

VI

HIV-2/SIV Proteins

Contents

VI-1	Introduction	377
VI-2	Annotated features	378
VI-3	Sequences	379
VI-4	Alignments	381
VI-4.1	Gag	381
VI-4.2	Pol	385
VI-4.3	Vif	392
VI-4.4	Vpx	394
VI-4.5	Vpr	395
VI-4.6	Tat	396
VI-4.7	Rev	397
VI-4.8	Env	398
VI-4.9	Nef	404

VI-1 Introduction

The number of HIV-2 and SIV sequences is growing much more slowly than for HIV-1, particularly for sequences of complete or near-complete genomes. In selecting the set of sequences to present here, we eliminated some sequences derived from the same sample, patient, or isolate. We further omitted a few that were very closely related. Within each of the protein alignments, we have included a slightly different set of sequences, drawing on both genome-length and subgenomic fragments in order to best represent viral diversity. Thus for HIV-2, we selected A and B genomes and subgenomic fragments that best represent the diversity of the virus in that region. We also included a single AB CRF and both genomes available from other HIV-2 groups (G and U).

In these alignments, as in the database, we label sequences according to their native host. Thus, transfers of SIVsmm into humans have produced HIV-2 groups A through G. Unintentional transfers of SIVsmm into captive macaques are named according to the receptive host (Rhesus macaques = mac, Stump-tailed macaques = stm, and Pig-tailed macaques = mne). However, intentional, experimental cross-species transfers of virus remain labeled as being from the original host (e.g., HIV-1 injected into a chimpanzee is labeled HIV-1, not SIVcpz; and SIVsmm injected into a Rhesus macaque is labeled SIVsmm, not SIVmac).

VI-2 Annotated features

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

Feature	Protein	Location	Page
Gag p15 start	Gag	1	381
p15	Gag	135	381
p27	Gag	136	381
p27	Gag	364	383
p2	Gag	365	383
p2	Gag	381	383
p8	Gag	382	383
p8	Gag	433	383
p1	Gag	434	383
p1	Gag	447	383
p6	Gag	448	383
PTAP motif	Gag	458-461	383
PSAP in HIV-2 B, U	Gag	476-479	383
p6 end	Gag	511	384
Gag end	Gag	511	384
Pol p15 start (-1 from Gag)	Pol	1	385
p15	Pol	67	385
protease	Pol	68	385
protease	Pol	166	386
p51 RT start	Pol	167	386
D catalytic site	Pol	276	386
DD catalytic site	Pol	351	387
p51 RT end	Pol	605	388
p15 RNase H start	Pol	606	388
p15 RNase H end	Pol	725	389
p31 Integrase start	Pol	726	389
p31 Integrase end	Pol	1020	391
Pol end	Pol	1020	391
Vif start	Vif	1	392
Vif end	Vif	215	393
Vpx start	Vpx	1	394
Vpx end	Vpx	113	394
Vpr start	Vpr	1	395
Vpr end	Vpr	102	395
Tat start	Tat	1	396
exon 1 end	Tat	99	396
exon 2 start	Tat	100	396
Tat end	Tat	131	396
Rev start	Rev	1	397
exon 1 end	Rev	24	397
exon 2 start	Rev	25	397
Rev end	Rev	108	397
Env start	Env	1	398
signal peptide end	Env	22	398
gp120 start	Env	23	398
V1	Env	113-167	398
V2	Env	169-211	399
V3	Env	311-344	399

Feature	Protein	Location	Page
V3 tip	Env	328-334	400
V4	Env	402-432	400
V5	Env	472-477	400
gp120 end	Env	525	401
gp41 start	Env	526	401
Env end	Env	880	403
Nef start	Nef	1	404
R17Y mutation	Nef	17	404
max HIV-1 similarity	Nef	153-182	404
premature stop in Mac239	Nef	93	404
normal Nef end	Nef	264	405

VI-3 Sequences

Sequences included in the HIV-2/SIV protein alignments.

Name	Accession	Proteins	Author	Reference
MAC.US.x.239	M33262	All	Kestler, H.	<i>Science</i> 248 (4959):1109-1112 (1990)
A.CI.88.UC2	U38293	All	Barnett, S.W.	<i>Virology</i> 222 (1): 257-61 (1996)
A.DE.x.BEN	M30502	All	Kirchhoff, F.	<i>Virology</i> 177 (1):305-311 (1990)
A.DE.x.PEI2_KR_KRCG	U22047	All	Kraus, G.	<i>ARHR</i> 14 (1); 65-77 (1998)
A.GH.x.GH1	M30895	All	Hasegawa, A.	<i>ARHR</i> 5 (6):593-604 (1989)
A.GM.87.D194	J04542	All	Kuehnel, H.	<i>Proc Natl Acad Sci USA</i> 86 (7):2383-2387 (1989)
A.GM.x.ISY_SBL_6669_85	J04498	All	Franchini, G.	<i>Proc Natl Acad Sci USA</i> 86 (7):2433-2437 (1989)
A.GM.x.MCN13	AY509259	All	Schmitz, C.	<i>J Virol</i> 78 (4):2006-2016 (2004)
A.GM.x.MCR35	AY509260	All	Schmitz, C.	<i>J Virol</i> 78 (4):2006-2016 (2004)
A.GW.86.FG_clone_NIHZ	J03654	All	Zagury, J.F.	<i>Proc Natl Acad Sci USA</i> 85 (16):5941-5945 (1988)
A.GW.87.CAM2CG	D00835	All	Tristem, M.	<i>J Gen Virol</i> 72 (PT 3):721-724 (1991)
A.GW.x.ALI	AF082339	All	Reeves, J.D.	<i>J Virol</i> 73 (9); 7795-804 (1999)
A.GW.x.MDS	Z48731	All	Becker	Unpublished
A.IN.07.NNVA	EU980602	All	Gurjar, S.R.	<i>JAIDS</i> 52 (3); 329-35 (2009)
A.IN.95.CRIK_147	DQ307022	All	Santhosh, C.V.	<i>ARHR</i> 24 (10); 1315-7 (2008)
A.SN.85.ROD	M15390	All	Clavel, F.	<i>Nature</i> 324 (6098):691-695 (1986)
B.CI.88.UC1	L07625	All	Barnett, S.W.	<i>J Virol</i> 67 (2):1006-1014 (1993)
B.CI.x.20_56	AB485670	All	Takekawa	Unpublished
B.CI.x.EHO	U27200	All	Rey-Cuille, M.A.	<i>Virology</i> 202 (1):471-476 (1994)
B.GH.86.D205_ALT	X61240	All	Dietrich, U.	<i>Nature</i> 342 (6252):948-950 (1989)
B.JP.01.KR020_IMCJ_KR020_1	AB100245	All	Kusagawa, S.	<i>ARHR</i> 19 (11):1045-1049 (2003)
G.CI.92.Abt96	AF208027	All	Brennan, C.A.	<i>ARHR</i> 13 (5):401-404 (1997)
AB.CM.03.03CM_510_03	EU028345	All	Yamaguchi, J.	<i>ARHR</i> 24 (1):86-91 (2008)
H2_01_AB.CI.90.7312A	L36874	All	Gao	Unpublished
H2_01_AB.JP.04.04JP_NMC307	AB499693	All	Ibe, S.	<i>JAIDS</i> 54 (3):241-247 (2010)
H2_01_AB.JP.07.07JP_NMC716	AB499694	All	Ibe, S.	<i>JAIDS</i> 54 (3):241-247 (2010)
H2_01_AB.JP.08.08JP_NMC842	AB499695	All	Ibe, S.	<i>JAIDS</i> 54 (3):241-247 (2010)
U.FR.96.12034	AY530889	All	Diamond, F.	<i>ARHR</i> 20 (6):666-672 (2004)
MAC.US.x.17EC1	AY033233	All	Anderson, M.G.	<i>Virology</i> 195 (2):616-626 (1993)

Name	Accession	Proteins	Author	Reference
MAC.US.x.17EFR	AY033146	All	Flaherty, M.T.	<i>J Virol</i> 71 (8):5790-5798 (1997)
MAC.US.x.1937	AY611495	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.2065	AY611493	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.251_1A11	M76764	All	Marthas, M.L.	<i>J Med Primatol</i> 18 (3-4):311-9 (1989)
MAC.US.x.251_32H_PJ5	D01065	All	Rud, E.W.	(in) Brown, F, Chanock, RM and Ginsberg, HS(Eds); VACCINES 92: MODERN APPROACHES TO NEW VACCINES INCLUDING PREVENTION OF AIDS: 229-235 ; Cold Spring Harbor Laboratory Press, New York (1992)
MAC.US.x.251_BK28	M19499	All	Hirsch, V.	<i>Cell</i> 49 (3):307-319 (1987)
MAC.US.x.80035	AY611486	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.81035	AY599200	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.85013	AY611490	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.87082	AY600249	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.92050	AY603959	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.92077	AY599201	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.93057	AY611492	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.93062	AY607704	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.95058	AY611494	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.95086	AY607703	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.95112	AY588946	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.96016	AY607701	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.96020	AY611488	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.96072	AY611491	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.96081	AY597209	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.96093	AY611489	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.96114	AY588945	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.96123	AY611487	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.96135	AY607702	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.97009	AY599199	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.97074	AY599198	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.MM142_IVMXX	Y00277	All	Chakrabarti, L.	<i>Nature</i> 328 (6130):543-547 (1987)
MAC.US.x.SMM142_patent	BD131285	All	Alizon, M.	Patent: JP 2002030099-A 2 29-JAN-2002; INSTITUT PASTEUR
MAC.US.x.r80025	AY576480	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MAC.US.x.r90131	AY576481	All	O'Connor, D.H.	<i>J Virol</i> 78 (24):14012-14022 (2004)
MNE.US.82.MNE_8	M32741	All	Kimata, J.T.	<i>J Virol</i> 72 (1):245-256 (1998)
MNE.US.x.MNE027	U79412	All	Kimata, J.T.	<i>J Virol</i> 72 (1):245-256 (1998)
SMM.SL.92.SL92B	AF334679	All	Chen, Z.	<i>J Virol</i> 70 (6):3617-3627 (1996)
SMM.US.x.F236_H4	X14307	All	Hirsch, V.M.	<i>Nature</i> 339 (6223):389-92 (1989)
SMM.US.x.H9	M80194	All	Courgnaud, V.	<i>J Virol</i> 66 (1):414-419 (1992)
SMM.US.x.PBJ14_15	L03295	All	Dewhurst, S.	<i>Nature</i> 345 (6276):636-640 (1990)
SMM.US.x.PBJA	M31325	All	Dewhurst, S.	<i>Nature</i> 345 (6276):636-640 (1990)
SMM.US.x.PBJ_143	M80193	All	Courgnaud, V.	<i>J Virol</i> 66 (1):414-419 (1992)
SMM.US.x.PBJ_6P6	L09212	All	Novembre, F.J.	<i>J Virol</i> 67 (5):2466-2474 (1993)
SMM.US.x.PGM53	AF077017	All	Novembre, F.J.	<i>J Virol</i> 72 (11):8841-8851 (1998)
SMM.US.x.SME543	U72748	All	Hirsch, V.	<i>J Virol</i> 71 (2):1608-1620 (1997)
STM.US.89.STM_37_16	M83293	All	Novembre, F.J.	<i>Virology</i> 186 (2):783-787 (1992)

	Gag p15 start	p15_p27
MAC.US.x.239	MGVRRNSVLSGKKADELEKIRLRPNKGGKMYMLKHVVWAANELDRFLAESLLENKEGCKILSVLAPLVTGSENLKSLYNTVCVWICIHAAEEKVKHTEEAQIVORHLVVEVETGTTETMPKTSRPTAPSSGRGGNYPVQ01.GGNYVHLPLSPRTLNAWVKLIEEKKFGAE	169
A.CI.88.UC2	-A---R-----G---R---I-----R---N---S-----T---D-----F-----L-----D-----KL---A---A---K---N---P---K---F---A---V-----V-----V-----	169
A.DE.x.BEN	-A---R-----V---G---R---I-----K-----S-----R---D-----F-----L-----D-----KLA---A---A---K---N---P---KR-----A---V-----V-----V-----	169
A.DE.x.PE12_KR_KRCG	-A-S---R---V-----G---R---I-----GK-----S-----IT---D-----F-----L-----D---G-----A---A---ADK---S---A---P-----R---F---TG---I---V-----V---D-----	170
A.GH.x.GH1	-A---R-----V---S---R---I-----K-----S-----K---E-----F-----L-----D-----KLA---A---A---K---NI---P---K---F---A---I---V-----V---D-----	169
A.GM.87.D194	-A---R-----V---G---R---I-----T-----S-----I---E-----F-----T-----L-----D---R---G-----E---A---K---NI---P---K---F---V-----V-----V-----	169
A.GM.x.ISY_SBL_6669_85	-AK---R-----G---R---I-----T-----S-----I---E-----F-----T-----L-----D---R---G-----E---A---K---NI---P---K---F---V-----V-----V-----	169
A.GM.x.MCN13	-A---K---T-----G---R---I-----S-----R---T---G-----F-----D---G-----AA---A---K---N---P---K---F---V---T---V-----V-----V-----	169
A.GM.x.MCR35	-A---K---T-----G---R---I-----S-----R---T---D-----F-----D---G-----AA---A---K---N---P---K---F---V---T---V-----V-----V-----	169
A.GM.86.FG_clone_NIHZ	-A---R-----V---G---K---I-----S-----R---T---D-----F-----D---G-----AA---A---K---N---P---K---F---V---A---T---I---G-----V-----	167
A.GW.87.CAM2CG	-A---R-----V---G---C---I---V---E---Y---T---S---Q---H---T---M-----F-----Y---L---D---K---A---S---A---D---K---K---A---P-----A---V---D-----	164
A.GW.x.ALT	-A---R---R-----R---G---C---I---V---E---Y---T---S---Q---H---T---M-----F-----Y---L---D---K---A---S---A---D---K---K---A---P-----A---V---D-----	164
A.GW.x.MDS	-A-S---R---V---G---R---I-----S-----R---K---E-----F-----V---V-----D---G---R---AL---A---A---K---D---P---K---S-----T---V---D-----V-----	169
A.IN.07.NNVA	-AK---R-----R---G---RV---T-----S-----T---D-----F-----F-----D---K---A---S---A---D---K---K---A---P---EQ---F---V---A---T---V---D-----	169
A.IN.95.CR1K_147	-AKS---R-----R---G---R---I---K---T-----S-----T---D-----F-----F-----D---K---A---S---A---D---K---K---A---P---EG---F---R---T---V---V-----	169
A.SN.85.ROD	-A---R-----R---G---R---I---K---T-----S-----T---D---M-----F-----F-----D---K---A---S---A---D---K---K---A---P---EG---F---R---T---V---V-----	169
B.CI.88.UC1	-A-S---T---V---G---R---C---II---V---Y---T---S---H---T---M-----F-----Y---L---D---K---A---S---A---D---K---K---A---P-----A---M-----V-----	164
B.CI.x.20_56	-A-G---T---V---G---R---C---I---V---E---Y---R---GS---R---RK---G-----F-----Y---L---D---K---A---S---A---D---K---K---A---P-----A---V---D-----	164
B.CI.x.EHD	-A-G---T---V---G---R---C---I---V---E---Y---R---GS---R---RK---G-----F-----Y---L---D---K---A---S---A---D---K---K---A---P-----A---S-----V-----	164
B.GH.86.D205_ALT	-A-G---T---V---G---C---I---V---Y---T---S---H---T---E-----F---I---F---L---D---K---A---S---A---D---K---A---NK---P-----L---A-----V-----	164
B.JP.01.KR020_IMCJ_KR020_1	-A-G---T---V---G---C---I---V---Y---T---S---H---T---E-----F---Y---L---V---D---D---K---A---S---A---AN---K---A---K---P---R---V---A---T-----V-----	164
G.CI.92.Abt96	-A-S---T---V---G---R---I---V---Y---T---S-----S-----L---Q-----V-----KV---A---I---P-----V-----V-----	169
AB.CM.03.03CM_510_03	-A-S---T---V---G---R---I---V---Y---T---S-----S-----K-----F---F---L---D---K---A---S---A---AD---A---K---TA---NK---P-----A---V-----V-----	163
H2_01_AB.CI.90.7312A	-A-G---T---V---G---R---I---N---S---T---M-----F---L---K---D---KLA---S---D---K---T---ADK---A---T---R---S---V---A---V-----V---L-----	164
H2_01_AB.JP.04.04JP_NMC307	-A-S---T---V---GAR---R---I---X---K---I---V-----F---L---Q---D---KLA---S---AAA---K---A---TADK---A---T---X---V---A---V-----V-----	164
H2_01_AB.JP.07.07JP_NMC716	-A-S---T---V---GAR---R---I---T---M---V-----F---L---Q---D---KLA---S---AAA---K---A---TADK---A---T---X---V---A---V-----V-----	164
H2_01_AB.JP.08.08JP_NMC842	-A-G---T---V---G---RX---I---T---M---V-----F---X---XL---K---D---X---KLA---S---AAV---K---A---TANK---A---T---X---V---A---V---X---XX-----	164
U.FR.96.12034	-A---#---V---V-----A-----L-----A---E---A---K---A---X---P---GR-----A-----V-----V-----	168
MAC.US.x.17EC1	-----X-----X-----V-----M-----	169
MAC.US.x.17EFR	-----X-----X-----V-----M-----	169
MAC.US.x.1937	-----X-----X-----V-----M-----	169
MAC.US.x.2065	-----X-----X-----V-----M-----	169
MAC.US.x.251_1A11	-----A---I-----	169
MAC.US.x.251_32H_PJ5	-----G-----	169
MAC.US.x.251_BK28	-----G-----	169
MAC.US.x.80035	-----X-----	169
MAC.US.x.81035	-----X-----	169
MAC.US.x.85013	-----R-----	169
MAC.US.x.87082	-----R-----	169
MAC.US.x.92050	-----R-----	169
MAC.US.x.92077	-----R-----	169
MAC.US.x.93057	-----R-----	169
MAC.US.x.93062	-----R-----	169
MAC.US.x.95058	-----X-----	169
MAC.US.x.95086	-----X-----	169
MAC.US.x.95112	-----V-----	169
MAC.US.x.96016	-----L-----	169
MAC.US.x.96020	-----L-----	169
MAC.US.x.96072	-----L-----	169
MAC.US.x.96081	-----M-----	169
MAC.US.x.96093	-----M-----	169
MAC.US.x.96114	-----M-----	169
MAC.US.x.96123	-----M-----	169
MAC.US.x.96135	-----M-----	169
MAC.US.x.97009	-----T-----	169
MAC.US.x.97074	-----T-----	169
MAC.US.x.MM142_IVMXX	-----G-----M-----A-----F-----T-----	169
MAC.US.x.SNM142_patent	-----G-----M-----A-----F-----T-----	169
MAC.US.x.r80025	-----X-----M-----A-----F-----T-----	169
MAC.US.x.r90131	-----X-----M-----A-----F-----T-----	169
MNE.US.82.MNE_8	-----G-----A-----A-----V-----T-----	169
MNE.US.x.MNE027	-----G-----A-----A-----V-----T-----	169
SMM.SL.92.SL92B	-A-G---E---V---G-R---II---R---S---D---R---A---M---FS---V---L---M---D---KT---S---S---A---KL---AQ---P---V---N---T---V-----	167
SMM.US.x.F236_H4	-A---E---V---G-R---II---R---S---D---R---A---M---FS---V---L---M---D---KT---S---S---A---KL---AQ---P---V---N---T---V-----	169
SMM.US.x.H9	-A---E---V---G-R---II---R---S---D---R---A---M---FS---V---L---M---D---KT---S---S---A---KL---AQ---P---V---N---T---V-----	169
SMM.US.x.PB114_15	-A---G---R---K---I---R---L---ADR---A---P---V---T---V-----	166
SMM.US.x.PBJA	-A---G---R---K---I---R---L---ADR---A---P---K---T---V-----	169
SMM.US.x.PBJ_143	-A---G---R---K---I---R---L---ADR---A---P---K---T---V-----	169
SMM.US.x.PBJ_6P6	-A---G---R---K---I---R---L---ADR---A---P---K---T---V-----	169
SMM.US.x.PGMS3	-A---G---R---K---I---R---L---ADR---A---P---K---T---V-----	169
SMM.US.x.SME543	-A---G---R---K---I---R---L---ADR---A---P---K---T---V-----	169
STM.US.89.STM_37_16	-A-S---V---G---S---IT---E---F-----V---K---ANK---A---P---V-----V-----	169

MAC.US.x.239	VVPGFQALSEGCTPYDINQMLNCVGDHQAAMQIIRDIINEEAADWLQHPQP.AP0Q.GQLREPSGSDIAGTSSVDEQIQWVQRQNPPIVGNIIYRRWIQLGLQKVRMYPNTNILDVKQGPKEPFQSYVDTRFYKSLRAEQTDAAVKNMNTQTLLIQNANPDCKLVLK	336
A.CI.88.UC2	-----Q-----E-----A-I-G-LP-A-D-R-----T-----P-V-----I-----V-----S-----A-P-----V-----	337
A.DE.x.BEN	-----S-I-G-LP-A-D-R-----T-----P-V-----I-----K-----I-----	337
A.DE.x.PE12_KR_KRCG	-----E-----V-----I-G-LP-A-D-R-----T-----F-A-V-----I-----	337
A.GH.x.GH1	-----D-----A-I-G-LP-A-D-R-----T-----P-V-----I-----	338
A.GM.87.D194	-----E-----V-----I-G-LP-A-D-R-----T-----E-----E-----V-----I-----S-----	337
A.GM.x.ISY_SBL_6669_85	-----E-----V-----I-G-LP-A-D-R-----T-----E-----E-----V-----I-----S-----	336
A.GM.x.MCN13	-----E-----V-----I-G-LP-A-D-R-----T-----F-P-V-----I-----	337
A.GM.x.MCR35	-----E-----V-----I-G-LP-A-D-R-----T-----F-P-V-----I-----	337
A.GW.86.FG_c1one_NIHZ	-----E-----VA-I-G-LP-A-D-R-----T-----F-P-V-----I-----IN-----	335
A.GW.87.CAM2CG	-----E-----AN-I-G-LP-A-D-R-----T-----F-A-V-----I-----S-----	337
A.GW.x.ALI	-----E-----VA-I-G-LP-A-D-R-----T-----F-PR-V-----I-----	337
A.GW.x.MDS	-----E-----A-I-G-LP-A-D-R-----T-----F-A-V-----I-----	337
A.IN.07.NNVA	-----E-----V-----I-G-LP-A-D-R-----T-----E-----F-A-V-----I-----	337
A.IN.95.CR1K_147	-----R-----E-----A-I-G-LP-A-D-R-----T-----F-A-V-----I-----	337
A.SN.85.ROD	-----E-----E-----V-----I-G-LP-A-D-R-----T-----E-----F-P-V-----I-----	337
B.CI.88.UC1	-----E-----Q-----I-G-LP-A-D-R-----T-----A-----V-----I-----	332
B.CI.x.20_56	-----E-----Q-----I-G-LP-A-D-R-----T-----A-----V-----I-----S-----	332
B.CI.x.EHO	-----E-----E-----Q-----S-G-MP-A-D-R-----T-----E-----P-----V-----I-----	332
B.GH.86.D205_ALT	-----E-----E-----Q-----S-G-MP-A-D-R-----T-----E-----A-----V-----I-----	332
B.JP.01.KR020_IMCJ_KR020_1	-----I-----E-----D-----Q-----S-G-MP-A-D-R-----T-----E-----P-----V-----I-----	332
G.CI.92.Abt96	-----L-----E-----E-----X-----Q-G-PP-A-X-T-----TIE-TH-----X-----X-----X-----X-----P-----	337
AB.CM.03.03CM_510_03	-----E-----D-----Q-----S-G-LP-A-D-R-----T-----E-----G-----I-----I-----	331
H2_01_AB.CI.90.7312A	-----I-----E-----V-----E-----Q-----V-G-LP-A-D-R-----T-----E-----G-S-----	332
H2_01_AB.JP.04.04JP_NMC307	-----I-----E-----V-----E-----Q-----V-G-LP-A-X-D-R-----TPE-----G-S-----	332
H2_01_AB.JP.07.07JP_NMC716	-----I-----E-----V-----E-----Q-----I-G-LP-A-D-R-----T-----E-----G-SV-----	332
H2_01_AB.JP.08.08JP_NMC842	-----I-----X-----X-----E-----X-----V-----X-----E-----Q-----I-GXLP-AX-D-R-----T-----E-----G-SV-----	332
U.FR.96.12034	-----E-----E-----D-----T-----NQ-G-PP-A-R-----T-----E-----P-----I-----	336
MAC.US.x.17EC1	-----	336
MAC.US.x.17EFR	-----C-----V-----	336
MAC.US.x.1937	-----	336
MAC.US.x.2065	-----	336
MAC.US.x.251_1A11	-----	336
MAC.US.x.251_32H_PJ5	-----	336
MAC.US.x.251_BK28	-----	336
MAC.US.x.80035	-----	336
MAC.US.x.81035	-----	336
MAC.US.x.85013	-----A-----V-----	336
MAC.US.x.87082	-----	336
MAC.US.x.92050	-----	336
MAC.US.x.92077	-----	336
MAC.US.x.93057	-----I-----V-----	336
MAC.US.x.93062	-----X-----	336
MAC.US.x.95058	-----I-----X-----	336
MAC.US.x.95086	-----	336
MAC.US.x.95112	-----	336
MAC.US.x.96016	-----	336
MAC.US.x.96020	-----	336
MAC.US.x.96072	-----	336
MAC.US.x.96081	-----	336
MAC.US.x.96093	-----	336
MAC.US.x.96114	-----	336
MAC.US.x.96123	-----	336
MAC.US.x.96135	-----	336
MAC.US.x.97009	-----	336
MAC.US.x.97074	-----	336
MAC.US.x.MM142_IVMXX	-----S-----L-----Q-----	336
MAC.US.x.SNM142_patent	-----S-----L-----Q-----	336
MAC.US.x.r80025	-----	336
MAC.US.x.r90131	-----	336
MNE.US.82.MNE_8	-----I-----E-----E-----Q-----	336
MNE.US.x.MNE027	-----E-----E-----E-----Q-----	336
SMM.SL.92.SL92B	-----E-----E-----E-----RGQQ-A-PA-G-----	337
SMM.US.x.F236_H4	-----E-----E-----E-----G-LP-A-R-----	337
SMM.US.x.H9	-----X-----E-----E-----X-----G-IP-A-X-----TX-----X-----X-----S-----	337
SMM.US.x.PB114_15	-----E-----E-----E-----G-IP-P-R-----T-----	337
SMM.US.x.PBJA	-----E-----E-----E-----G-IP-P-R-----T-----	337
SMM.US.x.PBJ_143	-----X-----E-----E-----X-----G-IP-A-X-----TX-----X-----X-----	337
SMM.US.x.PBJ_6P6	-----E-----E-----E-----G-IP-P-R-----T-----S-----	337
SMM.US.x.PGMS3	-----E-----E-----E-----G-IP-A-D-R-----T-----	337
SMM.US.x.SME543	-----E-----E-----E-----G-LP-A-R-----T-----	337
STM.US.89.STM_37_16	-----L-----E-----KE-----M-----PG-LP-A-----PE-----V-----T-----A-PS-----R-----	338

	p27_p2	p2_p8	p8_p1	p1_p6	PTAP motif	PSAP in HIV-2 B, U	
MAC.US.x.239	GLGVNPTLEEMLTACQGVGGQKARLMAEALKEALAPVPIPFMAAQQRGPRKPIKWCNCGKEGHSARQCRRAPRRQGCWKCGKMDHVMACPDRAQFLGLGPWGKKPRNFPMQVHOGMLMPTAPPEDPAVDLLKNYMLQKQORE	482	
A.CI.88.UC2	480	
A.DE.x.BEN	480	
A.DE.x.PET2_KR_KRCC	480	
A.GH.x.GH1	481	
A.GM.87.D194	480	
A.GM.x.ISY_SBL_6669_85	479	
A.GM.x.MCN13	480	
A.GM.x.MCR35	478	
A.GW.86.FG_clone_NIHZ	480	
A.GW.87.CAM2CG	480	
A.GW.x.ALI	480	
A.GW.x.MDS	480	
A.IN.07.NNVA	480	
A.IN.95.CR1K_147	480	
A.SN.85.ROD	480	
B.CI.88.UC1	497	
B.CI.x.20_56	496	
B.CI.x.EH0	495	
B.GH.86.D205_ALT	497	
B.JP.01.KR020_IMCJ_KR020_1	497	
G.CI.92.Abt96	485	
AB.CM.03.03CM_510_03	496	
H2_01_AB.CI.90.7312A	497	
H2_01_AB.JP.04.04JP_NMC307	497	
H2_01_AB.JP.07.07JP_NMC716	497	
H2_01_AB.JP.08.08JP_NMC842	497	
U.FR.96.12034	497	
MAC.US.x.17EC1	482	
MAC.US.x.17EFR	482	
MAC.US.x.1037	482	
MAC.US.x.2065	482	
MAC.US.x.251_1A11	482	
MAC.US.x.251_32H_PJ5	482	
MAC.US.x.251_BK28	482	
MAC.US.x.80835	482	
MAC.US.x.81035	482	
MAC.US.x.85013	482	
MAC.US.x.87082	482	
MAC.US.x.92050	482	
MAC.US.x.92077	482	
MAC.US.x.93057	482	
MAC.US.x.93062	482	
MAC.US.x.95058	482	
MAC.US.x.95086	482	
MAC.US.x.95112	482	
MAC.US.x.96016	482	
MAC.US.x.96020	482	
MAC.US.x.96072	482	
MAC.US.x.96081	482	
MAC.US.x.96093	482	
MAC.US.x.96114	482	
MAC.US.x.96123	482	
MAC.US.x.96135	482	
MAC.US.x.97009	482	
MAC.US.x.97074	482	
MAC.US.x.MM142_IVMXX	482	
MAC.US.x.SNM142_patent	482	
MAC.US.x.r80025	482	
MAC.US.x.r90131	482	
MNE.US.82.MNE_8	482	
MNE.US.x.MNE027	482	
SMM.SL.92.SL92B	483	
SMM.US.x.F236_H4	483	
SMM.US.x.H9	483	
SMM.US.x.PB114_15	483	
SMM.US.x.PBJA	483	
SMM.US.x.PBJ_143	483	
SMM.US.x.PBJ_6P6	483	
SMM.US.x.PGMS3	483	
SMM.US.x.SME543	483	
STM.US.89.STM_37_16	483	

p6 end
Gag end†

MAC.US.x.239 KORESREKPYKE.....VTEDLLHLNLSLFGGDQ*
A.CI.88.UC20-R---VTEDLLHLEQGETPRRE.A-----K-
A.DE.x.BEN0-R---VTEDLLHLEQGETPHRE.E-----K-
A.DE.x.PET2_KR_KRCGQKMR---VTEDLLHLEQGETPHKE.A-----K-
A.GH.x.GH10-R---VTEDLLHLEQGKAPHRE.A-----K-
A.GM.87.D1940-R---VTEDLLHLEQGETPHRG.A-----K-
A.GM.x.ISY_SBL_6669_850-OR---VTEDLLHLEQGETPHRE.T-----N-
A.GM.x.MCN130-R---VTEDLLHLEQGEAPCRE.T-----K-
A.GM.x.MCR350-R---VTEDLLHLEQGEAPYRE.T-----K-
A.GW.86.FG_clone_NIHZ0-R---VTEDLLRFEQGETPCRE.T-----K-
A.GW.87.CAM2CG0-OR---VTEDLLHLEQGEAPCRE.T-----K-
A.GW.x.ALI0-R---VTEDLLHLEQGETPHKE.R-----K-
A.GW.x.MDS0-R---VTEDLLHLEQGKTPCKE.T-----T-
A.IN.07.NNVA0-R---VTEDLLHLEQGEAPCKE.T-----K-L-
A.IN.95.CR1K_1470-R---VTEDLLHLEQGGAPCKE.M-----K-
A.SN.85.ROD0-R---VTEDLLHLEQGETPYREPP-----K-
B.CI.88.UC1R-----E-
B.CI.x.20_56R-----E-
B.CI.x.EHDN-R-----E-
B.GH.86.D205_ALT-R-----E-
B.JP.01.KR020_IMCJ_KR020_1G-----E-
G.CI.92.Abt96Q-N-----XX---S---D-
AB.CM.03.03CM_510_03-R-----E-
H2_01_AB.CI.90_7312AKQGR-----E-
H2_01_AB.JP.04_04JP_NMC307KQGR-----E-
H2_01_AB.JP.07_07JP_NMC716KQ-R-----E-
H2_01_AB.JP.08_08JP_NMC842KQ-R-----E-
U.FR.96.12034N-R-----N-
MAC.US.x.17EC1-----E-
MAC.US.x.17EFR-X-----E-
MAC.US.x.1937-----E-
MAC.US.x.2065-----E-
MAC.US.x.251_1A11M-----E-
MAC.US.x.251_32H_PJ5-----E-
MAC.US.x.251_BK28-----E-
MAC.US.x.80835-----E-
MAC.US.x.81035K-----E-
MAC.US.x.85013-----E-
MAC.US.x.87082-----E-
MAC.US.x.92050-----E-
MAC.US.x.92077-----E-
MAC.US.x.93057-----E-
MAC.US.x.93062-X-----E-
MAC.US.x.95058-----E-
MAC.US.x.95086-----E-
MAC.US.x.95112-----E-
MAC.US.x.96016-----E-
MAC.US.x.96020-----E-
MAC.US.x.96072-----E-
MAC.US.x.96081-----E-
MAC.US.x.96093-K-----E-
MAC.US.x.96114-----E-
MAC.US.x.96123-----E-
MAC.US.x.96135-----E-
MAC.US.x.97009-----E-
MAC.US.x.97074-----E-
MAC.US.x.MM142_IVMXXG-----E-
MAC.US.x.SNM142_patentG-----E-
MAC.US.x.r80025-----E-
MAC.US.x.r90131-----E-
MNE.US.82.MNE_8KR-----E-
MNE.US.x.MNE027KR-----A---E-
SMM.SL.92.SL92BR-SR-----E-
SMM.US.x.F236_H4N-R-----E-
SMM.US.x.H9N-R-----E-
SMM.US.x.PB114_15N-R-----E-
SMM.US.x.PBJAN-R-----E-
SMM.US.x.PBJ_143N-R-----E-
SMM.US.x.PBJ_6P6N-R-----E-
SMM.US.x.PGMS3N-R-----E-
SMM.US.x.SME543N-R-----E-
STM.US.89.STM_37_16KT-----E-

510
522
522
522
523
523
521
521
521
520
522
522
522
522
522
523
523
522
521
521
522
522
522
522
522
522
522
522
511
511
511
511
511
507
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
511
507
511
511
511
511
511
508
508
508
508
508
508
508
508
508
508
508

	Pol p15 start (-1 from Gag)	p15 protease	
MAC.US.x.239	FFRPWSMGKEAPQFPHGSSASGADANCS	PRGPPSCGSAKELHAVGQAAERKAERKOREALQ.GGDRGFAAPQFSLWRRPVVTAHIEGQVPEVLLDGTADDISVTGIELGPHYTPKIVGGIGGFINTKEYKNVEI
A.CI.88.UC2	---DGLT---	L-R-P-S---	T-ST-SRSGS-PVR-IF-A-EK-GAEGETIOGGDG.RLTAPRAG-DTS-R---
A.DE.x.BEN	---VGPT---	S-L-RDP-P---	T-STSGRS-S-TVG-IV-AREK-GAEGETIOGGDG.GLAAPRAE-DTS-R---
A.DE.x.PET2	---D-P---	S-L-RDP-PA---	T-ST-SR-SRP-R-VL-AREE-AENETIOGGDR.GLTAPRTR-DTT-R---
A.GH.x.GH1	---DG---	L-R-P-S---	T-ST-SRS-S-TGIV-A-ER-GAEGETIOGGDG.RLTAPRAGKSTS-R---
A.GM.87.D194	---DGPT---	L-R-P-S---	T-ST-NRS-S-PVG-IV-AREK-AEGETIOGGDG.GLTAPRAG-D-P-R---
A.GM.x.ISY	---A-T---	L-R-PKFA-NT-ST-N-S-S	PTG-V-AREKT-AETKTIOQRSDR.GLAASRAR-DTT-RD-L-T---
A.GM.x.MCN13	---DGPP---	L-R-P-PA-NT-ST-SRS-S	PTG-IV-ARKKTK-TERETIOQRSDR.GLTAPRAG-GTM-RD-L-LK---
A.GM.x.MCR35	---DGPP---	L-R-P-PA-NT-ST-SRS-S	PTG-IV-ARKKTK-TERETIOQRSDR.GLTAPRAG-STI-RD-L-LK---
A.GW.86.FG	---D-PL---	G-L-R-P-PA-NT-ST-I-S-S	PTG-IV-ARKK-KGAERETIOGSDR.GLTAPRAG-DTM-D-L---
A.GW.87.CAM2CG	---D-PL---	R-P-ST-NT-ST-I-S-S	TG-IV-AREK-GAETETIOGGDR.GLTAPRTR-GPM-DN-L---
A.GW.x.ALI1	---A-P---	S-L-RNP-SA-INT-ST-SRA-S	P-GAVY-A-EK-KAERETIOGGDG.GLTAPRAG-DTT-R-L---
A.GW.x.MDS	---D-PL---	L-R-PGSA-NT-ST-SRS-S	PTG-IV-AREK-GAERETIOGDR.GLAAPRAGKDTM-DN---
A.IN.07.NNVA	---DGPP---	L-R-PNPA-NT-ST-N-S-S	PIG-IV-AREK-GAERETIOGGDR.GPTAPRAG-ST-R-N-RL---
A.IN.95.CR1K.147	---DKPP---	S-L-R-P-ST-NT-SISN-S-S	PTG-IV-AREK-GAERETIOGGDR.GPTAPRAG-GT-R-N-L---
A.SN.85.ROD	---TGPL---	L-R-P-SA-T-ST-S-S-S	TG-IV-AREKT-AERETIOGSDR.GLTAPRAGDTI-GATN-L---
B.CI.88.UC1	---VRTL---	S-L-DP-S-TI-T-D-	R-HDTSGDITIC-PC-SSSGDAEKLHEDGETAEREP-T---
B.CI.x.20.56	---IRTL---	S-L-P-EG-TI-T-D-	NT-HDTSGGDICTPC-SSSGDAEKLHAAREAREP-T---
B.CI.x.EHO	---VRPL---	S-L-#GP-DS-I-A-DE	IRHDTSGDICTPC-SSRGDAKLHATREAEAGE-T---
B.GH.86.D205	---VRTL---	S-L-DP-S-TI-T-DE	R-HDTSGDITIC-PC-SSSGDAEKLHADGETTEREP-T---
B.JP.01.KR020	---I-PL---	S-L-CSP-TP-DT-I-TTNE	REHDTSGGNAIC-PC-SSSGDVEGLHAAREKAERG-G-T---
G.CI.92.ABT96	---V-TL---	S-L-DP-P-S-SISTTD-SRPTE	---X-EE-KG.....AEKQT---XX-L---
AB.CM.03.03CM.510.03	---VRTL---	S-L-DPG-S-TI-T-DE	RXHDTSGDITIC-PC-SSSGDAEKLHADGETTEREP-T---
H2.01.AB.CI.90.7312A	---A-TL---	S-L-P-DS-I-A-DEH-R	ODTSGDITIC-PC-SSSGDAEKLHATREAEAREAG-L---
H2.01.AB.JP.04.04JP.NMC307	---A-TL---	S-L-R-P-GG-I-A-D-H	REODTSGDITIC-PC-SSSGDAKLHATGEAAREEA-T---
H2.01.AB.JP.07.07JP.NMC716	---V-TL---	S-L-R-P-GG-I-A-D-N	REODISESDITIC-PC-SSSGDAKLHATGEATEREA-T---
H2.01.AB.JP.08.08JP.NMC842	---X-TL---	S-L-R-P-GG-I-T-D-H	REODISESDITIC-PC-SSSGDAKLHATGEAAREEA-T---
U.FR.96.12034	---A-PL---	K-S-L-P-PA-T-P-SPSSRT-S	DPDPSPPSSSGS-KELHATGQE.....TKGE-TI---
MAC.US.x.17EC1	---X-X---		
MAC.US.x.17EFR			
MAC.US.x.1937			
MAC.US.x.2065			
MAC.US.x.251.1A11			
MAC.US.x.251.32H.PJ5			
MAC.US.x.251.BK28			
MAC.US.x.80035			
MAC.US.x.81035			
MAC.US.x.85013			
MAC.US.x.87082			
MAC.US.x.92050			
MAC.US.x.92077			
MAC.US.x.93057			
MAC.US.x.93062			
MAC.US.x.95058			
MAC.US.x.95086			
MAC.US.x.95112			
MAC.US.x.96016			
MAC.US.x.96020			
MAC.US.x.96072			
MAC.US.x.96081			
MAC.US.x.96093			
MAC.US.x.96114			
MAC.US.x.96123			
MAC.US.x.96135			
MAC.US.x.97009			
MAC.US.x.97074			
MAC.US.x.MM142			
MAC.US.x.SNM142_patent			
MAC.US.x.r80025			
MAC.US.x.r90131			
MNE.US.82.MNE.8			
MNE.US.x.MNE027			
SMM.SL.92.SL92B			
SMM.US.x.F236.H4			
SMM.US.x.H9			
SMM.US.x.PBJ14.15			
SMM.US.x.PBJA			
SMM.US.x.PBJ.143			
SMM.US.x.PBJ.6P6			
SMM.US.x.PGMS3			
SMM.US.x.SME543			
STM.US.89.STM.37.16			

	protease_p51 RT start		D catalytic site										
MAC.US.x.239	NLLTALGMSLNFP	IAKVEPVKALPKG	DGPKLQWPLSK	EIKVAIVEICEK.MEKDGL	EEAAPTNPYNTPTFAIKKKDK	NKWRMLDFRELN	RVTDQFTEVQLG	IPHPAGLAKRKRITVLDIGDAYFSIPLDDEEF	FRQYTAFTLPSV	NNAEPGKRYI	KVLPQGWK	GSP	324
A.CI.88.UC2	-I	-L	-V	-RI	-I	-IT	-R	-T	-VE	-K	-E	-E	324
A.DE.x.BEN	-I	-L	-V	-I	-I	-T	-R	-T	-E	-K	-E	-E	324
A.DE.x.PET2	-I	-L	-V	-D	-I	-I	-VR	-T	-E	-K	-RE	-E	324
A.GH.x.GH1	-I	-L	-V	-I	-I	-T	-R	-T	-E	-K	-E	-E	324
A.GM.87.D194	-I	-AT	-L	-V	-LD	-I	-T	-R	-T	-E	-K	-RE	324
A.GM.x.ISY	-I	-L	-V	-I	-I	-T	-QR	-TR	-E	-K	-RE	-E	324
A.GM.x.MCN13	-I	-L	-V	-I	-I	-IM	-R	-T	-E	-K	-RE	-E	324
A.GM.x.MCR35	-I	-L	-V	-I	-I	-IM	-R	-T	-E	-K	-RE	-E	324
A.GW.86.FG	-I	-L	-V	-I	-I	-IM	-R	-T	-E	-K	-RE	-E	324
A.GW.87.CAM2CG	-I	-L	-V	-I	-I	-IM	-R	-T	-E	-K	-RE	-E	324
A.GW.x.AL1	-I	-L	-V	-I	-IE	-R	-R	-T	-E	-K	-RE	-E	324
A.GW.x.MDS	-I	-L	-V	-I	-I	-IM	-R	-T	-E	-K	-RE	-E	324
A.IN.07.NNVA	-I	-L	-V	-I	-I	-K	-R	-T	-E	-K	-RE	-E	324
A.IN.95.CR1K.147	-I	-V	-L	-V	-I	-I	-IM	-R	-T	-E	-K	-RE	324
A.SN.85.R0D	-I	-L	-V	-I	-I	-IM	-R	-T	-E	-K	-RE	-E	324
B.CI.88.UC1	-I	-NT	-T	-V	-I	-K	-IR	-L	-K	-E	-E	-E	324
B.CI.x.20.56	-I	-NT	-T	-V	-I	-K	-IR	-L	-K	-E	-E	-E	324
B.CI.x.EH0	-I	-NS	-T	-V	-RI	-Q	-E	-IR	-L	-K	-E	-E	324
B.GH.86.D205	-I	-NT	-T	-V	-I	-K	-IR	-R	-L	-K	-E	-E	324
B.JP.01.KR020	-I	-NT	-T	-V	-I	-K	-IR	-L	-K	-E	-E	-E	324
G.CI.92.ABT96	-I	-VK	-T	-V	-I	-K	-IR	-L	-K	-E	-E	-E	324
AB.CM.03.03CM.510.03	-I	-NTX	-T	-V	-XX	-EX	-X	-X	-IR	-RX	-L	-XK	324
H2.01.AB.CI.90.7312A	-I	-NT	-T	-V	-I	-K	-IR	-L	-K	-E	-E	-E	324
H2.01.AB.JP.04.04JP.NMC307	-I	-XSX	-M	-L	-X	-Q	-IR	-L	-L	-E	-E	-E	324
H2.01.AB.JP.07.07JP.NMC716	-I	-S	-M	-L	-X	-Q	-IR	-L	-L	-E	-E	-E	324
H2.01.AB.JP.08.08JP.NMC842	-I	-S	-M	-L	-X	-Q	-IRX	-L	-L	-E	-E	-E	324
U.FR.96.12034	-E	-G	-EAL	-IGG	-R	-L	-E	-K	-E	-E	-E	-E	324
MAC.US.x.17EC1													324
MAC.US.x.17EFR													324
MAC.US.x.1937													324
MAC.US.x.2065													324
MAC.US.x.251.1A11													324
MAC.US.x.251.32H.PJ5													324
MAC.US.x.251.BK28													324
MAC.US.x.80035													324
MAC.US.x.81035													324
MAC.US.x.85013													324
MAC.US.x.87082													324
MAC.US.x.92050													324
MAC.US.x.92077													324
MAC.US.x.93057													324
MAC.US.x.93062													324
MAC.US.x.95058													324
MAC.US.x.95086													324
MAC.US.x.95112													324
MAC.US.x.96016													324
MAC.US.x.96020													324
MAC.US.x.96072													324
MAC.US.x.96081													324
MAC.US.x.96093													324
MAC.US.x.96114													324
MAC.US.x.96123													324
MAC.US.x.96135													324
MAC.US.x.97009													324
MAC.US.x.97074													324
MAC.US.x.MM142													324
MAC.US.x.SNM142.patent													324
MAC.US.x.r80025													324
MAC.US.x.r90131													324
MNE.US.82.MNE.8													320
MNE.US.x.MNE027													320
SMM.SL.92.SL92B	-I	-A	-V	-Y	-V	-YT	-K	-E	-M	-R	-Q	-T	317
SMM.US.x.F236.H4	-M	-L	-X	-I	-T	-E	-R	-I	-I	-E	-R	-I	320
SMM.US.x.H9	-M	-L	-X	-I	-T	-E	-R	-I	-I	-E	-R	-I	320
SMM.US.x.PB114.15	-M	-L	-X	-I	-T	-E	-R	-I	-I	-E	-R	-I	320
SMM.US.x.PB1A	-M	-L	-X	-I	-T	-E	-R	-I	-I	-E	-R	-I	320
SMM.US.x.PB143	-M	-L	-X	-I	-T	-E	-R	-I	-I	-E	-R	-I	320
SMM.US.x.PB1.6P6	-M	-L	-X	-I	-T	-E	-R	-I	-I	-E	-R	-I	320
SMM.US.x.PGMS3	-M	-L	-V	-I	-T	-E	-R	-I	-I	-E	-R	-I	320
SMM.US.x.SME543	-M	-L	-V	-I	-T	-E	-R	-I	-I	-E	-R	-I	320
STM.US.89.STM.37.16	-V	-T	-I	-E	-K	-E	-K	-I	-R	-V	-G		320

		DD catalytic site																						
MAC.US.x.239	AIFQYTMRHVLEPFKAMPDVLVQYMDDILIASDRDTELDHDRVVLQSKELLNSIGFSTPEEKFKQDPPFQWMGYELWPTKWLKQKIELPQREIWTVDNIQKLVGVLNMAAQIYPGIKTKHLRLIRGKMTLVEEVQWTEMAEAEYENKIIISOEQEGCYQEGKPLEA	494																						
A.CI.88.UC2	-----QI-----I-I-----G-K-L--GL-----Q-K-----I-----L-L-R--D---H-E-E-----	507																						
A.DE.x.BEN	-----L-I-----G-K-L--GL-D-----C-----LQ-KDI-----S-----L-L-----Y-E-E-----	507																						
A.DE.x.PET2	-----H-Q-----I-I-----G-T-L--GL-D-----YK-C-----LQ-K-V-----L-L-----Y-E-E-----	507																						
A.GH.x.GH1	-----H-Q-----I-I-----G-K-L--GL-D-----L-----LQ-K-I-----K-----L-L-----Y-E-E-----	507																						
A.GW.87.D194	-----FM-QI-----I-I-----G-K-L--GL-D-----L-----LQ-K-I-----K-----L-L-----S-E-E-----	507																						
A.GW.x.ISY	-----O-----I-I-----K-L--GL-D-----Y-----LQ-K-V-----K-----L-L-----H-E-E-----	507																						
A.GM.x.MCN13	-----O-----L-III--K-L--NL-D-----YH-----LQ-KDV-----R-K-----L-L-----H-E-E-----	507																						
A.GM.x.MCR35	-----O-----L-III--K-L--NL-D-----YH-----LQ-KDV-----R-K-----L-L-----H-E-E-----	507																						
A.GW.86.FG	-----O-----E-III--K-L--GL-D-----YR-----LQ-K-V-----L-L-----R-K-----H-E-E-----	507																						
A.GW.87.CAM2CG	-----O-----E-III--K-L--NL-D-----YR-----LQ-K-V-----L-L-----R-K-----H-E-E-----	507																						
A.GW.x.ALI	-----H-QI-----O-I-I--K-L--GL-D-----YK-G-----LQ-K-V-----H-----L-L-----H-E-E-----	507																						
A.GW.x.MDS	-----O-----O-III--K-L--L-D-----YR-----LQ-K-V-----L-L-----G-----H-E-E-----	507																						
A.IN.07.NNVA	-----KI-----O-III--S-P-K-L--GL-D-----YY-----LQ-K-V-----L-L-----L-----H-E-E-----	507																						
A.IN.95.CR1K.147	-----O-----O-III--K-L--GL-D-----Y-----LQ-K-V-----L-L-----L-R-----H-E-E-----	507																						
A.SN.85.ROD	-----H-Q-----K-III--L--GL-D-----YH-----LQ-K-I-----L-L-----L-R-----H-E-E-----	508																						
B.CI.88.UC1	-----S-K-D-----S-III--S-----S-L--DM-----K-----KR-----LQ-EK-V-----L-F-R-I-K-----L-LQ-E-----S-K-V-----	508																						
B.CI.x.20.56	-----GK-D-----D-I-----V-----S-L--Q-DM-----K-----K-----LQ-EK-V-----S-LF-RMI-K-----L-LQ-E-----S-K-V-----	508																						
B.CI.x.EHD	-----AK-D-----N-II-----V-----S-L--NM-----K-----K-----LQ-EK-V-----L-F-R-I-K-----L-LQ-E-----S-K-V-----	506																						
B.GH.86.D205	S-C-S-K-D-----S-III--S-----S-L--DM-----K-----K-----LQ-EK-V-----A-K-RV-----L-LQ-E-----S-K-RV-----	507																						
B.JP.01.KR020	-----SVGK-D-----D-I-----V-----S-L--D-----K-----K-----LQ-E-R-----L-F-R-I-K-----L-LQ-E-----S-K-I-----	508																						
G.CI.92.ABT96	-----H-K-D-----E-XIXI--V-----Q-L--KX-D-----Y-----X-K-----N-XEK-----X-----L-N-XK-X-----L-LQ-X-X-A-K-XE-X-----	490																						
AB.CM.03.03CM.510.03	-----S-K-D-----S-II-----S-----S-L--NM-----X-XX-----R-----X-K-----LQ-EK-V-----L-F-X-R-M-X-----L-LQ-X-E-----S-K-V-----	508																						
H2.01.AB.CI.90.7312A	-----S-K-D-----R-N-II--V-----S-K-S-L--DM-----K-----QI-E-V-----L-F-R-I-K-----L-LQ-MQ-----G-S-K-V-----	508																						
H2.01.AB.JP.04.04JP.NMC307	-----S-K-D-----N-II-----X-----S-L--DM-----K-----QI-E-V-----L-F-I-----L-LQ-MQ-----E-X-S-K-XV-----	508																						
H2.01.AB.JP.07.07JP.NMC716	-----K-D-----N-II-----V-----S-L--NM-----K-----QI-E-V-----L-F-I-----L-LQ-MQ-----E-S-K-V-----	508																						
H2.01.AB.JP.08.08JP.NMC842	-----S-K-D-----N-II--XX-S-----S-X-X-DM-----X-X-----K-----QI-EX-V-----L-F-I-X-----L-LQ-MQ-X-X-E-X-S-K-XV-----	508																						
U.FR.96.12034	T-----NL-----E-AVII--N-----QL-S-L--KT-----D-----HK-----K-----E-DV-----V-----K-K-----FA-----A-R-EEN-----	504																						
MAC.US.x.17EC1	-----L-----L-----X-----X-----	494																						
MAC.US.x.17EFR	-----L-----L-----X-----X-----	494																						
MAC.US.x.1937	-----L-----L-----X-----X-----	494																						
MAC.US.x.2065	-----L-----L-----X-----X-----	494																						
MAC.US.x.251.1A11	-----N-----L-----L-----L-----	494																						
MAC.US.x.251.32H.PJ5	-----L-----L-----L-----L-----	494																						
MAC.US.x.251.BK28	-----L-----L-----L-----L-----	490																						
MAC.US.x.80035	-----L-----L-----L-----L-----	494																						
MAC.US.x.81035	-----L-----L-----L-----L-----	494																						
MAC.US.x.85013	-----L-----L-----L-----L-----	494																						
MAC.US.x.87082	-----L-----L-----L-----L-----	494																						
MAC.US.x.92050	-----L-----L-----L-----L-----	494																						
MAC.US.x.92050	-----L-----L-----L-----L-----	494																						
MAC.US.x.92077	-----L-----L-----L-----L-----	494																						
MAC.US.x.93057	-----L-----L-----L-----L-----	494																						
MAC.US.x.93062	-----L-----L-----L-----L-----	494																						
MAC.US.x.95058	-----L-----L-----L-----L-----	494																						
MAC.US.x.95086	-----L-----L-----L-----L-----	494																						
MAC.US.x.95112	-----X-----L-----L-----L-----	494																						
MAC.US.x.96016	-----L-----L-----L-----L-----	494																						
MAC.US.x.96020	-----L-----L-----L-----L-----	494																						
MAC.US.x.96072	-----L-----L-----L-----L-----	494																						
MAC.US.x.96081	-----L-----L-----L-----L-----	494																						
MAC.US.x.96093	-----L-----L-----L-----L-----	494																						
MAC.US.x.96114	-----L-----L-----L-----L-----	494																						
MAC.US.x.96123	-----L-----L-----L-----L-----	494																						
MAC.US.x.96135	-----L-----L-----L-----L-----	494																						
MAC.US.x.97009	-----L-----L-----L-----L-----	494																						
MAC.US.x.97074	-----XX-----L-----L-----L-----	494																						
MAC.US.x.MM142	-----L-----L-----S-----L-----	490																						
MAC.US.x.MM142_patent	-----L-----L-----S-----L-----	490																						
MAC.US.x.r80025	-----L-----L-----L-----L-----	494																						
MAC.US.x.r90131	-----X-----X-----L-----L-----	494																						
MNE.US.82.MNE.8	-----H-N-----V-----L-----L-----K-----	490																						
MNE.US.x.MNE027	-----H-N-----V-----L-----L-----K-----	490																						
SMM.SL.92.SL92B	-----A-Q-----L-----L-G-N-GLT--KM-T-LRDM--NL-----D-----N-L-----L-Y-K-----EK-R-----N-KM-----L-LA-R-N-----R-R-DE-----	487																						
SMM.US.x.F236.H4	-----XH-X-N-----I-----L-----L-G-----N-----L-L-----L-Y-K-----EK-R-----N-----P-----P-----I-----	490																						
SMM.US.x.H9	-----XH-X-N-----I-----L-----L-G-----N-----L-L-----L-Y-K-----EK-R-----N-----P-----P-----I-----	490																						
SMM.US.x.PB114.15	-----H-N-----I-----L-----L-G-----S-----L-----L-L-----L-Y-K-----EK-R-----N-----P-----P-----I-----	490																						
SMM.US.x.PBJA	-----XH-X-N-----I-----L-----L-G-----S-----L-----L-L-----L-Y-K-----EK-R-----N-----P-----P-----I-----	490																						
SMM.US.x.PBJ.143	-----H-N-----I-----L-----L-G-----S-----L-----L-L-----L-Y-K-----EK-R-----N-----P-----P-----I-----	490																						
SMM.US.x.PBJ.6P6	-----H-N-----I-----L-----L-G-----S-----L-----L-L-----L-Y-K-----EK-R-----N-----P-----P-----I-----	490																						
SMM.US.x.PGMS3	-----H-N-G-----I-----L-----L-G-----S-----L-----L-L-----L-Y-K-----EK-R-----N-----P-----P-----I-----	490																						
SMM.US.x.SME543	-----N-----I-----L-----L-G-----S-----L-----L-L-----L-Y-K-----EK-R-----N-----P-----P-----I-----	490																						
STM.US.89.STM.37.16	-----NI-----R-----I-----L-----NL-----DV-----K-----A-----A-----Q-R-D-----	490																						

p51 RT end p15 RNase H start

MAC.US.x.239	TVIKSQDNQWSYKIHQEDKILKVG.KFAKIKNTHNGVRLLAHVIOKIGKEAIVWIGQVPKFHLPVEKQDVMQWTDYVQVTWIPEWDFISTPPLVRLVFNLVKDP	IEGEEYYTDGSCNKQSKGKAGYITDRGKDKVKVLEQTTNQQAELFALMALTDSGPKANIIV	663
A.CI.88.UC2	-IQ-----T-----E-----Y-----Q-V-----L-----RI-----RET-----DN-----V-----T-----G-----P-T-----F-----R-----V-----R-----R-----A-----A-----V-----		676
A.DE.x.BEN	-IQ--GH--T-----E-----Y-----Q-V-----L-----RI-----RET-----DN-----V-----T-----G-----P-A-----F-----R-----V-----R-----R-----V-R-----A-----V-----		676
A.DE.x.PET2	-IQ-D--T-----GE-----Y-----Q-V-----L-----RI-----RET-----DN-----D-----V-----A-----P-----F-----R-----R-----R-----RI-----A-----T-----V-----		676
A.GH.x.GH1	-IQ-N--T-----A-----GERV-----Y-----Q-V-----L-----RI-----RET-----DN-----V-----T-----G-----P-A-----F-----R-----R-----V-----R-----R-----S-----A-----A-----V-----		676
A.GM.87.D194	-IQ-N--T-----A-----GERV-----Y-----Q-V-----L-----R-----R-----T-----DN-----V-----T-----G-----P-T-----F-----R-----V-----R-----R-----S-----A-----A-----V-----		676
A.GM.x.ISY	-IQ-D--T-----V-----GE-----Y-----Q-V-----L-----RI-----RET-----DN-----D-----V-----A-----P-A-----F-----R-----R-----R-----RI-----A-----V-----V-V-----		676
A.GM.x.MCN13	-IQ-D--T-----V-----GE-----Y-----Q-V-----L-----RI-----RET-----DN-----D-----V-----A-----PRT-----F-----R-----V-----R-----RM-----A-----A-----V-----		676
A.GM.x.MCR35	-IQ-D--T-----V-----GE-----Y-----Q-V-----L-----RI-----RET-----DN-----D-----V-----A-----PRT-----F-----R-----V-----R-----RM-----A-----A-----V-----		676
A.GW.86.FG	-IQ-D--T-----V-----GE-----G#-Y-----#-----Q-V-----L-----RI-----RET-----DN-----D-----V-----A-----GE-VP-----F-----R-----R-----R-----I-----A-----V-----		675
A.GW.87.CAM2CG	-IQ-D--T-----E-----M-----H-----K-----Q-V-----L-----*-----RI-----R-----RET-----DN-----D-----V-----A-----G-----P-T-----F-----R-----V-----R-----I-----A-----V-----		675
A.GW.x.ALI	-IQ-D--T-----GE-----E-----M-----H-----K-----Q-V-----L-----RI-----R-----RET-----DN-----D-----V-----A-----L-A-----F-----P-----R-----V-----R-----I-----A-----V-----		675
A.GW.x.MDS	-IQ-D--T-----V-----GE-----Y-----V-----A-----Q-V-----L-----RI-----RET-----D-----D-----V-----A-----P-A-----F-----R-----V-----K-----R-----A-----V-----		676
A.IN.07.NNVA	-IQ-D--T-----V-----GE-----Y-----V-----A-----Q-V-----L-----RI-----RET-----D-----D-----V-----A-----P-A-----I-----R-----V-----R-----R-----A-----V-----		676
A.IN.95.CR1K.147	-IQ-D--T-----V-----GE-----Y-----V-----A-----Q-V-----L-----RI-----RET-----D-----D-----V-----A-----P-T-----F-----R-----V-----R-----R-----S-----A-----V-----		676
A.SN.85.ROD	-IQ-D-E--T-----E-----Y-----V-----I-----Q-V-----L-----RI-----RET-----DN-----D-----V-----A-----G-----P-A-----F-----R-----V-----R-----K-----A-----V-----		677
B.CI.88.UC1	-IQ-NLA--T-----GNR-----Y-----V-----V-----L-----EI-V-----RET-----D-----V-----AY-----L-KV-----RA-----V-----R-----AL-Q-----QV-----		677
B.CI.x.20.56	-IQ-NLA--T-----GNRT-----Y-----V-----V-----L-----EI-M-----RET-----D-----V-----AY-----L-KV-----RA-----V-----R-----AL-Q-----QV-----		677
B.CI.x.EHD	-IQ-NLA--T-----G-----Y-----V-----V-----L-----EI-M-----RET-----D-----V-----I-AY-----L-V-----A-----V-----P-----AL-Q-----QV-----		675
B.GH.86.D205	-IQ-NLA--T-----GN-V-----Y-----V-----V-----L-----EI-V-----RET-----D-----V-----I-AY-----L-R-----RT-----V-----A-----AL-Q-----E-QV-----		676
B.JP.01.KR020	-IQ-NLA--T-----G-R-R-----Y-----V-----V-----L-----T-----EI-M-----RET-----D-----V-----I-AY-----L-V-----RT-----V-----A-----AL-Q-----QV-----		677
G.CI.92.ABT96	-IQ-NL--T-----V-----GNR-----Y-----V-----X-----L-----KL-F-----R-T-----D-----V-----H-----Y-----E-L-OT-----T-X-----V-----X-----A-----Q-----V-----X-----		659
AB.CM.03.03CM.510.03	-IQ-NLA--T-----GXRM-----Y-----V-----XX-----V-----L-----EI-I-----XRET-----D-----X-----V-----I-AY-----XK-----A-----V-----X-----AL-XQ-X-----QV-----		677
H2.01.AB.CI.90.7312A	-IQ-NLA--T-----GN-V-----Y-----V-----I-----V-----L-----I-M-----RET-----D-----V-----AY-----L-KV-----RT-----V-----R-----P-----AL-Q-----QV-----		677
H2.01.AB.JP.04.04JP.NMC307	-IQ-XLA--T-----G-----V-----Y-----V-----I-----I-----M-----RET-----D-----X-----V-----AY-----L-KV-----F-----RV-----V-----R-----P-----X-AL-Q-----QV-----		677
H2.01.AB.JP.07.07JP.NMC716	-IQ-NLA--T-----G-V-----Y-----V-----S-----I-----I-----M-----RET-----D-----V-----AY-----L-KV-----F-----RV-----X-----VX-----R-----P-X-----X-AL-Q-----QV-----		677
H2.01.AB.JP.08.08JP.NMC842	-IQ-NLA--T-----X-V-----Y-----V-----X-----I-----IXM-----RET-----D-----X-----V-----AY-----X-L-KV-----F-----RV-X-----VX-----R-----P-X-----X-AL-Q-----QV-----		677
U.FR.96.12034	-L-N-----T-----G-RT-----V-----V-----V-----KI-M-----RE-----V-----VLD-----V-----AY-----E-L-----L-----RA-Q-----V-----R-----P-----A-----E-----V-----VT-----		672
MAC.US.x.17EC1			663
MAC.US.x.17EFR			663
MAC.US.x.1937		R-----X-----X-----	663
MAC.US.x.2065		X-----	663
MAC.US.x.251.1A11	-N-----	N-----	663
MAC.US.x.251.32H.PJ5		K-----	663
MAC.US.x.251.BK28			663
MAC.US.x.80035			659
MAC.US.x.81035			663
MAC.US.x.85013			663
MAC.US.x.87082			663
MAC.US.x.92050			663
MAC.US.x.92077			663
MAC.US.x.93057			663
MAC.US.x.93062			663
MAC.US.x.95058			663
MAC.US.x.95086			663
MAC.US.x.95112		X-----	663
MAC.US.x.96016			663
MAC.US.x.96020			663
MAC.US.x.96072			663
MAC.US.x.96081			663
MAC.US.x.96093			663
MAC.US.x.96114			663
MAC.US.x.96123			663
MAC.US.x.96135		X-----	663
MAC.US.x.97009			663
MAC.US.x.97074			663
MAC.US.x.MM142			659
MAC.US.x.SNM142_patent			659
MAC.US.x.r80025			663
MAC.US.x.r90131			663
MNE.US.82.MNE.8	-N-----		659
MNE.US.x.MNE027	-N-----	K-----K-----V-----R-----A-----	659
SMM.SL.92.SL92B	-L-N-----G-R-----V-----I-----N-V-----SL-----KT-F-----RE-----D-----A-----I-----K-----V-----I-----RN-----V-----E-----LP-----A-----Q-L-L-----K-----PS-V-VVT-----		656
SMM.US.x.F236.H4			659
SMM.US.x.H9	P-----S-----V-----I-----E-----R-----XEI-----V-----V-----E-----Q-A-----F-V-----R-----R-----V-----R-----A-----L-----YL-----A-----		659
SMM.US.x.PBJ14.15			659
SMM.US.x.PBJA			659
SMM.US.x.PBJ.143	P-----V-----R-----XEI-----V-----V-----E-----Q-A-----F-V-----R-----R-----V-----R-----A-----L-----YL-----A-----		659
SMM.US.x.PBJ.6P6			659
SMM.US.x.PGMS3			659
SMM.US.x.SME543	-V-D-----V-----V-----V-----L-----E-----RE-----I-----D-----V-----E-----Q-A-----F-V-----R-----R-----V-----R-----A-----L-----YL-----A-----		659
STM.US.89.STM.37.16	-V-D-----I-----RE-----A-----V-----E-----L-----T-----F-----V-----R-----V-----N-----A-----A-----A-----V-----		659

p15 RNase H end_p31 Integrase start

MAC.US.x.239	DSQYVMGIITGCPTESESRLVNQIIEEMIKKSEIYVAWPAHKGIGGQIEDHLVSGIRQVLF.L.EKIEPAQEHEDKYHSNVKELVFKGLPRIVARQIVDTCDKCHOKGEAIIHQGANSDLGTWQMDCTHLEGGKIIIVAVHVSAGFIEAEVIPQETGRQTALFLLKLAG	832
A.CI.88.UC2	-----VASQ-----I-----D-----EAV-----I-----V-----E-----I-----IH-----I-----QL-----N-----AQ-----Q-----V-----AEI-----V-----S-----S-----	845
A.DE.x.BEN	-----VA-Q-----N-----I-----EAV-----V-----V-----E-----II-----TH-----I-----LL-----NS-----AQ-----Q-----V-----AEI-----V-----Y-----S-----S-----	845
A.DE.x.PET2	-----VA-Q-----K-----EAV-----V-----V-----E-----M-----SH-----K-----L-----N-----AQ-----Q-----V-----DAE-----S-----S-----	845
A.GH.x.GH1	-----VV-Q-----I-----D-----EAV-----V-----R-----E-----M-----TH-----I-----QL-----N-----AQ-----Q-----V-----AEI-----V-----S-----S-----	845
A.GW.07.D194	-----VA-Q-----N-----I-----D-----EAV-----V-----V-----E-----I-----TH-----I-----QL-----N-----AQ-----Q-----V-----AEI-----V-----S-----S-----	845
A.GM.x.ISY	-----V-Q-A-----I-----K-----EAV-----V-----R-----E-----G-----AH-----N-----L-----N-----AQ-----Q-----V-----AEI-----V-----S-----S-----	845
A.GM.x.MCN13	-----VA-Q-----I-----EAV-----V-----V-Q-----E-----M-----SH-----K-----L-----N-----AQ-----Q-----V-----DAE-----S-----S-----	845
A.GM.x.MCR35	-----VA-Q-----I-----EAV-----V-----V-Q-----E-----M-----SH-----K-----L-----N-----AQ-----Q-----V-----DAE-----S-----S-----	845
A.GW.86.FG	-----VA-Q-----N-----I-----EAV-----V-----V-----E-----I-----SH-----I-----K-----L-----N-----AHVQ-----V-----AE-----V-----S-----S-----	844
A.GW.07.CAM2CG	-----VA-Q-----N-----I-----EAV-----V-----V-----E-----T-----CH-----D-----I-----QL-----N-----ADYQ-----V-----AEV-----V-----I-----S-----S-----	844
A.GW.x.AL1	-----VA-Q-----N-----I-----EAV-----V-----V-----E-----M-----SH-----N-----L-----N-----AQ-----Q-----V-----AE-----V-----I-----S-----S-----	844
A.GW.x.MDS	-----VV-Q-----N-----I-----EAV-----V-----V-----E-----M-----SH-----Q-----L-----N-----AQ-----Q-----V-----A-----V-----S-----S-----	845
A.IN.07.NNVA	-----VA-Q-----N-----EAV-----V-----V-----E-----N-----SH-----K-----L-----N-----AQ-----Q-----V-----AE-----V-----S-----S-----	845
A.IN.95.CR1K.147	-----VA-Q-----NKI-----EAV-----V-----V-----E-----SH-----R-----K-----L-----N-----AQ-----Q-----V-----AE-----V-----S-----S-----	845
A.SN.85.ROD	-----SASQ-----KI-----EAV-----V-----V-----E-----SH-----I-----N-----L-----NS-----AQ-----Q-----V-----AE-----V-----S-----S-----	846
B.CI.88.UC1	-----VA-Q-----T-----P-----EA-----G-----R-----L-----V-----E-----G-----H-----I-----QL-----K-----NS-----Q-----V-----AE-----V-----S-----S-----	846
B.CI.x.20.56	-----VA-Q-----T-----D-----PI-----K-----EA-----G-----L-----V-----E-----G-----H-----I-----QL-----K-----NS-----Q-----V-----AE-----V-----S-----S-----	846
B.CI.x.EHD	-----VAAQ-----T-----PI-----RE-----EK-----G-----L-----V-----I-----E-----N-----H-----I-----QL-----NS-----Q-----V-----AE-----V-----S-----S-----	844
B.GH.86.D205	-----AAQ-----T-----PI-----AK-----EAV-----G-----L-----V-----I-----E-----G-----H-----I-----QL-----K-----NS-----Q-----V-----A-----V-----S-----S-----	845
B.JP.01.KR020	-----AAQ-----T-----PI-----KT-----V-----EA-----G-----L-----V-----HI-----N-----E-----G-----H-----I-----QL-----K-----NS-----Q-----V-----AE-----V-----S-----S-----	846
G.CI.92.ABT96	-----X-----AXQ-----T-----P-----K-----X-----EXL-----G-----L-----V-----X-----X-----E-----F-----TH-----I-----QL-----K-----NS-----Y-----H-----S-----Q-----V-----AE-----V-----V-----S-----S-----	828
AB.CM.03.03CM.510.03	-----AAX-----T-----PI-----XT-----X-----EA-----G-----L-----V-----X-----X-----G-----X-----H-----X-----QL-----K-----NX-----Q-----X-----X-----V-----AE-----V-----X-----X-----S-----S-----	846
H2.01.AB.CI.90.7312A	-----VA-Q-----T-----PI-----K-----EA-----G-----L-----V-----E-----G-----H-----I-----QL-----K-----NS-----Q-----V-----AE-----V-----S-----S-----	846
H2.01.AB.JP.04.04JP.NMC307	-X-----A-Q-----T-----SI-----R-----EAV-----G-----L-----V-----E-----G-----H-----HL-----K-----NS-----X-----Q-----V-----AE-----V-----X-----X-----S-----S-----	846
H2.01.AB.JP.07.07JP.NMC716	-----A-Q-----T-----SI-----R-----EAV-----G-----L-----V-----E-----G-----H-----HL-----K-----NS-----Q-----V-----AE-----V-----S-----S-----	846
H2.01.AB.JP.08.08JP.NMC842	-X-----A-Q-----T-----SI-----K-----EAV-----G-----L-----X-----V-----E-----G-----X-----HL-----K-----NS-----Q-----V-----AE-----V-----X-----S-----S-----	846
U.FR.96.12034	-----I-----A-Q-----T-----P-----SK-----L-----R-----EAV-----G-----L-----V-----E-----Q-----H-----I-----QL-----K-----N-----HV-----Q-----M-----V-----TEV-----V-----V-----S-----S-----	841
MAC.US.x.17EC1	-----V-----	832
MAC.US.x.17EFR	-----V-----	832
MAC.US.x.1937	-----V-----	832
MAC.US.x.2065	-----V-----	832
MAC.US.x.251.1A11	-----V-----	832
MAC.US.x.251.32H.PJ5	-----V-----	832
MAC.US.x.251.BK28	-----V-----	828
MAC.US.x.80035	-----V-----	832
MAC.US.x.81035	-----V-----	832
MAC.US.x.85013	-----V-----	832
MAC.US.x.87082	-----V-----	832
MAC.US.x.92050	-----V-----	832
MAC.US.x.92077	-----V-----	832
MAC.US.x.93057	-----V-----	832
MAC.US.x.93062	-----V-----	832
MAC.US.x.95058	-----V-----	832
MAC.US.x.95086	-----V-----	832
MAC.US.x.95112	-----V-----	832
MAC.US.x.96016	-----V-----	832
MAC.US.x.96020	-----V-----	832
MAC.US.x.96072	-----V-----	832
MAC.US.x.96081	-----V-----	832
MAC.US.x.96093	-----V-----	832
MAC.US.x.96114	-----V-----	832
MAC.US.x.96123	-----V-----	832
MAC.US.x.96135	-----V-----	832
MAC.US.x.97009	-----V-----	832
MAC.US.x.97074	-----V-----	832
MAC.US.x.MM142	-----V-----	828
MAC.US.x.SNM142_patent	-----V-----	828
MAC.US.x.r80025	-----V-----	832
MAC.US.x.r90131	-----V-----	832
MNE.US.82.MNE.8	-----V-----	828
MNE.US.x.MNE027	-----V-----	828
SMM.SL.92.SL92B	-----LN-----Q-----S-----D-----DI-----A-----QLVQ-----EAV-----IG-----N-----V-----R-----S-----D-----AQ-----YNI-----QL-----K-----NA-----N-----Q-----T-----AEV-----V-----R-----I-----S-----	825
SMM.US.x.F236.H4	-----A-Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828
SMM.US.x.H9	-----VA-Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828
SMM.US.x.PB114.15	-----VA-Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828
SMM.US.x.PBJA	-----VA-Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828
SMM.US.x.PBJ.143	-----VA-Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828
SMM.US.x.PBJ.6P6	-----VA-Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828
SMM.US.x.PGMS3	-----A-Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828
SMM.US.x.SME543	-----VA-Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828
STM.US.89.STM.37.16	-----Q-----K-----EA-----X-----E-----V-----E-----#K-----E-----K-----X-----L-----K-----L-----X-----V-----XX-----G-----X-----S-----S-----	828

Accession	Sequence	Position
MAC.US.x.239P-T.....V-QS.....S.....T.....L-R-N-K-Q.....V-A.....R-RO	1015
A.CI.88.UC2P-T.....V-QS.....S.....T.....L-R-N-K-Q.....V-A.....R-RO	1015
A.DE.x.BENP-T.....V-QS.....S.....T.....L-R-N-Q.....F.....D-V-A-I.....R-RR	1015
A.DE.x.PE12P-T.....I-QS.....S.....TM.....LHA-L.....F.....D-V-A-I.....R-RR	1015
A.GH.x.GH1S-T.....I-QS.....S.....TI.....L-R-N-Q.....D-V-A-I.....R-RO	1015
A.GM.87.D194P-T.....I-QS.....S.....TI.....L-R-N-K-Q.....D-V-A-I.....R-RO	1015
A.GM.x.ISYT.....V-QS.....E.....TM.....V-V.....L-A-L.....F-N-Q.....D-V-I.....R-PRO	1015
A.GM.x.MCN13T.....V-QS.....S.....T.....T.....LHA-L.....F.....D-V-I.....R-RO	1015
A.GM.x.MCR35T.....V-QS.....S.....T.....T.....LHA-L.....F.....D-V-I.....R-RO	1015
A.GW.86.FGT.....V-Q.....S.....T.....I.....LHA-L.....F.....D-V-E.....R-RO	1014
A.GW.87.CAM2CGT.....V-Q.....S.....T.....S.....L-A-L.....F.....D-V-II.....R-RO	1014
A.GW.x.AL1S-V.....T.....V-QS.....S.....T.....L-T-L.....P.....D-V-I.....R-RO	1013
A.GW.x.MDST.....I-QS.....S-D-T.....S.....A.....L-A-L.....F.....D-V-I.....R-ROD	1015
A.IN.07.NNVAT.....V-QS.....S.....T.....S.....A.....L-A-L.....F.....D-V-II.....R-RO	1015
A.IN.95.CR1K.147T.....I-QS.....S.....T.....S.....L-A-L.....F.....D-V-II.....R-RO	1015
A.SN.85.RODT.....I-QS.....S.....TI.....I.....S.....L-A-L-D.....F.....LV-II.....R-RO	1016
B.CI.88.UC1T-D-A-I-Q.....D-V-I-V-A.....V.....F-A-L-Q-Q.....LV-E-I.....RH	1016
B.CI.x.20.56T-D-A-I-Q.....E.....D-V-I-V-T.....IV.....F-A-L-Q-Q.....D.....I-E-I.....RH	1014
B.CI.x.EHOT-D-A-I-Q.....E.....D-V-I-V-T.....IV.....F-A-L-Q-Q.....D.....I-E-I.....RH	1014
B.GH.86.D205T-PS-V-Q.....L-D-V-I-V-T.....IV-A.....F-A-L-Q-Q.....I-E-I.....RH-G	1015
B.JP.01.KR020T-D-A-I-V-Q.....D-V-I-V-A.....IV-A.....L-T-KL-Q.....F.....I-E-I.....N.....1016	1016
G.CI.92.ABT96T.....V-QA.....S-K-K-D-I.....I.....T.....I-E.....998	998
AB.CM.03.03CM.510.03	X.....T-D-V-Q.....S-L-X-X-I-V-X-T.....V.....F-A-L-Q-QX.....X-E.....RH-G	1016
H2.01.AB.CI.90.7312AT-D-V-Q.....D-V-I-V-A.....V.....F-A-L-Q-Q.....I-E-I.....RH	1016
H2.01.AB.JP.04.04JP.NMC307T-D-V-QS.....D-V-L-V-A.....V.....F-A-L-Q-Q.....V-E.....RH	1016
H2.01.AB.JP.07.07JP.NMC716T-D-V-QS.....D-V-L-V-A.....V.....F-A-L-Q-Q.....V-E.....RH	1016
H2.01.AB.JP.08.08JP.NMC842T-D-V-QS.....D-V-L-V-A.....V-YF-T-L-Q-Q-X.....V-E.....RH-G	1016
U.FR.96.12034T.....V.....E-TM-A.....I-S-L.....AD-LI-T.....R.....1011	1011
MAC.US.x.17EC11002	1002
MAC.US.x.17EFRX-X.....X.....V.....1002	1002
MAC.US.x.1937S-T.....1002	1002
MAC.US.x.20651002	1002
MAC.US.x.251.1A11V.....1002	1002
MAC.US.x.251.32H.PJ5V.....1002	1002
MAC.US.x.251.BK281002	1002
MAC.US.x.808351002	1002
MAC.US.x.810351002	1002
MAC.US.x.850131002	1002
MAC.US.x.870821002	1002
MAC.US.x.920501002	1002
MAC.US.x.920771002	1002
MAC.US.x.930571002	1002
MAC.US.x.93062X.....1002	1002
MAC.US.x.950581002	1002
MAC.US.x.950861002	1002
MAC.US.x.951121002	1002
MAC.US.x.960161002	1002
MAC.US.x.960201002	1002
MAC.US.x.960721002	1002
MAC.US.x.96081X.....1002	1002
MAC.US.x.960931002	1002
MAC.US.x.961141002	1002
MAC.US.x.961231002	1002
MAC.US.x.961351002	1002
MAC.US.x.970091002	1002
MAC.US.x.97074X.....1002	1002
MAC.US.x.MM142998	998
MAC.US.x.SNM142_patent1002	1002
MAC.US.x.r800251002	1002
MAC.US.x.r901311002	1002
MNE.US.82.MNE.8M.....L.....V.....998	998
MNE.US.x.MNE027M.....L.....L.....L-T-YLN-Q-Q.....A.....VI-E.....RQ	998
SMM.SL.92.SL92BK-T.....L-V-QS.....DL-KN-K-E-L-A.....IV.....L-T-YLN-Q-Q.....A.....VI-E.....RQ	995
SMM.US.x.F236.H4T.....X-Q-X-X.....T.....I-I.....V.....E.....998	998
SMM.US.x.H9T.....X-Q-X-X.....T.....I-I.....V.....E.....998	998
SMM.US.x.PB114.15T.....X-Q-X-X.....T.....I-I.....V.....E.....998	998
SMM.US.x.PBJAT.....X-Q-X-X.....T.....I-I.....V.....E.....998	998
SMM.US.x.PBJ.143T.....X-Q-X-X.....T.....I-I.....V.....E.....998	998
SMM.US.x.PBJ.6P6T.....X-Q-X-X.....T.....I-I.....V.....E.....998	998
SMM.US.x.PGMS3T.....X-Q-X-X.....T.....I-I.....V.....E.....998	998
SMM.US.x.SME543T.....X-Q-X-X.....T.....I-I.....V.....R.....E.....998	998
STM.US.89.STM.37.16V-T.....Q.....T-D-T-V.....K-L.....V.....V.....E.....998	998

	p31 Integrase end Pol end	
MAC.US.x.239	VDSSSHMEDTGEAR.EVA.....*	1019
A.CI.88.UC2	L--P-L-GAR-DG.....	1033
A.DE.x.BEN	L--P-L-GAR-DG.-M.....C	1033
A.DE.x.PE12	-----L-G-R-DG.....	1033
A.GH.x.GH1	L-----L-GAR-EDG.....	1033
A.GM.87.D194	L-----L-GAR-DG.....	1033
A.GM.x.ISY	M-G-L-GAR-DG.-M.....	1033
A.GM.x.MCN13	L-G-L-GAR-DG.-M.....	1033
A.GM.x.MCR35	L-G-L-GAR-DG.-M.....	1033
A.GW.86.FG	M-G-L-GAR-DG.-M.....	1032
A.GW.87.CAM2CG	L-----L-GAR-NG.....	1032
A.GW.x.ALI	L-GP-L-GAR-DG.....	1033
A.GW.x.MDS	M-GP-L-G-R-DG.....	1033
A.IN.07.NNVA	---GP-L-G-R-DG.....	1033
A.IN.95.CR1K_147	M-GP-L-G-R-DG.....	1033
A.SN.85.ROD	M-G-L-GAR-DG.-M.....	1034
B.CI.88.UC1	L-CGTD---RQ---M-0SGQVPEA-	1042
B.CI.x.20_56	L-CGADV---RQ---M-0SGQVPEA-	1042
B.CI.x.EHD	L-C-ADV---MQ---QSN.....	1035
B.GH.86.D205	L-C-AD---RQ---M-QSD.....	1036
B.JP.01.KR020	L-C-AN---Q---M-QPSQISET-	1042
G.CI.92.ABT96	----TN---RQTG.....	1016
AB.CM.03.03CM_510_03	L-C-AD---RQ-S.-M-QSDQIP..	1040
H2_01_AB.CI.90.7312A	L-C-TDV---RQ---M-0SSQVSEA-	1042
H2_01_AB.JP.04.04JP_NMC307	L-CGTDV---RQ---M-0SGQVPEA-	1042
H2_01_AB.JP.07.07JP_NMC716	L-CGADV---RQ---M-0SGQVPEV-	1042
H2_01_AB.JP.08.08JP_NMC842	L-CGTD---RQ---M-0SQVPEA-	1042
U.FR.96.12034	L----NL-GAEKV.-M-LPDQTP#V-	1036
MAC.US.x.17EC1	-----	1020
MAC.US.x.17EFR	-----	1020
MAC.US.x.1937	-----	1020
MAC.US.x.2065	-----	1020
MAC.US.x.251_1A11	M-----	1020
MAC.US.x.251_32H_PJ5	-----	1020
MAC.US.x.251_BK28	-----	1016
MAC.US.x.80035	-----	1020
MAC.US.x.81035	-----	1020
MAC.US.x.85013	-----	1020
MAC.US.x.87082	-----	1020
MAC.US.x.92050	-----	1020
MAC.US.x.92077	-----	1020
MAC.US.x.93057	-----	1020
MAC.US.x.93062	-----	1020
MAC.US.x.95058	-----	1020
MAC.US.x.95086	-----	1020
MAC.US.x.95112	-----	1020
MAC.US.x.96016	-----	1020
MAC.US.x.96020	-----	1020
MAC.US.x.96072	-----	1020
MAC.US.x.96081	-----	1020
MAC.US.x.96093	-----	1020
MAC.US.x.96114	-----	1020
MAC.US.x.96123	-----	1020
MAC.US.x.96135	-----	1020
MAC.US.x.97009	-----	1020
MAC.US.x.97074	-----	1020
MAC.US.x.MM142	M-----	1016
MAC.US.x.SNM142_patent	M-----	1016
MAC.US.x.r80025	-----	1020
MAC.US.x.r90131	-----	1020
MNE.US.82.MNE_8	-----	1016
MNE.US.x.MNE027	-----	1016
SMM.SL.92.SL92B	MG--AS---Q*KNSGAL.....	1013
SMM.US.x.F236_H4	L-G-L-----	1016
SMM.US.x.H9	L-G-L-----	1016
SMM.US.x.PB114_15	L-G-L-----	1016
SMM.US.x.PBJA	L-G-L-----	1016
SMM.US.x.PBJ_143	L-G-L-----	1016
SMM.US.x.PBJ_6P6	L-G-L-----	1016
SMM.US.x.PGMS3	L-GP-L-----	1016
SMM.US.x.SME543	L-G-L-----	1016
STM.US.89.STM_37_16	---G-L---R--G.....	1016

Vif start

MAC.US.x.239	MEEEKRWIAVPTWRIPE.RLERWHSLIKYLKYKTKDLOKQVYVPHFKVGAWMTCSRVIPLQEGSHLEVGQYWHLTPEKGLWSTYAVRITWYSKNFWDVTPNYADILLHSTYVFCPTAGEVRRAIRGEQLLSCCRFPRAHKYQVPSLQYLALKKVVS..DVRSQGENPT	167
A.CI.88.UC2	---G-S-V---V-G-M---V---R---EG-R-H-----KGE-I-A-N---SHS-L-TEK---DC-I-C-A---K---NY-Q--A---V-Q.QND-P-RKG-A	168
A.DE.x.BEN	---DRN-V---V-G-M-K-A-V---R---EE-R-H-----GK-I-A-N---SH-L-TEK---DC-I-S---K---NY-Q--A---V-Q.QND-P-RKGTA	168
A.DE.x.PE22_KR_KRCG	---GE-V---V-G-M-K-V---HR-EG-R-H-----GN-I-A-N---S---TER---DC-S-I-S---K---NY-Q--RSK-L-F-V-Q.QNG-P-KNST	168
A.GH.x.GH1	---G-N---V-GR-M---V---R---EE-R-H-----KGE-I-A-N---SHS-L-TEK---DC-S-I-S---K---NY-Q--V---V-Q.QND-P-RKGTA	168
A.GM.07.D194	---G-N-V---V-G-M---V---R---EE-R-H-----EGE-I-A-N---SHS-L-TEK---DC-S-I-S---K---NY-Q--A---V-Q.QNG-P-RKGAA	168
A.GM.x.ISY_SBL_6669_85	---DQG---V-G-M-K-V---R---EO-R-H-----KGN-I-A-N---S-S-M-EG---DC-T-I-S---KS---NY-Q--SK---F-V-Q.QNDKP-RD-T	168
A.GM.x.MCN13	---G-N-V---V-G-M-K-V---R---ED-H-----KGN-I-A-N---S-S-M-TER---DC-S-I-S---K---NY-Q--RA---F-V-Q.QND-P-RNGTP	168
A.GM.x.MCR35	---G-N-V---V-G-M-K-V---R---ED-H-----KGN-I-A-N---S-S-M-TER---DC-S-I-S---K---NY-Q--RA---F-V-Q.QND-P-RNGTP	168
A.GM.86.FG_c1one_NIHZ	---G-V-I-V-G-M---V---R---E-H-----K-N-I-A-N---SHS-TEK---DC-T-I-S---K---KY--RS---F-V-Q.QND-P-RNSA	168
A.GW.07.CAM2CG	---G-S-V---V-G-M-K-V---R---G-EQ-H-----RGD-R-I-A-N---S-M-TEK---DC-T-I-S---K---KY--RS---F-V-Q.QND-P-RDRT	168
A.GW.x.ALI	---G-S-V---V-G-M-K-V---R---E-H-----GR-I-A-N---S-M-TEK---DC-S-I-G-S---K---NY-Q--RS---F-V-Q.QNGK-RNST	168
A.GW.x.MDS	---G-T-V---V-G-M-K-V---R---EE-R-H-----KGN-I-A-N---S-L-TEK---DC-S-I-T---A---K---NY-Q--S-T---F-V-Q.QNG-P-RDST	168
A.IN.07.NNVA	---G-V---V-G-M-K-V---R---E-H-----NNN-N-I-A-N---S-L-TEK---DC-S-I-G-A---KV---NY-Q--S-T---F-V-Q.QND--RDST	168
A.IN.95.CR1K_147	---D-V---KV-G-M-K-V---R---E-H-----KDN-I-I-A-N---E-S-S-TE---DCT-S-I-A---KT---NY-Q--RT---F-V-Q.QND--RDGT	168
A.SN.85.ROD	---D-V---V-G-M-K-V---R---E-H-----KGN-I-A-N---E-S-S-TE---DC-V-I---K---NY--RA---F-V-Q.QND-P-RDST	168
B.CI.88.UC1	---G-N-V---G---K-V---HR-E-0-S-H-----K-EAY---N---R-F-S---L-KRS-Y---DV-Q-GS-S-N---KT-Y-NY-S-EG---F-R-IOEGKDG---SA	169
B.CI.x.20_56	---N-V---G---V---HR-E-0-S-H-----K-AY-I---N---R-L-S---L-RES-Y---DV-Q-GS-S-N---KT-Y-NY-S-DR---F-R-OEGKNG---SA	169
B.CI.x.EHD	---N---C---V---R-0-S-H-----K-AY-I---N---R-F-S---L-ERS-Y---DV-R-GS-S-N---KT-H-NY-S-TG---F-R-OEGKNG---ST	169
B.GH.86.D205_ALT	---D-V---G---K-V---R-GE-0-S-H-----I-NK-AW---N---R-F-S---L-ER-Y---DV-Q-GS-S-N---KT-Y-NY-S-EG---F-R-OEGKNG---SA	169
B.JP.01.KR020_IMCJ_KR020_1	---N---G---K-V---HR-E-EQ-T-H-----K-AY-I---N---R-F-A---L-K-S-Y---DV-H-GS-F-N---KT-Y-NY-S-RG---F-R-OEGGNP---SAA	169
G.CI.92.Abt96	---G---K---F-F---AV-H-----TKEA-I---N---Q---L-TRK-Y---ET-Q-GS-D---I---NY-T-R---F-Q-QK.GHG-K-SX	168
AB.CM.03.03CM_510_03	---D-V---K-G-A---GE-0-S-H-X-----DKXAX-X---N---R-L---L-XER-Y---DV-Q-GS-S-N---KT-Y-NY-S-EG---F-R-IOEGKNG---K-T	169
H2_01_AB.CI.90.7312A	---G-N-V---K-G-K-V---HR-GE-0-S-H-----KG-AC---N---R-F-S---L-EES-Y---DV-Q-G-S-N---KT-Y-NY-S-EG---F-R-OEGKNG---SA	169
H2_01_AB.JP.04.04JP_NMC307	---N-V---G---K-V---HR-RX-0-S-H-----K-RAW-I---N---R-F-S---L-EES-Y---DV-Q-G-S-N---KT-Y-NY-SX-ER---F-R-OEGKNG-K-GA	169
H2_01_AB.JP.07.07JP_NMC716	---N-V---G---V---GE-0-S-H-----KDRAW---N---R-F-S---L-EE-Y---DV-Q-G-S-N---KT-Y-NY-S-ER---F-R-OEGKNG-K-GA	169
H2_01_AB.JP.08.08JP_NMC842	---D-V---G---K-X---HR-GE-0-S-X-H-----XXKAWI-X---N---X-F-S---L-TES-Y---DV-Q-G-S-N---KT-Y-NY-S-XR---F-R-OEGKNGKXX-XA	169
U.FR.96.12034	---N-V---V-R---H-#-N-E-H-----KK-A-N---S-L-TRG---DC-Q-GS-S---N-S-RQ---F-RALQ--G-KR	166
MAC.US.x.17EC1	-----	167
MAC.US.x.17EFR	-----	167
MAC.US.x.1937	-----	167
MAC.US.x.2065	-----	167
MAC.US.x.251_1A11	-----	167
MAC.US.x.251_32H_PJ5	-----	167
MAC.US.x.251_BK28	-----	167
MAC.US.x.80035	-----	167
MAC.US.x.81035	-----	167
MAC.US.x.85013	-----	167
MAC.US.x.87082	-----	167
MAC.US.x.92050	-----	167
MAC.US.x.92077	-----	167
MAC.US.x.93057	-----	167
MAC.US.x.93062	-----	167
MAC.US.x.95058	-----	167
MAC.US.x.95086	-----	167
MAC.US.x.95112	-----	167
MAC.US.x.96016	-----	167
MAC.US.x.96020	-----	167
MAC.US.x.96072	-----	167
MAC.US.x.96081	-----	167
MAC.US.x.96093	-----	167
MAC.US.x.96114	-----	167
MAC.US.x.96123	-----	167
MAC.US.x.96135	-----	167
MAC.US.x.97009	-----	167
MAC.US.x.97074	-----	167
MAC.US.x.MM142_IVMXX	-----	167
MAC.US.x.SNM142_patent	-----	167
MAC.US.x.r80025	-----	167
MAC.US.x.r90131	-----	167
MNE.US.82.MNE_8	-----	167
MNE.US.x.MNE027	-----	167
SMM.SL.92.SL92B	-----	167
SMM.US.x.F236_H4	-----	167
SMM.US.x.H9	-----	167
SMM.US.x.PB114_15	-----	167
SMM.US.x.PBJA	-----	167
SMM.US.x.PBJ_143	-----	167
SMM.US.x.PBJ_6P6	-----	167
SMM.US.x.PGMS3	-----	167
SMM.US.x.SME543	-----	167
STM.US.89.STM_37_16	-----	167

	Vif end	
MAC.US.x.239	WKQWRDRNRRLMAKONSRGDKQGGKPPPTKGANFPLGAKVLGILA*	214
A.CI.88.UC2	R----HW---V-R-DY-SLE-----APR-H--V---E----	216
A.DE.x.BEN	R----HW---V-REDH-SL-G-SE-SAPR-H--V---E----	216
A.DE.x.PE12_KR_KRCG	R-R-SNYW--F-L-RKDG-H-----SE-AS-Y--V---E----	216
A.GH.x.GH1	R----HW---V-R-DY-SL-----SE-SAPR-H--V---E----	216
A.GW.87.D194	R----HW---V-R-DY-SL-G-SE-SAPR-H--V---E----	216
A.GM.x.ISY_SBL_6669_85	R----NY---L-R-DG-SH-----SE-AQ-Y--V---E----	216
A.GM.x.MCN13	R----Y---QL-R-DG-SH-----SES-APR-Y--V-E-E----	216
A.GM.x.MCR35	R----Y---QL-R-DD-SH-----SES-APR-Y--V-E-E----	216
A.GW.86.FG_c1one_NIHZ	R-R-G-Y---R-D--Y-----SES-PTR-H-----E-E----	216
A.GW.87.CAM2CG	R----Y---L-R-D--SY-----SES-AP-Y--V---E----	216
A.GW.x.AL1	R----Y---L-R-D--L-----ES-AP-H--V---E----	216
A.GW.x.MDS	R-R--Y-A---L-R-D--SH-----ES-APR-Y--V---E----	216
A.IN.07.NMVA	R-R--Y---L-REDG-SH-----SES-AQ--H--V---E----	216
A.IN.95.CR1K_147	R-R--Y---L-R-DG-SH-----ES-AQ-TH-----E-E----	216
A.SN.85.ROD	R-R--Y---L--D--SH-----SSES-PRTY--V-E-E----	216
B.CI.88.UC1	R-R-N---SI-L-RK-NNRAQ-GSSQ-LAPRTH-----E-----	217
B.CI.x.20_56	R-R-G-S---I--RK-TSRTQ-GSSQFAPRTH-----E-----	217
B.CI.x.EHO	R-R-NS---I--RD-I-TSQ-GSSQSLAQ-TY-----E-----	217
B.GH.86.D205_ALT	R-R-NS---SI-L-RK-NNRAQ-GS-Q-FAPRTH-----E-----R	217
B.JP.01.KR020_IMCJ_KR020_1	R-R-N---SI--RK-NNRAQ-SQCSSAQ-TY-----E-----R	217
G.CI.92.Abt96	R-R-G---I--RK-XR-Q-DSSQSF-Q-Y-----E-----	216
AB.CM.03.03CM_510_03	RE-R-N---SI-X-RK-NNRAQ-GSSQ-LAPRTH-----E-----	217
H2_01_AB.CI.90_7312A	R-R--S---I--RK-NSRTO-GSSQ-FAPRTH-----E-----	217
H2_01_AB.JP.04.04JP_NMC307	R-R--I--L-RK-NNRAQ-GSSQ-F-P-TH-----E-----	217
H2_01_AB.JP.07.07JP_NMC716	R-R--IQL-RK-NNRAQ-GSSQ-F-P-TH-----E-----	217
H2_01_AB.JP.08.08JP_NMC842	R-R--ITX-RK-NNRAQ-GSSQSSAP-TH-----E-----	217
U.FR.96.12034	R-FR-NC---FQL-RK-CERYQ-GSSATSSQ-TY-----A----	214
MAC.US.x.17EC1		215
MAC.US.x.17EFR		215
MAC.US.x.1937		215
MAC.US.x.2065		215
MAC.US.x.251_1A11		215
MAC.US.x.251_1A11		215
MAC.US.x.251_32H_PJ5		215
MAC.US.x.251_BK28		215
MAC.US.x.80035		215
MAC.US.x.81035		215
MAC.US.x.85013		215
MAC.US.x.87082		215
MAC.US.x.92050		215
MAC.US.x.92077		215
MAC.US.x.93057		215
MAC.US.x.93062		215
MAC.US.x.95058		215
MAC.US.x.95086		215
MAC.US.x.95112		215
MAC.US.x.96016		215
MAC.US.x.96020		215
MAC.US.x.96072		215
MAC.US.x.96081		215
MAC.US.x.96093		215
MAC.US.x.96114		215
MAC.US.x.96123		215
MAC.US.x.96135		215
MAC.US.x.97009		215
MAC.US.x.97074		215
MAC.US.x.MM142_IVMXX		215
MAC.US.x.SNM142_patent		215
MAC.US.x.r80025		215
MAC.US.x.r90131		215
MNE.US.82.MNE_8		215
MNE.US.x.MNE027		215
SMM.SL.92.SL92B		217
SMM.US.x.F236_H4		215
SMM.US.x.H9		215
SMM.US.x.PBJ14_15		215
SMM.US.x.PBJA		215
SMM.US.x.PBJ_143		215
SMM.US.x.PBJ_6P6		215
SMM.US.x.PGMS3		215
SMM.US.x.SME543		215
STM.US.89.STM_37_16		215

	Vpx start	Vpx end
MAC.US.x.239	MSPDRERIPPGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVWORSWEYWHDEQGMSPSYVKYRYLCLIQKALFMHCKKGCRCLEGEHGAGG.WRP.GPPPPPPGLA*	112
A.CI.88.UC2	-T---V-----D-I-AL-----R-----L-S-T---M--M-I-L--T--G--P-----V-	113
A.DE.x.BEN	-T---V-----D-I-AL-----R-----L-S-T---M--M-I-L--T--G--P-----V-	113
A.DE.x.PE12 KR KRCC	-A---KTV--G--V---A-E-M-A-----R---DL-Q-T---R-M-Y-M-I-V--T--G--P-----V-	113
A.GH.x.GH1	-T---V-----D-I-AL-----R-----L-S-T---M--M-I-L--T--G--P-----V-	113
A.GM.07.D194	-A---V-----D-I-AL-----R-----L-S-T---M--M-I-L--T--G--P-----V-	113
A.GM.x.ISY SBL 6669 85	-TN--T-----E-D-D-A-----R---R-T---M--V-F--T-R--P-----S--V-	113
A.GM.x.MCN13	-T---TV-----E-N-D-A-----R---O-T---M--T-R-T-G--P-----V-	113
A.GM.x.MCR35	-T---TV-----E-N-D-I-A-----R---O-T---M--T-R-T-G--P-----V-	113
A.GW.86.FG c1one NIHZ	-T---TV-----E-A-D-A-----H-----R---T-T---M-MY-A-RDGT-G#-PE--S#-V-	112
A.GW.07.CAM2CG	-T---TV-----E-A-D-A-----R---O-T---M--V-F-R--T--G--P-----V-	113
A.GW.x.AL1	-AN--TV-----E-A-D-AL-----R---O-T---M--MYT-FM--T--G--P-----S--V-	113
A.GW.x.MDS	-A---TV-----EQ-A-D-A-----R---T-T---M--MYI-GR--T-R--P-----V-	113
A.IN.07.NNVA	-T---TV-----A-E-A-----G---A-T---M--M-I-ARR--T--G--G--S--L--VW-	113
A.IN.95.CR1K 147	-T---TV-----A-D-A-----G---A-T---M--T-I-VRR--T--G--P-----S--V-	113
A.SN.85.R0D	-T---TV-----A-D-A-----R---E-T---I---VY-VR--T--R--P-----V-	113
B.CI.88.UC1	KM---V---D---D-E-IT-L-V-----C-A-RE-S-T---L-M--M-V-YT---Q--P--S--	113
B.CI.x.20 56	KM---V---D---V-A-E-LA-L-K-----K-A-R---I-T---L-M--NYI-IA--T-RR--P--Q--	113
B.CI.x.EH0	KM---V---V-----E-T-L-HL-V-----K-A-RE-I-T---L-M--M-I-FA--G-R--P--Q--	113
B.GH.86.D205 ALT	KM---V---D---V-A-E-IT-L-V-----A-RE-I-T---L-M--M-V-YT---Q--P--S--	113
B.JP.01.KR020 IMCJ KR020 1	EM---V---D---E-IT-L-V-----A-E-IT-I-T---M--Y-LAR--A-R--P-----S--	113
G.CI.92.Abt96	DM---X-E-D-----E-XT--I-----R---A-T---L-M--M-V---T-QK--P--Q--	113
AB.CM.03.03CM 510 03	KM---V---D---X-E-IT-L-V-----C-A-R---S-T---L-M--M-V-YT--T-RQ--X--S--	113
H2 01 AB.CI.90.7312A	KM---V---D---V-A-E-IV-L-----A-R-D-I-T---L-M--M-V-FA--T-RG--P--R--	113
H2 01 AB.JP.04.04JP NMC307	KM---V---D---D-E-IT-L-V-----A-RE-S-XT---L-M--M-V-FS--T-RR--P--Q--	113
H2 01 AB.JP.07.07JP NMC716	KM---V---D---E-N-E-IT-L-----A-RE-S-T---L-M--M-V-FS--T-RR--P--Q--	113
H2 01 AB.JP.08.08JP NMC842	RMX-X-XV-D---X-E-IT-L-----A-RE-T-A---L-M--M-V-FS--T-RR--P--Q--	113
U.FR.96.12034	-G---E-V--S-E-RD--A-Q-----Q-R-E-T---Q--P--S--I-	113
MAC.US.x.17EC1	-----	113
MAC.US.x.17EFR	-----	113
MAC.US.x.1937	-----	113
MAC.US.x.2065	-----	113
MAC.US.x.251 1A11	-----	113
MAC.US.x.251 32H P35	-----V-----Q-----M-----	113
MAC.US.x.251 BK28	-----Q-----M-----	113
MAC.US.x.80035	-----X-----X-----	113
MAC.US.x.81035	-----X-----X-----	113
MAC.US.x.85013	-----X-----X-----	113
MAC.US.x.87082	-----	113
MAC.US.x.92050	-----	113
MAC.US.x.92077	-----	113
MAC.US.x.93057	-----	113
MAC.US.x.93062	-----	113
MAC.US.x.95058	-----	113
MAC.US.x.95086	-----	113
MAC.US.x.95112	-----	113
MAC.US.x.96016	-----	113
MAC.US.x.96020	-----	113
MAC.US.x.96072	-----	113
MAC.US.x.96081	-----	113
MAC.US.x.96093	-----	113
MAC.US.x.96114	-----	113
MAC.US.x.96123	-----	113
MAC.US.x.96135	-X-----	113
MAC.US.x.97009	-----	113
MAC.US.x.97074	-----	113
MAC.US.x.MM142 IVMXX	-----Q-T-----Q-T-----	113
MAC.US.x.SNM142 patent	-----Q-T-----Q-T-----	113
MAC.US.x.r80025	-----	113
MAC.US.x.r90131	-----Q-----	113
MNE.US.82.MNE 8	-----	113
MNE.US.x.MNE027	-----	113
SMM.SL.92.SL92B	-T-----HN-AL-QT-Q-----R-C--V--Y--A--VQ-M--M-Q-FR--T-R--SQ--T-	113
SMM.US.x.F236_H4	-----K-----A-----R-----M-E-T-----V-----E-----T-----	113
SMM.US.x.H9	-XX-X-----XX-X-D-H-X-A-X-----R-X-----M-V-T-X-----E-----T-----	113
SMM.US.x.PB114 15	-----D-D-----A-----R-----Y-M-V-T-----M-----E-----E-----	113
SMM.US.x.PBJA	-----D-D-----A-----R-----M-V-T-----M-----E-----E-----	113
SMM.US.x.PBJ 143	-XX-X-----XX-X-D-H-X-A-X-----R-X-----M-V-T-X-----E-----T-----	113
SMM.US.x.PBJ 6P6	-----D-D-----A-----R-----M-V-T-----M-----E-----E-----	113
SMM.US.x.PGMS3	-----E-----G-A-----R-----R-M-E-T-----V-R-----E-----S-----	113
SMM.US.x.SME543	-----E-----G-A-----R-----R-M-E-T-----V-R-----E-----S-----	113
STM.US.09.STM 37 16	-----A-E--H--D--A-----R-----PG-----S-R-----	113

	Vpr start	Vpr end	
MAC.US.x.239	MEER...PPE.NEGPQREPWDEWVVEEELKEEALKHFDPRLLLTALGNHYNRHGDTELEGAGELIRILQRALFMHFRGGCIHSRIGOPGGGNPLSAIPPSRSM*		101
A.CI.88.UC2	...#WK-LG-DC#...*I---*R---I-D-Y---N---R-K---S-L-V-GSR---TRRRT-CP-T-TP-G-H-		83
A.DE.x.BEN	-T-APTEF...DGT-R-DLGS-D-I-T-R-I---R---I-YY-H---R-KT---V-A-NR---*TRRRT-CP-A-TP-A-H-		105
A.DE.x.PE12_KR_KRCG	-T-APAEF...D-T-P-G-G---IGI-R-R---R---T---Y-CA---S-R-NV---V-L-A-KI---TR-ET-F---TP-G-Q-		106
A.GH.x.GH1	-T-APTEF...DGT-R-LGGD-IRI-G-I---I---Y-HS---P-R---V---V-L-A-NR---S-TRRRT-FP-A-TP-G-Y-		106
A.GM.87.D194	-T-APTEF...DGT-R-LGST-T-K-I---C-I---Y---R---V---V-I-A-DR-K-TRRA-CP-A-TP-G-H-		106
A.GM.x.ISY_SBL_6669_85	-T-APAEF...DGT-P-G-I-I-R-I---R---Y---T---R---V---T-A-G-R---R-R---TP-N-Q-		106
A.GM.x.MCN13	-T-APTEF...DGT-P-G-II-I-RKI-K---Y-HT---R-NV---R-A-RL---T-RT-FP-TSTP-T-Q-		106
A.GM.x.MCR35	-T-APTEF...DGT-P-G-II-I-RKI-K---Y-HT---R-NV---R-A-RL---T-RT-FP-TSTP-T-Q-		106
A.GW.86.FG_clone_NIHZ	-T-APTEL...DRT-P-G-A-I-I-R-IE---R---RY-T---R---V---A-A-G---TR---TP-G-H-		106
A.GW.87.CAM2CG	-T-APTEL...DGT-P-G-I-I-RDI---R---G---A---R---V---T-A-N---TR---TP-R-*		105
A.GW.x.ALI1	-T-APTEF...A#-GT-PGS#-I-I-R-I---R-M-I-GY-T---R-R-NA---T-A-GR-V---TR-R---TP-N-Q-		105
A.GW.x.MDS	-T-APTEL...DGT-P-G-I-I-R-IT---R-RS---R-T---R---V---T-A-G---TR---TP--Q-		106
A.IN.07.NNVA	-T-APTEL...DRT-P-G-I-I-R-I---R---RY-T-Y-P---R---V---T-A-G---R---TP--R-		106
A.IN.95.CR1K_147	-T-APTEL...DGT-P-G-I-I-R-I---R---IY-D-Y---R---V---T-A-G---R---T-TP--R-		106
A.SN.85.ROD	-A-APTEL-V.DGT-L-G-II-I-R-I---I-KY-T---R-KV---T-A-G---TR---TP-N-Q-		106
B.CI.88.UC1	-A-AAPEL...S---E-ED-M-I-Q-R---F-S---A---K---L---A-Q---*--R-		105
B.CI.x.20_56	-A-AAPEL...D---E-ED-I-Q-R---F-S---A---K---L---O---*G-Q-		105
B.CI.x.EHO	-A-AAPEL...DKN---EQ-D-I-Q-R---F-S---N---A---K---L---O---*G-Q-		105
B.GH.86.D205_ALT	-A-AAPEL...N---E-IG-I-I-Q-R---F-S---A---K---L---A-Q---S---T---P*G-R-		105
B.JP.01.KR020_IMCJ_KR020_1	-T-AAPEL...D---E-ED---Q-R---F-G---IA---L---O---V---*--QK		105
G.CI.92.Abt96	-A-E...I---D-A---I---Y-D---K---L---D-RC-GNA---T---GVF-		103
AB.CM.03.03CM_510_03	-A-AXXXI...S---Q-E-RXG-I---I-Q-X---X-F-S-Y---A-VK---L---A-R---X---P*G-Q-		105
H2_01_AB_C1_90_7312A	-A-AASET...D---E-ED-I-Q-R---F-S---A---K---L---O---T-EG-Q-		106
H2_01_AB_JP_04_04JP_NMC307	-A-AAPEL...D-T-G-E-EDI-Q-R---F-S---A---K---L---O---PGG-*		105
H2_01_AB_JP_07_07JP_NMC716	-A-AAPEL...N-G-E-EDI-Q-R---F-S---A---K---L---O---V---PGGL*--		105
H2_01_AB_JP_08_08JP_NMC842	-A-AAPEL...D-S-XG-E-EDI-Q-R---F-S-Y---A---K---L---O---PGG-*		105
U.FR.96.12034	-A-I...DGA-G-R-I---VR---Y---K---I-L-R-Q---SSR---TTV---*GV-		101
MAC.US.x.17EC1			102
MAC.US.x.17EFR			102
MAC.US.x.1937			102
MAC.US.x.2065			102
MAC.US.x.251_1A11			101
MAC.US.x.251_32H_PJ5			102
MAC.US.x.251_BK28			101
MAC.US.x.80035			102
MAC.US.x.81035			102
MAC.US.x.85013			102
MAC.US.x.87082			102
MAC.US.x.92050			102
MAC.US.x.92077			102
MAC.US.x.93057			102
MAC.US.x.93062			102
MAC.US.x.95058			102
MAC.US.x.95086			102
MAC.US.x.95112			102
MAC.US.x.96016			102
MAC.US.x.96020			102
MAC.US.x.96072			102
MAC.US.x.96081			102
MAC.US.x.96093			102
MAC.US.x.96114			102
MAC.US.x.96123			102
MAC.US.x.96135			102
MAC.US.x.97009			102
MAC.US.x.97074			102
MAC.US.x.MM142_IVMXX			102
MAC.US.x.SM1142_patent			102
MAC.US.x.r80025			102
MAC.US.x.r90131			102
MNE.US.82.MNE_8			102
MNE.US.x.MNE027			102
SMM.SL.92.SL92B			101
SMM.US.x.F236_H4			90
SMM.US.x.H9			102
SMM.US.x.PB114_15			102
SMM.US.x.PBJA			102
SMM.US.x.PBJ_143			102
SMM.US.x.PBJ_6P6			102
SMM.US.x.PGMS3			102
SMM.US.x.SME543			102
STM.US.89.STM_37_16			102

	Tat start	exon 1 end	exon 2 start	Tat end	
MAC.US.x.239	METPLREQNSLESSNRSSCISEADASTPESANLGEELISQLYRPLEACYNTCYCKKCCYHCQFCFLKKGGLGICYEQSRKRRRTPKKAKANTSSASNKPISNRTRHCOPEKAKKTEVKEAVATAPGLGR*				130
A.CI.88.UC2	-----KAP-S-K-Y-P-P-T-WEVAAQ-L-KQ-L-L-A-H-----T-P-----SF-L-----W-VRKGG-----RT-THPP-TPD-S-IQ-GDSR-T-KQ---P-TP---TS-----				131
A.DE.x.BEN	-----KAP-S-KPY-P-T-R-VTAQ-L-KQ-L-L-A-H-P-T-K-----R-SF-L-S-----S-RKGR-----R-T-TPSP-PD-S-T-GDS-T-EQ-K-S-AT-V-TC---Q-----				131
A.DE.x.PET2_KR_KRCG	-----SKAP-S-M-C-P-T-Q-VKSO-L-KQ-RL-----Q-N-P-----L-----R-KGR-----H-S-----D-S-T-GNS-T-KQTK-P-T-LE-R---Q-----				131
A.GH.x.GH1	-----H-KAP-S-Y-P-T-QGVTAQ-L-KQ-L-L-A-H-----T-S-----Q-SF-L-----W-ARKSR-----R-T-THS-----D-S-T-GDS-T-EQ-K-T-TMV-TCS---S-----				131
A.GM.87.D194	-----K-P-S-Y-P-T-R-VTAQ-R-KQ-L-L-A-H-----T-S-----Q-S-L-----W-ARQGR-----R-T-THPP-D-S-T-GDS-T-KQ-K-P-TT-VS-C---H-----				131
A.GM.x.ISY_SBL_6669_85	-----KAP-S-G-Y-P-RT-Q-VA-Q-L-Q-----T-N-K-F-G-F-L-N-----DRKGR-S-T-HS-P-D-S-T-GNS-T-KQ-K-LGTTLEAD-----				131
A.GM.x.MCN13	-----KAP-S-M-Y-P-T-R-VGSO-L-KQ-L-H-P-N-K-G-F-L-N-----DRKGR-----T-HS-----D-S-T-GNS-----KQ-K-L-TTLE-R-----				131
A.GM.x.MCR35	-----KAP-S-M-Y-P-T-R-VGSO-L-KQ-L-H-P-N-K-G-F-L-N-----DRKGR-----T-HS-----D-S-T-GNS-----KQ-K-L-TTLE-CS-----				131
A.GW.86.FG_clone_NIHZ	-----KAP-S-C-P-RT-Q-VA-Q-L-RQ-----T-S-----D-L-Q-----W-DRKGR-----T-HP-----D-S-T-NS-----KQ-K-L-AT-E-DL-----				131
A.GW.87.CAM2CG	-----K-P-S-G-C-P-RT-Q-L-K-----Q-E-D-S-----R-L-Q-----W-DRKGR-----H-S-----D-S-T-NS-A-KQ-K-L-AT-E-DL-----				131
A.GW.x.AL1	-----K-PGS-MPY-P-T-Q-VAQ-L-Q-----T-N-----L-N-----W-DRKGR-S-I-HS-----D-S-T-NS-EKQ-K-L-TTLG-DC-P-S-----				131
A.GW.x.MDS	-----I-KAP-S-Y-P-HT-Q-VA-Q-L-RQ-----N-N-R-----I-N-----W-RKGR-----T-HP-----D-S-T-GNS-T-KQ-K-L-AT-E-DL-P-----				131
A.IN.07.NNVA	-----KVP-S-C-PF-HT-Q-VA-Q-VQE-----FQ-----D-S-----F-EL-Q-----DRKGR-----R-T-HP-TPD-S-T-NS-T-EQ-K-L-TT-D-DL-P-----				131
A.IN.95.CR1K_147	-----KAP-S-C-PFLHT-Q-VA-QG-----AS-----L-L-Q-----R-KGR-----T-THPTLT-D-S-T-WNS-----E-K-L-TT-V-DLDP-----				130
A.SN.85.ROD	-----KAP-S-K-C-PF-RT-Q-VA-Q-L-RQ-----T-N-S-----R-M-N-----R-KGR-----T-THP-PTPD-S-T-GDS-T-KQ-K-AT-E-DT-P-----				131
B.CI.88.UC1	-----I-Q-S-K-S-P-ST-PVVN-QG-D-Q-----D-K-----L-----W-D-HS-K-SS-R-VTA-----DESL-AN-GDS-T-KQ-TK-TKGL-DL-P-----				129
B.CI.x.20_56	-----I-SQ-----RF-----ST-GEQ-----QGLD-Q-----N-----F-----L-----W-D-HS-K-SS-R-VTA-----PDESLS-TS-GDS-T-KQ-TK-EKGI-DL-P-----				129
B.CI.x.EHO	-----I-K-S-N-SGH-ST-GV-N-QGLD-R-----K-S-----S-L-----S-K-SS-R-TT-----P-ESL-A-GDS-T-KQ-KE-TTR-DL-P-S-----				129
B.GH.86.D205_ALT	-----I-Q-S-K-----P-ST-PVVN-QGLD-Q-----K-D-----L-----D-S-----SA-R-TTAP-PD-SL-A-GDS-T-KQ-KE-TTGT-D-P-----				129
B.JP.01.KR020_IMCJ_KR020_1	-----GI-Q-----F-S-----ST-EG-N-RGLD-Q-----R-K-----L-----D-HS-K-SS-R-VTAPT-----ESL-T-A-DG-A-KQ-KE-TTRT-D-----S-----				129
G.CI.92.Abt96	-----S-----K-CR-L-TF-ETVDA-GLEGTQA-----Y-----S-K-----L-N-V-----P-R-----T-SF-----ES-A-NX-K-K*-K-TK-----DL-----				129
AB.CM.03.03CM_510_03	-----I-QG-S-K-S-P-ST-PVVD-QG-DSQ-----X-K-N-K-----X-----L-R-----W-D-HS-K-R-R-TTA-P-PD-S-T-DS-A-KQ-KE-TTGT-DL-P-----				127
H2_01_AB.CI.90_7312A	-----I-QG-----RF-S-----ST-VVN-QGLD-Q-----H-R-N-----L-----D-S-K-SS-RN-TA-----D-SV-T-NS-T-KQ-KE-TTG-DL-P-S-----				129
H2_01_AB.JP.04.04JP_NMC307	-----Q-G-SKF-K-H-ST-VVN-QGLD-Q-----G-N-K-A-----L-N-----S-K-SS-TA-----D-SV-TS-GNS-T-KQ-KE-ATR-DL-S-----				129
H2_01_AB.JP.07.07JP_NMC716	-----Q-G-SKF-K-H-ST-VVN-QGLD-Q-----G-N-K-A-----L-N-----HP-K-SS-RTA-----D-SV-TS-GNS-T-KQ-KE-ATR-DL-S-----				129
H2_01_AB.JP.08.08JP_NMC842	-----Q-G-SKF-K-H-ST-VVN-QGLD-Q-----G-N-K-A-----L-N-----D-S-K-SS-R-TA-----D-SV-TS-GNS-TT-KQ-KE-TTR-D-----S-----				129
U.FR.96.12034	-----I-K-S-S-K-CK-Q-ST-G-VG-QG-AA-----S-K-----L-----P-----S-V-YS-AP-----SL-G-NS-KEK*-KAL-T-----DL-----				128
MAC.US.x.17EC1					131
MAC.US.x.17EFR					131
MAC.US.x.1937					131
MAC.US.x.2065					131
MAC.US.x.251_1A11					131
MAC.US.x.251_32H_PJ5					131
MAC.US.x.251_BK28					131
MAC.US.x.80835					131
MAC.US.x.81035					131
MAC.US.x.85013					131
MAC.US.x.87082					131
MAC.US.x.92050					131
MAC.US.x.92077					131
MAC.US.x.93057					131
MAC.US.x.93062					131
MAC.US.x.95058					131
MAC.US.x.95086					131
MAC.US.x.95112					131
MAC.US.x.96016					131
MAC.US.x.96020					131
MAC.US.x.96072					131
MAC.US.x.96081					131
MAC.US.x.96093					131
MAC.US.x.96114					131
MAC.US.x.96123					131
MAC.US.x.96135					131
MAC.US.x.97009					131
MAC.US.x.97074					131
MAC.US.x.MM142_IVMXX					131
MAC.US.x.SNM142_patent					131
MAC.US.x.r80025					131
MAC.US.x.r90131					131
MNE.US.82.MNE_8					131
MNE.US.x.MNE027					131
SMM.SL.92.SL92B					127
SMM.US.x.F236_H4					117
SMM.US.x.H9					129
SMM.US.x.PB114_15					129
SMM.US.x.PB1A					129
SMM.US.x.PB1_143					129
SMM.US.x.PB1_6P6					129
SMM.US.x.PGM53					131
SMM.US.x.SME543					129
STM.US.89.STM_37_16					129

	Rev start	exon 1 end	exon 2 start	Rev end
MAC.US.x.239	MSNIHEREEELRKLRLIHLHLLQTNPYPTGPGTANQRORRKRWRRRRQWLLALADRIYSFPDPPTDPLDLAIQQLONLAIESIPDPPTN...TPEALCDPTEDS.RSPQD*			107
A.CI.88.UC2	-CEKAD---QRE---R-----YR-----G---N-R-R-Q-LR-----KLHTA---A-SS-W---H-G-T-REL---D...L-...SOSNO.GLAET-			104
A.DE.x.BEN	-ERAD-G-OGK-LR-----0---S---N-R-R-Q-LR-V---NKLCV---S-R---H-R-T-OEL---D...L-...SNSNO.GLAET-			104
A.DE.x.PET2_KR_KRCG	-NGRAD-G-ORQK-Q-R-----0-L-R---N-R-R-KOH-R-V---NS-T---A-S-R---R-G-T-OEL---D...L-...SSESTNNNO.GLAETY			108
A.GH.x.GH1	-HEKADG--QE---R-----H-0---S---N-R-R-Q-FR-V---KLHTI---S-R---G-T-OEL---D...L-...SNSNO.GLAET-			104
A.GM.87.D194	-RORAD-G-GEK--R-----I-----0-L-R---N-R-R-E-K-I---T---A-P-OT---G-T-OTL---D...L-...SNSNO.GLAET-			104
A.GM.x.ISY_SBL_6669_85	-TERAD-GV-RK--R-----0-L-S---N-R-R-KO-R-V---K-T---A-S-GRT---H-E-T-ODL---S...F-SSEGTDSNO.-LAKN-			97
A.GM.x.MCN13	-TGRAD---RK-----0-L-S---N-R-R-KO-R-V---K-T---A-S-RT---H-E-T-ODL---S...F-SSEGTDSNO.-LAKN-			108
A.GM.x.MCR35	-TGRAD---RK-----0-L-S---N-R-R-KO-R-V---K-T---A-S-RT---H-E-T-ODL---S...F-SSEGTDSNO.-LAKN-			108
A.GW.86.FG_c1one_NIHZ	-TERAD-G-ORK--R-----0---S---N-R-R-KO-R-I---S-T---A-S-R---H-G-T-ODL---H...L-...SPESTNSNO.-LAEA-			108
A.GW.87.CAM2CG	-TERAD-G-ORK--R-----0---S---N-R-R-KO-R-I---S-T---A-S-RT---H-E-T-ODL---H...L-...SPESTNSNO.-LAEA-			101
A.GW.x.ALI	-TERAG-D-ORK--R-----0---S---N-R-R-Q-R-IV---FT---ASS-R-V-H-G-T-ODL---D...L-...SSESANNO.GLAET-			108
A.GW.x.MDS	-NERAD-G-ORK--R-----0---S---N-R-R-Q-R-I---S-T---A-S-OTV---G-T-ODL---H...L-...SO.-PAAT-			101
A.IN.07.NNVA	-TERAG-GPERK--R-----0---S---N-R-R-Q-I-I---S-T---A-S-V-OT---G-T-OEL---H...L-...SO.-LAEA-			101
A.IN.95.CR1K_147	-NERAD-G-ORK--R-----0---S---N-R-R-Q-L-I-T---S-T---ANS-V-OT---G-T-OEL---H...L-...SO.-LAEI-			101
A.SN.85.ROD	-NERAD-G-ORK--R-----0---S---N-R-R-KO-R-I---S-T---A-S-OT---H-G-T-OEL---H...L-...SO.-LAEI-			101
B.CI.88.UC1	-T.T-KD-Q-G--L-----0T---S---N-R-R-K-GL-I---R-LS-S-EE---V-R-E-T-DL-N-S-...TT-AF			95
B.CI.x.20_56	-T.T-KD-Q-G--L-----0A---S---N-R-R-K-GL-I---R-S-S-EE---V-R-E-T-DL-N-S-...TT-AF			95
B.CI.x.EHO	-N.A-RD-Q-G--L-----0---S---N-R-R-K-GL-I---HPL-S-EG---R---I-KDL-N-S-...TA-AS			95
B.GH.86.D205_ALT	-T.A-GD-Q-E--L-----0---S---N-R-R-K-GL-I---H-STA-AEE-N---R---TV-DL-N-P-LNOS-...TT-AP			98
B.JP.01.KR020_IMCJ_KR020_1	-T.T-KD-Q-G--L-----0---S---N-R-R-K-GL-I---Q-HPL-S-EE-N---K-T-DL-T-S...I-...TA-TS			95
G.CI.92.Abt96	---L.E-----F---0-X---N-R-R-K-KO-L-I---X-ARE---S-T-ODX-S-A...V-KN-Q-PSN*-TSPAR			106
AB.CM.03.03CM_510_03	-T.T-R-.E-E--L-----H---N-R-R-K-GL-I---R-SAA-AEE-N---R---TV-DL-N-S-...TT-AP			93
H2_01_AB.CI.90.7312A	-T.A-KD-Q-GI--L-----0---S---N-R-R-K-GL-I---HPL-NS-AEE-RR---TV-DL-N-S-...TT-AP			95
H2_01_AB.JP.04.04JP_NMC307	-N.A-KD-Q-G--L-----0A---S---N-R-R-K-GL-I---HPL-S-AEE-RR---DL-N-S-...TT-AP			95
H2_01_AB.JP.07.07JP_NMC716	-N.T-KD-Q-G--L-----0A---S---N-R-R-K-GL-I---HPL-T-AEE-RR---T-DL-N-S-...TT-AP			95
H2_01_AB.JP.08.08JP_NMC842	-T.A-KD-Q-G--L-----0A---R---N-R-R-K-GL-I---HPL-NS-AEE-VRR---T-DL-N-S-...TI-GP			95
U.FR.96.12034	-P.-GVQR-----N-R-R-Q-L-I---K-T---S-V---G---ODL-E-AV...V-HQ-...TSAP-			99
MAC.US.x.17EC1	-N-----R-----T-----X-----I-X-----			108
MAC.US.x.17EFR	-XX-R-----R-----X-----X-----I-X-----			108
MAC.US.x.1937	-S-----R-----X-----I-X-----K-----			108
MAC.US.x.2065	-S-----R-----X-----I-X-----K-----			108
MAC.US.x.251_1A11	-S-----R-----X-----I-X-----K-----			108
MAC.US.x.251_32H_PJ5	-S-----S-----R-----I-X-----G-N-----			108
MAC.US.x.251_BK28	-S-----S-----R-----I-X-----KG-----			108
MAC.US.x.80035	-X-----R-----X-----X-----			108
MAC.US.x.81035	-X-----R-----X-----X-----			108
MAC.US.x.85013	-X-----R-----X-----X-----			108
MAC.US.x.87082	-X-----R-----X-----X-----			108
MAC.US.x.92050	-X-----R-----X-----X-----			108
MAC.US.x.92050	-X-----R-----X-----X-----			108
MAC.US.x.92077	-X-----R-----X-----X-----			108
MAC.US.x.93057	-X-----R-----X-----X-----			108
MAC.US.x.93062	-X-----R-----X-----X-----			108
MAC.US.x.95058	-X-----R-----X-----X-----			108
MAC.US.x.95086	-X-----R-----X-----X-----			108
MAC.US.x.95112	-X-----R-----X-----X-----			108
MAC.US.x.96016	-X-----R-----X-----X-----			108
MAC.US.x.96020	-X-----R-----X-----X-----			108
MAC.US.x.96072	-X-----R-----X-----X-----			108
MAC.US.x.96081	-X-----R-----X-----X-----			108
MAC.US.x.96093	-X-----R-----X-----X-----			108
MAC.US.x.96114	-X-----R-----X-----X-----			108
MAC.US.x.96123	-X-----R-----X-----X-----			108
MAC.US.x.96135	-X-----R-----X-----X-----			108
MAC.US.x.97009	-X-----R-----X-----X-----			108
MAC.US.x.97074	-X-----R-----X-----X-----			108
MAC.US.x.MM142_IVMXX	-RS-TG-R-----R-----X-----R-----I-----N-----A-----			108
MAC.US.x.SNM142_patent	-RS-TG-R-----R-----9-S-----KR-Q-----I-----#-----N-#-----A-----			107
MAC.US.x.r80025	-R-----R-----R-----R-----R-----R-----R-----R-----			108
MAC.US.x.r90131	-R-----R-----R-----R-----R-----R-----R-----R-----			108
MNE.US.82.MNE_8	-S-AE-R-----R-----S---N-R-R-Q-----F---V---N---V-----I-I-H---NP-----			108
MNE.US.x.MNE027	-S-AE-R-----R-----S---N-R-R-Q-----F---V---N---V-----I-I-H---SP-----			108
SMM.SL.92.SL92B	-Q-P-----R-----I-----A---S---N-R-R-K-I-IV-----A-DI-----FDS-S-Q-L-E-T---V-T-R-QSA*-AC-KR			105
SMM.US.x.F236_H4	-ST-----R-----F-----0-----R-R-R-Q-----I-----V-----G-EL---S-A-P-N-VAKSP*-H-N-			105
SMM.US.x.H9	-SN-----R-----F-----XX-----R-R-R-XO-----I-X-----A-----G-XL-N-AS---A-P-K-AA-SP*-H-S-			105
SMM.US.x.PB114_15	T-SN-G-----R-----S-----D-----R-R-R-Q-----I-----V-----R-----EL-N-AS---A-P-K-AA-SP*-H-S-			105
SMM.US.x.PBJA	-SN-----R-----R-----D-----R-R-R-Q-----I-----V-----R-----EL-N-AS---A-P-K-AA-SP*-H-S-			105
SMM.US.x.PBJ_143	-SN-----R-----XX-----R-R-R-XO-----I-X-----A-----G-XL-N-AS---A-P-K-AA-SP*-H-S-			105
SMM.US.x.PBJ_6P6	-SN-----R-----D-----R-R-R-Q-----I-----V-----R-----EL-N-AS---A-P-K-AA-SP*-H-S-			105
SMM.US.x.PGMS3	-SSAG-R-----F-----0-----R-R-R-Q-----I-----A-S---V---D-SL-N---S-V-----L-NP---R-			105
SMM.US.x.SME543	-ST-----R-----F-----0-----R-R-R-Q-----I-----N-----G---DL---S-A-T-K-AAKS*-H-N-			105
STM.US.89.STM_37_16	-DQ-----OF-----Q-----N-R-R-Q-N-----N-----AS---V---G-S-QDL---P---L-KD-Q-TA-N*-R-TR			105

Env start	signal peptide end_gp120 start	V1
MAC.US.x.239	MGCLGNQLLIAILLLSVY.GIYCT.LYVTVFYGVPWARNATIPLFCA ¹ TKNRD ² TWGT ³ QCLPDNGDYSEVALNVTESFD ⁴ AWNNTVTEQAIEDV ⁵ WOLFETS ⁶ IKPCVKLSPL ⁷ CI ⁸ TRMCK ⁹ NS ¹⁰ ETDRWGLTKSITTTASTTST...TA.SAKV.DMVNETSSIAO ¹¹ DNCT.G	159
A.CI.88.UC2	-EPGR--AV--T-AC.L-K.Q--V--S--I--D-Q-IP--A--D--R--T-VA-N-P...VTGNNTNA-AKP-AARP...TNP-YLT..II-S-T-VGA--	156
A.DE.x.BEN	-EPGR--FVW--T-AC.LV-S.Q--I--K-S--I--D-Q-II--A--D--V--H--T-VA-N-SR...VQGNITTPNRTSS...SRPPTSAA ¹² SI--N-ENNT-A	157
A.DE.x.PET2_KR_KRCG	IMDSR-IV--T-AC.L-A.Q--I--K-S--R--E--I--D-Q-IP--A--D--V--N--V--T-VO-E-S...TSTESSNS-SEG-VPEI..L--T--TNS-S.D	148
A.GH.x.GH1	KM-GKSL-CV-S-A-A-LV-Q--V--S--I--K-D-Q-IT--A--D--V--S--T-VA-S-S...TINNTT...GSTTGM ¹³ S..ET..PS-YS	145
A.GM.87.D194	-EPGR--V--T-AC.L-K.Q--I--S--I--D-Q-IT--A--D--R--T-VA-N-I...TSG-TA...PSPNIT..IID-N-T-GDN	145
A.GM.x.ISY_SBL_6669_85	TMSGKI--V-F-T-AC.L-K.Q--I--V-K-S--I--D-Q-IP--A--D-I--V--N--T-VA-N...ASTESAVA...TSP-SGRD.M--I-D-DP-QLN-S	146
A.GM.x.MCN13	-MGGR--V--T-TC.L-N--I--S--I--D-Q-IT--A--D--N--T-VA--N...TDARNITTT...PT-A-SPRT-I-PVTEIS-N-RAN-S	152
A.GM.x.MCR35	-MGGR--V-A-TC.L-N--I--S--I--D-Q-IT--A--D--N--T-VA--N...TDARNITTT...PT-A-SPRT-I-PVTEIS-N-T-RAN-S	152
A.GM.86.FG clone NIHZ	-KGSK--V-A-A-L-H-K-QF--I--S--I--D-Q-IT--A--D--V--N--T-VA-N...TRNMTT...W-GRT...DTONITII-D-H-ARA	145
A.GW.87.CAM2CG	-EPGR--V--T-AC.L-R.QQ--I--K-S--I--D-Q-IP--A--D-I--V--N--T-VA-K-I...STSDITM...TRT...TPS-TAKEAPISDW-P-RTN-S	151
A.GW.x.ALI	-MSSR--VT--A-AC.LV-K.Q--I--K-S--I--D-Q-IP--A--D--V--N--T-VA-K-SN...TSTESTIT...SPSPST...LKPL..I-SDP--PR-	130
A.GW.x.MDS	-TRKMH--T-AC.L-KPOO--I--S--I--D-E-I--A--D--V--S--T-VA-N-SRLSDSASTRNITTTNA-A-IIR..ATTI..PRNTTIS-N-P-RA-S	164
A.IN.07.NNVA	-TRERA--VV--T-TC.L-K.QQ--I--K-S--I--D-Q-I--A--D--V--N--T-VA-E-KN...ETGNATT...TAPT-T...PEINETTE..IS-NT-VRG-S	153
A.IN.95.CR1K_147	-AHE-T--F-T-AC.L-K.Q--I--K-S--R--I--D-Q-IV--A-A-D--N--T-VA-V--T...STT.TTT...TAPTST...AGSTATPKPI--NT-MHAN-S	154
A.SN.85.ROD	...M--A-AC.LV-Q--T-K--R--I--D-Q-IT--A--D--H--T-VA-K-SS...TESSTGN-T-KTST...TTPTDQEQEIS-DTP-ARA-S	154
B.CI.88.UC1	-AHTS-H-F-LL-I--GFLGHK.KN--I--V--T--V--T-ISV-I--A--VD-S--L--T-VA--S...NNTG-N-TTKPI-T..PITTTKPS ¹⁴ ENLL-D-P-KN-T-P	155
B.CI.x.20_56	-AHIHSH--SL-I--GCMC.KQ--I--V--A--V--S--T-ISV-I--A--VD-N--L--T-VA-N-ST...NNTR-NNTAST-T...GNSTTPIV...AIP-VKAN-S	152
B.CI.x.EHO	-AHVN-Y-VTL-I-I-GYMG.KNF--I--K-S--R--V--D-T-IO-I--A--D--D-TK-S--L--T-VA-V-K...W-SASKET--SA..SLRS-TO_TLL-D-K-QN-S-A	153
B.GH.86.D205_ALT	-AYFSSR-P-L-IGISGFVC.KQ--I--V--I--T--V--T-IR-I--A--D--Q-VD-R--T-VA-N-S...ETNPGN-S-T-I..KPTTTSRGLKTI--DP-KN-S	155
B.JP.01.KR020_IMCJ_KR020_1	-AHINKY-FA-L-I-GFMG..EDF--I--K--V--I--V--D-D-I-I--A--D--VA-N--T-VA-N-SN...LNTTAN-TTANNIT..ASKNI...S-L-GND--RD	151
G.CI.92.Abt96	-AY--TX--T-GF...K--I--SV--L--A--D--V--N--T-VA--T-NK--GTT-V-PA...VPTTKM ¹⁵ VTAEL-S-Q-LMY	159
ABM.CM.03.03CM_510_03	-AYSSSC-P-TL--ICGFTX..KQ-X--I--X-XX-X--I--D-Q-IT--A--X--R--T-XXVA-N-TT...EATPON...T...RNTT...IIEDNDP--SRN	145
H2_01_AB_C1_90_7312A	KM-GK-L-FV-S-A-A-L-K--I--D-Q-I--A--V--S--T-VA-S-S...TAT-PP--NNT...TTEPTTGGPEI--FP-MRT	155
H2_01_AB_JP_04_04JP_NMC307	KM-GK-L-FV-S-A-AC.L-S.KK--V--S--Q--V--D-Q-I--A--D--V--N--T-VA-N-TN...AT-S-NITR ¹⁶ PX-I..DN-ITTE..I-X--QR-S	152
H2_01_AB_JP_07_07JP_NMC716	KM-GK-L-F-S-T-S-L-SA.KK--V--S--Q--V--D-Q-I--A--D--V--N--T-VA-N-T...N-TKS-N-N..GN-VNITGTTLLI--P-PR-S	148
H2_01_AB_JP_08_08JP_NMC842	KMWGK-L-F-S-A-AC.L-S.KK--V--S--Q--V--D-Q-I--A--D--V--N--T-VA-N-T...N-MT..N-T..NTPT-STGGGLI--P-PR-S	148
U.FR.96.12034	...L-C-MACV..Q--I--N--L--A--V--S--T-A-S-T-N-GTAA...SPAMQESSKVI--EP-RNNS-S	156
MAC.US.x.17EC1	...G...V...M...R...A	159
MAC.US.x.17EFR	...G...M...R...A	159
MAC.US.x.1937	...X-XG-X...M...R...A	159
MAC.US.x.2065	...G...M...X...A	159
MAC.US.x.251_1A11	...G...M...X...A	159
MAC.US.x.251_32H_PJ5	...G...Q...E...L...E...K...L...P-AP...A...I...N...I...VH	158
MAC.US.x.251_BK28	...G...Q...L...E...E...K...L...P-AP...A...I...N...I...TH	158
MAC.US.x.80035	...X...G...M...X...A	161
MAC.US.x.81035	...G...M...X...A	159
MAC.US.x.85013	...G...X-X...M...X...A	159
MAC.US.x.87082	...G...M...R...A	159
MAC.US.x.92050	...G...M...R...A	159
MAC.US.x.92077	...G...M...R...A	159
MAC.US.x.93057	...G...Q...M...X...A	159
MAC.US.x.93062	...G...M...R...A	159
MAC.US.x.95058	...G...X...M...R...A	159
MAC.US.x.95086	...G...Q...L...E...E...K...L...P-AP...A...I...N...I...X	159
MAC.US.x.95112	...G...M...R...A	159
MAC.US.x.96016	...G...M...R...A	159
MAC.US.x.96020	...P...X...G...Q...L...E...E...K...L...P-AP...A...I...N...I...X	159
MAC.US.x.96072	...P...X...G...Q...L...E...E...K...L...P-AP...A...I...N...I...X	159
MAC.US.x.96081	...G...X...M...R...A	159
MAC.US.x.96093	...G...X...M...R...A	159
MAC.US.x.96114	...G...X-X...M...R...A	159
MAC.US.x.96123	...G...X-X...M...R...A	159
MAC.US.x.96135	...G...M...R...A	159
MAC.US.x.97009	...G...M...R...A	159
MAC.US.x.97074	...G...M...R...A	159
MAC.US.x.MM142_IVMXX	...G...D...M...K...X...T...TAK-VETR-I...P-VWH	162
MAC.US.x.SNM142_patent	...#A...G...I-I...Q...D...L...E...E...K...S...T...TAK-VETR-I...P-VWH	160
MAC.US.x.r80025	...G...D...L...E...E...K...S...T...TAK-VETR-I...P-VWH	159
MAC.US.x.r90131	...G...N...M...R...A	159
MNE.US.82.MNE_8	...F-A-G-I-Q...V-R...D-L-I-E...E...H...T-K...K...S...AP-AI...PT..K-EAIKV--N-P-NH	159
MNE.US.x.MNE027	...F-A-G-I-Q...V-R...D-L-I-E...E...H...T-K...K...S...AP-TK...T..TT-EIEV--N-T-VNR	160
SMM.SL.92.SL92B	-A-P-LH--D-F-LG.TW-A-Q--I-I-Q--V--D-L-I-A-D--N--T-A-K-N...RAA-S-P-T-SP..LTAA-PSGEEI--D-M-TKNN-S	164
SMM.US.x.F236_H4	-L-V-LE-C-V-Q--I-K-S--I--D-Q-I--A--D--N--T-A-T...GNAG-TTA..IT..TATPSVAENVL-SNP-KNNS-A	161
SMM.US.x.H9	-LS-X-ASG--V-Q--I--V--R--D-L-I-A-D--N--T-A-T...GXPAP-TTQ-T-TOAS-TPT-PITAKV--DSDP-RSN	161
SMM.US.x.PB114_15	-L-ASG--V-Q--I--V--R--D-L-I-A-D--N--T-A-T...GTPAP-TTQ-T-TOAS-TPT-PITAKV--DSDP-KIN	166
SMM.US.x.PBJA	-L-ASG--V-Q--I--V--R--D-L-I-A-D--N--T-A-T...GTPAP-TTQ-T-TOAS-TPT-PITAKV--DSDP-KIN	166
SMM.US.x.PBJ_143	-L-X-ASG--V-Q--I--V--R--D-L-I-A-D--N--T-A-T...GXPAP-TTQ-T-TOAS-TPT-PITAKV--DSDP-RSN	166
SMM.US.x.PBJ_6P6	-L-ASG--V-Q--I--V--R--D-L-I-A-D--N--T-A-T...GTPAP-TTQ-T-TOAS-TPT-PITAKV--DSDP-KIN	166
SMM.US.x.PGMS3	-L-A-G--V-Q--I-K-S--R--D-L-I-R...N--T-A-T...NEIP-TT-T-TKSPK-EA..ITAKVI-SDP-SNN	164
SMM.US.x.SME543	-L-V-LE-C-V-Q--I-K-S--R--D-L-V-I-A--D--N--T-A-T...GRAE-TTAK-T..STTTTTVPKVI-GD--KNNS-A	163
STM.US.89.STM_37_16	-A-P--AC.LT--Q--L-I--A--D--N--T-A--N--K--GKTV-VTP...AAA-TKPEL--VSNN	159

	V1	V2	V3	
MAC.US.x.239	LEQEOMISCKFNMTGLKRDKKKEYNETWYSADLVCE	.Q.GNNTGNESSRCYMHNCN	TSVIEQESCDKHVDAIRFRYKCAPPGYALLRCDNYSGFMPKCSKVVVSSSTRMMETQTSTWFGFNGTRAENRTYIYWHGRDNRITISLNKYNYLTKCKRRPGNKTVLPVTIM	325
A.CI.88.UC2	-GD-G-VN	-EQ-I-G-TD-D-V-D	STNKGT-TI-R	-K-SMK-T-A-N-P-AA-H-H-K-V-I-L
A.DE.x.BEN	-GY-E-MQ-E	-K-EQ-RR-KD-LE-V-D	NTT-AG-T-R	-I-K-M-F-E-T-AA-R-K-V-I-L
A.DE.x.PE12_KR_KRCG	-GS-EVVD-R	QL-PQQ-S	K-V-D.TTN-TSRKC	-T-M-LC-E-P-AAT-TH-H-K-S-V-I-LR
A.GH.x.GH1	-GK-EIVN-Q-Y	-EQ-Q	K-V-SMNTKD-K-N	-T-K-F-E-A-T-SIH-K-V-I-L
A.GM.87.D194	-GK-EVVE-E	-EQ-RK-DA-R-V-D	KTNIGT-T-R	-K-MK-F-E-AA-H-K-V-I-L
A.GM.x.ISY_SBL_6669_85	-RE-D-VE-Q	-EL-Q-S	K-V-SDNSDTRKRC	-T-M-FV-E-N-A-T-P-L-K-IL-E-V-I-L
A.GM.x.MCN13	-GE-EVNV-Q	-E-Q-S	K-V-GN-TDTC	-T-M-F-A-N-AAT-F-IH-K-V-I-L
A.GM.x.MCR35	-GK-EVNV-Q	-E-Q-A	K-V-GN-TTSTC	-T-M-F-A-N-AAT-F-IH-K-V-I-L
A.GW.86.FG_clone_NIHZ	-KE-E-D-Q-S	-E-R-Q-T-A	K-V-D-NNTSSOSKC	-T-M-F-A-N-AAT-K-NF-H-K-V-I-F
A.GW.87.CAM2CG	-E-KIVK-H	-E-Q-P	S-V-DNSTDQTTNETTC	-T-M-F-I-K-A-N-A-T-K-H-S-Y-V-I-L
A.GW.x.ALI	-GD-E-VN-R	-Q-P-Q	K-V-PNPTTNOTRC	-T-M-F-D-I-A-N-AAT-H-H-K-V-I-L
A.GW.x.MDS	-E-E-VK-Q	-OI-Q	S-V-G-EMNNTNQSRRC	-S-K-TM-Q-N-A-T-K-AS-SLH-K-V-I-L
A.IN.07.NNVA	-RE-EIVI-Q	-I-P-R	S-V-PDSTTSTR_NNSSSGQHMR-WYMRD	-DM-F-E-N-T-AA-S-H-SIY-K-V-I-H
A.IN.95.CR1K_147	-GE-EVNV-E	-V-P-T	R-VD-PDSTTNSRKC	-T-A-N-A-T-SIH-K-T-I-L
A.SN.85.ROD	-GE-ET-N-Q	-E-Q	K-V-TMNSTNQTC	-T-A-N-A-T-SLH-K-T-I-KQIML
B.CI.88.UC1	IGL-NTVD-Y	-R-E-Q-KD-EK-E-N	GNST-STI-RT	-SL-SL-I-HN-T-M-SK-H-I-I
B.CI.x.20_56	IGL-DVN-T	-RQ-ER-Q-D-RR-E	G-T-RT	-SL-SL-I-HN-T-M-SK-H-I-I
B.CI.x.EHO	IGL-E-D-Q-K	-ES-Q-KD-KQ	KGTRSN-E-K-IKT	-I-SL-F-F-K-N-LY-K-S-H-K-M-V-IRTV
B.GH.86.D205_ALT	-GE-EIMO-N-S	-R-EL-Q-KD-E-E-N	NTR-KY-T-IRT	-TI-SL-FF-N-A-S-EK-T-SIH-K-V-IRTV
B.JP.01.KR020_IMCJ_KR020_1	I-F-N-VT-R	-V-EP-T-D-AE-VQ-N	G-G-HIKT	-I-QL-F-FM-S-N-A-KS-V-K-IH-V
G.CI.92.Abt96	IQ5-S-VG	-QR-Q-QX	S-KSE-R	-R-F-A-N-T-IX-I
AB.CM.03.03CM_510_03	-N-E-VT-N-S	-EQ-R	Q-V-D.KNNSSETLNKT	-R-K-TM-T-Q-N-AA-I-H-K-V-I-L
H2_01_AB.CI.90.7312A	-GE-E-VQ-Q	-E-T-Q-S	SNNASD-R-D	-T-J-F-M-SK-IH-K-V-I-L
H2_01_AB.JP.04.04JP_NMC307	-GE-E-VQ-Q-D-I	-E-R	K-V-G-EANG_NI	-T-T-F-E-T-N-H-K-V-I-L
H2_01_AB.JP.07.07JP_NMC716	-GE-GIVQ-Q-D	-E-R	K-V-TNNGTNRIT-D-K	-T-T-F-D-KSN-H-K-V-I-L
H2_01_AB.JP.08.08JP_NMC842	-GE-GIVQ-Q-D	-Q-R	K-V-TNNGTN-T-N	-T-T-F-M-D-N-V-S-H-K-V-I-L
U.FR.96.12034	PLV-R	-R-Q-K-H-I	TSES-K	-K-N-T-RG-S-Y
MAC.US.x.17EC1		N		K
MAC.US.x.17EFR		N		K
MAC.US.x.1937				X
MAC.US.x.2065		X		
MAC.US.x.251_1A11		G		S
MAC.US.x.251_32H_PJ5		G		S-D
MAC.US.x.251_BK28		T		S-D
MAC.US.x.80835		T		I
MAC.US.x.81035		T		T
MAC.US.x.85013				
MAC.US.x.87082		X		D
MAC.US.x.92050		X		X
MAC.US.x.92077				
MAC.US.x.93057		X-X		X
MAC.US.x.93062				
MAC.US.x.95058				
MAC.US.x.95086		X		NX
MAC.US.x.95112				
MAC.US.x.96016		X		
MAC.US.x.96020				
MAC.US.x.96072		X		
MAC.US.x.96081				
MAC.US.x.96093				
MAC.US.x.96114				
MAC.US.x.96123				
MAC.US.x.96135				
MAC.US.x.97009		X		X
MAC.US.x.97074				
MAC.US.x.HM142_IVMXX		X		X
MAC.US.x.HM142_patent		P		S
MAC.US.x.r80025		P		C-D-C-N-R
MAC.US.x.r90131		X		X
MNE.US.82.MNE_8		P		S-E
MNE.US.x.MNE027		P		R-S-ED-K-N-SK-N
SMM.SL.92.SL92B	I-P-G-Q	-Q-RQ	R	G-ES-Y-A-N
SMM.US.x.F236_H4	P-G	N	R-I	SANES-K-H-S-L-A-N
SMM.US.x.H9	P-V	R	R	NS-GNETD-K-X-A-N-T
SMM.US.x.PB114_15	P-V	R	R	NS-NETD-K-S-A-N-T
SMM.US.x.PBJA	P-V	R	R	NS-NETD-K-S-A-N-T
SMM.US.x.PBJ_143	P-V	R	R	NS-ENETD-K-X-A-N-T
SMM.US.x.PBJ_6P6	P-V	R	R	NS-NETD-K-S-A-N-T
SMM.US.x.PGMS3	P-V-R	K	R-I	STTSR-K-S-A-N
SMM.US.x.SME543	P-G	I	R-I	PAN-S-K-Q-S-A
STM.US.89.STM_37_16	-E-SLVG	-R-S-I	-V-E-R	-L-T-A-N-V-A-N-S

gp120 end_gp41 start

MAC.US.x.239	ITMSAEVAELYLRLGDKYKLEITPIGLAPTDVKRYTGGTSRNKRGVFLGFLGFLATAGSAMGAASLTLTAQSRLL.AGIVQQQQQLLDVVKKQQLLRLTVWGTKNLQTRVTAIEKYLKDKQAQLNAGWCAFRQVCHTTVPW...PNASLTPKWNNTWQEWKRV	644
A.CI.88.UC2	-F-----I-----F-SE-SSA.PA-----L-----S-----I-----A-----S-----T-----V-D--R--M--KQ-	634
A.DE.x.BEN	-F-----I-----F-OR-SSA.PV-----R-----S-----M-----A-----H-----S-----Y--L--E-IV-D-M--KQ-	629
A.DE.x.PET2_KR_KRCG	-F-----I-----F-SE-SSA.PO-----V-----S-#-#-M-----A-----R--R--S-----Y--L--E-IV-D-M--KQ-	622
A.GH.x.GH1	-F-----I-V-----F-RE-SSA.PV-----M-----A-----S-----V-D--S-D-M--KQ-	622
A.GW.87.D194	-F-----I-V-----F-KE-SSA.PV-----M-----A-----S-----V-D--D--M--KQ-	620
A.GM.x.ISY_SBL_6669_85	-F-----V-----F-AE-SSA.PG-H-L-----T-A-G-----F-R-----M-----A-----R-----V-DT--E-M--HI-	623
A.GM.x.MCN13	-F-----V-----F-AE-SSA.PM-----M-----A-----S-----V-DT--E-M--GI-	624
A.GM.x.MCR35	-F-----V-----F-AE-SSA.PM-----M-----A-----S-----V-DT--E-M--GI-	624
A.GW.86.FG_c1one_NIHZ	-F-----A-----F-SE-SSA.HG-HT-----M-----A-----S-----V-DT--D-M--O-	621
A.GW.87.CAM2CG	-F-----D-----F-SO-SSA.PG-P-A-----T-V-T-----M-----I-----A-----R-----V-N--K-D-D-M--O-	631
A.GW.x.ALI	-F-----D-----F-SER-SSA.PR-----M-----A-----S-----V-N--K-D-D-M--O-	627
A.GW.x.MDS	-F-----D-----F-PE-SSA.HG-H-----T-A-----S-----V-S--E-D-M--O-	641
A.IN.07.NNVA	V-F-G-D-----V-----F-PE-SFA.SR-OT-----D-----S-P-----M-----A-----A-----E-DT--V-N-K-M-CK--O-	629
A.IN.95.CR1K_147	-F-----V-----V-----F-SE-SS-#-G-HT-L-E-----S-P-----M-----A-----A-----V-D--D-K-M--*--O-	632
A.SN.85.ROD	-F-----V-----V-----F-KE-SSA.HG-HT-----V-S-----P-----M-----A-----Q-----R-----V-D--A-D-D-M--KQ-	630
B.CI.88.UC1	-----G-----F-ET-SSA.P-----M-----L-M-----T-----S-----A-----L-----E-T--D-E-M--O--KR-	633
B.CI.x.20_56	-----G-----F-NT-SSA.P-R-M-----L-----S-----A-----L-----E-N-TM--O--E-	623
B.CI.x.EHO	-V-----S-----K-----F-SI-SSV.P-----M-----S-----A-----L-----V-E--K-D--M--O--O-	628
B.GH.86.D205_ALT	-SV-----S-----R-----F-R-SSV.KP-----M-----M-----T-----S-----PV-----M-----A-----N-F--M--O--KQ-	632
B.JP.01.KR020_IMCJ_KR020_1	-S-----S-----S-----F-S-SSA.P-----P-----S-----L-----S-----A-----KNF--N-D-M--O--NQ-	621
G.CI.92.Abt96	-IF-----F-----F-N-SSV-PK-----M-----T-----S-----R-----S-----M-----DALGA-KT-E-O--M--KQI	649
AB.CM.03.03CM_510_03	-F-----I-----F-TEG-SSA.PG-----S-----A-----X-----S-----X-----V-E--N-D-M--O--E-	622
H2_01_AB.CI.90.7312A	-F-----I-V-----F-SE-SSA.PG-H-----T-A-----S-----V-D--D-D-M--O--KQI	634
H2_01_AB.JP.04.04JP_NMC307	-F-----I-V-----F-AE-SS.PG-H-----T-A-----S-----M-----A-----S-----X-----V-D--D-D-M--O--KQI	628
H2_01_AB.JP.07.07JP_NMC716	-F-----I-V-----F-EQ-SS.PG-H-A-----T-----S-----M-----A-----S-----V-D--A-D-D-M--O--KQI	626
H2_01_AB.JP.08.08JP_NMC842	-F-----I-V-----F-EQ-SSD.XX-H-----X-----T-AX-X-----S-----M-----A-----S-----X-----X-XX--D-M--O--TOI	625
U.FR.96.12034	-S-----R-----F-P-SSV-P-----S-----I-DT--N-D-M--E-	639
MAC.US.x.17EC1	-----T-----I-----	644
MAC.US.x.17EFR	-----T-----X-----XNX--R-D-----	644
MAC.US.x.1937	-----G-----X-----	644
MAC.US.x.2065	-----N-R-----L-----D-----	643
MAC.US.x.251_1A11	-----D--D-----	643
MAC.US.x.251_32H_PJ5	-----D--D-----	646
MAC.US.x.251_BK28	-----D--D-----	644
MAC.US.x.80035	-----X-----I-----	644
MAC.US.x.81035	-----N-----X-----	644
MAC.US.x.85013	-----X-X-----X-----	644
MAC.US.x.87082	-----X-----X-----	644
MAC.US.x.92050	-----X-----I-----	644
MAC.US.x.92077	-----H-----X-----	644
MAC.US.x.93057	-----G-----X-----	644
MAC.US.x.93062	-----X-----X-D-----	644
MAC.US.x.95058	-----N-----X-----	644
MAC.US.x.95086	-----X-----X-X-----	644
MAC.US.x.95112	-----X-X-----D-----	644
MAC.US.x.96016	-----H-A-----D-----	644
MAC.US.x.96020	-----H-A-----D-----	644
MAC.US.x.96072	-----H-A-----D-----	644
MAC.US.x.96081	-----X-----E-----	644
MAC.US.x.96093	-----G-----X-----	644
MAC.US.x.96114	-----G-----X-----	644
MAC.US.x.96123	-----G-----N-----S-----	644
MAC.US.x.96135	-----X-----S-----	642
MAC.US.x.97009	-----G-----V-----S-----H-----D-----	647
MAC.US.x.97074	-----N-----V-----S-----H-----D-----	645
MAC.US.x.MM142_IVMXX	-----N-----V-----S-----H-----X-X-----	644
MAC.US.x.SNM142_patent	-----N-----V-----S-----H-----X-X-----	644
MAC.US.x.r80025	-----N-----V-----S-----H-----X-X-----	644
MAC.US.x.r90131	-----N-----V-----S-----H-----X-X-----	644
MNE.US.82.MNE_8	-----N-----N-----N-----	646
MNE.US.x.MNE027	-----N-----N-----N-----	647
SMM.SL.92.SL92B	-----M-H-R-S-K-----S-----L-----D-V-D-M--K-	651
SMM.US.x.F236_H4	-----I-----S-R-T-A-----V-S-----S-----L-----D-V-D-M--K-	650
SMM.US.x.H9	-----F-----S-----T-A-----S-----S-----X-----X-----DT--X--MX--KQ-	651
SMM.US.x.PB114_15	-----AF-----S-----T-A-----V-S-----S-----X-----X-----DT--N--M--KQ-	654
SMM.US.x.PBJA	-----AF-----S-----T-A-----A-----S-----S-----R-----DT--N--M--KQ-	654
SMM.US.x.PBJ_143	-----F-----S-----T-A-----V-S-----S-----X-----X-----DT--X--MX--KQ-	656
SMM.US.x.PBJ_6P6	-----F-----S-----T-A-----V-S-----S-----A-----A-----DT--N--M--KQ-	654
SMM.US.x.PGMS3	-----S-----R-----T-A-----V-RS-----S-----R-----V-N--M--Q-	653
SMM.US.x.SME543	-----R-----R-----T-A-----S-----H-----S-----D--V-N-D-M--G-	652
STM.US.89.STM_37_16	-A-----N-----TS-T-----T-----S-----D--V-D--M-----	647

MAC.US.x.239	DFLEENITALLLEEAIQIQEKNMYELQKLNWSDFGWNFDLASWIKYIQYGVYIVVGVILLRIVYIVQMLAKLRQGYRPFVSSPPSYFOOTHIQDPAALPTREGKERDGGEGGNSWSPWQIEYIHFILRQLIRLLTLWLFNSCRTLLSRVYQILOPI...LQRLSATLQR	811
A.CI.88.UC2	RY-A-SOS-----T-----I-A-A-V-L-SRF-K-----G-L-I-HT-RGO-AN-ET-G-A-DAS-YDF--P-N-QL-HL-T---G-Y-I-D---ANSPTRRLL...S-N....TA	797
A.DE.x.BEN	RY-A-SOS-----IL-----T-V-----H-----I-A-A-V-L-SRF-K-----G-L-I-HK-RGO-AN-T-E-V-GDS-YDL--P-N-VQ--HL-T---IG-YNI-D---KNSPTRRLL...S-S....TA	792
A.DE.x.PE12_KR_KRCG	RD-A-SRS-Q-----T-----VII-I-A-V-L-SR-K-----G-I-I-HK-WEO-D-ETDE-A-NST-D---P-A-----G-Y-V-KD---SFPT-L...F-S-QRA-TT	789
A.GH.x.GH1	RY-A-SOS-Q-----T-----I-G-A-----L-SR-K-----G-L-I-HT-RGO-AN-T-E-DRDD-YDL*-P-N-----HL-T---G-YK-D---TNSPTHRL...S-N....TA	784
A.GM.07.D194	HY-A-SOS-Q-----T-----I-G-A-----L-SR-K-----G-L-I-HT-RGO-AN-ET-E-A-DOS-FGL--PLN-Q--HL-T---G-YNS-G---KNSPTRRLL...S-S....TA	783
A.GM.x.ISY_SBL_6669_85	R-A-SES-Q-----T-----M-----I-IVV-----V-----SR-K-----G-I-I-HK-WEO-D-ET-E-V-NV-SR--P-A-----L-----R-YNS-D	779
A.GM.x.MCN13	RD-A-S00-Q-----T-----II-IVV-----SR-K-----G-L-I-HK-WEO-A-ET-E-V-NV-D-L*-P-R--H---AG-YNI-N---ISLT-R-V...F-S-QRA-TA	791
A.GM.x.MCR35	RD-A-S00-Q-----S-----II-IVV-----SR-K-----G-L-I-HK-WEO-A-ET-E-V-NV-D-L*-P-R--H---G-YNI-N---ISLT-R-A...F-S-QRA-TA	790
A.GW.86.FG_c1one_NIHZ	RY-A-SOS-Q-----T-L-FT-VR-----V-IVA-----V-----SR-K-----G-I-I-HK-QEO-A-ET-E-V-SN-DR--P-A-----L-----G-YNI-D---ISP...F-S-QRA-TA	788
A.GW.87.CAM2CG	RY-A-SOS-----L-----T-----N-----T-----S-----S-----G-----I-----I-----HK-QEO-AS-ET-E-V-GN-DR--P-A-----L-----G-YNI-N---IF-T...H	793
A.GW.x.ALI	RY-A-SEQ-R-----T-----L-TA-V-----I-IVA-----V-----SR-K-----G-I-I-HK-QEO-----G-ET-E-V-DN-DRL--P-A-L---HL-A---IG-Y-I-D---ISP...FRS-QRA-TT	794
A.GW.x.MDS	RY-A-SOK-----Q-----I-----I-V-----V-----SR-K-----G-I-I-HK-REO-D-ET-E-V-GN-DR--P-A-----L-----R-YNS-D---IFLT...F-N....	801
A.IN.07.NNVA	RH-A-SOS-Q-----I-----I-----I-A-V-----V-L-CSF-K-----G-I-I-HK-QEO-D-----T-E-V-DN-YR--P-A-----L-----R-YNS-KD---SFLT...-N....	789
A.IN.95.CR1K_147	RY-A-SOS-Q-----S-----C-----I-V-----V-----SR-K-----G-I-I-HK-LEO-A-----I-E-V-DN-YR*-P-A-----HL-----R-YNS-KD---SFLT...-N....	791
A.SN.85.ROD	RY-A-SKS-Q-----T-----V-----L-I-A-A-----V-----SR-K-----G-I-I-HK-RGO-AN-ET-E-SN-DRY--P-A-----HL-----R-Y-I-D---SFLT-L...Y-N....	790
B.CI.88.UC1	N-DA-----R-----FT-MA-RL-L-V-A-L-V-----M-----R-K-----T-IP-RKHGQ-AN-ET-DE-NE-AVR-----A-----RN-I-YNG-N-LKTS...-A....	792
B.CI.x.20_56	HY-DA-----H-----FT-MA-RL-L-V-----L-V-----VI-----R-K-----T-IP-RKHGQ-AN-ET-DE-RE-DYR-----A-----RN-I-YNG-N-L-TS...-A....	782
B.CI.x.EHO	R-DA-K-----Q-----I-S-----FT-MA-RL-L-----I-IVV-A-----I-----R-K-----T-IP-RK-RGO-AN-ET-EG-NNE-YR-----P-----RD-I-Y-G---KTF-T...-V....	787
B.GH.86.D205_ALT	H-DA-----I-----HL-L-----A-LVV-----V-----R-K-----T-IP-RK-RGO-AN-ET-EG-ND-DYR-----R-----RN-I-YNG---LKTF--HQ-STN...-F....	794
B.JP.01.KR020_IMCJ_KR020_1	R-D-----K-----V-----E-----K-----Q-----S-----FT-A-----I-L-VI-----LVV-----L-----R-K-----V-IP-HKGQGR-AS-ET-EG-DDEYR-----R-----RN-I-YNS-S-LKIF-P-LV...-F....	810
G.CI.92.Abt96	N-D-R-----R-----T-----V-VYL-L-V-A-----V-----V-I-GR-----X-----V-I-RK-QEX-K-EI-GES-NK-YR-----X-----GN---Y---A---IX-T--L...F-I-R--A	786
AB.CM.03.03CM_510_03	RD-A-SOI-Q-----S-----T-----R-----A-VV-X-A-----XSR-K-----SG-V-I-HT-QEO-S-ET-EG-ND-DYR-----V-----L-----RN-I-YNG---LKTS...-F....	781
H2_01_AB.CI.90_7312A	RD-A-SES-Q-----D-----S-----X-----V-----F-----I-IV-----I-GR-R-----X-----G-L-IR-HK-QEO-AN-ET-EG-ND-YR-----RN-I-YDG---LKTF-T...-A....	793
H2_01_AB.JP.04.04JP_NMC307	RY-A-SES-Q-----D-----S-----X-----V-----F-----I-IV-----I-LXSR-R-----X-----G-L-IR-HKHREO-AN-ET-EG-SDE-YR-----RN-I-YGG---KTF--LA...-A....	787
H2_01_AB.JP.07.07JP_NMC716	RY-A-SVS-Q-----D-----S-----T-----F-----I-IVA-----I-V-GR-K-----G-L-IR-HKHREO-AN-ET-EG-SDE-YR-----L-----RN-I-YEG---KTF-T...-A....	785
H2_01_AB.JP.08.08JP_NMC842	R-A-SKS-Q-----T-----L-----E-IVA-----V-----L-GR-R-----G-L-IR-HKHREO-DN-ET-EG-NDE-YR-----L-----RN-I-YEG---KTF--LA...-A....	784
U.FR.96.12034	NY-----Q-----A-----N-----I-----T-----V-VYL-L-V-A-----I-----V-L-GS-K-----H-VE-IP-R-QEQ-KGEI-EGA-DS-YR-----R-----H-----Y-S-D---IC-S-L...F-SIR...-	801
MAC.US.x.17EC1	-----N-----V-----G-----	811
MAC.US.x.17EFR	-----N-----X-----G-----*	810
MAC.US.x.1937	-----X-----G-----X-----X-----	811
MAC.US.x.2065	-----X-----G-----	811
MAC.US.x.251_1A11	-----X-----G-----Q-----R-----	809
MAC.US.x.251_32H_PJ5	-----I-V-----*-----T-----G-----A-----A-----R-----	810
MAC.US.x.251_BK28	-----X-----G-----	812
MAC.US.x.80035	-----X-----G-----	811
MAC.US.x.81035	-----X-----G-----	811
MAC.US.x.85013	-----X-----G-----	811
MAC.US.x.87082	-----X-----G-----X-----	811
MAC.US.x.92050	-----X-----G-----L-----F-----	811
MAC.US.x.92077	-----X-----G-----	811
MAC.US.x.93057	-----X-----G-----X-----	811
MAC.US.x.93062	-----X-----G-----	805
MAC.US.x.95058	-----X-----V-----H-----G-----	811
MAC.US.x.95086	-----X-----G-----	811
MAC.US.x.95112	-----X-----G-----	811
MAC.US.x.96016	-----X-----G-----	811
MAC.US.x.96020	-----T-----X-----G-----L-----X-X-----	811
MAC.US.x.96072	-----T-----X-----G-----	811
MAC.US.x.96081	-----T-----G-----X-----	811
MAC.US.x.96093	-----T-----G-----	811
MAC.US.x.96114	-----X-----G-----X-----	811
MAC.US.x.96123	-----X-----G-----	811
MAC.US.x.96135	-----X-----G-----	811
MAC.US.x.97009	-----X-----G-----	811
MAC.US.x.97074	-----X-----G-----	809
MAC.US.x.MM142_IVMXX	-----A-----T-----I-----I-----R-----*-----T-----K-----K-----G-----G-----A-----F-----R-----	813
MAC.US.x.SNM142_patent	-----A-----T-----I-----I-----R-----*-----T-----K-----K-----G-----G-----A-----F-----#-R-----	810
MAC.US.x.r80025	-----X-----G-----	811
MAC.US.x.r90131	-----X-----G-----	811
MNE.US.82.MNE_8	-----R-----*-----R-----Q-----K-----T-----G-----GS-----A-----F-----F-----T-----	812
MNE.US.x.MNE027	-----R-----R-----Q-----K-----T-----G-----DS-----A-----F-----F-----T-----	814
SMM.SL.92.SL92B	E-A-QM-RL-----T-VR-----FL-I-IV-----V-----SR-----H-I-R-QE-AK-T-GE-N-YRL-----I---YN-LTR-A-A-N-QL...C---EIS-P	818
SMM.US.x.F236_H4	N-A-Q-----X-----XX-XX-X-TX-----L-L-G-----V-----R-----A-V-IP-HKGQEP-K-E-G-DR-SR-----S-DW-L-S...V...S-T	817
SMM.US.x.H9	N-A-OS-----T-----I-----T-----L-L-G-----V-----R-----A-V-VP-TGOE-K-E-G-GR-R-----X-----S-DW-L-XC...V...S-R	818
SMM.US.x.PBJ14_15	N-A-OS-----T-----I-----T-----L-L-G-----V-----R-----A-V-IP-TGOE-K-E-G-GR-R-----S-DW-L-NC-T...V...S-R	821
SMM.US.x.PBJA	N-A-OS-----T-----I-----T-----L-L-G-----V-----R-----A-V-IP-TGOE-K-E-G-GR-R-----S-DW-L-NC-T...V...S-R	821
SMM.US.x.PBJ_143	N-A-XX-----X-----XX-XX-X-TX-----L-L-G-----V-----R-----AXVX-IP-XTGQE-K-E-G-XR-R-----X-----S-DW-L-XC...V...S-R	823
SMM.US.x.PBJ_6P6	N-A-OS-----T-----I-----T-----L-L-G-----V-----R-----A-V-IP-TGOE-K-E-G-GR-R-----S-DW-L-NC-T...V...S-R	821
SMM.US.x.PGMS3	D-A-QA-----I-----T-----T-----L-L-VG-----V-----I-----R-----V-V-IP-HKGQEO-K-E-G-DR-R-----N---Y-L-I...-V...F...-	817
SMM.US.x.SME543	A-Q-----V-----T-----R-----L-L-VG-----V-----R-----P-A-V-IP-HK-QEP-K-E-GE-DR-SR-----S-DW-L-I...-V...R-----	819
STM.US.89.STM_37_16	A-Q-----V-----T-VR-----L-I-LVM-VA-M-L-R-K-----CR-IP-HKGQEO-K-T-EG-DR-IN-----T-----V-----YN-F-AC---I-T-H-T...F-I-RI---	814

	Env end	
MAC.US.x.239	IR.EVLRTELTYLQYGSYFHEAVQAVWRSATETLAGAWGD.LWETLRRGGRWILAIIPRRIRQGLELTL..*	
A.CI.88.UC2	--DW--LKAA----CEWIO--F--IA--T--R-----RG--CKAVQ--I--G--V-----A--IA--	879
A.DE.x.BEN	--DW--LKAAQ----CEWIO--F--FA--T--R-----W--AA--I--G--V-----A--A--	866
A.DE.x.PE12_KR_KRCG	--DW--LTIA----CEWIO--VL--VLA--T--R-----S--R--G--GAMG--I--G--V-----A--A--	861
A.GH.x.GH1	--DW--LKAA----CEWIO--F--FAK--T--R-----S--G--CAAVQ--V--G--V-----A--IA--	858
A.GM.87.D194	--DW--LKAA----CEWIO--FR--FA--T--R-----I--RG--CAAQ--I--G--V-----A--IA--	853
A.GM.x.ISY_SBL_6669_85	L--DW--LKAA----CEWIO--F--LA--V--T--R-----TS--GRS--GA--G--I--G--V-----A--IA--	852
A.GM.x.MCN13	--DW--DAA----CEWIOG--F--FA--A--T--R-----T--R--GA--Q--I--G--V-----A--IA--	848
A.GM.x.MCR35	--DW--DAA----CEWIOG--F--FA--A--T--R-----T--R--GA--Q--I--G--V-----A--IA--	860
A.GW.86.FG_clone_NIHZ	--DW--LKAA----CEWIO--F--LA--T--R-----GR--RA--Q--I--G--V-----A--A--	859
A.GW.87.CAM2CG	L--DW--PKVAF----CEWIO--F--AA--A--R-----CR--V--GM--Q--I--G--V-----A--A--	857
A.GW.x.AL1	--DW--LKAA----CEWIO--FR--FA--I--R-----TNT--R--GAVQWV--R--V-----A--IA--	860
A.GW.x.MDS	L--DO--LRTAF----CEWIO--F--FA--V--T--R-----TS--CRG--G--G--I--G--V-----A--IA--	863
A.IN.07.NNVA	L--DW--LRTAF----CEWIO--F--AA--T--G-----S--CRS--G--G--I--G--V-----A--A--	870
A.IN.95.CR1K_147	L--DW--LKTAF----CEWIO--F--FA--T--GK-----S--GRN--CG--VE--I--G-----A--VA--	858
A.SN.85.ROD	L--DW--LRTAF----CEWIO--F--AA--A--T--R-----CRG--RV--E--I--G-----A--IA--	860
B.CI.88.UC1	...P--LS--A----I--W--Q--I--AT--A--R-----NTGRA--KA--TAEA--I-----A--	859
B.CI.x.20_56	...P--LS--A----I--W--Q--I--AT--A--G-----N--GRA--A--TAGA--I--V-----A--	858
B.CI.x.EHO	...P--LPPA--R--I--W--Q--I--AA--A--G-----S--ART--S--GV--AAGE--I-----A--A--	853
B.GH.86.D205_ALT	...P--LPVA----I--W--Q--LR--AA--ATG-----S--GET--A--AA--A--I-----A--	860
B.JP.01.KR020_IMCJ_KR020_1	...Q--LP--A----IGW--Q--I--AT--A--G-----TNTGRA--A--TA--A--V-----A--	860
G.CI.92.Abt96	...H--L--AA--FS--FRWLQ--CT--AT--A--Q-----TST--RA--K--G--V--G-----	846
AB.CM.03.03CM_510_03	...P--LPPA----IGW--Q--LR--AA--ATR-----S--GET--A--GAA--A-----A--	885
H2_01_AB.CI.90.7312A	...P--LLFA----IGW--Q--AAGATG-----STGRT--A--TA--G--I--V-----A--	847
H2_01_AB.JP.04.04JP_NMC307	...P--LP--AH--I--W--Q--I--AA--A--G-----TSTRT--A--TA--G--I-----A--	859
H2_01_AB.JP.07.07JP_NMC716	...P--LL--AH--I--W--Q--I--AA--A--G-----S--SRT--I--A--TA--G--I-----A--	853
H2_01_AB.JP.08.08JP_NMC842	...PF--VP--AH--I--W--Q--I--AA--A--G-----TN--SRT--I--A--AA--G--I--X--X--	851
U.FR.96.12034	...R--HL--IA-----Q--K--F--FGKA--R--SRTGRE-----G--V--LR-----F--A--	850
MAC.US.x.17EC1Y-----V-----	868
MAC.US.x.17EFRY-----V-----	880
MAC.US.x.2065	-X-----X-----X-----X--X-----X-----	879
MAC.US.x.251_1A11N-----VD-----*--A-----	881
MAC.US.x.251_32H_PJ5Q-----VG-----	880
MAC.US.x.251_BK28	V-----G-----R-----	877
MAC.US.x.80835Y-----V-----	879
MAC.US.x.81035	-X-----X-----X-----X-----X-----	881
MAC.US.x.85013X-----X-----	880
MAC.US.x.87082A-----	880
MAC.US.x.92050A-----	880
MAC.US.x.92077X-----X-----X--X-----	880
MAC.US.x.93057C--Q-----	880
MAC.US.x.93062C--Q-----	874
MAC.US.x.95058X-----X-----X-----	880
MAC.US.x.95086X-----X-----X-----	880
MAC.US.x.95112X-----X-----X-----	880
MAC.US.x.96016I-----Y-----F-----	880
MAC.US.x.96020I-----Y-----F-----	880
MAC.US.x.96072X-----X-----X-----	880
MAC.US.x.96081X-----X-----X-----	880
MAC.US.x.96093X-----X-----X-----	880
MAC.US.x.96114X-----X-----X-----	880
MAC.US.x.96123X-----X-----X-----	880
MAC.US.x.96135X-----X-----X-----	880
MAC.US.x.97009X--X-----AO-----E--A--Q-----	880
MAC.US.x.97074X--X-----AO-----E--A--Q-----	878
MAC.US.x.MM142_IVMXX	#--L-----Q-----#Q-----E#--A--Q-----	882
MAC.US.x.SNM142_patent	#--L-----Q-----#Q-----E#--A--Q-----	878
MAC.US.x.r80025Q-----VA-----G--V-----	880
MAC.US.x.r90131	V-----Q-----VA-----G--V-----E-----	880
MNE.US.82.MNE_8	...-LV-R-AG-IR--N--I--C-EA--Q-AIV--L--I--G-V-G-A-----M-N--	881
MNE.US.x.MNE027	V--I--I--IA-----R--Q-----W-KL-R--S--I--G-V-G-----	883
SMM.SL.92.SL92B	A--I--V--I-----R--LQ--AXXW-KFXR--S--R--G-V-G--XX-----	887
SMM.US.x.F236_H4	A--I--V--I-----R--LQ--AXXW-KFXR--S--R--G-V-G--XX-----	886
SMM.US.x.H9	A--I--V--I-----R--LQ--AXXW-KFXR--S--R--G-V-G--XX-----	887
SMM.US.x.PB114_15	A--I--V--I-----R--LQ--AXXW-KFXR--S--R--G-V-G--XX-----	890
SMM.US.x.PBJA	A--I--V--I-----R--LQ--AXXW-KFXR--S--R--G-V-G--XX-----	890
SMM.US.x.PBJ_143	A--I--V--I-----R--LQ--AXXW-KFXR--S--R--G-V-G--XX-----	892
SMM.US.x.PBJ_6P6	A--I--V--I-----R--LQ--AXXW-KFXR--S--R--G-V-G--XX-----	890
SMM.US.x.PGMS3	V--I--I--I-----CRW--Q--A--ARG--S--R--S--G-V-V-----A--	886
SMM.US.x.SME543	V--I--I--I-----CRW--Q--A--ARG--S--R--S--G-V-V-----A--	888
STM.US.89.STM_37_16	...-V-LGAA----CIWIQ--A--A--A--G-----S--GR-----G--V--R--G-----	883

Nef start	R17Y mutation	premature stop in Mac239	max HIV-1 similarity	
MAC.US.x.239	MGGAI5MRRSRPSGDLRQLLRARGE	TYGRLLGVEVDGYSQSPGLDKGLSSLSLC	.EGQKYNOGYMNTFWR.NPAEEREKLAYRKQNMDDIDE*DDD.LVGVSVR.PKVPLRTMSYKLAIDMSHFIEKGGLEGIIYSARRHRILDIYLEKEEIGIIPWDQDYTSGPG	163
A.CI.88.UC2	--ASG-KKL-KQ-OG-E	-----G.G-V-QCSASGGES--Q--SGREOK-P	-----Q-R-D-----T-M-GQ-ES-KQ--V-SD-N-Q--I-T-SR--P-T--V-----MF--RD-----L-----V-----N-H	164
A.DE.x.BEN	--ASG-KKL-KH-RG-E	-----D.G-KQRDASGGE--FOEESGREON-P	-----Q-Q-E-S-----T-Q-DL-Q--V-SD--I-P-T-R-R-E-T--V-----Q-MF--R-----N-H	164
A.DE.x.PET2_KR_KRCG	--ASG-KKC-SLOG-E	-----C.GQWDGASGE-L-FOE-SGR-QNLP	-----R-Q-DF-----T-AG-GTL-KQ--V-ADN-N-I-P-T-R-A-T--V-I-LN--D-M--E-----M-----N-H	164
A.GH.x.GH1	--ASG-KKH-KH-OR-E	-----H.G-G-VUQCNASGGE--QE-SG-QK-P	-----Q-R-DF-----T-I-GQ-KL-KQ--V-SS--P-T-R-A-T--V-----R-D-MF--RD-----L-----N-H	163
A.GM.87.D194	--ASG-KK-EH-OG-E	-----G.G-VKQRNASGGE--QE-SGREOK-P	-----Q-Q-EF-----T-AIQ-NS-KQ--V-SD--P-M-R-E-T-----F--RE-----N-H	164
A.GM.x.ISY_SBL_6669_85	--ASG-KK--RG-QE	-----G.AC-G-WD-S-G--FHE-SGREOKLP	-----R-Q-DF-----T-T-K-ES-Q--V-SD--DT-SR--A-T-RM-V--DL-D--M--E-----LF-----N-H-L	164
A.GM.x.MCN13	--ASG-KE--OG-E	-----C.GHCSGSGGE-L-QEESGREOR-P-Y	-----R-Q-DF-----T-K-G-EL-KQ-R-V-SD--K-T-R-ELTHR--V-----M--E-----K-----A-N-H	164
A.GM.x.MCR35	--ASG-KE--OG-E	-----C.GHCSGSGGE-L-QEESGREOR-P-Y	-----R-Q-DF-----T-K-G-EL-KQ-R-V-SD--K-T-R-ELTHR--V-----M--E-----K-----A-N-H	163
A.GM.86.FG_clone_NIHZ	--ASG-KK-K-LOG-OE	-----Q--C-GRCN-SGG-L-HE-SGREON-P	-----R-Q-DFV-----T-A--EL-KQ--V-LD--Q-FP-T-R-P-TF--V-----LF--Q--L-D-----A-N	147
A.GW.87.CAM2CG	--ASG-KK--LOG-OE	-----AG-C-ECYMAL-GESLR-QE-S-REON	-----R-Q-DF-----A-A-GK-N-Q--V-LD--S--PAT-R--T--V--L-R--MF-E-----A-N	164
A.GW.x.ALI	--ASG-KK-G-LOG-E	-----QTP--C-GCSCSGSGE--Q-SGR-QKLP	-----R-Q-DF-----T-T--EL-KQ--V-LD--S--PAT-R--T--V--L-R--MF-E-----A-N	163
A.GW.x.MDS	--ASG-KK--RG-OE	-----AG.AC-GHWL-LGGE-R-OE-S--OR-P	-----Q-Q-DF-----T-T-G-D5-Q--V-SD--Y-T-RT--A-T-V-----L-F-E-----N-H	164
A.IN.07.NNVA	--ASG-KKH--PRG-E	-----AG.AC-GRDALGGE-RFOEESGR-Q-P	-----RQNH-E-ADR--#I-ST-PNDW-GE-K#-V-SD--P-T-SRTQ-E-T-F--L-D--M--E--RE-I-L-----N	163
A.IN.95.CR1K_147	--ASG-KKH-K-RG-E	-----QA--CVGQWK-LGGE--QE-S--OR-P	-----Q-Q-DF-----T-T-KG-DL-Q--V-SD--L-R--T-R--T-RM--L--MF-E-----N-H	136
A.SN.85.ROD	--ASG-KKH--PRG-QE	-----AG.AC-GYWN-SGGE-RFQE-S-REOK-P	-----R-Q-Q-DF-----K-D-A-NL-Q--V-SD--Q-R--T--P-THR--L-TR--MF-E--K-N-----A-N-H	164
B.CI.88.UC1	--S-G-KKQ-KQORG-E	-----TOE--P-K-SEGRQRQS-R--S-D-N-P--RN	-----RN-----A-RA-GGG--Q-DT-S-DNE--Y-PNR--S-T-M-----E--T-N--VSG-N-Y	146
B.CI.x.20_56	--S-G-KKQ-KQORG-E	-----T-E-P-K-SE-RRGOS-R--ES-D-N-P-Y--RM	-----RM-----POGQKGGG--Q-DK-NS-D-E--AY-PS--A-T-V--L-----LF-E--TH-N--VSG-N-H	146
B.CI.x.EHD	--S-G-KKQ-KQORG-E	-----RG-PR-ESS--ROERSL-Y--S--D-N-P--D	-----D-----TLGA-GGG--Q-DS-ED-E--R-PGV--P-TF--V--L-----E--F-E--K--TH-N--VSG-N-H	146
B.GH.86.D205_ALT	--S-G-KK-ERQOG-EK	-----VPER-P--S--RRGOS-R--ES-D-N-P--N	-----N-----ARGA-GGG--Q-DA-ES-D-E--ALCKT-I--P-T--V--Q--M--E--T-F-N--VSG-N-H	147
B.JP.01.KR020_IMCJ_KR020_1	--S-G-KKQ-KQRRG-E	-----QTQEG-P--SE-QRQOS-R--ES-D-N-P--RS	-----RS-----A-GAQGGG--Q-GIE-S-DNE--Y-PNR--P-T--W-Q--M--E--L-T-N--VSG-N-H	146
G.CI.92.Abt96	--S-G-KK-APQQLG-K--Q	-----P-K-WEGL-E--YREESG--P	-----P-Y-----T-A--X-N-V-SD-N--P-Y--V-----E--M-Q--N-A	164
ABM.CM.03.03CM_510_03	--S-G-KK-EQQRG-E	-----V-ER-P-K-S-GRRERS-R--EES-D-N-P--N	-----N-----ALGA-GGG--Q-A-ES-D-E--AY-PSR--P-T-T--L--D--E--T-F-N--VSG-N-H	145
H2_01_AB.CI.90_7312A	--S-G-KKQ-KQORG-E	-----Q-G-P--S-GRRGES--EES-R-N-P--T	-----T-----A-GA-GGG--Q-DIEED-DNE--Y-POV--P-T-----F-E--T-F-N-R-VGG-N-Y	146
H2_01_AB.JP.04_04JP_NMC307	--S-G-KKQ-K-QRG-E	-----PAG-P--SEGRRGES--EES-D-N-P--N	-----N-----A-GA-GGG--Q-DIEED-D-E--AY-PN--A-T-----F-E--T-F-N--VSG-N-H	146
H2_01_AB.JP.07_07JP_NMC716	--S-G-KKQ-K-PRG-E	-----PAG-P--SE-RRGES-R--ES-D-N-P--N	-----N-----T-GA-GGG--Q-DIEED-D-E--AY-PN--S-T-----F-E--T-F-N--VSG-X-N	146
H2_01_AB.JP.08_08JP_NMC842	--S-G-KKQ-K-LRE-E	-----TPAG-P--S-GRRGES--EXS-D-NXP--N	-----N-----A-GA-GGG--Q-DIEED-DNE--AY-PN--E-T--V--F-E--T-F-N--VSG-N-H	146
U.FR.96.12034	--S-KKH-KR-K-E-A	-----N--W-LG-FELYQ-SG-D-RF-EAQ-G-SE-EF	-----T-TGKD-Q-QQ--V-E-N--A--R-A-T--V--F--D-F-Q--N--N	166
MAC.US.x.17EC1				164
MAC.US.x.17EFR				164
MAC.US.x.2065				164
MAC.US.x.251_1A11				164
MAC.US.x.251_32H_PJ5				164
MAC.US.x.251_BK28				164
MAC.US.x.80035				164
MAC.US.x.81035				164
MAC.US.x.85013				164
MAC.US.x.87082				164
MAC.US.x.92050				164
MAC.US.x.92077				164
MAC.US.x.93057				164
MAC.US.x.93062				164
MAC.US.x.95058				164
MAC.US.x.95086				164
MAC.US.x.95112				164
MAC.US.x.96016				164
MAC.US.x.96020				164
MAC.US.x.96072				164
MAC.US.x.96081				164
MAC.US.x.96093				164
MAC.US.x.96114				164
MAC.US.x.96123				164
MAC.US.x.96135				164
MAC.US.x.97009				164
MAC.US.x.97074				164
MAC.US.x.MM142_IVMXX				163
MAC.US.x.SNM142_patent				162
MAC.US.x.r80025				164
MAC.US.x.r90131				164
MNE.US.82.MNE_8				164
MNE.US.x.MNE027				164
SMM.SL.92.SL92B				165
SMM.US.x.F236_H4				164
SMM.US.x.H9				162
SMM.US.x.PBJ14_15				162
SMM.US.x.PBJA				162
SMM.US.x.PBJ_143				162
SMM.US.x.PBJ_6P6				162
SMM.US.x.PGM53				164
SMM.US.x.SME543				164
STM.US.89.STM_37_16				164

	max HIV-1 similarity	normal Nef end	
MAC.US.x.239	IRYPKTFGLWKLVPVNSDEAEQ...DEE.HYLMHPAQTSQDDPWGEVLAWKFDPTLAYTYEAYVRYPEEFSGKSLGSEEEVRRRLTARGLL..NMADKK.....ETR*		262
A.CI.88.UC2	V--MC-----M-Q-----D-TNC-----RY-IH-T-V-R-NSM--E-K-FTL-----H---P-K-WKAK-K-I.....PYSE.....		258
A.DE.x.BEN	V--MY-----S-EL-Q-----ED-ANC-V---RH-EH-T-V-Q-SM--N-K-FTL-----H---P-K-WKAK-K-I.....PYSE.....		258
A.DE.x.PET2_KR_KRCG	V--F-----D-POGE-----D-C-L-T---GS--H-T-M-R---R-E-T-FN-----Y---P-K-WKAK-K-I.....PFS.....		254
A.GH.x.GH1	V--MC-----D-Q-----D-TN-T---RH-EH-T-L-R---D-K-FILH-----H---P-K-WKAK-K-I.....PYS.....		256
A.GM.87.D194	T--MY-----DI-Q-----EV-TNC-V---RY-EH-T-V-R--M-S-K-FILH-----H---P-K-WKAK-K-I.....PYS.....		258
A.GM.x.ISY_SBL_6669_85	V--MF-----T-PO-G-----T-TLC--S-V-RF---H-T-V---M-HE-TFILH-----H--ME-DDWKAK-K-I.....PFS.....		257
A.GM.x.MCN13	--MF-----D-TRQE-----DGT-C-L---RF--H-T-I---#---HE-K-FILH-----H---P-DWKAK-K-I.....PFS.....		257
A.GM.x.MCR35	--MF-----D-TRQE-----DGT-C-L---RF--H-T-I---#---HE-K-FILH-----H---P-DWKAK-K-I.....PFS.....		214
A.GW.86.FG_clone_NIHZ	V--MF-----DT-Q-G---DTET-T-T-C-L---RH-WH-T-V---SM--K-FTL-----HN--P-K-WKA-K-I.....PFS.....		181
A.GW.87.CAM2CG	V--MF-----D-PO-G-----T-T-C-L-V---RH--TH-T-V-R--K-HD-K-FILH-----H---P-D-WKAK-K-I.....PFS.....		261
A.GW.x.AL1	T--C-----DLPQDE---NT-T-C-V---I-KF-OH-T-V-R-SL--E-T-FKL-----HQ--P--WKA-K-I.....PFN.....		263
A.GW.x.MDS	--MFH-----PQ-G-----T-T-C---KF-RH-T-V-R---D-V-FRLH-----H---P--WKA-K-I.....PFN.....		256
A.IN.07.NNVA	V--MF-----D-PO-G-----T-T-C-V---KF--H-T-V-E---L-S---FI-----H---P--WKA-K-I.....PFS.....		256
A.IN.95.CR1K_147	V--MF-----D-PO-G-----T-T-C-V---KF--H-T-V-E---L-S---FI-----H---P--WKA-K-I.....PFS.....		172
A.SN.85.ROD	V--MF-----D-PO-G-----T-T-C-V---KF--H-T-V-E---L-S---FI-----H---P--WKA-K-I.....PFS.....		257
B.CI.88.UC1	--R-----DIPE-E-R---GA-TSC-V---I-S--IH-T--R--L-HD-V-FN-----YQ--P-K-WKA-K-I.....PTD.....		238
B.CI.x.20_56	V--F-----I-MIA-P---E-T-C-V---A--HE-T-V-Q-SL--D-V-FS-F---YQ--MP-K-WKAK-R-I.....PTE.....		239
B.CI.x.EHO	V--F-----I-MIA-P---E-T-C-V---A--HE-T-V-Q-SL--D-V-FS-F---YQ--MP-K-WKAK-R-I.....PTE.....		239
B.GH.86.D205_ALT	--Y-----E-PAATR---E-T-C---I-S--IH-T-I-Q-SL--D-V-FN-F---YQ--P--WKA-K-I.....PTD.....		241
B.JP.01.KR020_IMCJ_KR020_1	V--F-----IPE-D.K---EA-TSC---A-P--VH-T-I-Q-SL--D-I-FN-----YQ--P-K-WKV-K-I.....PTD.....		239
G.CI.92.Abt96	--X-----T-----T-C-V---P--ET-----D-R-FILH-----RW--P-AVWKEK-KQ-----PIE.....		257
AB.CM.03.03CM_510_03	--Y-----ETPT-R---G-T-C---S--YH-T-V---L-HD-V-FN-F---YQ--P-K-WKAK-K-I.....PTD.....		239
H2_01_AB.CI.90.7312A	--H-----E-AAVTR---E-T-C-V---AA--H-T-V-Q-SL--SD-FN-F---YQ--P-K-WKA-K-I.....PTE.....		240
H2_01_AB.JP.04.04JP_NMC307	X--Y-----XE-AAAT---E-T-C-V---A--H-T-V-Q-XX-S--FN-F---YQ--XP-K-WAK-K-I.....PTE.....		240
H2_01_AB.JP.07.07JP_NMC716	X--Y-----XTAAT---E-TXC-V---A--H-T-V-Q-SH-S--FK-X---YQ--XP-K-WAK-K-I.....PTE.....		240
H2_01_AB.JP.08.08JP_NMC842	V--F-----E-AAATP---E-T-C-V---S--H-T-V-Q-SL-X-S--FN---YQ--MP-K-WKA-K-I.....PTE.....		240
U.FR.96.12034	T-W-MC-----E-DAN-K---R-L-VGS---CEE-H--A-V---SS-S-Q-FIKC-----K-----PVKNC.....		261
MAC.US.x.17EC1			264
MAC.US.x.17EFR			264
MAC.US.x.1937			264
MAC.US.x.2065			264
MAC.US.x.251_1A11			264
MAC.US.x.251_32H_PJ5			264
MAC.US.x.251_BK28			264
MAC.US.x.80035			263
MAC.US.x.81035			264
MAC.US.x.85013			264
MAC.US.x.87082			264
MAC.US.x.92050			264
MAC.US.x.92077			264
MAC.US.x.93057			264
MAC.US.x.93062			264
MAC.US.x.95058			264
MAC.US.x.95086			264
MAC.US.x.95112			264
MAC.US.x.96016			264
MAC.US.x.96020			264
MAC.US.x.96072			264
MAC.US.x.96081			264
MAC.US.x.96093			264
MAC.US.x.96114			264
MAC.US.x.96123			264
MAC.US.x.96135			264
MAC.US.x.97009			264
MAC.US.x.97074			264
MAC.US.x.MM142_IVMXX			264
MAC.US.x.SM1142_patent			263
MAC.US.x.r80025			262
MAC.US.x.r90131			264
MNE.US.82.MNE_8			264
MNE.US.x.MNE027			264
SMM.SL.92.SL92B			267
SMM.US.x.F236_H4			263
SMM.US.x.H9			262
SMM.US.x.PBJ14_15			262
SMM.US.x.PBJA			262
SMM.US.x.PBJ_143			262
SMM.US.x.PBJ_6P6			262
SMM.US.x.PGMS3			264
SMM.US.x.SME543			264
STM.US.89.STM_37_16			264