QSG.GT.-N-.-K-RQ-RRRWT-ILQ-----FLY-DTPP-RDPLQEALSN-QQ-TLSDLPE-PVNPFNSPSSFAVDRS\$	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-TPDLAI-QLQGLAIEK.LPN-PASAPEPLKDAES\$	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-AIFGSRTAALSDSL-QLQISD\$	77
SYK-KE51GRER...-DTQEQLLRTLLR-AQQ...LA---LPQ-.PSRA.--N-N-YRQLQA-RLYVQQ--F--IARRSAETLEQSFQELQITD\$	81
TAN.UG.-.TAN1S-D-----IIQ--RI--H.S-Q-----E-----R--FQQ-QR--AA-S--FIAQQRDPSPGSGESLAAAFDQLVLDNQ--VIETLPDPPQEPHD-SSTA\$	103
VER.DE.-.AGM3P-----S-----LL-L-A--NK.N-----V..E-----R--RQAE-LRA----WHSR-EEQLVQ-IDQLV-DQQH-AIQQLPD-PSSS\$	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.-.TYO1P-----P-----V-L-WL--S-----SG.E-----R--RQ-QD-IRV-V--LQ-QVYAVDRLEAEQHLAIQQ.....LPD-P-SA\$	84

PLV protein alignment: REV

635

H1B.FR.83.HXB2	..GSPQIL...VESPTVLESQTKES\$	116
CPZ.CD.-.CPZANTQSQHPLD.NNQNS\$	118
CPZ.GA.-.CPZGABGNYSILGKGAKNS\$	124
GRV.ET.-.GRI_677	CSSSGBGCBSLGTGTCYCPIRLSGSHQSKKSAARP\$	147
GSN.CM.99.CM166	QSDSLH\$	124
GSN.CM.99.CM71	LY\$	120
LST.CD.88.SIVlhoest447	SSESLLQQTGGKRSDECELELEDQKEQKRHISSCKKPTGTKEEMEIQIRKKNWVSQ\$	173
LST.CD.88.SIVlhoest485	SLESQQRMDGGKRSDECALELEDQKEQKRHISSCBRHQTGTKEKMEIQIQREKKNWVSQ\$	173
LST.CD.88.SIVlhoest524	SLANPQOIPGGECSENCELDQEQEQRRAIKNS\$	148
LST.KE.-.SIVlhoest	PLADQQQMDGGKRSDESLELAQEQEMQKQRTVIEH\$	148
SUN.GA.98.SIVSUN	PLADPQRLAGFAPYSYEQDQERVQNOQGESIIVSEGGK\$	150

PLV proteins



PLV proteins

	- 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.WEAU160	-...QS.-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.84.84ZR085	-...QS.-Q-L--VA-VL--I-.....F--.Y..RRIK...R-R-...-W-ID-----E-----E--K---T---HHA.PWD--DD..		82
H1F1.BE.93.VI850	-...SY.-LAIG-AA-----I-.....Y--.Y..-V...R-R--..NK-YK-----E--A---AA-G---PFI.PGDIN-....		81
H1G.SE.93.SB6165	-...QS.-V-L--V-----F-A-----F--.Y..REIR....-R..GK-D-----VT-----DFD.PWVG-....		81
H1H.CF.90.056	-...YI.-GLGIGALVVTFI-A.VI----Y--.Y..-V...R-....-IE-G-.....-K-M---HLN.LGY-AD....		80
H1J.SE.93.SE7887	-...I-.Q-A--VAF--IF--GM---Y--.Y..-R-....R-R-...-K-ID-----AD--R-PHD.LWN-ND....		81
H1K.CM.96.MP535	-...VS.-A-SIVALVVALILA..I-....Y--.Y..R-V....R-...-NW-ID-----E--A---AD.IG-L-HLI.LGNI-....		80
H1N.CM.95.YBF30	-...L.SLGFIA-GA-VSI..VI--ALLYR-.Y..-IK...L-E--KHIRQ.....E-----A-W-DGDEEWLVTLSSSKL-QGNWV-		83
H1O.BE.87.ANT70	-...HH.RDLL--IIISAL-FIN.VIL-GFILRK..Y..LEQK.EQDRKEREILE--R-....IRD--DY--N-EE-QEVM--LSHGFDPMPFE....P-		85
H1O.CM.91.MVP5180	-...HQ.ENLL-LIA-SALCLIN.VLI-LFNLRI..Y.LVQRKQDRRE-EIL..E--R--K-....IRD--DY--NEEQQEVME-IHSHGFANPMPFE....		85
H101_AE.TH.90.CM240	-...S--V-----S--A--V..I-....R-R-.....VK-----E--D--AK-....DFD.PWVG-....		80
H102_AG.NG.-.IBNG	-...Q-.T-T--V--V--F-A-----S--Y--.Y..R-IR.....D-----ST..M---YEY.I-DN-....		81
H103_AB.RU.97.KAL153_2	-...QSLAIAAIVALVVVGI.-AIVVGSIVF.IEYR-I-....R-R-.....ID-----E--Q-A...M---HLV.PWDA-D....		78
H104_cpx.CY.94.CY032	-...LF.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-BEADTYYLGSFGFANPVYREGDE.....		83
CPZ.CM.-.CAM3	-...LT.W-QIGLI-IGIEI-I-...-A-G-AFK-.W..R-GKEENRA.QKI.RT-IE--IS.....Q-----EDQRQLDREIHVYGFDPMPFD....W-		85
CPZ.CM.98.CAM5	-...LI.W-QIGLIA-GIE-IV-...-G-AYKK..W..REIKEERNRT.QRI.YN-AE--I.....Q-----N-EE-AQLDN-LHANGFDPMPFDW....P-		86
CPZ.GA.-.CPZGAB	-...L.-VGLVLILVGL.IAWN.-CI-GYIIKW.GY..RRYK..RHRLETEI.E--NLIL-.....N-EE--RLEQ-IHNY-HNNHFANPMPFD--		88
CPZ.GA.88.GAB2	-...LSMVAIGLI-IGTL-VIN-...-G---SV.Y..-RWK..RH-EEQRI.ID-IIK-T-.....EDK-TLAT-LHNNGFDPMPFE-R...I-		88
CPZ.TZ.-.TAN1	-IKIVVGSVSTNVIGILC-LLILIG.GLLIGI--RR.E.LERERQ...H-R.VL.E--A.-RLS.....IDSGVEEDEE...FNWNNFDPHNYNPRDW...I-		83
CPZ.US.85.CPZUS	-...LN.WFEIGLIA-GIEG--V.VII-GL-ARL..W..RQIKIKENT-QE-Q.NL-E--I.....E-----EE--TLAK-LSSLLELDNPRI.....V-		84
DEN.CD.-.CD1	-...L.KIKLGPPIEYVCL-FAVITWIAAIG--YLA.Y..RAYK....SYREELRYR-RLWSI.....DSGYESSQEDP.....		64
GSN.CM.99.CN166	-...H-AAVW-WGAAI-TFIY-CVALLALYLAWDK.W..V-GK.....PKPTQVAV--L.....IEDEEDSGIYD-ASSETGFNGFA-PGFEV.....		78
GSN.CM.99.CN71	-...SAAALW-WGAAV-TFIYFCLAIFALYLAWDK.W..I-GK.....PKIP.VAV--L.....VED-EESGIFE-ASSE.PNAYGFA-PGFEV.....		76
MON.CM.99.L1	-...N.YWWSLVAITYSLILI..LPVAAWAW..WRYY-ITKRFKRIDQEIQRLIQIHE-R.....HDSGVDTESESEQHEETH..GFV-PVFNDDFGEW..V-		87
MON.NG.-.NG1	-...N.TYWI-AISVWCVIV-L..APC-LYLL..YQSY-EHKRVSFEQ.DFQRLVQXYQEXDSGYEDEDEHDNSFDNPLFDGDDPQW-XFRKQIMGHSILWSAG--		98
MUS.CM.01.1085	-...N.YWYL-AVIVTGIYFVIA-FAFVLAQY...RWC-PKKVEVSVI.....L.....LE-GD-DSGIFEDA-DDMAES-HHAFANPAPEQ.....		76





signal peptide \/ gp120

SAB.SN.-.SAB1C	M...KLLTVL.....LWLSGCWSLWVLV.....QY.VTVFYGIPVWKNS...VQAFCK.....TPNTNLWASTNCIP.DDEPEGT.IAEVPIPNITEKFDANKNRNPLVGOAESNIHL	95
H1B.FR.83.HXB2	-RVKE-YQHLW.....R-GWRWGTMLLGLMLM..ICSATEKLW.---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	...KVMEKKKRDNWLSLIIITIIITII-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-PPGKPVQNMMPNVTEAFDLF.....K-SFS-EV...WI	88
DEB.CM.99.CM40RLIQ.L.LLLILGLVIGVTQKQKQKQ.--L--Y--V---VDAK...-DL--T.....ANSSESG--V-A-L-.HALVREEVPMNPVTQ-FNAFD...-P...IEEQLWQDMT	95
DEB.CM.99.CM5LSLLIIIA-LTLVGAAPTK.....-QW--Y--V---H-AT...-GL--T.....ANSSESG--V-A-L-.HAMVREQVPMNPVSYFDSFS...-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.CM171	...MRKVGf.....KM-VVSLT.LMALLI..PGSRSGNW.T---V---RDAK...PPL--ASDADITSNEPG-I-I--A-L-.S-PSPAEVPL.NITEEFNIY.....K-YM-DEVRDDMVS	105
LST.CD.88.447	...NCPGLI.....LLII-RVQG.....EKTCK.---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-LNVTESPDAWN.....TVTE--IEDVWQ	92
MND-1.GA.-.MNDGB1	...KCPGLF.....ILCLAXLVGIIG.....K.....-V---EAK...THLI-A.....-D-SS-VT---SLPDYDE.VEIPD-KENFTGL.....IREN-IVYQAWH	89
MND-2.-.-.5440	...LRY.....LRYIVLGIIVS-IV...GE...-W...-Y--T-K-HKAR...THL--A.....-D-NSF-VT-S-V-.SLLHYEEQHIPNITE-F-GPI.....EENBIVTQAWG	88
MND-2.CM.-.MND302ENV	...LRY.....IILGIIIVGLGL...GN...-W...-Y--T-K-HKAE...THL--A.....-E-DS--VT-S-V-.SLLHYEEQIPNITW-F-GPM.....EENEVVMQAWG	85
MND-2.CM.98.CM16	...LRH.....LITAIIVIIGIGI...GK...-W...-Y--T-K-HPAR...THL--A.....-D-NS--VT-S-V-.SLLHYEEQLIPNITE-F-VPI.....QNEVVRQAWG	86
MND-2.GA.-.M14	...TQG--QNLNQSMMRKLMLSTLCLM...MTLVNGSNW.T--Y--V---PAT...PPL--ASDPNYGSKEAG-N-LASS-L-.T-PTPQSLYL.NITEEFNAY.....QNYMVBEMVEDMK	108
MON.CM.99.L1	...IQGMRMRMKTIIIVLTILC-MMGTL.TSGSXSGNKS.....-V--Y-PAT...PPL--ASDPXYGSKEAG-N-LASS-L-.S-PVPRALPL.NITESFNAY.....NYMVBEMVEDIK	109
MON.NG.-.NG1	0
MNE.US.-.MNE027	...GC-GNQ.....LIAILFLSAYGIY...CI.....-V-A-R-AT...IPL--V.....-R-RDT-GT-Q-L-.NDDYSE.L-LNITESFDAWE.....TVTE--IEDVWH	92
MUS.CM.01.1085	...PRMTWLRASHTTL-TLLSSSIWGR-W.....-V---DAT...PPL--ASDANIANREPG-V-IT-A-L-.T-PAPQVQL.NISGEKFDIY.....KNYMVDQMTDDVR	103
RCM.GA.-.GAB1	...D.....K-VIVLIV-IGIIL.VQGSQK.P--I---V---R-T...-PM-V.....-D-QS-GTL---EGGISPE.VSINVSEKFDAW.....N-S-YE--KD-VWN	93
RCM.NG.-.NG411	...A.....ISI-IIIGLLGIII...VGEK--...-V---T...-PM-A.....-D-AQS-GTL---EGGISPE.VPNVSEKFDAW.....E-S-YE--K--VWH	91
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.....IGFSLLG-.IAI.....-T-K-EPAV...-PLI-A.....SA-NS-VT-S-L-.LQTYAE.VPITGLEENFTG.....ISNN-IVQQAQW	87
SYK.KE.-.KE51	...KLFNIIISVVLGIVLSLIGIVARLE.....-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	0
VER.KE.-.9063	...KP--GIV.....IAII.GISS.-L...SER.-W...-M.....-T-R---T-----HDYTE.VPLNITEPFEAWA-R.....A--G---	93
VER.KE.-.AGM155	...T-F-GIF.....IV-GI.GIGI.GI...STKQ-W.I---V-----M.....-T-R---T-----HDYTE.VPLNITEPFEAWA-R.....A--G---	94
VER.KE.-.TYO1	...RYTIIT.....GIIVI.GIGI.VL...SK...-W.I-----M.....-T-S---T-----HDYTE.VPLNITEPFEAWG-R.....IA--A---	92
VER.KE.-.VER266E	0

PLV protein alignment: ENV

	↓- * * * V1 loop ^^^ - - ^^ V2 loop *	
SAB.SN.-.SAB1C	...LFESTLKPCKLSPMCIKMNCYRLEGGAAATTSPTSTARPEVVS.....VGFNDSVIEQEMEKEQAMN.CS...FAMAGYRRDVKKNY.STVWDDQEVVCEE.GREKSN..	194
H1B.FR.83.HXB2	...-WDQS-----T-L-VSLK-T...DLKNDTNTN-SSGR.....MIMEKGE.....IK-....-N1STSI-GKVQKEYAFFYKL.DIIPIDNDT.....	188
CPZ.CD.-.ANT	...-QOSH-----T-----TGYN-TPT-P...-T-ST.....-TPKTTTTPIVDGM-.LQE.-N...-NQSTGFK-K-QKMKAIFYKG.DLMKCDNN-TN...	189
CPZ.GA.-.CPZGAB	...-WDQS-----T-L-VTLQ-SKANFS.....QAKNLT.....NQTSSPPL.....MK-....-NVTTEL--K--QVY-LFYVE.D--NLGNENN.....	187
H2A.DE.-.BEN	...-TSI-----T-L-VA--S-.VQ-NT--PN-R--SSTTSRPPTS.....AA-I-NETSNCIENNT.-A...GLGYEEMMQCEF-MKGLEQ-KRRYKDT.WYLEDVVC	192
BAB.TZ.85.2010E	0
COL.CM.-.CGU1	ITQTTL-QR-R--A--TAY-AP-I-TKVN.....RTENGT-TVAPT.....TNNASADWDES.....NWKEYPWY--R.MNST-PLLK-R-ELELGFVSE-LT-LGN..KNDS...	184
DEB.CM.99.CM40	SL...YKQSF-----T-Y-VS-Q-IKTSTNTP-NTST-T-TIATTKTTTDWSENIIMTQ.....YW--....F-VSGPYR-KK-KSSAV.WL-DDIQW	185
DEB.CM.99.CM5	SL...YRQSF-----T-Y-VR--SSTTTTTTAA-TTIKT-TDWSGENKTVTQ.....YW--T.....F-VTGPYR-KK-MQSAV.WL-DDIQW	172
DEN.CD.-.CD1	...-YKQ-F-----T-V-VR--T-K-DNSTNAVTV-S.....TTTTTTTIKPTTTEWYGGNKTTIFLK.-N...-N-TTGKFRD-RVMYEAHCRI-D-M-DRET.....	191
DRL.-.-.FAO	ALSSMIDAV-----IN-Y-V--E-SD-...KIVNAA.....STTPTT.....TTVATTVNTSVA-TTTANDTL-I-V-.NT...EPVMEVN-VC-Y-VTGLCR-CK-EIT-N.F-YDDVVC	193
GRV.-.-.GRI2E	0
GRV.-.-.GRI3E	0
GRV.ET.-.GRI_677	...-Q-MR-----I-----S-VE-N-T-T-KATTTAT-TMT.....TPCQN.....CSTEQ--G--AE-P-S-.T.....I--Q-----.-MT-Y--L--NN.KTGSEK..	188
GSN.CM.99.CN166	...-NQA-----T--VG-Y-NLNTST.SEPSTT..PR.....PPNVSTTQWGSWGGENGTPQ.P.LY-....-NQTTEF--Q--QMY-LF-RE.DIME-TH-...	197
GSN.CM.99.CN71	...-NQA-R-----T--VR-K-KLPNTST--P-TT-A-T.....SPSSPTTPWG-NGG.NGTGQP..IY-....-NQTTEF--K-RQMY-LF-K-.DIMRAQD-.....	199
LST.CD.88.447	AMTSMVDTIM----VN-Y-VR-K-EGEGI-TTKA-KTK-...STTPMPCFITEKTTTTTKRSTTGGP-...LIEDI-.-IKVNNETTEIN-ECRY-VTGLCR-CRTEIKQS.F-YDETTT	203
LST.CD.88.485	AMTSMVDTIM----IN-Y-VR-K-TGGGI-PK--P--IGK.....TTTPMPCFIAETTTTTTK-NTTMES...LIEDT-.-TRWLNETTEIN-EC-Y-VTGLCR-CK-EIKQS.F-YDETTT	202
LST.CD.88.524	AMTSMVDAIM----T-IN-Y-VR-K-TGEGE-KKR--TTT-P.....TTTPMPCFITESTTTTTKSSGTBQPS...LIEDT-.-TRWLNETTEIN-ECSY-VTGLCR-CK-EIKQS.F-YDETTT	205
LST.KE.-.lho7	AMNAMVD-IM----IN-Y-VR-Q-GEVTKTPT--PK.....TTTQMPCFINEQV.....TVKNPNET.RL-EDL-.-TRGLNETTERNAECQY-VTGLCR-CRTEIKQS.F-YDDVTC	195
MAC.US.-.239	...-TSI-----L--T-R-NKS-TDRWGL-KS..I-TTASTT-TTA.....SAKVDVNETSSCIAQD-.T...GLEQEQMISC-F-MTGLKR-KKKEYN-T.WYSADLVC	194
MND-1.GA.-.MNDGB1	AMGSMIDTI-----IN-Y-V--Q-QET-NVS--AK-I.....TTP.....TTTSTVASSTEIYLD..VD-N...NT...BEKVERNHCYRY-ITGLCR-SK-EIVTN.F-GDDVVC	188
MND-2.-.-.5440	AISSMIDAV-----R-T-Y-V--K-TEGQNETEQA-AK.....TTTPVP.....TTTTPSTTTSSSTNKTTTPVLVVEKQNN...TTTQON-VC-F-TTGLCR-CKLEIE-N.F-YEDVTC	196
MND-2.CM.-.MND302ENV	0
MND-2.CM.98.CM16	AISSMIDAV-----T-Y-V--Q-Q-KGKDL-TPIPTT.....STTTTTTTT...KTVANTT-LDIDT-NTE...TTTQON-VC-F-TTGLCR-CKLEIE-N.F-YEDVTC	184
MND-2.GA.-.M14	AISSMIDAV-----T-Y-V--Q-Q-KG-NSPTS.....TTTPMP.....TTTTPTTTPAVTGLEVTQON.....NE...STTQEN-VC-F-TTGLCR-CKLEIE-S.F-YEDVTC	182
MON.CM.99.L1	S...-SQA-----T--VR-L-VEVNTVSNAS-T-AP.....STPTPWGNWGG.N.GTGQP.VYN-....FNQTEFRD--KQMYSLFWKEDIMK--G.....	195
MON.NG.-.NG1	S...-SQA-----T--VR-R-QXINESTVXPTTXX.....XATTPWGSWGG.N.GTGQP.LYN-....FNQTEFRD--KQMYSLFWKEDIMQDKD.....	196
MNE.US.-.MNE027	...-TSI-----T-L--T-K-NKS-TDKWGL-KS--T--A-TTKTTTT.....KEIB-VNENSTCVNRD-.T...GLEQEQMISC-F-MTGLKR-KKREYN-T.WYSADLVC	195
MUS.CM.01.1085	S...I-NQA-----T--V--K-AKVNYTRS-GGN.....STGTTAPPTTPWGSWGNDSAGQP.-FN-....FNQTEFRD--RQMYSLFFVDDLMRDSSES-N...	200
RCM.GA.-.GAB1	...-YD-----R--L--T--SAIN-SWDG.....IP...TSAPPTTTKTTTQRTIGVEKECTAGNET.-E...EVQDADVMSCEFAVAGLKR-EKHKYNDT.WYSRDLWC	191
RCM.NG.-.NG411	...-YD-----R-T--T-Y-TA-N-SWDG-ATQ-P.....TTTTRATRTTKEIGKDCYANQTE-...SVPDADVMNCEFTVAGLKR-EKMK-NDT.WYSRDLWC	189
SAB.-.-.C5	0
SAB.-.-.SAB3E	0
SAB.-.-.SAB4E	0
SMM.SL.92.SL92B	...-TSI-----T-L-A-K-NKN-TDRWGL-RAA-T-SS-TTT-PLTAASPSGE.....E.IVNDT-SCTKNN-...GLEQEQMIGCQF-MTGLKR-QKQYN-T.WYSRDLVC	199
SMM.US.-.H9	...-TSI-----T-L-T-R-NKS-XXRWGL-GXPAP-TTQTSTPPSPPIIA.....K...VNDSDPICIRSN-.T...GLEQEQMVSC-F-MTGLKR-KKREYN-T.WYSRDLVC	196
SMM.US.-.PGM53	...-TSI-----T-L-T-R-NKT-TDRWGL-GTKNELPTTTTTTTKSPKAEAITA.....K...NESDPCISNN-.T...GLEQEQMVSCR-F-MTGLKR-KKREYN-T.WYSRDLVC	201
STM.US.-.STM	...-TSI-----T-L-A-R-NKN-TDKWGL-GKTVT-VT-TAAAAAT.....KPELVNETSSCVSNN-.T...GLEEESLVGC-F-MTGLKR-KKREYN-T.WYSSDLIC	194
SUN.GA.98.L14	AMTSMVDAIM----IN-Y-V--K-PPKPTTSSNSTVKS-CDYWTITTTAKTTTQTSSST.....S-TASTTTPMPLDW-.TDT.ENI-ESNKVC-Y-VTGLCR-CKTE-EQN.F-DTEVTC	202
SYK.KE.-.KE51	DL...-LQADR-----RTL-VEDKSSSGLNTTII-PTTTTNKAAAPPWAGDNSTET.....RY-....F-MTGGFK-KKQYNAF.FYKSDIMR	184
SYK.KE.-.SYK173	AL...-LQANR-----A--R-L-TLDNSP-TS.....TPTTSPPTTPNETWWDGNSP.T...RF-....F-LTGGFK-KKQYRAF.FYKDLDMK	177
TAN.-.-.TAN17E	0
TAN.-.-.TAN40E	0
TAN.-.-.TAN49E	0
TAN.CF.-.AGMB05	0
TAN.CF.-.AGMTB14	...-T-----R--T--PSPTSSS-TTRT-R.....NPCPGTTNDECNSH...LVTNS--F-NSSM-.T...M-Q--T-.NST-Y-A-LM--P.ESK-...	154
TAN.UG.-.TAN1	...-T-----T-----TK-TST-P-SST-TS-STT.....DPCPNIDECCNAT...LVTNS-DY-NSSI-...K-.NST-Y--L--K.ENNTT...	197
VER.-.-.VER1E	0
VER.-.-.VER2E	0
VER.DE.-.AGM3	...-Q-----L--S-VE-NSSEP--PK--TASTTNITASTTT.LPCVQNKSTV.LESC-ETI--K-LNE-P-S-.T...V-Q-K-.V--N-A-IM-KK-.NN...	205
VER.ET.-.VER385E	0
VER.KE.-.9063	...-Q-----L--S-VE-NSSRRNP-PT-A--TKA.PKTG...DPCIKSTNNVNLQPC-A-L--E-L-E-A-S-.T...I-Q-----V--N-A-IY-KN.SSSNS...	205
VER.KE.-.AGM155	...-Q-----L-----VE-K-S-TS-PAT--TAGTK.....LPCVRNKTD.SNLQSC-TI--K-ND-A-S-.T...I-Q-----V--N-A-IF-KR.ST.SH-...	200
VER.KE.-.TYO1	...-Q-M-----L-----VE-NSTRERA-T-T-TPKSTG.....LPCVGTSGE.NLQSC-A-I--R--D-P-S-.T...V-Q---Y-V--N-A-IY-KN.KT..NS...	199
VER.KE.-.VER266E	0



	V2 loop		
SAB.SN.-.SAB1C	..AT...HTV.....GCYMIHCNTSVIKEADCKTYWDTFRLRYCAPAGYALLRCADTDYSYGHK..ACRNVTVSACTRLINTTSTVSTGIGINGSYVA.NRTEIW.....QKNGNSNDSVIIRLNR	299	
H1B.FR.83.HXB2T.....SYKLTS-----TQ--P-VSFEPIPIH-----F-I-K-NKTFN-TGP..-T--STVQ--HG-RPV---QLLL---LAE.EEVV-R.....SVNFTDNAKT-IVQL	288	
CPZ.CD.-.ANTCY-LW---TT-TQS-E-STFEPIPIH-----I---E-E-FT-VGM..-K--S-VH--HG-SPM-A-WLLL--T-QTNTSVVMN.....GRKN..ESVLVRFQKE	288	
CPZ.GA.-.CPZGABTYRI-N---TA-TQ--P--SFEPIPIH-----F-I-K-N-K-F--KGK..-T--STVH--HG-KPV-T-QLL---LAE.GNITVR.....VENK-KNTDVMIVQL	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.85.2010ENSI.....RAT-KD-ANYVTQV--M-IV-PV-TGF--AP--M-----D-KKWD-TGA..-N---AVS--HEF-I--MSHVLV-A-KEL.SDWAKD.....REGVWKNDSGT-EYYWF	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-ETSFRFKP-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-ETSFRFKP-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	79	
GRV.-.GRI3EL-V-----K-N-K-N-FAP.K-K--S-VH-----IT---L---R-E.-----EK---T--K-K	79	
GRV.ET.-.GRI_677D-----D-----L-V-----K-N-K-R-FAP.K-K--S-VH-----IT---L---R-E.-----G-D--T--K-K	292	
GSN.CM.99.CN166NQS.....Y-IRN---Y-TQK-V-SSFQVPVPIH---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.CN71NDT.....DY-I-N---YVTQK-V-SFQVPVPIH---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.DDEL-PLTP.KKM.GNLLLEAKFVYKVG	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.PKKMGDLNGAKFVYKVG	312	
LST.KE.-.lho7	SGERENR.....T--T--D-I-TQD-N-GVMQAYF-L-----M---NEQLNFSK...-E-I-ATP--GYMLSS--SFF-F--TNHT.RDEL-PLT.PNKMEDLNGAKFVYKVG	305	
MAC.US.-.239	EQGN.NTGNES.....R--N---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.LIRCLIG-V-P--VM--YNEKLNNNKL...-S-ISAVQ--QHLVA---SFF-F--TMHK.EGEL-PIDDKYRGPPEFHQRKFVYKVG	295	
MND-2.-.5440	...-KLNKTGSATNSTEPY.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.-.MND302ENVS---TQ-----TQD-N-AST-EIKF-L---P--V---LERLNVSK...-T-I-AVQ--QPLPA-T--MC-SS-TKHD.YNEL-QTNT.KKG-EEFHDHYVY-VDE	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-EEFHDHYVY-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...VNPQKEFHDHYVY-V-K	293	
MON.CM.99.L1SNGSHY-ILN---Y-TQ--E-SNYEPVP-H---P-----D-PAFT-QGS...-S--SAVT--HA-QPI-A-WFQL-STGN..PN.TTV.....MM-KQK-E-IVV--AK	298	
MON.NG.-.NG1NTTDDYY-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.01.1085DSY-ITN---Y-TQ--E-SSFQVPVIQ-----S--K-N-ANFT-QGE..-N---ARH--HG-LPL-A-WLQL--T-LB.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..-V-W-Y.....QRRQSNRTV-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.-.C5K-K-----E-----NTA	0	
SAB.-.SAB3EK-K-----E-----NTA	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-GIMQAYF-L-----M---DEKLNATK...-K-I-ATP--NYMTS---SFF-F--TRHK.EDEL-PIN...NKVGGQAEYVWVKVAA	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	78	
TAN.-.TAN40EL-----F--K-R--N-A-FG..I-N--S-VS--G-M---I-SAF-L---QAE.---H-----K-K	78	
TAN.-.TAN49EL-----F-I-K-K-N-T-YG..I---S-VS--G-M---SAF-L---QAE.---HRIN-----K-K	78	
TAN.CF.-.AGMB05NSR.....D-----E-----L-----AV-I-K-K-N-T-FGT...-S-VS--G-M---SAF-L---QAE.---H-V-----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	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	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---HE.---Q-----R-H-V-----LI-K	79	
VER.DE.-.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-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-----F--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-----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	

PLV proteins

	^ ^ ^ ^ - ^ ^ ^ ^	V3 loop	-	^	^ ^	*					
SAB.SN.-.SAB1C	YFNLITRCRRPGNKTVLPVTIMA	.GLVF.HS.....	.QKYNTRLKQAWCWF	.GGNWRGAWKEVKETIVRLPPKKYSGTN.....	.DTNKIFLQQRQWGDPESEFFFFNCQGE		395				
H1B.FR.83.HXB2	NTSVE-N-T--N-N-RKRIR-QR-	-PGRAVFTI.....	-G-I.GNMR--H-NISRAK-NNTL-QIASKLRQFGNKNKTII.....	-K-SSG----	IVTHS----	381			
CPZ.CD.-.ANT	FE---T-I----R--RNLQ---	-PGMTFFYNV.....	.EIATGDTRK-F-TVNKTL-EQ-RNKTEHVLAEHKK-VDNK--	A-TIWTFF-D---	VKVHW----	384			
CPZ.GA.-.CPZGAB	VEAVSLN-H---N-RGE-Q---	-PGMTFFYNI.....	.ENNVGDTRS-Y-KIN-TT-NRTVE---	KALATSSNRTAAN.....	ITLNRASG----	VTHHM----	379		
H2A.DE.-.BEN	-Y--M--K-----I-L-S-----	-PI--PR---R---R--E-MQ---Q-L-QH-RY-GINDT.....GKINFTKPGAGS----	VA-MWT--R--	393			
BAB.TZ.85.2010E	HY--VT-K-----I-----M--	R---H-.N-----R---E---RE--K--KDR-Q-----	-AQ--Y---F---AANLW-----		130			
COL.CM.-.CGU1	PKDIALG-I-R--SSHRNLNTAN.	-AK-YYE.....	.LIPYSKGIYGR-Q-.VPMT-QN-KN-TQFAIEIK-NLTAWLERISRKNI.....	ITPRNGNRTS--	AT-T-VI-HRL	391		
DEB.CM.99.CM40	KY--LT-V-----YRAIHMAT.--	S-YTT.....	.FIQRL-I-R-H-RL-N-S-AN-T--MRQK-LEIFG-ANRTN.....	NLTIHYPK--R-VQSVW-Q-H--	393			
DEB.CM.99.CM5	KY--SLD-V-----YRA-HMAT.--	S-YTT.....	.PIPRL-I-R-H-RM.R---T--V--IRSRLKEIYNVTKITV.....	HYPG--R-VQNTW-Q-H--	375			
DEN.CD.-.CD1	SSVIKLT-T---V-LKGLNMA--	I-M..RAT.....	.IIPRRAIR-A-KIV---TRTRQ-IEKAVKKEEFNINGT.....	V-WDVP-VG---TSSAR-Q-R--	389			
DRL.-..FAO	KW--K-I---K--RSIISTPSAS.--	L-Y-G.....	.LEPGKN--RGM-K-.E-Q-GK-LGSLAKELKSVSGSIWRN-T.....NSTCISLKGKKIQDRNVTGCMKQL-	VDNYTTK--LA--	NLMLL-G--	432		
GRV.-..GRI2E	FY--V-----	H-.Q-D-H-----QEVKKA...-NLT.....	DD-K-R-----	AAN-W-S---	169		
GRV.-..GRI3E	FY--V-----I-----R-	K-----HS.Q---L---R-K-KEE---	NLTEV.....	KIEN-H-R-----	AAN-W---H--	172	
GRV.ET.-.GRI_677	FY--V-----	H-.Q-D-K-----R-EVKKVKNLTEVSIE.....	N-H-R-I-----	AN-W---	385	
GSN.CM.99.CM166	DYH-NLT-I-----IRNLQ---	-AGMTFFYSQ.....	.VIVGGNTRK-Y-RLDPPR-NR-IR-AMKAMNEHWEN-TG.....	RN--QIRWTSEPK--L-VQTHW-Q---		397		
GSN.CM.99.CM71	DYQ--LT-I-----LKNLQ---	-AGMTFFYSQ.....	.IIVGGDTRK-Y-KLNHTK-DI-IRQAMKAMKNHWE-INND-T.....	PNK-QIRWTSEPK--L-VQTHW-Q---		402		
LST.CD.88.447	KWG-I---I-K--RSEVSTISST.--	L-YYG.....	.LEHGS--RL-Q-K-.E-Q-GRMPHNGLNMLK--NATAMNY-HGN..	CIDNTKPC.....	GRQLKGLPIANMT-RG-NLAT-MLMHT-GE-		422			
LST.CD.88.485	KWG-I---I-K--RSEVSTISST.--	L-YYG.....	.LEHGS--RL-Q-K-.E-Q-GRMPHNGLNMLK--NATAMNY--	GO.CPEGKQPC.....	GRHLKGLPIANMT-RGEGLAT-MLMHT-GE-		421			
LST.CD.88.524	KWG-I---I-K--RSEVSTISST.--	L-YYG.....	.LEHGS--RL-Q-K-.E-Q-GRMPHNGLNMLK--NETAMNYKE.GN.	CSTPGKPC.....	GRQLKGLPIANMT-RGV-LAT-MLMHT-GE-		424			
LST.KE.-.lho7	KWG-I---I-K--RSEVSTISST.--	YL-YYG.....	.LEHGS--RL-Q-K-.E-Q-GRMPHNGLNMLK--NAEAMNY-EGTGTCDKSKKTTTC.....	GRKLKGLPIANMT-HGA-LAT-MLMHT-GE-		419			
MAC.US.-.239	-Y--MK-----S-----	PI-D-P-----K-KD-I---Q---KH-RYTGTNNT.....	KINLTAPGG.---	VT-MWT--R--	398		
MND-1.GA.-.MNDGB1	KYG-K-E-H-K--RS-VSTPSAT.--	L-Y-G.....	.LEPGKN--KGM-T-.K-R-GL-LWSLAKELNK-NDSIKVNQTCNFTSTGEEENKQNTDKQEFKCIK-L-	DNYTTS--	RAA-MMMT---				417		
MND-2.-..5440	KWG-QV---K--RSIISTPSST.--	L-Y-G.....	.LEPGKN-RKKG-QL.E-K-Q-LHLSLSELRKINDSI-KDNH.....	NMTCKSSNNK...KNTTGCCHKL-IS-SESTVK-E-GA-TIMLL-G--					432		
MND-2.CM.-.MND302ENV	KY--QVV---K--RS-ISTPSAT.--	I-YSG.....	.LEPGKN--KGM-QL.K-Q-GI-MHDLAIELRKIDSSIWRNVT..	KGYCKGLKRK...ENRTGCALK-I-VSDYTTK-E-GA-TIMLL-G--					411		
MND-2.CM.98.CM16	KWG-KVV-D-K--RSIVS-PSAS.--	I-Y-G.....	.LEPGRN--KGM-RL.I-Q-GR-LNALSRLEKKNVNASI-R-F..	ASGPCRGINNRG..VVNKTGCALK-IEVSNYTTI-E-GA-TIMLL-G--					411		
MND-2.GA.-.M14	HLHVN-T-I-----IRNLQ---	-AGMTFFYSQ.....	.LIVGGNTRK-Y-KVNKTQ-ET-LQA-H-AVKTWE--NN.....	GTNV-TISWRFPQ--K-VQTHW-----		396		
MON.CM.99.L1	TLSVN-T-I-----LRNLQ---	-AGMTFFYSQ.....	.IIVGGNTRK-Y-KVNKTQ-DA-LHAAXA--XXEWES--NN.....	NVTVA-ITWRYQPK--N-VQTHW-----		398		
MON.NG.-.NG1	-N--MK-----S-----	PI-E-P---R-.E--KE-I---Q---KH-RYTGTNNT.....	KINLTAP-G.---	VT-MWT--R--	399		
MNE.US.-.MNE027	N-RVNLT-I-----IRNLQ---	-AGMTFFYSQ.....	.LIVDGDTRR-Y-KINKTQ-DP-LRQAMLAMRKHWHGRLQ.....	KQNHSSINETKIKIRWRSEPK--Q-VQTHW--G--		406		
MUS.CM.01.1085	FY--SVT---S-R--KGISLAT.-	VFI.....	.LRVEK-P-G--R-.E--TD-----RVKTTKGYRGTSNT.....	DKIKIRTVYG--D-ARY-WL--N--		395		
RCM.GA.-.GAB1	H---VT---S-R--KGISLAT.-	VFI.....	.LRVQK-P-G--R-.D-D-K-----A-NKTKGY-G.....	TRQPE-ITIRSVYG--D-AKC-WL--D--		389		
RCM.NG.-.NG411	0		
SAB.-..C5	H-----	174		
SAB.-..SAB3E	-Y-----	174		
SAB.-..SAB4E	174		
SMM.SL.92.SL92B	-Y---K-----S-----	PI-E-PR-----K--E-MQ---K---KH-RYTGTNDT.....	RKINLTAPGG.---	VT-MWT--R--	401	
SMM.US.-.5440	-X---M-----S-----	PI-E-P-----E-KE-IR---XXL-KH-RY-GTNXX.....	XQI-LTAPGG.---	VT-MWT--R--	401	
SMM.US.-.PGM53	-Y-----S--I-----	PI-E-P-----S--E-IQ---KH-RYTGTNET.....	KKINLTAPGG.---	VT-MWT--R--	405	
STM.US.-.STM	-Y--MS-----S-----	PI-E-P-----E---I-----L-KH-RYTGTNDT.....	AKIR-VAPGG.---	VT-MWT--R--	397	
SUN.GA.98.L14	KWG-V-Q-I-K--RSQVSTISST.--	L-YYG.....	.LEPGSK-RL-Q-K-.V-Q-GR-FATLGKLLRQVE-LANMSNGC.TF.DNITKTCRFT.NGTDFFKKMIKFK..	PMEKHGA-AAT-MLMHT-GE-					425		
SYK.KE.-.KE51	KLK-QLV-E-----ESIKNIQLA--	FFLPVI...QGRL-TGKAA-R-F-KV.K-E-GKFFEQ-HNESIKV.W-NVTS-	SWRSQPQ--L-VRTHW-Q-G--	384	
SYK.KE.-.SYK173	ALR-Q-I-E-----ESIKNIQLA--	YFLPVI...QGKL-TGRDA-R-F-RV.T--TEFF-Q-H-QATKT.W-NVTN.....	TTWRSQPQ--L-VRTHW-Q-G--	377	
TAN.-..TAN17E	FY---K-----	R---H-.Q-K-KE---R---K-KE--R--	K--W-R-----	AANIW-----	174
TAN.-..TAN40E	HYQ-K-K-I-----	K-R--H-.Q-D-K-----R---K-KE-Q--K.....	N-TE-W-R-----	AANIW-----	174
TAN.-..TAN49E	-YK-GVT-----	K-R--H-.H-D-K-----R---K-RDR.....	N--Q-W-R-----	AANIW-----	174
TAN.CF.-.AGMB05	HYQ---V---K-----M---	Q---K-R--H-.Q-D-K-----R---TK-KV-Q---	N-KQ-W-R-----	AANIW-----	111
TAN.CF.-.AGMTB14	HY---L-----	K-R--H-.Q-D-K-----R---K-KH-Q---	K--W-R-----	AASIW---L-	353
TAN.UG.-.TAN1	HYK-K-V-----	Q---K-R--H-.Q-D-K-----R---R-K--E-KE-R---	N-RQ-W-S-----	AANIWL-----	396
VER.-..VER1E	HY-----V-----	K-R--H-.K-D-----R-I-QK--E--KER-Q-----	KQ-----	ASNLW---H--	175
VER.-..VER2E	A---VT-----	R---H-.Q-----E-IRQE-TK--H-R---	E-----F---	AANLW-----	175
VER.DE.-.AGM3	HY---VT-K-----	R---R---H-.Q-----NE--K--KDR-Q-----	EE-Y---LF---	AANLW-----	405
VER.ET.-.VER385E	HY---T-K-----	R---H-.Q---H---I-QE--K--KDR-R---	EQ-Y-T---	AANLW-----	175
VER.KE.-.9063	-Y--VT-K-----	K-R--H-.Q-R--E-----E-IK--KDR-R---	E--Y---F---	AANLW-----	405
VER.KE.-.AGM155	HY---VT-K-----	R---H-.Q---K---Q-E--K--KER-Q-----	F-----	AANLW-----	400
VER.KE.-.TY01	HY--SVT-----	MK-R--H-.E-----R---QK--E--KDR-K---	N-EH-Y---	ASNLW-----	399
VER.KE.-.VER266E	HY--VT-K-----	L-R--H-.Q---K-----E--K--KER-R---	KR-----	TANLW---H--	175



	-	^ ^ ^	V4 loop	* ^ ^ ^ *	-	* ^ ^ ^	-	V5 region	^ ^ ^ ^ ^	
SAB.SN.-.SAB1C	FFYCKMDWFLNLYLNKS		VDP	DHNCAKNNTPKPCWQRTYVPC	HIRQVVDWYTLSSKTYAPPREG		HLECNSTATALYVELNYSK		NRTNVTLSPOQIESI	493
H1B.FR.83.HXB2	---NSTQLF-STWFN		TWSTEGSNNTGSDT		ITL--R-K-II-M-QKVG-AM	---	IS...QIR-S-NI-G-LLTRDGGN		SNNSEIFRPGGDMRDN	478
CPZ.CD.-.ANT	---DITPWF-ATYT		GNLITNGAL		IAH-R-K-I--H-GIV-GI-LA	---	R...NVS-T-SI-GIML		EGQIY...NETVKVSPAARVA	468
CPZ.GA.-.CPZGAB	---NTSQIF..TD		ITNGII		IL--R---I-SS-MRVGRGI	---	IR...NIT---NI-G-LLTSDTPVT		N.NSG-L-FRPTGGNMKD	463
H2A.DE.-.BEN	-L--N-T---WVED-N		QTRR		N-...-K-II-T-HKVG-NV-L	---	E-A-E--V-SIIANIDIDKN		RTH--I-F-AEVAEL	475
BAB.TZ.85.2010E	-----I--LT...	VDANHN	TNRCNKTKG		TSGAPG--A	---	S-----V-L-----R--V-		GMTVELNYYN	235
COL.CM.-.CGU1	---NASSLWKHDS		PVMN-T--KL--S-V-HARIL-G	---	PG...Q--WEKQPVIAFMGTIEG		DNDGNGC		AYPAAPNFKHAL	472
DEB.CM.99.CM40	---NISKA-DL-LLQN		NTRNST		WSDKWL--R-N-L--T--VGQHI-L	---	K--E-K-S-HIS-FVFDVDHYNG		SITL-P-AD-RAV	481
DEB.CM.99.CM5	---NVSKAF-L-M--T		NNITWS		NN-LM--K-N-L--T-FKVGQHV-L	---	K--E-K-H-SVS--VF-VAHYGQ		NM-I-P-SD-KAV	461
DEN.CD.-.CD1	---NLTRMF-SSQYPV		NNS		NIINNQ--MV-N-K-L--T-H-VGMDI-L	---	A--EIT-T-NVSGII-DTDVHPNR		THITL--ADVRQV	478
DRL.-.-.FAO	Y-F-NVSKIWRTW--RT		SNVW		YP-AS--K-IID--ARVGR-I-M	---	VS-FNNEIR-TQDV-EMFF		IQKLETEETGDSN	526
GRV.-.-.GRI2E	-----I--I--ET...	VDADG	NKCNPK		GPGPG--V---A-----T---R	---	Q-----SV-----I--NE		SGPI-----R--	271
GRV.-.-.GRI3E	-----I---QT...	VDADG	KCSTSHK		KGQPG--AKK--A-----T	---	V-----I--N-----S		SGPI-----VR--	275
GRV.ET.-.GRI_677	-----I---RT		E-AEGTNRCTDK		G-PGP--V---A-----V---V	---	SV-----AID--N		SGPI-----VR--	488
GSN.CM.99.CN166	---NLSILFQLN--T		INSS.NIGN		ITSKYKQWLA-K---F-TQ-GYV	---	SI-L--Q...-IN-T-NI-G-LIDGAMYES		SINMTPSA	489
GSN.CM.99.CN71	---NLSVLFQFN--IT		AVDDTNINN		VTSKYKQWMA-R---F-TQ-GYV	---	RSI-L--Q...-IN-T-NI-G-LIDGAMYGH		SINMTPSA	496
LST.CD.88.447	M-F-NVTRIFQEW--N		S-KW		YPWAN--KS-ID--ASVG--I-L	---	TS-FNNRIR-THRV-EMWF		MEKPEPEENLGGNLS	516
LST.CD.88.485	M-F-NITRI-QEW--N		S-KW		YPWAN--KS-ID--A-VG--I-I	---	TS-FNNRIR-THRV-EMWF		EKWEPEENLGGNLS	515
LST.CD.88.524	M-F-NVTRIFQEW--N		S-KW		YPWAN--KSIID--A-VG--I-L	---	TS-FNNRIR-THRV-EMWF		MEKWEPEENLGGNLS	518
LST.KE.-.lho7	M-F-NVTRIFQEW--N		S-KW		YPWAN--KSIID--A-IG--I-L	---	TS-FNNRIR-THRV-EMFF		MEKWEPEENLGGNLS	513
MAC.US.-.239	-L---N---WVEDRN		TAN		QKP-EQHK-N-----II-T-HKVG-NV-L	---	D-T---V-S-IANIDWIDG		Q--I-M-AEVAEL	489
MND-1.GA.-.MNDGB1	M-F-NVTRIMRAW-DPN		EKKW		YP-AS-Q---I-D--MQVGR-I-L	---	TS-FNNRIR-THRV-EM-F		MQKIDSNETK	508
MND-2.-.-.5440	Y-F-NWTKIWKAW-S-Q		SSVW		YP-MSRN--I-G--HKVG--I-M	---	VS-FNNEIR-TNDV-EMFF		VQKTDDD	522
MND-2.CM.-.MND302ENV										0
MND-2.CM.98.CM16	Y-F-NWTKIWRW-DQN		SSVW		YPWMS-N--IID--HKVG--I-M	---	VS-FNNEIR-SNDV-EMFF		VQKTEEG	499
MND-2.GA.-.M14	Y-F-NWTRIKWTW--QN		SNVW		YP-MS-K---I-D--HKVGR-I-M	---	VS-FNNRIR-TNDV-EMFF		VQLVEG	500
MON.CM.99.L1	---NVNLFIRRTNK		T		GISPPDVN-KPNTTYHGWL	---	T--M-TQ-GYV-SI-L--K...-VQ-T-NI		--LITGELYQN	493
MON.NG.-.NG1	---NLSRLFXSSDPXK		NK		NETFESTFNVST-ASTAYNGEWLT	---	RL--L-TQ-GYI--SX-L--Q...-VQ-T-NI		--L-DGELYGS	498
MNE.US.-.MNE027	-L---N---WVED-N		LTG		TTQKPQEQHK-N-----II-T-HKVG-NV-L	---	D-T---V-S-IANIDWIDG		Q--I-M-AEVAEL	492
RUC.CM.01.1085	---NLSVLFQFV-HTE		INES		NIHTV-TKYHKDQWV-R---F-Q-GYV	---	SI-L--K...LVR-E-NI-G-LIDGAMYRD		SI-M-P-ANVLDAW	501
RCM.GA.-.GAB1	-L--LN---L---ET		G		T-NEKRKAPF--ITKMI	---	V-R-V-T--PD...A-K-SAQVSY-LADID		INDSE--I--AEVGDY	484
RCM.NG.-.NG411	-L---N-----T		EGNKNEARQ		AMP--ITKM-----V-R-V-T--PD	...	A-R--A-V-Y-LADID-TD		NMT-----AEVGD-	478
SAB.-.-.C5										0
SAB.-.-.SAB3E	-----Q...	VDPDH	NPCNK		TSD-K---N-A					272
SAB.-.-.SAB4E	-----Q...	VDPNH	NTCTP		KSD-R					272
SMM.SL.92.SL92B	-L---N---WVEDRN		TSSP		RWTQTQK-EQHK-N-----II-T-HRVG-NV-L	---	D-T---V-S-IANIDWIDN		E--I-M-AEVAEL	497
SMM.US.-.H9	-L---N---WVEXIK		NGS		RWTSQ-Q-ERY-KN-----X-II-T-HKVG-NVXL	---	D-T---V-S-IA-IDWING		K--I-M-AEVAEL	496
SMM.US.-.PGM53	-L---N---WVEDRN		TSGS		LWTQTKADEKK-N-----II-T-HRVG-NV-L	---	D-T---V-S-IADIDWIDG		Q--I-M-AEVAEL	501
STM.US.-.STM	-L---N---WIE-R-		TSE		MRDWNK-K-EQK-N-----I-T-HKVG-NV-L	---	Q...D-T---V-SIIANIDWTNN		E--I-A-AEVAEL	493
SUN.GA.98.L14	M-F-NLTRIFKVV-DTT		SNKW		YPWAN--KS-ID--ASVG--I-L	---	TS-FNNRIR-ANRV-EAWFT		ERVEDKVNNGSNISVVAFQPTNTLNQ	520
SYK.KE.-.KE51	---NVSKIFANVS-GY		AN-SNY		AKNL-LS-A---II-Y-GYVT-LM-L	---	T--IK-T-NI-VLTDIE-YPN		SQL-FAPTANV-DV	472
SYK.KE.-.SYK173	---NVSKLFANIT-GN		ASKNN		YASNLRLS-A---II--RYVR-LI-L	---	TA...IK-T-NV-VLTDIE-YPG		STL-F-PATANV-DV	465
TAN.-.-.TAN17E	---TP--V---D-ARRRKEDEVEG		NDCK		DPRK-FK-S-----S-----T-V-SMM-S-D-VN	---	M-----TANL-N-			275
TAN.-.-.TAN40E	---TP--MV-W--E-DSGKSVNVEG		NTCTP		EKDRL-KK--A---S-----T-V-SIIAS-D-N	---	M-----TANL-NL			277
TAN.-.-.TAN49E	---TP--V-W--E-AGGNYTDEVEG		NPCSNSTK		GGMTGSRK-LK--G--S-----S-----T-V-SMM-S-D	---	M-----TANL-N-			283
TAN.CF.-.AGMB05	---TP--V---E-ANGTNDLVEG		NKC-DRIS		GGLTGSRK-LK--A--S-----S-----T-V-SMM-S-D	---	M-----TANL-N-			220
TAN.CF.-.AGMTB14	---TP--V---E-SEGSFTDVEG		NRCSNITSG		GLTGS-RK-LK--GL--S-----LQ-R	---	R--V-SML-S-D--D--M-----TANL-N-			463
TAN.UG.-.TAN1	---TP--V-W--E-NSGRNVDEVEG		NNCTTG		KD--YK-----SI-----T-V-SMM-S-D	---	E-----TANL-N-			499
VER.-.-.VER1E	-----M--I--LT...	VDANH	NNCNSTQ		GKRPV--V---G--S-I---K-AR	---	T--V-GMA---IH			276
VER.-.-.VER2E	-----LT...	VDADH	NSCNGSRT		SPRAG--A---A--S-I---I	---	Q-T--V-GMT-----R			278
VER.DE.-.AGM3	-----RT...	VDPDH	NPCNGTKG		KGKAPG--A---A--S-I---R	---	Q-T--V-GMS-----T			508
VER.ET.-.VER385E	-----T...	WNADH	NOCDNSTK		TGRAPG--V---A--S-I---I	---	Q-T--V-GMT-----IND			278
VER.KE.-.9063	-----Q...	VDANH	NQCSNETK		KG-APG--V---A--S-I---I	---	T--V-GMT-----NN		Y-----G-	508
VER.KE.-.AGM155	-----LT...	VDADH	NHCNNAG		KGSRPG--V---A--S-I---I	---	Q-T--V-GMT-----NQ			503
VER.KE.-.TYO1	-----T...	WDAYH	NFCSSKK		KGHAPG--V---AY--S-I--S	---	Q-R--V-GMT-----			501
VER.KE.-.VER266E	-----LT...	VDADH	NECKNTSGTK		SGN-RAPG--V---A--S-I---I	---	T--V-GMT-----INQ			282

PLV proteins

	gp120 \ / gp41	*	
SAB.SN.-.SAB1C	WANELGDYKLVEIKPIGFAPTKVRRYTG...PERQKRVPF.VLG..FLGFLGAAGAAMGAAATALTQVSQQLLAGILQQQKNLL.AAVEQQQMLKLTIWGVKNLNARVTALEKYLEDQARLNIWGC	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-----	598	
BAB.TZ.85.2010E	--A---R-----T-----E-----G...Q-----A-----T-----D-	302	
COL.CM.-.CGU1	STL---R---KMRTTYV--DIK-SVNVNWHHG---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---W	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---W	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.-.-.SAB3EN-----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---W	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---W	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---W	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---W	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-.A--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-.A--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	



Accession	Sequence	Position
SAB.SN..SAB1C	RQVCHTTLVWKYNN.....TPDWENMTWQEWERQIEKYEANISRILEQAHEQEQKNDLSYQKLVSWSD.FWSWFDLTKWFGWMKIAIMVIAGIIVARVLLVIIGILRKRKGYAPLSSLP	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-IVFAVLS-VNRV-Q--S..L-FQ	718
CPZ.CD..ANT	KVT----P-NNSWVNFQTC.AKNSSDIQCI-----D-LVQNSTGQ-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-IAIYVQVQ-L-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-----IPFEL--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-IIMFVLM--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--QEFN..--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--QEFN..--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--QEFN..--M--T--QYI-Y-VLI-I---L-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--QEFN..--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-VHDLE--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	-----PNDS.....LV--N-----KKV-FL---TQM--E-RL-QE--MYEL--N--DV..GN-----S-VRYIQYGVFLVI--VLL-IVIYVQM-SRL-Q--R-VF-S-	735
SMM.US..H9	-----X-P-PNDT.....L--X-N-X---K-VNPL---TX--E-QI-QEX-MYEL--NXXDX.XGN-X---X-IRYIQYGVFLV-L-VGL-IVIYVQM-ARL-Q--R-VF-S-	735
SMM.US..PGM53	-----P-PNAS.....LV-N-N-----VDDL---TQA--E-QI-QE--MYEL--N--DI..GN-----S-IRYIQYGVFLV-L-VVGL-IVIYVQM-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-QEFN..--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---VDNLDTH-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--QEVNLDTH-DGL-RE-Q--QER-VHDLT--QE-DS.L-----S---FYL--GFY--GALVLL-LVFSV--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-YTVYSCIARV-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-YSCIGRV-Q--S..L-PQ	745
VER.KE..AGM155	K-----P-TW.....E-N---L---K---GL-G--TKQ---R---E---A---SD--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---S-----PS--LNIL--MGFL--V--GL-L-YTVY-CIVRV-Q--V..L-PQ	738



PLV proteins

SAB.SN.-.SAB1C	SSH\$. . . QIHIHLKQDPKKEEQQSGGLNSDSRSTWQREFLRHLCHRLITWLRNLTWSFSTIFSNLHR.	CLQD. IQ. QRTRQL. 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-GRKDRIW-ALQHGFFA-L. WVDLTSIQWIYQICRTCLLNWLWAVLQHL. CRI. TF-LCNHLENN-ST	804
CPZ.GA.-.CPZGAB	TLIPV. . -REQGR. . . . LGEID-G-GEQDR-R-VRLVEGCLP-I. WDDLRLNLGIW-YQSLTSLACNVWRQKLTGL. HLILH.SL. RLL-E-LCLL. GGIIQ	799
H2A.DE.-.BEN	PGYLQQIH--K.DRGO.-AN-GTEEDV-GD-GYDLWP-PINYVQF-I-L-TRL-IG-YNICRDL- -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. LQQLLW-YQTFRNCIWLIRRGFELLQEGIYRLSPALW-R.LR.	LLGE-F-V-G. GYLQ	815
DEB.CM.99.CM5	P. K-VEVLQQLGEGGDK-GEVKSRTWTSRPGFLRLIWD. I-QLLTW-YQTLRNCIWLIRRGFELLQETIYRLSPAFW-R.LR.	LLGEHL-VVW. GYCO	795
DEN.CD.-.CD1	IR.PQKAAN. . . . QPEEG-G-S-D. . . . G-RLKWWQYPRGFLSII. WEDLSQLLW-YLTLRNL.VFLIRT.	ACSRIWD. FVH-HSLILLGYFQ-	810
DRL.-.-.FAO	FKDAYR.PETTNY.HR--R-KGEEQDRGEW.NIRSEPS. . . . KPGSSKAWSEETVGTWLKE-RGYIW-KNVKAV.	EYGWS. ELQEAGRGC.	847
GRV.ET.-.GRI_677	IHIHSS.AERPDNNG--DRGG-SSS-KLIRLQEE-S-PS.RINNWNLNFKSCS--IR-WCYNICLTL-IF. I-TA. VGY-Q-	800
GSN.CM.99.CN166	ILIPQAVERP. . . . -AT--GVGAQGVKVR-VR-LTGFSS-L. WDDLRLNLVIWIYQILASLAWTLRGI.	G. KF-WEQL-KAIQKAI. QOAR.QLRE. VA-RVIA	823
GSN.CM.99.CN71	ILTPQQVERDP. . . . -GTRKD-GAQQ-VR-VR-LTGFSS-L. WDDLRLSLVIWIYQILATLVWTLQRI.	G. KF-WEQL. QKAF. QOARQLGEV. A-RAAA	828
LST.CD.88.447	AYVEQDCKWEKRENQE----G-IEKDKE-I.YINLEQYKKESSTPPNVNDWSEPLQD-L-V-LLKWKAFGILLLSLVNQPLSWLGHVILFFQHGQHL.RK.	ACS-RV. VENAQ	863
LST.CD.88.485	AYVEQDYKWEKGENPE----G-IEKDKEDI.YINLEQCKKESIPPVNDWSEPLQD-L-V-LLKWKAFGILLLSLVNQPLSWLGHVILFFQHGQHL.WK.	ASSGWM. VENAQ	862
LST.CD.88.524	AYVEQDFKWEENGESQE----G-TEKDRE-I.YINLEQCKKESFRPPVNDWSEPLQD-L-V-LLKW-KAWGILLTLVWQPLTWLGHVILVLSYHHGQHL.W-	TLF-F. VESAO	865
LST.KE.-.lho7	VYIEQDYKWEKKEENQE--R--K-ADTETI.YINLEQCKKESRPLVNDWSEPLQD-LLV-LLKW-KEGGILLLSLVNQPLSWLGHVILVLSYHHGQHL.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-EEPDKQ.-TKSDSS. KVEFGKPKSKEQIRDWLKT-RGYVW-KNLQAV.	IEYGW. ELKTAGGKIF.	818
MND-2.GA.-.M14 VLKGDYLRPHNLK--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.	LQVRSQR. LPPLLHIRLLQLWESLRRLLAYC	826
MON.NG.-.NG1	I. . . PTQVPREP. . . . -A-EETG-G-GEQGRNRXTV-ARGFLTII. WTDCLDLVWTFQLLRDSGLTIYRS.	LQVRSRSH. LIPLLRDLCL-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. WIQLIHLPLSRIWQLSQICRSCSSIIPOS.	LRWILAKIQYGW.	798
RCM.NG.-.NG411	PCYH. . Q--P-.RNREEQPIK-GEKG-SGEGGHHK-YNWQRGYF. YI-ILRPIELCRRLYQICSNLSVVIYQSLQSIIPSRILLHLRWIGAKLQDGW.	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-YNNLRLA-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-SCRDWLLRXCO.	ILQPVLSQSLRT--R.AR. EVI-VE. ITYLQ	831
SMM.US.-.PGM53	PVYVQQIP--K.GQE-Q.-T--G-E-D--DRGGN--WP--I-YIHF-IRQ--RL-TW-FNCR-YLLRIYQ.	ILQPVFQRL---R.-R. EVV-TE. ISYLQ	832
STM.US.-.STM	P-CRQQIP--K.GQE-Q.-T--GTEEG--DRGGIN-WP--I-YTHF-IRQ-VRL-TW-YNN-RACL-RIYQ.	TLHPTFORISRI--R.-R. EVV-LG. AYLQ	827
SUN.GA.98.L14	AYVEQDWLQE. . . TCPK-TDK--EETEKERIYINLEQSKKESLPPWTVDWDEPLRD-LLV-LLKW-KAWGIVLAQNIYHLLSFLWHLTLTSPFHGQRL.W-	TLRGW. GSHLI	864
SYK.KE.-.KE51	QP. TLYHKAPA-PEEGIETDNDYKPSVSTFSREPLRQSL.A. EAQKLW-RVSESCRSRLIRGLVIAW. GFIS	785
SYK.KE.-.SYK173	TQ. GRKDPGKPADEEEG-GDREGLNVSTFSRESLRQSL.E. AGQLWRTVCSSFRSLIRQLTITW. 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.-.GEGDKRKNSESEPQKESGTAEWKSNWCKRLTNWCSISSIWLWYNSCLTL-VH.	L-SA. FQYIQ-	823
VER.KE.-.9063	IHIHPW.KGQPDNAGE.-.BEGG-RDNDSSD-WQKSGTRQKRGWSKRLTNWLSSTFS-WLYN-FLTL-IH.	L-GA. WQYIQ-	823
VER.KE.-.AGM155	IHIHPW.KGQPDNAGE.-.BEGGRT-KSKSTHSQKEFGGRDKRTSWCRLTNWF-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-IEVVQ.....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-DTIIWAGQEIWEWLGIAQVAIQGLV.....WGG-NLL-I-A-I--GA-LLL--	850
DEN.CD.-.CD1	-LQELQ TGL..RDLGTSAIQQGR-TAE-VL-ALT...RAA-E-V-I---I--G--IVL--	864
DRL.-.-.FAOTLLQEVAQRIRWREGRLGL-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	KIASWSWNKIRNRNRDISA...HARNL-LGQK...KRW-FRFRGRSGFPS-TTETAL-	915
LST.CD.88.485	KIARWSWNKIRNRNRDISA...HARDI-LGQK...KRW-FRFRGRSGFPS-TTETAL-	914
LST.CD.88.524	KTASWIRNKSRNRNGELSKT...NGTNI-LGQNR...RRW-FRFR..SGLPS-TTETTL-	916
LST.KE.-.lho7	KIQSW-REKCRNRNGQLSST...DRKNIQLGKK...KRW-LRFGGRSGISS-ATETAL-	912
MAC.US.-.239	YGWSYPHE-VQAVWRSATETLAG-WGDLWETL.....RGW-WIL-I---I--G--LTL-	879
MND-1.GA.-.MNDGB1	SLLVSFA-ISWRAGYTGWLQ-NRIPTGC-E-II....AAGTCIWHIL--I--SA--PL--	879
MND-2.-.-.5440QVLQGFQRLWSRGYQLGL-GVRGAA...AFG-GIWNI---I--GA-ALL--	890
MND-2.CM.98.CM16KVLQGYAQGLWSRGHQWGL-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--XXWKFRET LA--WRDLWETLG...RVG-GIL-IXX-I--G--LTL-	886
SMM.US.-.PGM53	YGCWRFQE-VQAAWR-RGETLA--WRDSWETLG...RVG-W-L-I---I--G--LALL-	887
STM.US.-.STM	YGCWRFQE-VQAAWR-RGETLA--WRDSWETLG...RVG-RIG-I---I--G--LTL-	882
SUN.GA.98.L14	QATSRIRN-CRTSRER-SSQKAR-RTFSLGRK...WRPKWNRTRGS-IPS-TTETTL-	919
SYK.KE.-.KE51	YGV T--E--I--GRE-AHQMVAIWQALLAYA.....RVAEN-A-L---L--G--IYL--	840
SYK.KE.-.SYK173	YGFN--I-----GREIRDVVAIWIQAIYA-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



PLV proteins

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--E---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-RGDYGGKQRDASGGEYS-FQEESGREQNSP-CEGQQYQQ.....GEYMN-PWRNPATERQKDLRYRQONMDDVSD-DDL..I-V..-T-RV-R-E-----MS	120
COL.CM.-.CGU1	--SLL-CLNI-QQPQTRGGVLRGLRFGIYWYPGVQAFRFLRHFRGGY-LRRVAGVERLNI-ED.....GYMTVPKVLNCPCLCYRHPDVVYHN-EPE-VEDV-.AGCY..-MSR--V-QA-I--LV-IS	122
DEB.CM.99.CM40	---N--IPS-GQVGSFGSGSRGLLRWYRDLSEQEQQFSECLLES-REQSS--TEFW.....GSP.....RREIKCTNKQQQDLQDT-A.....QVSI-DP---M--YS	103
DEB.CM.99.CM5	---N--TPSPGQGGKSGSGKGLLRWYKDLSEGAETFSPFLQES-KEQSCY-TE.....EPLYA.....DPHREIKCTQRAQECYD--HGI--A--RV-I-EP---M--YS	105
DEN.CD.-.CD1	--S-N--ASGIWGLAP-NKEGRQRKWYRLLPEQQEKWSPDGS-RA-KL--TRLL-HDE.....DEETFVDAEEI IQSSPALDPQHPRIH-YVVDHD----.---.---RR--H-P--DM--MS	122
DRL.-.FAO	---QN--KREEAYARFYKR--RGYGAKGGNLDYHQLEPSEPLLRKWQ-FP-GLDKA-RSSST.....SSLDTTHMQAAAAEKD...VTLPTEEEQPS-E---.---.---C-QR---EP---DLV--S	118
GRV.ET.-.GRI_677	--SSN--R--QGL-K--RG--GK-GADWVLLSDPLIGQSSTVQECCGA-KK-WGKGKMPD.....GR-LQEGDTFDEW-D----.---.---Q-RV---Q-----V--S	104
GSN.CM.99.CN166	--S-N----QE-STALLSS-GTQQ.....RPYFTLVDEY-ENFWLSPDASD-GRRYYLTEEPKPKRGS-L-Y.EPSC----RV---DP---VMV-LS	92
GSN.CM.99.CN71	--S-N----SQE-STALLSSHGTGQ.....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-PSADSWRVLRR--AGSGTRAEGTEESYQKLM-ELG.....WDKTEE-G-SDSE.....QR--CSS---TL--LS	80
LST.KE.-.lho7	--NAFGRPAADGWKTLRRF-AGSRNAEGTEDSYRALIEKTSN.....WAKRRDG-S-S-E.....-K-QR--CSP---TL--LS	81
MAC.US.-.239	---AT-MRRS-P-GD-RQR-LR-RGETYGRLLGEVEDGYS-SPGGL-KG-SSL-CEGQYQ.....G-YMNTPWNRPAEEREKLYRQONMDDIDE\$-DDL..-V.S---KV---T-S-----MS	119
MND-1.GA.-.MNDGB1	--SSQ--KRSEAWRYS-A--LVGGPVTPDGYKQIESSQGAEK.....QSLLRG-AYGTYSEGLDKVQNDPLTKDEKDLDTQQ-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	-----EEKY-KYYKAM-RGYGAEGTNGDYQQLHASEPLLGALSTSQEEDFRE-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-STPLLGSSSGRMRYFMLEDDYGEQSWLSPDAS-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-CEGQYK.....G-FMNTPWKNPAGEREKLYRQONMDDIDE--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-RWRFLMLTT-GEYVVRWHTLLDQWPWCAEGSGRASRDFVIRGGIT-E.....TQASIDDIDWYE-T-DTL..---.---K-QV---P-S-----MS	103
RCM.NG.-.NG411	-----ASAAGLRRWRGLMLTT-GEDYARFAETLQDQPRCAEBSGRASRDFLTRGGPTIE.....TQQSVDAIDWYE-T-DTL..---.---QV-I-P-S-----LS	103
SMM.SL.92.SL92B	---IT-SKPA-KHGD-RK--LS-RGDSYGRLLWGLEEGSQ-SQGG-S-RA-N...-CLTRDGY...YNVGEFMNTPWRNPATERQKQYRQOCMDDIDED-D-L..A-V.V-W-KV-----S-----MS	121
SMM.US.-.H9	---VT--K-RXAGGN-XER-L--RGETYGRLEWGLEGEYS-SXDASGKG-SSL-CEPQKYCE.....G-FMNTPWNRNPATEGAKLYRQONMDDVDN--L.....-C--S-RV-V-I-----MS	118
SMM.US.-.PGM53	--AAG--K-S-QRGG-GE--L--RGETHGKLEWGLEEDGYS-SRGLERDWNHL-FEGQYSE.....G-FMNTPWNRNPAREKLYRQONMDDVD-D-D-L..I-V.S-H-KV-----S-----MS	120
STM.US.-.STM	--ASG--K-RKQHG-RER-LR-RGETYGRLEWGLEEGSGPSQAS-KG-NSH-CEPQRYNE.....G-FMNTPWKNPAEASAKLEYRQONMDDVDE--DNL..-V.A-H-RV---E-----LS	120
SUN.GA.98.L14	--NAFGRPSEVGVWRTLFR--AGSGTRAEPAGREYHRLRR-EV.....EPLV.....SAENGGNGI-Q-E-E-....QR--CKP---QL--LS	84
SYK.KE.-.KE51	--SQ--K-PSRWDEK-RTRWVFPKPYSPMPDELLRMSQP-HEDF-RGWRSTLT.....EPI.....DPKRDIFS...-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-D-....RV---Q-----V--S	107
VER.KE.-.9063	--LGN--P-HKKHVS--HA-HKTO.....TK-GL-----I-T--AV--YD-G--KS---KRNGNM-.PDGRRLQE-D....EWDEWSD-D-....K-RV---Q-----V--S	107
VER.KE.-.AGM155	--LG--P-HKKQ-TI-RA-HATR.....TR-GL-----I---TL---D-G--KS---KRNGNM-.PEGRRLOD-D....WDEWSD-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	HFLKEKGGLEGIYYSERKKILDLYALNEWGIVDGWQNYTDGPG.....TRYPKCFWCKFLVPVLDSEAKNSE..NHC..LLHPAQVAYEDD.....AW.KETLVWKFDPPLAVDY.VAVRLHP	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---MWMF-TQ--FPD-----EE-----V---L-R-----ESPPD.....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-----ICQHGM-----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--BED-A.N.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---D-DRYEGQA-----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.-.lh07	--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-LGIEQEX.....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-LGIEQEX.....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-KSGPQLDELQ	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-NEIEGYLVDE-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--NEIEGLEYE-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...IEW-S.....P-.-----N-----Q-.NPDSFKD	214
VER.DE.-.AGM3	-----D--NQ--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-BYF	219



SAB.SN.-.SAB1C	EQVPSA.QG\$	226
H1B.FR.83.HXB2	-YFKNC.\$	205
CPZ.CD.-.ANT	-YFRD\$	198
CPZ.GA.-.CPZGAB	-YYKD\$	205
H2A.DE.-.BEN	-EFGHKSLPEKEWKAKLKARGIPYSE\$	257
COL.CM.-.CGU1	-M-D\$	223
DEB.CM.99.CM40	REHTRRCMFPRK\$	227
DEB.CM.99.CM5	PEHNRRCMFPRK\$	230
DEN.CD.-.CD1	QTSQHN.NYCQRAKPRK\$	248
DRL.-.-.FAO	-EFGHVKSIVVEEPG.....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.-.lh07	AKIQNTAFAPDCRKR\$	207
MAC.US.-.239	-EFG-KSGLSEEEVRRRLTARGLLNMDKKETR\$	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	QEERKRRLTANRIL\$	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	-EFGWKSGLTEEEVERRLANKPKPKQKMKADKKETS\$	266
SMM.US.-.H9	-EFG-QSGLSKEEVQRRLTARGLLMADKKETS\$	261
SMM.US.-.PGM53	-EFG-KSGLSEEEVKRRLTARGLLMADKKETS\$	263
STM.US.-.STM	-EFG-KSGLPKEEVERRLTARGLLMADKKETS\$	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

648

PLV protein alignment: NEF