

**V**

---

**HIV-2/SIVsmm proteins**

---

Introduction .....	532
Table of HIV-2/SIVsmm protein alignments .....	533
Gag .....	537
Pol .....	542
Vif .....	551
Vpr .....	553
Tat .....	554
Rev .....	555
Vpx .....	556
Env .....	557
Nef .....	565

## Construction of the HIV-2/SIV Amino Acid Alignments

### Classification of the sequences

Classification of HIV-2 is straightforward in the sense that the subtype classification of this virus is more or less standardized. SIVs derived from sooty mangabey (SIV-smm), macaques (SIV-mac) and HIV-2s intermingle in phylogenetic trees, however, due to multiple cross-species transmissions (1,2) which complicates the story. Classification of the other SIVs is even more complicated. As the tree in section HIV-1/HIV-2/SIV of this compendium shows, SIVcpz's are the most related to HIV-1, so included in HIV-1 alignments. The sequences from four subspecies of African Green Monkey, (SIVagm) are included in one alignment due to their distinctive clustering. Included with the SIVagms are representatives of some of the very diverse SIVs, SIVlhoest, SIVsun, and SIVsykes. The HIV-2/SIVsmm/SIVmac are also included as a separate alignment in this Compendium, as they form their own phylogenetic group.

### References

1. Hahn, B., G. Shaw, K. De Cock, P. Sharp, 2000. AIDS as a zoonosis: scientific and public health implications. *Science* **287**:607–614.
2. Chen, Z., P. Telfier, A. Gettie, P. Reed, L. Zhang, D. D. Ho, and P. A. Marx, 1996. Genetic characterization of new West African simian immunodeficiency virus SIVsm: geographic clustering of household-derived SIV strains with human immunodeficiency virus type 2 subtypes and genetically diverse viruses from a single feral sooty mangabey troop. *J Virol* **70**:3617–27.

Table 1: Table of HIV-2/SIVsmm protein Alignments

Name	Accession	Region	Author	Reference
H2AB.CI.7312A-1	L33088	POL	Gao, F	J Virol 68:7433-7447 (1994)
H2AB.CI.7312A-2	L33092	POL	Gao, F	J Virol 68:7433-7447 (1994)
H2AB.CI.7312A	L33079	GAG	Gao, F	J Virol 68:7433-7447 (1994)
H2AB.CI.7312A	L33094	ENV	Gao, F	J Virol 68:7433-7447 (1994)
H2A.-.CBL21	U05350	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.-.CBL22	U05351	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.-.CBL23	U05352	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.CI.UC2	U38293	ENV GAG NEF POL REV TAT VIF VPX	Barnett, SW	Virology 222(1):257-261 (1996)
H2A.DE.BEN	M30502	ENV GAG NEF POL REV TAT VIF VPR VPX	Kirchhoff, F	Virology 177(1):305-311 (1990)
H2A.DE.PE12	U22047	ENV GAG NEF POL REV TAT VIF VPR VPX	Talbott, R	PNAS USA 90(9):4226-4230 (1993)
H2A.GH.GHI	M30895	ENV GAG NEF POL REV TAT VIF VPR VPX	Hasegawa, A	ARHR 5:593-604 (1989)
H2A.GM.CBL24	U05353	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.GM.D194	J04542	ENV GAG NEF POL REV TAT VIF VPR VPX	Kuehnel, H	PNAS USA 86(7):2383-2387 (1989)
H2A.GM.ISY	J04498	ENV GAG NEF POL REV TAT VIF VPR VPX	Franchini, G	PNAS USA 86(7):2433-2437 (1989)
H2A.GW.ALI	AF082339	GAG NEF POL REV TAT VIF VPR VPX	Azevedo-Pereira, JM	Unpublished (1998)
H2A.GW.CAM1	U05359	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.GW.CAM2CG	D00835	ENV GAG NEF POL REV TAT VIF VPR VPX	Tristem, M	J Gen Virol 72:721-724 (1991)
H2A.GW.CAM2env	U05354	ENV	Tristem, M	J Gen Virol 72:721-724 (1991)
H2A.GW.CAM3	U05355	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.GW.CAM4	U05356	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.GW.CAM5	U05357	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.GW.CAM6	U05358	ENV	Breuer, J	J Gen Virol 76(Pt 2):333-345 (1995)
H2A.GW.FG	J03654	ENV GAG NEF POL REV TAT VIF VPR VPX	Zagury, JF	PNAS USA 85(16):5941-5945 (1988)
H2A.GW.HIV-2.ALI	L25445	ENV	Costa Taveira, N	ARHR 10:223-224 (1994)
H2A.GW.MDS	Z48731	ENV GAG NEF POL REV TAT VIF VPR VPX	Becker, M	Unpublished (1995)

Table 1: Cont.

Name	Accession	Region	Author	Reference
H2A.-.HIV2JAU1	L28935	VIF	Jacinto, A	Unpublished (1994)
H2A.SN.A1871	U81840	VPR	Stivahtis, GL	J Virol 71(6):4331-4338 (1997)
H2A.SN.A1871	U81847	VPX	Stivahtis, GL	J Virol 71(6):4331-4338 (1997)
H2A.SN.A.2057	U81842	VPR	Stivahtis, GL	J Virol 71(6):4331-4338 (1997)
H2A.SN.A.2057	U81849	VPX	Stivahtis, GL	J Virol 71(6):4331-4338 (1997)
H2A.SN.A.640	U81838	VPR	Stivahtis, GL	J Virol 71(6):4331-4338 (1997)
H2A.SN.A.640	U81845	VPX	Stivahtis, GL	J Virol 71(6):4331-4338 (1997)
H2A.SN.H2A7924A	U81836	VPR	Stivahtis, GL	J Virol 71(6):4331-4338 (1997)
H2A.SN.HIV2ST	M31113	ENV GAG NEF POL REV		
		TAT VIF VPR VPX	Kumar, P	J Virol 64(2):890-901 (1990)
H2A.SN.ROD	M15390	ENV GAG NEF POL REV		
		TAT VIF VPR VPX	Clavel, F	Nature 324(6098):691-695 (1986)
H2B.CI.EHO	U27200	ENV GAG NEF POL REV		
		TAT VIF VPR VPX	Rey-Cuille, MA	Virology 202(1):471-476 (1994)
H2B.CI.UC1	L07625	ENV GAG NEF POL REV		
		TAT VIF VPR VPX	Barnett, SW	J Virol 67:1006-1014 (1993)
H2B.-.GH-2	D10458	ENV	Kawamura, M	Virology 188:850-853 (1992)
H2B.GH.D205	X61240	ENV GAG NEF POL REV		
		TAT VIF VPR VPX	Kreutz, R	ARHR 8(9):1619-1629 (1992)
H2C.LR.22381G	L33077	GAG	Gao, F	J Virol 68:7433-7447 (1994)
H2C.LR.2238POLB7	M87138	POL	Gao, F	Nature 358:495-499 (1992)
H2C.LR.HIV2238E	M87118	ENV	Gao, F	Nature 358:495-499 (1992)
H2D.LR.FO784	L33083	GAG	Gao, F	J Virol 68:7433-7447 (1994)
H2D.LR.FO784	M87069	ENV	Gao, F	Nature 358:495-499 (1992)
H2D.LR.FO784	M87115	NEF	Gao, F	Nature 358:495-499 (1992)
H2D.LR.FO784PA	M87110	POL	Gao, F	Nature 358:495-499 (1992)
H2D.LR.FORTC2	M87111	POL	Gao, F	Nature 358:495-499 (1992)
H2E.SL.PA	L33087	GAG	Gao, F	J Virol 68:7433-7447 (1994)
H2G.CI.ABT96	AF208027	ENV GAG NEF POL REV		
		TAT VIF VPR VPX	Brennan, CA	ARHR 13(5):401-404 (1997)
MAC.-.MM132ZL	L22809	ENV	Kodama, T	J Virol 67(11):6522-6534 (1993)
MAC.-.MM316S	L22790	VIF VPR VPX	Kodama, T	J Virol 67(11):6522-6534 (1993)
MAC.-.MM316ZQ	L22814	ENV	Kodama, T	J Virol 67(11):6522-6534 (1993)
MAC.-.SEGA	L26938	ENV	Kodama, T	J Virol 67(11):6522-6534 (1993)

Table 1: Cont.

Name	Accession	Region	Author	Reference
MAC--SEGE	L26942	ENV	Kodama, T	J Virol 67(11):6522-6534 (1993)
MAC--SIU18034	U18034	ENV	Zhu, GW	J Neurovirology 1(1):78-91 (1995)
MAC--SIU18039	U18039	ENV	Zhu, GW	J Neurovirology 1(1):78-91 (1995)
MAC--SIU18046	U18046	ENV	Zhu, GW	J Neurovirology 1(1):78-91 (1995)
MAC.US.CTL7202	M61062	ENV	Burns, DPW	J Virol 65:1843-1854 (1991)
MAC.US.MAC251-1A11	M76764	GAG NEF POL REV TAT VIF VPR VPX	Planelles, V	ARHR 11:889-898 (1991)
MAC.US.MM142	M16403	ENV GAG NEF POL REV	Chakrabarti, L	Nature 328(6130):543-547 (1987)
MAC.US.MM239	M33262	TAT VIF VPR VPX ENV GAG NEF POL REV	Kestler, H	Science 248:1109-1112 (1990)
MAC.US.MM251	M19499	TAT VIF VPR VPX ENV GAG NEF POL REV	Franchini, G	Nature 328(6130):539-543 (1987)
MAC.US.MMPI1	M72323	TAT VIF VPR VPX NEF REV TAT	Colombini, S	PNAS USA 86(13):4813-4817 (1989)
MNE--SIU06277	U06277	ENV	Chackerian, B	J Virol 68(6):4080-4085 (1994)
MNE--SIU06352	U06352	ENV	Chackerian, B	J Virol 68(6):4080-4085 (1994)
MNE--SIU06377	U06377	ENV	Chackerian, B	J Virol 68(6):4080-4085 (1994)
MNE.US.MNE	M32741	ENV GAG NEF POL REV	Benveniste, RE	Unpublished(1990)
SMM--ENVVW25	X86727	TAT VIF VPR VPX NEF	Whatmore, AM	J Virol 69(8):5117-5123 (1995)
SMM.LR.SMMLIB1	M62651	POL	Marx, PA	J Virol 65:4480-4485 (1991)
SMM--NEFW61	X90853	NEF	Whatmore, AM	J Virol 69(8):5117-23 (1995)
SMM--P209C15	L20009	ENV	Hynes, NA	ARHR 9:803-806 (1993)
SMM.SL.SIU17646	U17646	GAG	Chen, Z	J Med Primatol 24(3):108-115 (1995)
SMM.SL.SIV smSL92f	U48815	GAG	Chen, Z	J Virol 70(6):3617-3627 (1996)
SMM--SMMPBJE	L03298	NEF	Dewhurst, S	Nature 345:636-640 (1990)
SMM--SMP209	L20008	ENV	Hynes, NA	ARHR 9:803-806 (1993)
SMM.US.MM32H	D01065	ENV GAG NEF POL REV	Rud, EW	J. Gen. Virol. 75(Pt 3):529-543 (1994)
SMM.US.SIU04982	U04982	TAT VIF VPR VPX ENV NEF REV TAT VPR VPX	Hirsch, VM Mullins, JI	J Virol 68:2649-2661 (1994) Unpublished (1992)
SMM.US.SIVSMMB670	M90048	ENV NEF	Cournaud, V	J Virol 66:414-419 (1992)
SMM.US.SIVSMMH9	M80194	ENV GAG NEF POL REV TAT VIF VPR VPX		

Table 1: Cont.

Name	Accession	Region	Author	Reference
SMM.US.SMFMF236	X14307	ENV GAG NEF POL REV TAT VIF VPR VPX	Hirsch, VM Li, Y	Unpublished Unpublished(1989)
SMM.US.SMMM7	M27256	POL		
SMM.US.SMMPBJ14-15	L03295	ENV NEF POL REV VIF VPR VPX	Dewhurst, S	Nature 345:636-640 (1990)
SMM.US.SMMPBJ14-441	M31325	GAG POL REV TAT VIF VPR VPX	Dewhurst, S	Nature 345:636-640 (1990)
SMM.US.SMM-PBJ-6P6	L09212	GAG NEF POL REV VIF VPR VPX	Novembre, FJ Khan, AS	J Virol 67:2466-2474 (1993) J Virol 65(12):7061-7065 (1991)
STM.US.STMAK2	X60667	GAG	Khan, AS	J Virol 65(12):7061-7065 (1991)
STM.US.STMAK3	X60668	ENV		
STM.US.STM	M83293	ENV GAG NEF POL REV TAT VIF VPR VPX	Novembre, FJ	Virology 186(2):783-787 (1992)

```

|<- Gag p17 matrix
MGARNSVLRGKkadeLEkiRLRPgGKkKkYrLkHiVWAANEldrFgLaESLLeSKeGCOKiItVldPlVPtGSEnKLSlfntVcViWClHAeEKVKDTEaK.?ivqRHLVvAeTgTaEKWpntSRPtaPpS
-----R--N-----KL-----
H2A.CI.UC2      128
H2A.DE.BEN     129
H2A.DE.PE12    129
-----V--GK--I-----G--Q-----D--S--A-----
H2A.GH.GH1     129
-----V--N--R--V-----K--E-----KLA--G--S-----
H2A.GM.D194    129
-----R--R-----Q-----R--K--E-----Y--T-----R--G--E-----I-----
H2A.GM.I5Y     129
-----R--V-----K-----R--K--E-----V--V-----G--Q--I--A--IE-----S-----
H2A.GM.AH1     129
-----V-----K-----K-----R--K-----I-----R--AL--A-----D-----
H2A.GM.CAM2CG  129
-----S--V-----V-----F--I-----KFAR-----
H2A.GM.FG      129
-----S--V-----V-----I--A-----OK-----K--T--S-----
H2A.SN.HIV2ST  129
-----R-----K-----M-----I-----G--R-----S-----S-----
H2A.SN.ROD     129
xANELDRFGLAENLLeSKeGCOKiItVlAPLMPtGSEnKLSlfntVcViWClHAeEKVKDTEaK.KLAQSHLVVDT...EKtPtADKPAATSS          91

MGARgSVLsGkKtDLeKvRLRPgKkKkYmLkHiVwAvNELdRfGLAeSLLeSkEgc?KiIkvLaPlVpTgSEnKLSlfntVcViFclHAeEKVKDTEaK.KIAQRHLAADT...EKWpAtSkPtaPpS
-----C--I-----K--V-----R--G-----R--I--G-----Y-----Y-----I-----Y-----R-----M--SK--T-----
H2B.CI.BHO     126
H2B.CI.UC1     126
H2B.GH.D205    126

xANELDRFALAEsLLeTKEGcOKiItVLEPlVpTgSEnKLSlyntVcViWClHAeEKVKHTEaRN.EVVERHLLAeTKNAEKMPAtSRPtaPpS          93
xANELDRFGLAeSLLenKEgcOKiIsvLaplVpTgSEnKLSlyntVcViWClHAeEKVKDTEgAK.QiVQRHLVvVETgTADKMPStSRPtaPpS          93
xARELDRFGLAeSLLenKEgcHRIItVLEPlMPtGSEnKLSlfnAVcVvYClHAaAVKVKDTEaKkHVQOHHLvAGGkTTEKlLPqSRPtaPpS          94
MGARSSVLSGKKADeLEKvRLRPgKkKYmLkHvVWAANELDRFGLAeSLLeSKeGCOKiIsvLaplVpTgSEnKLSlyntVcViWClHAeEKVKHTEaK.QVvQRHLVvVETgTTEKVPAtSRPIaPpS          129
MGARNSVLSGKKADeLEKIRLrP?GKKKYmLkHvVWAANELDRFGLAeSLLenKEGCOKiIsvLaplVpTgSEnKLSlyntVcViWClHAeEKVKHTEaK.QiVQRHLVvVETgTAEtMpkTSRpTAPpS          128
--V-----N-----G-----G-----M-----T-----F-----
MAC.US.MM142   129
MAC.US.MM251   129
MAC.US.WAC251-Ia11  129
MNE.US.MNE     129
MGARNSVLSGKKADeLEKIRLrPFGKkKYmLkHvVWAANELDRFGLAeSLLenKEGCOKiIsvLaplVpTgSEnKLSlyntVcViWClHAeEKVKHTEaK.QiVQRHLVvVETgTAEtMpkTSRpTAPpS          129
mgarnsvlsgkkadeleki.rlrpggkky?lkh?vwaaneLdrFGLAeSLLenKEGCOKiIsvLaplVpTgSEnKLSlyntVcViWClHAeEKVKHTEaK.QiVQRHLVvVETgTAD?mpatSRPtaPpS          125
x--K-----S--R--R--V-----F--M--N-----A--NE--Q-----
S--R--R-----S--R--R-----F--L--M-----L--M-----A--NE--Q-----
90
-----K--M--V-----K--M--I-----R--L-----K-----K-----S-----
129
-----V-----K--M--I-----K--M--J-----L-----J-----K-----R-----
129
-----E--V-----N-----K--M--V-----D-----L-----L-----K-----
129
MGARSSVLSGKKADeLEKvRLRPgKkKYmLkHvVWAANELDRFGLAeSLLeSKeGCOKiItVLEPlVpTgSEnKLSlfntVcViWClHAeEKVKHTEaK.QVvQRHLVvVETgTANKMPAtSRPtaPpS          129
MGARSSVLSGKKADeLEKvRLRPgKkKYmLkHvVWAANELDRFGLAeSLLeSKeGCOKiIItVLEPlVpTgSEnKLSlfntVcViWClHAeEKVKHTEaK.QVvQRHLVvVETgTADKMPAtSRPtaPpS          129
    
```





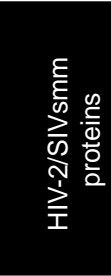
CONSENSUS_H2A	p24 capsid \ / x peptide / p9 NC	
H2A.CI.UC2	VPVGNiYRRWtIQlGLOKcVrMvNpTnIldIkOGPKepFoSvYvDRfYkSlRaEQtdPaVKNwMwTQTLlVQnaNpDCKlVlKGLGmNpTlEEmLlTAcOGvGpGqKaRlMaEAlKkEam?PaPiPFAAAQO..	381
H2A.DE.BEN	-----V-V-----S-----A-----L-T-P-----	385
H2A.DE.PE12	-----K-----I-----G-S-----	385
H2A.GH.GH1	I-----V-----LA-----	386
H2A.GM.D194	-----V-----I-----L-T-P-----	386
H2A.GM.ISY	-----V-----S-----L-T-----	385
H2A.GM.ALI	-----S-----I-----S-----R-----	384
H2A.GM.CAM2CG	-----S-----I-----T-----	385
H2A.GM.FG	-----N-----I-----T-----G-P-----	385
H2A.GM.MDS	-----V-----I-----V-A-----	383
H2A.SN.HIV2ST	-----S-----A-----I-----T-----	385
H2A.SN.ROD	-----S-----A-----I-----VIG-----	385
H2AB.CI.7312A	IPVGNiYRRWtIQlGLOKcVrMvNpTnIldVlKOGKpEPFoS	256
CONSENSUS_H2B	VPVGNiYRRWtIQlGLOKcVrMvNpTnIldIkOGPKepFoSvYvDRfYkSlRaEQtdPaVKNwMwTQTLlQnaNpDCKlVlKGLGmNpTlEEmLlTAcOGvGpGqKaRlMaEAlKkEAlTpaPiPFAAAQKa	381
H2B.CI.BHO	-----I-----	382
H2B.CI.UC1	-----S-----	382
H2B.GH.D205	-----V-----	382
H2C.LR.22381G	VPVGNiYRRWtIQlGLOKcVrMvNpTnIldVlKOGKpEPFoS	261
H2D.LR.FO784	IPVGNiYRRWtIQlGLOKcVrMvNpTnIldIkOGKpEPFoS	261
H2E.SL.PA	IPVGDiyRRWtIQlGLOKcVrMvNpTnIldVlKOGKpESFoS	261
H2G.CI.ABT96	IPVGNxYRRWtIxLGLQKcVrMxNpTnIldVlKOGKpEPFoSvYvDRfYkSxRaEQtdPaVKNwMwTQTLlQnaNpDCKlVlKGLGmNpTlEEmLlTAcOGxGpGqKaRlMaEAlKkEAlNpTlPFAAAQqKT	387
CONSENSUS_MAC	IPVGNiYRRWtIQlGLOKcVrMvNpTnIldVlKOGKpEPFoSvYvDRfYkSlRaEQtdPaVKNwMwTQTLlQnaNpDCKlVlKGLGvNpTlEEmLlTAcOGvGpGqKaRlMaEAlKkEAlAaVpPiPFAAAQqTG	385
MAC.US.MM142	-----P-----T-----A-----K-----	386
MAC.US.MM251	-----P-----T-----A-----K-----	386
MAC.US.MAC251-1A11	-----P-----T-----A-----K-----	386
MNE.US.MNE	IPVGNiYRRWtIQlGLOKcVrMvNpTnIldVlKOGKpEPFoSvYvDRfYkSlRaEQtdPaVKNwMwTQTLlQnaNpDCKlVlKGLGmNpTlEEmLlTAcOGvGpGqKaRlMaEAlKkEAlAaVpPiPFAAAQqKG	386
CONSENSUS_SMM	iPVGNiYrRwTlQlGLOKcVrMvNpTnIldVlKOGKpEPFoSvYvDRfYkSlRaEQtdPaVknWwTQTLlQnaNpDCKlVlKGLG?NpTlEEmLlTAcOGvGpGqKaRlMaEAlK?al??pPiPFAaVQqKG	377
SMM.SL.SIVsmSL92a	V-----V-----I-----x-----	276
SMM.SL.SIVsmSL92f	V-----I-----K-----G-----x-----	273
SMM.US.MM32H	-----x-----A-----V-----E-APV-I-----A--R-	386
SMM.US.SIVSMH9	-----x-----S-----M-----D-TQG-----D--x-----	387
SMM.US.SMMP236	-----x-----S-----M-----D-TQG-----D--x-----	387
SMM.US.SMMP236	-----x-----S-----M-----D-TQG-----D--x-----	387
SMM.US.SMMPBJ4-441	-----x-----S-----M-----D-TQG-----D--x-----	387
STM.US.STM	IPVGNiYRRWtIQlGLOKcVrMvNpTnIldVlKOGKpEPFoSvYvDRfYkSlRaEQtdPaVKNwMwTQTLlQnaNpDCKlVlKGLGmNpTlEEmLlTAcOGvGpGqKaRlMaEAlKkEAlVpPiPFAAAQqQG	388
STM.US.STMAK2	IPVGNiYRRWtIQlGLOKcVrMvNpTnIldVlKOGKpEPFoSvYvDRfYkSlRaEQtdPaVKNwMwTQTLlQnaNpDCKlVlKGLGmNpTlEEmLlTAcOGvGpGqKaRlMaEAlKkEAlVpPiPFAAAQqQG	387



CONSENSUS_H2A	EDLlhlEggetPhre?TEDLLHlnSIFGkDQ	p6 \	523
H2A.CI.UC2	-----R--A-----	-----	521
H2A.DE.BEN	-----R--E-----	-----	521
H2A.DE.PEI2	-----K--A-----	-----	521
H2A.GH.GH1	-----KA--A-----	-----	522
H2A.GM.D194	-----G--A-----	-----	521
H2A.GM.ISY	-----T-----N--	-----	520
H2A.GM.ALI	-----K--V-----	-----	521
H2A.GM.CAM2CG	-----A--C--T-----	-----	521
H2A.GM.FG	-----RF--A--C--T-----	-----	519
H2A.GM.MDS	-----K--CK--T-----	-----	521
H2A.SN.HIV2ST	-----F--Q--KQ--C--T-----	-----	521
H2A.SN.ROD	-----Y--PP-----	-----	522
CONSENSUS_H2B	EDLLHL.....NSLFGEDQ		521
H2B.CI.EHO	-----		519
H2B.CI.UC1	-----		521
H2B.GH.D205	-----		521
H2G.CI.ABT96	xxLLHL.....SSLFGDDQ\$		509
CONSENSUS_MAC	EDLLHL.....NSLFGGDQ		509
MAC.-.MM239	-----		510
MAC.US.MM142	-----		506
MAC.US.MM251	-----		506
MAC.US.MAC251-1A11	-----		510
MNE.US.MNE	EDLLHL.....NSLFGEDQ		506
CONSENSUS_SMM	EDLLHL.....NSLFGEDQ		511
SMM.US.MM32H	-----G--		510
SMM.US.SIVSMMH9	-----		507
SMM.US.SMM-PBJ-6P6	-----		507
SMM.US.SMMF236	-----		507
SMM.US.SMMPEJ14-441	-----		507
STM.US.STM	EDLLHL.....NSLFGEDQ		507
STM.US.STMAK2	EDLVHL.....NSLFG		510

HIV-2/SIVsmm protein alignment: POL

	-> Pol reading frame (-1 from Gag) p15			
	ktgllle?wg?rt?hgk?p??t?g?FFRfd?p?GkeapQIPIrGpSsaGAdTNSrpsstgtsgeiYAArekae?AE.?etiOrGDrGLtAprRagrdt?OrgrdRGlaAaPQFSLW		p15 peptide \ / Protease	98
CONSENSUS_H2A	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	90
H2A.CI.UC2	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	128
H2A.CI.UC2	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	90
H2A.DE.BEN	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	90
H2A.DE.PE12	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	90
H2A.GH.GH1	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	110
H2A.GH.GH1	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	110
H2A.GW.ALI	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	87
H2A.GW.CAM2CG	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	90
H2A.GW.DI94	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	128
H2A.GW.FG	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	90
H2A.GW.ISY	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	90
H2A.GW.MDS	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	110
H2A.SN.HIV2ST	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	110
H2A.SN.ROD	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	91
H2AB.CI.7312A1	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	0
H2AB.CI.7312A2	W-L-G--P-L--G--N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--		---	0
CONSENSUS_H2B	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
H2B.CI.BHO	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
H2B.CI.UC1	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
H2B.GH.D205	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
H2C.IR.2238POLB7	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
H2D.IR.FO784PA	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
H2D.IR.FORTC2	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
H2G.CI.ABT96	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
CONSENSUS_MAC	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
MAC.MM239	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
MAC.US.MAC251-1A11	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			70
MAC.US.MM142	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			110
MAC.US.MM251	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			111
MNE.US.MNE	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			111
CONSENSUS_SMM	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			111
SMM.IR.SMML1B1	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			76
SMM.US.MM32H	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			115
SMM.US.SIVSMMH9	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			115
SMM.US.SMM-PBJ-6P6	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			77
SMM.US.SMMF236	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			74
SMM.US.SMMI7	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			0
SMM.US.SMMPBJ14-15	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			109
SMM.US.SMMPBJ14-441	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			109
STM.US.STM	V-V-H--L-GK-R--G-I-A--S-N--IG--S--T-N--IN--A--A-AV--G--KR--R-A--G--T--TR-GPM-GD.N--			96



```

Protease \ / Reverse Transcriptase
CONSENSUS_H2A      RRPVVTAHIEGQPVVEVLLDTGADDSIVAGIEELGSNYSPKIVGGIGGFINTKEKVEIKVLRKRVrATIMTGDTPINIFGRNLLTALGMSINLFPVAKIePik??LKPGKDG?LrQWPLTRKkiEALKEI 225
H2A.CI.UC2         -----D---D---T---V-----R-----IT---R-K-----R-K-----V----- 240
H2A.DE.BEN        -----D---D---D---T-----D---D---R-K-----R-K-----R-K----- 240
H2A.DE.PE12       -----V---V---V---D---V---VD---VI---KV-----KV-----R----- 240
H2A.GH.GH1        -----V---V---Q---D---V---I---I---ED---K-----EVR---K-----R----- 240
H2A.GW.ALI        -----H---H---D---T-----E---G---K-----IM---R-----R----- 217
H2A.GW.CM2CG      -----F---D---R-N-----D---T-----E---G---K-----LD---VT---R-K----- 220
H2A.GW.DL94       -----H---D---R-N-----D---T-----E---G---K-----V-----IM---R-K----- 258
H2A.GW.FG         -----D---D---D---R-----D---R---K-----V---VT---K-----K----- 220
H2A.GW.ISY        -----H---H---V-----V-----E---G---R-----IM---K-----R----- 240
H2A.SN.HIV2ST     -----HV-----V-----R-----R-----IM---K-----K----- 240
H2A.SN.ROD        -----H---N-----E---K-----V-----IM---K-----K----- 221
H2AB.CI.7312A1    ----- 0
H2AB.CI.7312A2    ----- 0
CONSENSUS_H2B      RRPVVkAcIEGQsVEVLLDTGADDSIVAGIEELGSNYTPKIVGGIGGFINTKEKVEIEVYKVRAtiMTGDTPINIFGRNIIINTLGMTLNFPPVAKIePik??LKPGKDGPKIRQWPLSKKkiLALKEI 239
H2B.CI.BHO        -----R---T-----P-----V-----S-----R-----Q---E----- 240
H2B.CI.UC1        -----R-----V-----K-----E-----R-----R----- 241
H2B.GH.D205       ----- 0
H2C.LR.2238POLB7  ----- 0
H2D.LR.FO784PA    xRNLLTELGMSINLFPVAKIePikVRLKPGMDGPKLRQWPLSKKkiEALKEI 0
H2D.LR.FORTC2     ----- 51
H2G.CI.ABT96      RRPVVTAHIEGQPVVEVLLDTGADDSIVAGIEELGSNYSPKIVGGIGGFINTKEKVEIKVYKRIQATVMTGDTPINIFGRNIIIVKLGMSINLFPVAKIePikVRLKPGMDGPKIRQWPLSKKkiQALREI 200
CONSENSUS_MAC      RRPVVTAHIEGQPVVEVLLDTGADDSIVTGIELGPHYTPKIVGGIGGFINTKEKVEIEVLGKRiKGTIMTGDTPINIFGRNLLTALGMSINLPIAKVPEPKvaLKPGK?GPKLKQWPLSKKkiVALREI 239
MAC.MM239         -----I-----R-----F-----D-----V-----D----- 245
MAC.US.MM142      -----I-----R-----K-----Sp---D-----D----- 245
MAC.US.MM251     -----I-----K-----K-----T-----V-----D----- 241
MNE.US.MNE        RRPVVTAHIEGQPAEVLDDTGADDSIVAGIEELGPHYTPKIVGGIGGFINTKEKVNKIEVLGKRiKGTIMTGDTPINIFGRNLLTALGMSINLFPVAKIePikVRLKPGMDGPKLRQWPLSKKkiVALREI 241
CONSENSUS_SMM      RRP?VTAHIEGQPVVEVLLDTGADDSIVAGIEELGpNYTPKIVGGIGGFINTKEKGVKIKVLGKvIKGTIMTGDTPINIFGRNLLTANGMSINLFPVAKIePikVtLKPGKdGPKLrQWPLSKKkiIALREI 205
SMM.LR.SMML1B1    -----V---H---G---T---LH-----N---E---R-----L-----V---A-----V----- 245
SMM.US.MM32H      -----V---I-----I-----I-----X-----K----- 207
SMM.US.SIVSMMH9   -----I-----V-----SMM-PBJ-6P6-----E----- 204
SMM.US.SMMF236    -----I-----V-----SMMF236-----E----- 204
SMM.US.SMMF7      -----I-----I-----SMMF7----- 239
SMM.US.SMMPBJ14-15 -----I-----I-----SMMPBJ14-15----- 239
SMM.US.SMMPBJ14-44I -----I-----I-----SMMPBJ14-44I----- 226
STM.US.STM        RRPVVTAHIEGQPVVEVLLDTGADDSIVAGIEELGLQYTPKIVGGIGGFINTKEKVNIEVLGKRiKGTIMTGDTPINIFGRNLLTALGMSINLFPVAKIePikVtLKPGKdGPKIKQWPLSKKkiEALKEI
    
```



```

CONSENSUS_H2A
H2A.CI.UC2
H2A.DE.BEN
H2A.DE.PE12
H2A.GH.GH1
H2A.GW.ALI
H2A.GW.CAM2CG
H2A.GM.DL94
H2A.GM.FG
H2A.GM.ISY
H2A.GW.MDS
H2A.SN.HIV2ST
H2A.SN.ROD
H2A.CI.7312A1
H2A.CI.7312A2
CONSENSUS_H2B
H2B.CI.BHO
H2B.CI.UC1
H2B.GH.D205
H2C.IR.2238POLB7
H2D.IR.FO784PA
H2D.IR.FORTC2
H2G.CI.ABT96
CONSENSUS_MAC
MAC.MM239
MAC.US.MAC251-1A11
MAC.US.MM142
MAC.US.MM251
MNE.US.MNE
CONSENSUS_SMM
SMM.IR.SMMLIB1
SMM.US.MM32H
SMM.US.SIVSMH9
SMM.US.SMM-PBJ-6P6
SMM.US.SMMF236
SMM.US.SMM7
SMM.US.SMMPBJ14-15
SMM.US.SMMPBJ14-441
STM.US.STM
LEPPFRKANPVDVLIHQYMDDILIASDRDLEHDKVVLQKELLNGLGFSTPDEKFKQDPPYQWMMGYELWPTKWKLOKIQLPQKewVTWVNDIQKLVGVLNWAAQIYPGIKTKhICrLIRGKMTLIEEVQWTE
-----G-----T-----I-----
-----F-----C-----L-----S-----
-----V-----RT-----L-----L-----K-----
-----O-----K-----G-----H-----G-----
-----S-----I-----R-----N-----F-----I-----K-----
-----E-----I-----R-----K-----P-----
-----Q-----I-----R-----S-----R-----RN-----I-----
-----K-----I-----R-----H-----I-----L-----
LDPPFRANNDVTIIQYMDDILIVASDRDLEHDKVVSQKELLNMGFSTPEEKFKQDPPYQWMMGYELWPKWKLOKIQIPEREVWTVNDIQKLVGVLNWAAQIFPGIKTRHICRLIRGKMTLIEEVQWTE
-----V-----N-----N-----A-----
LDPPFRKANSdVIHQYMDDILIASDRDLEHDKVVSQKELLNMGFSTPEEKFKQDPPYQWMMGYELWPKWKLOKIQLPQKewVTWVNDIQKLVGVLNWAAQIYPGIKTKhICrLIRGKMTLIEEVQWTE
-----N-----T-----V-----N-----R-----A-----
LEPPFRKANSdVTLIQYMDDILIASDRDLEHDKVVLQKELLNGLGFSTPEEKFKQDPPYHMMGYELWPTKWKLOKIQELPQREDWTWVNDIQKLVGVLNWAAQIYPGIKTKhICrLIRGKMTLIEEVQWTE
LDPPFRKANEDxIXIQYMDDILIVASDRDLEHDKVVLQKELLNKxGFSTPDEKFKQDPPYQWMMGYEXWPKKWKLOKINLxKETWTVNDIQKLVGVLxWAAQIYPGIKTKhICrLIRGxMTLIEEVQWTE
LEPPFRKANPDVTLIQYMDDILIASDRDLEHDKVVLQKELLNSIGFSTPEEKFKQDPPYQWMMGYELWPTKWKLOKIQELPQRETFWTVNDIQKLVGVLNWAAQIYPGIKTKhICrLIRGKMTLIEEVQWTE
-----N-----S-----S-----
-----S-----S-----S-----
LEPPFRKANPDVTLIQYMDDILIVASDRDLEHDKVVLQKELLNSIGFSTPEEKFKQDPPYQWMMGYELWPTKWKLOKIQELPQKETFWTVNDIQKLVGVLNWAAQIYPGIKTKhICrLIRGKMTLIEEVQWTE
LEPPFRKANPDVTLIQYMDDILIASDRDLEHDKVVLQKELLNSIGFSLPEEKFKQDPPYQWMMGYELWPTKWKLOKIQELPQRETFWTVNDIQKLVGVLNWAAQIYPGIKTKhICrLIRGKMTLIEEVQWTE
-----V-----N-----N-----
-----S-----G-----S-----
SMM.US.SMMPBJ14-15
SMM.US.SMMPBJ14-441
LEPPFRANPDVTLIQYMDDILIASDRDLEHDKVVLQKELLNGLGFSTPEEKFKQDPPYQWMMGYELWPTKWKLOKIQELPQRDVWTVNDIQKLVGVLNWAAQIYPGIKTKhICrLIRGKMTLIEEVQWTE

```

HIV-2/SIVsmm protein alignment: POL

```

CONSENSUS_H2A                                613
H2A.CI.UC2                                    630
H2A.DE.BEN                                    630
H2A.DE.PE12                                    630
H2A.GH.GH1                                     630
H2A.GW.ALI                                     607
H2A.GW.CM2CG                                  609
H2A.GW.DL194                                  648
H2A.GW.FG                                      610
H2A.GW.ISY                                    610
H2A.GW.MDS                                    630
H2A.SN.HIV2ST                                 630
H2A.SN.ROD                                    611

LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    332
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    0
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    627
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    630
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    631
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    631
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    0
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    0
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    441
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    590
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    629
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    635
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    635
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    631
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    631
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    631
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    595
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    22
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    635
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    597
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    597
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    594
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    24
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    629
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    616
LAEAELQENKIIILSQEQEGSYKQEGVPLEATVQKNLANQWTYKIHQGNKVLKVGKYAKVKNTHTNGVRLAHVVQK    616

```



Reverse Transcriptase \ / Rnase H			
CONSENSUS_H2A	LV9GPlpCaETfYTDGScNzQSkEGKAGYVTDGrDrKv7vLEqTtNOQAELFaFamaltDSGPK7NIIVDSQYVMGivagQPTESE?riVnqIIIEEMIKKEaiYVAVVPAHKGI GGNQEVVDHLVSOQIRQ		740
H2A.CI.UC2	-----T-----R-----A-----V-----S-----D-----V-----I-----		760
H2A.DE.BEN	-----K-----K-----V-----R-----A-----V-----N-----V-----		760
H2A.DE.PE12	-----K-----E-----R-----I-----A-----S-----K-----L-----T-----L-----		760
H2A.GH.GH1	-----K-----R-----R-----R-----V-----V-----D-----		760
H2A.GW.ALI	-----K-----L-----K-----K-----L-----V-----A-----N-----		737
H2A.GW.CAM2CG	-----T-----T-----R-----R-----K-----I-----A-----N-----		739
H2A.GW.DL94	-----T-----T-----R-----R-----K-----K-----S-----A-----V-----D-----		778
H2A.GW.FG	-----E-----V-----I-----I-----K-----R-----K-----A-----N-----		740
H2A.GW.ISY	-----K-----K-----I-----K-----K-----R-----I-----V-----V-----V-----		740
H2A.GW.MDS	-----K-----L-----K-----K-----R-----K-----A-----V-----N-----		760
H2A.SN.HIV2ST	-----K-----L-----K-----R-----R-----L-----A-----V-----A-----		760
H2A.SN.ROD	-----K-----K-----R-----K-----K-----K-----S-----S-----S-----		741
H2AB.CI.7312A1			332
H2AB.CI.7312A2		xMGIVAGQPTETESPIVNKIIEEMIKKEAIYVGWVPAHKGLGGNQEVVDHLVSOQIRQ	57
CONSENSUS_H2B	LVKDPLRgVETfYTDGScNzrAsKEGKAGYVTDGRKDKVvLEqTtNOQAELFAFALALqDSgPQVNIIVDSQYVMGivagQPTESEPIV?2IIEEMIKKEaiYVGWVPAHKGLGGNQEVVDHLVSOQIRQ		755
H2B.CI.BHO	-----K-----K-----P-----P-----T-----T-----N-----Q-----		760
H2B.CI.UC1	-----K-----R-----T-----T-----I-----I-----G-----G-----L-----N-----		761
H2B.GH.D205	-----R-----T-----T-----T-----E-----E-----A-----K-----		761
H2C.IR.2238POLB7		xLGIAGQPTETESPLVNRILIEELIRKEAVYISWVPAHKGLGGNQEVVDHLVSOQIRQ	57
H2D.IR.FO784PA		xIGIVAGQPTESSEKLVNQIIIEEMIKKEAIYIAWVPAHKGI GGNQEVVDHLVSOQIRQ	57
H2D.IR.FORTC2		LVKBPLeKtETfYVDGScHRQSkEGKAGYVTDGRDREKTRALQQTtNOQAELFAFLLALQDSgPKANIIVDSQYVMGivagQPTESEKLVNQIIIEEMIKKEAIYIAWVPAHKRI GGNQEVVDHLVSOQIRQ	571
H2G.CI.ABT96		LVKBPLeQTEfYTDGScNKTSKxKAGYVTDGRKDKVxVLEqTtNOQxELFAFAMALQDSgPKVNIIVDSQYVMxIIAxQPTESEPLVNIIEEIKKExLKYVGWVPAHKGI GGNQEVVDxLVSOQIRQ	720
CONSENSUS_MAC	LVKDPiEGEETfYTDGScChKQSkEGKAGYITDRGKDKVvLEqTtNOQAELFAFLLMALTDGSPKANIIVDSQYVMGIIIGCPTESSESLVNIIEEMIKKSEIYVAVVPAHKGI GGNQEI DHLVSOQIRQ		759
MAC.US.MAC251-1A11	-----V-----S-----S-----K-----K-----K-----		765
MAC.US.MM142	-----V-----S-----S-----S-----T-----T-----		761
MAC.US.MM251	-----V-----S-----S-----S-----T-----T-----		761
MNE.US.MNE	LVKDPIKGEETfYVDGScNKQSkEGKAGYITDRGKDKVvLEqTtNOQAELFAFLLMALADSGPKANIIVDSQYVMGIIIGCPTESSESLVNIIEEMIKKTEIYVAVVPAHKGI GGNQEI DHLVSOQIRQ		761
CONSENSUS_SMM	LVKepIggaeTfYVDGScNzrqsREGKAGYVTDGRgrdkaKlLEqTtNOQAELFAFALALADSGPKANIIVDSQYVMGivagQPTESEsr1VnQIIIEEMIKKEaiYVAVVPAHKGI GGNQEVVDHLVSOQIRQ		725
SMM.IR.SMMLT1B1	-----Lx-----V-----E-----K-----V-----A-----E-----E-----		152
SMM.US.MM32H	-----D-----K-----E-----Y-----T-----K-----V-----V-----		765
SMM.US.SIVSMMH9	-----R-----R-----R-----R-----G-----T-----		727
SMM.US.SMM-PBJ-6P6	-----R-----R-----R-----R-----G-----T-----		727
SMM.US.SMMF236	-----D-----D-----E-----I-----G-----I-----R-----S-----I-----F-----		724
SMM.US.SMMH7	-----D-----D-----E-----I-----G-----I-----R-----S-----I-----F-----		153
SMM.US.SMMPBJ14-15	-----D-----D-----E-----I-----G-----I-----R-----S-----I-----F-----		759
SMM.US.SMMPBJ14-441	-----D-----D-----E-----I-----G-----I-----R-----S-----I-----F-----		759
STM.US.STM	LVKBPLeGTEfYVDGScNRQSkEGKAGYITDRGNKVKALeQTTNOQAELFAFAMALADSGPKANIIVDSQYVMGIIIGCPTESSEKLVNQIIIEEMIKKEAIYVAVVPAHKGI GGNQEVVDHLVSOQIRQ		746

HIV-2/SIVsmm proteins

Rnase H \ / Integrase					
CONSENSUS_H2A	VLFLFK . IEPAQEEHkYHenvkELshKfgiP?LVARQIVNTcagcQQKGEaIHGQVnaeIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQESGRQTALFLLKLaSRWPITHLHtdNGANfTSQEVK				868
H2A.CI.UC2	II I - Q -	I - V			889
H2A.DE.BEN	II T - L - S -	I - V - Y -			889
H2A.DE.PE12	L - K -	D -			889
H2A.GH.GH1	M - T - C -	I - V -			889
H2A.GW.ALI	L - N -	V - I -			866
H2A.GW.CM2CG	T - C - D - O -	Y -			868
H2A.GW.D194	I - T - C -	I - V -			907
H2A.GW.FG	I - K -	HV -			869
H2A.GW.ISY	A - L - N -	V -			869
H2A.GW.MDS	L - C -	D -			889
H2A.SN.HIV2ST	L - K - T -	V -			889
H2A.SN.ROD	N - S -				870
H2AB.CI.7312A1					332
H2AB.CI.7312A2	VLFLFK . IEPAQEEHkYHGNVKELVHKFGIPQLVAKQIVNSCDKQQKGEaIHGQVnaeIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfTSQDVK				186
CONSENSUS_H2B	VLFLFK . IEPAQEEHkYHGNVKELVHKFGIPQLVAKQIVNSCDKQQKGEaIHGQVnaeIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfTSQGVK				884
H2B.CI.BHO	I - - - - N - - - - R -	S -			889
H2B.CI.UC1	L - - - - - L -	V -			890
H2B.GH.D205	- - - - -	D -			890
H2C.IR.2238POLB7	VLFLFK . IEPAQEEHkYHNSAKELTHKFGIPNLVAKQIVNTCHVQQKGEaIHGQVnaDLgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfTSQEVK				186
H2D.IR.FO784PA	VLFLFK . IEPAQEEHkYHNSIKELVFKFGLPKLVARQIIDTCdChRkGEaIHGQVnaeIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfTSQEVK				186
H2D.IR.FORTC2	VLFLFK . IEPAQEEHkYHNSIKELVFKFGLPKLVARQIIDTCdChRkGEaIHGQVnaeIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfTSQEVK				658
H2G.CI.ABT96	VLFLKk . IEPAQEEHkYHNSVKElTHKFGIPQLVAKQIVNSVHSQQKGEaIHGQVnaeIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfTSQEVK				849
CONSENSUS_MAC	VLFLFK . IEPAQEEHkYHNSVKElVFKFGLPKLVARQIIDTCdChQChQKGEaIHGQVnaSDIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfASQEVK				887
MAC.MM239	- - - - - Y - I -	A -			894
MAC.US.MAC251-1A11	- - - - -	I -			894
MAC.US.MM142	S - - - - I -	V -			890
MAC.US.MM251	- - - - -	V -			890
MNE.US.MNE	VLFLFK . IEPAQEEHkYHNSVKElVFKFGLPKLVARQIIDTCdChQChQKGEaIHGQVnaSDIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfASQEVK				890
CONSENSUS_SMM	VLFLkK?IEPAQEEHkYHNSVKElVFKFGLPKLVARQIIDTCdChQChqkGEaIHGQVnaeIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPqETGRQTALFLLKLaSRWPiThLHtdNGANfTSQEVK				854
SMM.IR.SM11B1	I - - - - Y - I -	V -			281
SMM.US.MM32H	- - - - D -	I - R - SD -			894
SMM.US.STVSMH9	- - - - K - x -	L - x - xx -			856
SMM.US.SMM-PBJ-6P6	- - - - -	- G -			856
SMM.US.SMMF236	- - - - K - x -	- R - T - RI - I -			854
SMM.US.SMMF7	- - - - -	Q - - - - R - T - RI - I -			282
SMM.US.SMMPBJ4-15	- - - - -	- - - - -			888
SMM.US.SMMPBJ4-441	- - - - -	- - - - -			888
STM.US.STM	VLFLFK . IEPAQEEHkYHNSVKElVFKFGIPRLVAKQIVDTCDKChQKGEaIHGQVnaeIgtWQMDCTHLEgKi.IIVAVHVASGfLEAEVIPQETGRQTALFLLKLaSRWPITHLHtdNGANfTSQEVK				875

```

CONSENSUS_H2A
H2A.CI.UC2
H2A.DE.BEN
H2A.DE.PE12
H2A.GH.GH1
H2A.GW.ALI
H2A.GW.CAMCG
H2A.GH.DL94
H2A.GW.FG
H2A.GW.ISY
H2A.GW.MDS
H2A.SN.HIV2ST
H2A.SN.ROD
H2AB.CI.7312A1
H2AB.CI.7312A2
CONSENSUS_H2B
H2B.CI.BHO
H2B.CI.UC1
H2B.GH.D205
H2C.LR.2238POLB7
H2D.LR.FO784PA
H2G.CI.ABT96
CONSENSUS_MAC
MAC.US.MM239
MAC.US.MAC251-1A11
MAC.US.MM142
MAC.US.MM251
MNE.US.MNE
CONSENSUS_SMM
SMM.LR.SMMLIB1
SMM.US.MM32H
SMM.US.SIVSMH9
SMM.US.SMM-PBJ-6P6
SMM.US.SMMF236
SMM.US.SMM7
SMM.US.SMPEBJ14-15
SMM.US.SMPEBJ14-44
STM.US.STM
MVAWVVGIEQTFGVPYNPQSGVVEAMNHLKNQIDRIReQANTVETIVLMAVHCMNFKRRGGIGDMTPaERLI.NMI.tTEQEIQFLgqKnsk.lknFrVvYfREGRDQLWkGPgELLWkGGAViVKVGtdI
-----R--NF-K-Q--Y-----E-----A-----
-----R--NF--Q--Y-----E-----
-----H-----
-----R--NF--O--Y-----A-----
-----T--L-F--P--Y-----
-----S-----
-----R--NF-K-Q--Y-----A-----
-----V--V-----N--Q-----E-----
-----S-----V-A-----A-----
-----S-----D-----E-----L-----
MVAWVVGIEQTFGVPYNPQSGVVEAMNHLKNQIDRIReQANTVETIVLx
MaAWWi.GIEQTFGVPYNPqSGVVEAMNHLKNQIDRI.RDQAVSLFVVLMAat.HCMNFKRRGGIGDMTPaERLI.YNMITTEQEIQfFOQKnlkFQNFvVYyREGRDQLWkGPgELLWkGEGAVi.IKVGTEI
-----E-----
-----A-----L-----
-----L-----
MVAWVVGIEHTFGVPYNPQSGVVEAMNHLKNQIDRIReQANTMETIVLx
MVAWVVGIEQTFGVPYNPQSGVVEAMNHLKNQIDKIRDOANSVETIVLx
MVAWVVGIEQAFGVPYNPQSGVVEAMNHLKQIDKIRDOANSIETIVLMAVHCMNFKRRGGIGDMTPaERLI.INMITTEQEIQfOQTKNSKFKFNFRVYyREGRDQLWkGPgELLWkGEGAVi.IKVGTEI
MVAWVVGIEHTFGVPYNPQSGVVEAMNHLKNQIDRIReQANSVETIVLMAVHCMNFKRRGGIGDMTPaERLI.NMITTEQEIQfOQSKNSKFKFNFRVYyREGRDQLWkGPgELLWkGEGAVi.LKVGTDI
-----S-----T-----
-----
MVAWVVGIEHTFGVPYNPQSGVVEAMNHLKNQIDRIReQANSMETIVLMAVHCMNFKRRGGIGDMTPaERLI.NMITTEQEIQfOQSKNSKFKFNFRVYyREGRDQLWkGPgELLWkGEGAVi.LKVGTDI
-----N-----K-D-----T-----
-----H-----N-----V-----I-----
-----x-----x-----I-----VS-----
-----D-----H-----K-----R-----N-----E-----V-----VS-----
MVAWVVGIEQTFGVPYNPQSGVVEAMNHLKQIDRIReQANTVETIVLMAVHCMNFKRRGGIGDMTPaERLI.NMITTEQEIQfOQSKNSKFKFNFRVYyREGRDQLWkGPgELLWkGEGAVi.VKVGTDI
-----N-----K-----D-----
-----x-----x-----I-----VS-----
-----D-----H-----K-----R-----N-----E-----V-----VS-----
Integrase \

```

CONSENSUS_H2A	KVPPRRKAKIIrDYGgRqeLDsgshLFGaRE?dgeva?	1034
H2A.CI.UC2	-----SP-----	1055
H2A.DE.BEN	-----SP-----	1142
H2A.DE.PEI2	-----R-V-S-T-----	1055
H2A.GH.GH1	-----E-----	1055
H2A.GW.ALI	-----P-----X	1029
H2A.GW.CM2CG	-----S-----N	1034
H2A.GH.DL94	-----K-----M-----M	1073
H2A.GM.FG	-----P-----M-----M	1035
H2A.GM.ISY	-----DM-----P-----T-----	1035
H2A.GW.MDS	-----K-----M-----N-----	1055
H2A.SN.HIV2ST	-----K-----M-----N-----	1055
H2A.SN.ROD	-----M-----N-----	1036
CONSENSUS_H2B	KViPPRRKAKIIrhYGGGKeLDCsaDmEDITrQ.AREmaQS?	1052
H2B.CI.EHO	-----N-----V-----M-----N-----	1058
H2B.CI.UC1	-----N-----V-----M-----N-----	1064
H2B.GH.D205	-----GT-----GT-----GQVPEA	1059
	--V-----G-----G-----D	
H2G.CI.ABT96	KVPPRRKAKIIKDYGGGKEVDSSTNMEDTrQ.TGEVA\$	1015
CONSENSUS_MAC	KVPPRRKAKIIKDYGGGKE?DSSSHMEDTGE.AREVA	1052
MAC.-.MM239	-----V-----	1060
MAC.US.MAC251-1A11	-----M-----	1060
MAC.US.MM142	-----M-----	1056
MAC.US.MM251	-----V-----	1056
MNE.US.MNE	KVPPRRKAKIIKDYGGGKEVDSSTNMEDTGE.AREVA	1056
CONSENSUS_SMM	kvpprrkakiikdygggkeldsgshledtge.aveva	1020
SMM.US.MM32H	-----V-----S-----M-----	1060
SMM.US.STVSMH9	-----V-----S-----M-----	1022
SMM.US.SMM-PBJ-6P6	-----V-----S-----M-----	1022
SMM.US.SMMF236	-----V-----S-----M-----	1020
SMM.US.SMMPBJ4-15	-----V-----S-----M-----	1054
SMM.US.SMMPBJ4-441	-----V-----S-----M-----	1054
STM.US.STM	KVPPRRKAKIIKDYGGGKEVDSGSHLEDTrE.AGEVA	1041



CONSENSUS_H2A	vpk cds   ->				
H2A. . HIV2JAUL	KLLSCCNYPgAHksqVPSLQFLALVVVQQNd.rpQr??ttrKgwRrdyrrGIRvArqD?rs?kQrseppaprayFFGvAKVleILA				206
H2A.CI.UC2	-----K-----	-----KY-----	-----W-----	-----SGGH-----	-----K-----
H2A.DE.BEN	-----A-----	-----Y-----	-----HW-----	-----YLE-----	-----GK-----
H2A.GH.PEI2	-----R-----	-----K-----	-----A-----	-----EHL-----	-----G-----
H2A.GH.GH1	-----Y-----	-----L-----	-----SNW-----	-----PLK-----	-----CGH-----
H2A.GH.D194	-----A-----	-----Y-----	-----HW-----	-----YJ-----	-----G-----
H2A.GH.ISY	-----K-----	-----DN-----	-----N-----	-----L-----	-----G-----
H2A.GW.ALI	-----K-----	-----K-----	-----NS-----	-----L-----	-----S-----
H2A.GW.CAM2CG	-----K-----	-----R-----	-----DR-----	-----L-----	-----SY-----
H2A.GW.FG	-----K-----	-----R-----	-----S-----	-----GNSA-----	-----R-----
H2A.GW.MDS	-----T-----	-----G-----	-----DS-----	-----R-----	-----A-----
H2A.SN.HIV2ST	-----Y-----	-----G-----	-----DN-----	-----N-----	-----G-----
H2A.SN.ROD	-----R-----	-----RA-----	-----DS-----	-----R-----	-----L-----
CONSENSUS_H2B	KILSYNYPsAhEGVPSLQFLALRVVvOEGKGSOGESATRkQRrNsRrSrlrLArkNmraOQgSsQplAprTyFFGLAEVLGILA				
H2B.CI.EHO	-----H-----	-----T-----	-----I-----	-----N-----	-----G-----
H2B.CI.UC1	-----H-----	-----T-----	-----I-----	-----N-----	-----G-----
H2B.GH.D205	-----H-----	-----T-----	-----I-----	-----N-----	-----G-----
H2G.CI.ABT96	QILSCCNYPtAhKkQVPSLQFLALQVVQKH.GSKGESXTRkQRrGDNRRGIRMAKNSXRDDQDSSQSFtQgAYFFGLAKVLGILA\$				
CONSENSUS_MAC	qLLSCC:FpRAHkyQVPSLQYLALkVV.SdV.RSOGENPTWkQWRrDNRrRgLuRmAKQNSRgDKQRgKpTKGAnFFGLAKVLGILA				
MAC. . MM239	-----R-----	-----H-----	-----R-----	-----H-----	-----S-----
MAC. . MM316S	-----R-----	-----H-----	-----R-----	-----H-----	-----S-----
MAC.US.MAC251-1A11	-----R-----	-----H-----	-----R-----	-----H-----	-----S-----
MAC.US.MM142	-----R-----	-----H-----	-----R-----	-----H-----	-----S-----
MAC.US.MM251	-----K-----	-----R-----	-----R-----	-----R-----	-----S-----
MNE.US.MNE	QLLSCCRFPRAHkTQVPSLQYLALRVV.SVY.RSQRENPTWkQWRrDNRrRSLRMAKQNSRgDKQRgKpTKGVDFPGLAKVLGILA				
CONSENSUS_SMM	kLLSCCFpRAHknQVPSLQYLALtVV.ShV.rSQGenPTWkQWRrRnnRrgLrLrArQNsRrtnKQgssesfaeG7NFFGLAKVLGILA				
SMM.US.MM32H	-----R-----	-----Y-----	-----X-----	-----X-----	-----X-----
SMM.US.SIVSMMH9	-----R-----	-----Y-----	-----X-----	-----X-----	-----X-----
SMM.US.SMM-PBJ-6P6	-----D-----	-----D-----	-----K-----	-----M-----	-----K-----
SMM.US.SMMF236	-----D-----	-----D-----	-----K-----	-----M-----	-----K-----
SMM.US.SMMPBJ14-15	-----D-----	-----D-----	-----K-----	-----M-----	-----K-----
SMM.US.SMMPBJ14-441	-----D-----	-----D-----	-----K-----	-----M-----	-----K-----
STM.US.STM	KLLSCCRFTKAHKNQVPSLQYLALKVV.EHV.RSQRENTARkQWRrRNGSIRVATQNGRHkPRGSKPSTEGTDFPGLAKVLGILA				



HIV-2/SIVsmm  
proteins

CONSENSUS\_H2A  
H2A.CI.UC2  
H2A.DE.BEN  
H2A.DE.PEI2  
H2A.GH.GH1  
H2A.GM.D194  
H2A.GM.ISY  
H2A.GM.ALI  
H2A.GW.CAM2CG  
H2A.GW.FG  
H2A.GW.MDS  
H2A.SN.HIV2ST  
H2A.SN.ROD  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130  
130

<-| Vpr cds  
MetpKapeSL?syNEPpsctSeqdva?QeIakqGeeIlsOlyrPLaC2N?CyK7CcyhCQlCF1kKGLG1WY?RkgrRRRRtPKktK?ksssasD.KSISrTgmsqpkkOktlettvetd?glGr?  
---K---P---WE-A---L-A-H---T-P---K-SP---V---G---R---T-PP-TP---IQ-D-R-T---E-P---P-A-TS---\$  
---R---TA---R---TA-H---P-T-K---R-SP---S---E---R---TP-P-PP---D---T-E---S-A---V-TG---O  
---M-C---KS---PEI---RL---O---N---S---C-E---A---A---T---P---A---R---Q---\$  
---H---E---G-TA---L-H---H---T-S---Q-SP---R---T---D---T-E---T---MW-TCS---  
---E---R---TA-R---L-A-H---T-S---Q-S---N---C-D---R---T---P---V---SAC---H  
---G---R---T---N---L---A---N---C-D---S---I---A---P---T---G---L-A-P---  
---E-G---MP---V---G---ATT---L---L---Q---E-D-S---K---R---A---A---P---R---A---L---G-C-P---  
---E-C---R---T---R---R---T---S---K---D---Q---D---L---A---P---R---A---L---L---  
---I---N---E---H---T---R---R---N---N---R---I---N---E---A---T---P---T---A---L---P---\$  
---G---G---A---A-G-VSP-D---Y---Q---D---K---K---M---N---N---A---L---IG-P---\$  
---K-C---F---R---T---R---R---T---N---S---R---M---N---C---E---D---T---V---A---T---P---  
129  
129  
128  
128  
112

ME1PlgQEESLKSssepSSSTSEpVVntQGLDNGEEILLsOLYrPLKACgntCYCKKCCYHCQlCF1kKGLG1CYdrS..RKRSSkRAKtTasSAppd??SLSarTgDSOPtKKOkkEVETtg?TD1GpGR  
---K---N---GH---G-A---R---S---E---K---S---T---NMR---E---V---S---E---N---TK---K-L---  
-----N-----R---A---P---R---K-----T---P-----

H2G.CI.ABT96  
CONSENSUS\_MAC  
MAC.US.MM239  
MAC.US.MAC251-1A11  
MAC.US.MM142  
MAC.US.MM251  
MAC.US.MMP11  
MNE.US.MNE  
130  
130  
131  
130  
130  
130

METPSREQNSLKSRELSSCTFEETVDAPLEGTQAEILYLYRPLAECNSKCYCKKCCYHCQlCF1kKGLG1GVCYEOP.RRRRTPKTKANFSASN.ESISARTRXQPKKK\$  
METPLRQENSLESsneiSSciseAga?cPEsAnILEEILLsOLYrPLAECYNTCYCKKCCYHCQlCF1kKGLG1CYEgS?KRRRTPKKAKANTSSASN?kpIpmRtRhCQPeKAKKtVEKAVATAPGLGR  
---S---SGH---T---A---S---L---S---NR---S---S-----  
---Y---A---AL---L---Y---S---K---HR---S---I---L---K---A---E-----  
---L---T-----L-----L-----  
---L---T-----L-----L-----

H2B.CI.EHO  
H2B.CI.UC1  
H2B.GH.D205  
CONSENSUS\_SMM  
SMM.US.MM32H  
SMM.US.S1U049B2  
SMM.US.SIVSMNH9  
SMM.US.SMMF236  
SMM.US.SMMPBJ14-441  
STM.US.STM  
129  
129  
128  
128  
128  
128  
128  
129

METPLRQENSLESsEISEVda?tPEsAnILEEILLsOLYrPlEACYNkCYCKkCCHCQhCF1kKGLG1CYEQ??RRRTPKKtKANT?saSn?ksI?RaRncQPkKkEtVE?eVAtDlGLGR  
---R-----N-R-C---A---I---G---T---K---F---SRK---A---S---NRPI---N-I-H---E-A---R---A---AP---\$  
-----S---C---D---D---S---X---X---X---X---X---F---H---H---D---XX---D---XX---K---I---A---X  
-----C---C---VP-----D-----S-----H-----H-----PFP-----T-----R-----T-----A-----  
-----S-----D-----S-----Q-----S-----D-----R-----R-----A-----A-----A-----\$

H2A.GW.ALI  
H2B.CI.EHO  
SHIYIS\$ 136  
SNTSTRFAN\$ 139



Multiple sequence alignment of HIV-2/SIVsmm protein REV. The alignment shows conserved regions across various sequences including H2A, H2B, H2G, MAC, MNE, and STM. Specific residues are highlighted with asterisks to indicate conserved positions.

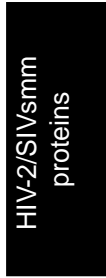


HIV-2/SIVsmm proteins

HIV-2/SIVsmm protein alignment: VPX

CONSENSUS_H2A	<-  vif cds	110
H2A.CI.UC2	MadPRetvPPGNSGEETiGeAFwLqRtVbAInREAVNH1PRELlFQVWQRswrYWHdeGMS?SYTKRYRLcImQkAmy?HfkrgrcTClGgghGpgGwrrp?GPPPPPpPGLV	112
H2A.DE.BEN	-----I-L-----L-S-----FI-L-K-----	113
H2A.DE.PE12	-----E-I-L-----A-----IFT-----W-EDM-RE-LEDQ-----	111
H2A.GH.GH1	-----G-V-----A-E-M-----DL-O-R-Y-FI-V-K-----	112
H2A.GM.D194	-----R-----I-L-----P-----VFI-----S-----	112
H2A.GM.I5Y	-----R-----I-L-----A-----I-V-I-K-----R-----S-----	112
H2A.GM.ALI	-----E-D-----R-----VFM-K-----R-E-----S-----	112
H2A.GM.CAM2CG	-----N-----L-----Q-----I-MK-----S-----	112
H2A.GM.FG	-----E-A-----Q-----VFI-----S-----	113
H2A.GM.MDS	-----E-A-----Q-----I-M-----M-A-DG-----x-----Sx-----	112
H2A.SN.A640	-----EQ-A-----H-----T-----I-GRK-----R-E-----	112
H2A.SN.A1871	-----I-----D-----E-V-----T-----	112
H2A.SN.A2057	-----I-----E-----V-----T-----	112
H2A.SN.HIV2ST	-----E-----L-----R-I-----I-M-C-----	112
H2A.SN.ROD	-----G-I-----I-----FI-S-----	112
	-----A-N-----E-----II-----V-M-VRK-----R-----	112
CONSENSUS_H2B	M.DPRERvPPGNSGEETiGeAFwLqRtVbAInREAVNH1PRELlFQVWQRswrYWHdeGMSiSYTKRYRLcImQkAMFvHyTKGCrClqEGHGFGWRS.GPPPPPpPGLA	110
H2B.CI.BHO	-----G-----E-T-LEH-----K-----C-----I-FA-----G--R-----	111
H2B.CI.UC1	-----I-----D-----A-----S-----	111
H2B.GH.D205	-----A-----	111
H2G.CI.ABT96	M.DPRERxPPENSDEETiGeAFwLqRtVbAInREAVNH1PRELlFQVWRRSWEYWHDEQMSASyTKRYRLcImQkAMFvHCKKGTCTlQKHGPGGWRQ.GPPPPPpPGLA\$	112
CONSENSUS_MAC	MSDPRERIPPGNSGEETiGeAFwLqRtVbEInREAVNH1PRELlFQVWQRswrYWHdeGMSpSyvKYRYLcIiQkAlFmHCKKGTCTlQKHGPGGWRP.GPPPPPpPGLA	112
MAC.MM239	-----	112
MAC.MM316S	-----	112
MAC.US.MAC251-1A11	-----	112
MAC.US.MM142	-----Q-T-----	112
MAC.US.MM251	-----Q-----M-----	112
MNE.US.MNE	MSDPRERIPPGNSGEETiGeAFwLqRtVbEInREAVNH1PRELlFQVWQRswrYWHdeGMSQsYvKYRYLcIiQkAlFmHCKKGTCTlQKHGPGGWRP.GPPPPPpPGLA	112
CONSENSUS_SMM	MsdPRERiPPGNSGEETiGeAFwLqRtVbEInRaAVNH1PRELlFQVWvRswEYWHdeGMSvSyTKRYLcIiQkAlFmHCKKGTCTlQKHGPGGWRP.GPPPPPpPGLA	111
SMM.US.MM32H	-----V-----E-N-----E-----Q-----Q-V-----EG-----	112
SMM.US.SIU04982	-----E-N-----G-----E-----V-----R-----E-----T-----	91
SMM.US.SITVSMH9	-xx-x-----xx-x-----H-x-----x-----x-----	112
SMM.US.SMM-PBJ-6P6	-----D-----D-----M-----M-----	112
SMM.US.SMMF236	-----K-E-----N-----E-----V-----	112
SMM.US.SMMPBJ4-15	-----D-----D-----Y-----M-----	112
SMM.US.SMMPBJ4-441	-----D-----D-----M-----M-----	112
STM.US.STM	MSDPRERIPPGNSGEETiGeAFwLqRtVbEInREAVNH1PRELlFQVWRRSWEYWHdeGMSpGSyvKYRYLcIiQkAlFmHCKKGTCTlQKHGPGGWRP.GPPPPPpPGLA	112

Consensus H2A, H2A.CI, H2A.DE, H2A.GH, H2A.GM, H2A.SN, H2A.SN.ROD, H2AB.CI.7312A, Consensus H2B, H2B.CI.EHO, H2B.CI.UCI, H2B.GH.D205, H2C.IR.HIV22238E, H2D.IR.FO784, H2G.CI.ABT96, Consensus MAC, MAC.F.MM239, MAC.F.MM132ZL, MAC.F.MM316ZQ, MAC.F.SEGA, MAC.F.SEGA, MAC.F.SIU18034, MAC.F.SIU18039, MAC.F.SIU18046, MAC.US.CTL7202, MAC.US.MM142, MAC.US.MM251, Consensus MNE, MNE.F.SIU06277, MNE.F.SIU06352, MNE.F.SIU06377, Consensus SMM, SMM.F.P209C15, SMM.F.SMP209, SMM.US.MM32H, SMM.US.SIU04982, SMM.US.SIVSMMB9, SMM.US.SIVSMMB670, SMM.US.SIVSMMF236, SMM.US.SIVSMMFJ4-15, STM.US.STM, STM.US.STMAK3



HIV-2/SIVsmm proteins

CONSENSUS\_H2A 183  
 H2A - CBL21 224  
 H2A - CBL21 224  
 H2A - CBL22 220  
 H2A.CI.UC2 227  
 H2A.DE.BEN 224  
 H2A.DE.FE12 216  
 H2A.GH.GH1 216  
 H2A.GM.CBL24 225  
 H2A.GM.ISY 215  
 H2A.GM.D194 213  
 H2A.GW.CAM1 223  
 H2A.GW.CAM3 223  
 H2A.GW.CAM2CG 225  
 H2A.GW.CAM2env 225  
 H2A.GW.CAM4 226  
 H2A.GW.CAM5 218  
 H2A.GW.CAM6 228  
 H2A.GW.FG 214  
 H2A.GW.HIV-2ALI 221  
 H2A.GW.MDS 235  
 H2A.SN.HIV2ST 218  
 H2A.SN.ROD 224  
 H2AB.CI.7312A 0  
 CONSENSUS\_H2B 184  
 H2B - GH-2 0  
 H2B.CI.EHO 223  
 H2B.CI.UC1 223  
 H2B.GH.D205 224  
 H2C.LR.HIV2238E 0  
 H2D.LR.FO784 0  
 H2G.CI.ABT96 200  
 CONSENSUS\_MAC 229  
 MAC - MM239 230  
 MAC - MM132ZL 230  
 MAC - MM316ZQ 230  
 MAC - SEGA 230  
 MAC - SEGE 230  
 MAC - SIU18034 228  
 MAC - SIU18039 228  
 MAC - SIU18046 228  
 MAC.US.CTL7202 230  
 MAC.US.MM142 233  
 MAC.US.MM251 232  
 CONSENSUS\_MNE 130  
 MNE - SIU06277 132  
 MNE - SIU06352 127  
 MNE - SIU06377 132  
 MNE.US.MNE 230  
 CONSENSUS\_SMM 217  
 SMM - P209C15 233  
 SMM - SMP209 232  
 SMM.US.MM32H 229  
 SMM.US.SI04982 232  
 SMM.US.SIVSMMB9 233  
 SMM.US.SIVSMMB670 229  
 SMM.US.SHMFP236 231  
 SMM.US.SNMPBJ14-15 236  
 STM.US.STM 229  
 STM.US.STMAK3 0



HIV-2/SIVsmm proteins

Table with 3 columns: protein identifier (e.g., CONSENSUS\_H2A, H2A\_C1, H2A\_C2), amino acid sequence, and position number (e.g., 408, 463, 459, 456, 458, 453, 454, 453, 463, 451, 459, 464, 464, 462, 456, 465, 452, 458, 472, 455, 461, 0, 403, 68, 459, 462, 463, 0, 0, 446, 473, 474, 474, 474, 474, 469, 472, 467, 474, 477, 476, 374, 378, 373, 378, 476, 458, 482, 481, 473, 481, 478, 484, 478, 89).



HIV-2/SIVsmm proteins

CONSENSUS\_H2A H2A...CBL21 657  
 H2A...CBL22 715  
 H2A...CBL23 711  
 H2A.CI.UC2 708  
 H2A.DE.BEN 718  
 H2A.DE.FE12 706  
 H2A.GH.GH1 706  
 H2A.GM.CBL24 715  
 H2A.GM.ISY 706  
 H2A.GM.D194 704  
 H2A.GW.CAM1 714  
 H2A.GW.CAM3 716  
 H2A.GW.CAM2CG 715  
 H2A.GW.CAM2env 717  
 H2A.GW.CAM4 716  
 H2A.GW.CAM5 709  
 H2A.GW.CAM6 718  
 H2A.GW.FG 705  
 H2A.GW.HIV-2ALI 711  
 H2A.GW.MDS 725  
 H2A.SN.HIV2ST 708  
 H2A.SN.ROD 714  
 H2AB.CI.7312A 150  
 CONSENSUS\_H2B H2B...GH-2 641  
 H2B.CI.EHO 321  
 H2B.CI.UC1 712  
 H2B.GH.D205 716  
 H2C.IR.HIV2238E H2C.IR.HIV2238E 151  
 H2D.IR.FO784 H2D.IR.FO784 150  
 H2G.CI.ABT96 H2G.CI.ABT96 703  
 CONSENSUS\_MAC MAC...MM239 525  
 MAC...MM132ZL 728  
 MAC...MM316ZQ 728  
 MAC.US.MM142 727  
 MAC.US.MM251 731  
 MNE.US.MNE 730  
 CONSENSUS\_SMM SMM...P209C15 710  
 SMM...SMP209 736  
 SMM.US.MM32H 735  
 SMM.US.SIU04982 727  
 SMM.US.SIVSMH9 735  
 SMM.US.SIVSMH670 732  
 SMM.US.SMMF236 734  
 SMM.US.SMMPBJ14-15 738  
 STM.US.STM 731  
 STM.US.STMAK3 342

OAgLNSWGCafRQVChTTtVpW...vNdsItPdWnMTWQeWbq?vryLeANISgsLEgAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP  
 ...T-Q-R-D-K-OI-T-GQ-RI-IVV-  
 ...R-E-GKI-D-Q-IL-V-V-VV-I  
 ...S-K-KO-E-IL-V-H-A-L-F  
 ...E-N-IV-OT-D-R-S-V-V-VV-VV-I  
 ...O-HO-EL-IV-A-I-L  
 ...E-HK1-F-E-M-V-V  
 ...KR-H-E-G-A-I-L  
 ...A-K-K-R-V-I  
 ...K-H-EM-N-T-S-V-A-I-F  
 ...A-E-E-L-N-T-S-V-V-VV-S  
 ...A-E-E-L-N-T-S-V-V-VV-I  
 ...LRT-D-L-S-R-F-II-VV-I-I  
 ...HKI-D-D-S-R-V-C-A-V  
 ...K-K-K-T-L-F-VR-V-V-I  
 ...N-K-D-Q-EQ-T-L-A-V-I  
 ...S-E-E-K-E-Q-I-V  
 ...T-T-N-RI-N-E-I-V  
 ...A-D-K-Q-K-I-L-AV-V  
 VNSMQEwQNMTWQEWKQIAFLedNITeLEAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP  
 PNSTLTPDWNMTWQEWKQIAFLedNITeLEAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP  
 pNctI2PdWnMTWQEWKq?nFtdANITaLLEEAQIQQEKNmYELQKLSMDVfGNWFD?TSW?YI?LGLY?YaG1?vLRIvYIvQMLARLRKGYRPVfSSP  
 ...D-K-Q-E-E-H-V-I-V  
 ...V-S-K-R-R-K-L-IK-H-V-I-V  
 ...T-E-R-R-Q-I-S-F-MA-R-I-I-V-AL-I  
 ...T-N-E-R-R-M-A-R-V-I-I-M  
 ...T-N-H-H-I-IK-H-I-V-V  
 VNSMQEwQNMTWQEWKQIAFLedNITeLEAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP  
 PNSTLTPDWNMTWQEWKQIAFLedNITeLEAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP  
 PNASLTPKwNNeTWQEWKQIAFLedNITeLEAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP  
 ...S-§-A-T-I-I-R  
 ...D-D-D-I-V  
 PNASLTPNwNNeTWQEWKQIAFLedNITeLEAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP  
 PngtIvP?WnNmtWQEWKqYdfLEanITqLLEEAQIQQEKNmYELQKLSMDiFGNWFDltSwiYIOYGLiVIGV?gLFRIvYvYQMLARLRKGYRPVfSSP  
 ...S-I-D-Q-S-P-V-V-V  
 ...S-I-D-Q-S-P-V-V-V  
 ...AS-T-D-D-K-E-A-A-K-K-Y-V-LL-I-K  
 ...E-N-S-x-K-N-xx-x-x-x-K-I  
 ...T-N-x-K-S-M-V-xx-xx-x-x-K-I  
 ...E-N-K-S-M-V-xx-xx-x-x-K-I  
 ...T-N-K-N-S-T-I-I  
 PNDSLVPDWNMTWQEWKQIAFLedNITeLEAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP  
 PNDSLVPDWNMTWQEWKQIAFLedNITeLEAQIQQEKNmYELQKLSMDVfGNWFDltSwiKyiQYgyvYvgiiaLFriVivvOmJSRlRKGyRPVfSSP





CONSENSUS_H2A	iGRGILlavPRRIRROGAEIALL	792
H2A.-.CBL21	---I-----L--\$	865
H2A.-.CBL22	---E-----L--\$	850
H2A.-.CBL23	-----L--\$	858
H2A.CI.UC2	-----L--\$	865
H2A.DE.BEN	-----L--\$	860
H2A.DE.FE12	V-----L--\$	857
H2A.GH.GH1	-----L--\$	852
H2A.GM.CBL24	-----L--\$	866
H2A.GM.ISY	-----L--\$	846
H2A.GM.D194	-----L--\$	851
H2A.GW.CAM1	-----L--\$	858
H2A.GW.CAM3	F---SI-----L--\$	860
H2A.GW.CAM2CG	-----L--\$	859
H2A.GW.CAM2ENV	---V-----L--\$	843
H2A.GW.CAM4	L-----L--\$	860
H2A.GW.CAM5	-----L--\$	849
H2A.GW.CAM6	-----L--\$	862
H2A.GW.FG	-----L--\$	856
H2A.GW.HIV-2ALI	V--R-----L--\$	862
H2A.GW.MDS	-----L--\$	869
H2A.SN.HIV2ST	-----L--\$	859
H2A.SN.ROD	-----L--\$	858
CONSENSUS_H2B	aAgaiIaIPRRIROGLLeLlL	461
H2B.-.GH-2	M---V-----R--\$	852
H2B.CI.BHO	---E-----A--\$	857
H2B.CI.UC1	---E-----A--\$	859
H2B.GH.D205	--R-----T--\$	854
H2G.CI.ABT96	VGRGILaIPRRIRROGLELTLlL\$	525
CONSENSUS_MAC	GGRWILaIPRRIRROGLElTLlL	879
MAC.-.MM239	-----L--\$	876
MAC.-.MM1322L	-----x	877
MAC.-.MM3162Q	-----L--\$	881
MAC.US.MM142	-----L--\$	880
MAC.US.MM251	-----L--\$	880
MNE.US.MNE	VGRWILaIPRRIRROGLELTLlL	857
CONSENSUS_SMM	vGRgILaIPrIRROGLELTLlL	887
SMM.-.P209CL15	-----V-----L--\$	886
SMM.-.SMP209	-----V-----A--\$	878
SMM.US.MM32H	G--W-----L--\$	886
SMM.US.SIU04982	-----F--\$	886
SMM.US.SIVSMMH9	-----xx-----L--\$	883
SMM.US.SIVSMMB670	---WV-----A--\$	885
SMM.US.SMMF236	-----L--\$	889
SMM.US.SMMPBJ14-15	-----L--\$	882
STM.US.STM	VGRrICaIPRRIRROGLELTLlL	

```

CONSENSUS_H2A                                <-| env cds                                ITR -->
H2A.CI.UC2                                   MGASGSKrS?psqgLRERLLrargg?cggq???sgEysqsgsgreQkspsCEGq?YgQgdfmmtPwrLPA?egek?lVkQQNMDDvDsdddd?lvqvpvtp.rvpLR?MTYkLAVdMshfIkeKgg
H2A.DE.PE12                                  L-KQ-----GYVR-CSA--S--G-----N-Q-ES-----N-Q-ES-----P-----P-----
H2A.DE.BEN                                  C-RSL-----ET--WDG-A--L-F--G-NL-----AGR-GT-----A-N-N--I-----A-----I-----LN--
H2A.GH.GH1                                   L-KH-R-----DGy-K-RDA--F-E-N--T-RO-D-R-----I-----R-E-----I-----A-----X--X--
H2A.GM.D194                                   H-KH-R-----H--GYVK-CMA--KG-----I-O-K-----I-S-----M-----A-----I-----
H2A.GM.D194                                   EH-----G-VVK-RVA--S-----A-I-Q-NS-----L-----M-----A-----I-----
H2A.GM.I5Y                                   R-R-Q-----QTP-ET--CSG--G-----L--L-----G--L-----R-----S-----SD-S-----A-----RM-----DL--D-----
H2A.GM.A1I                                   G-L-----A-T--ECYNALe--SLR--D--N-L-----R-----R-----E-----Q-----I-----L-----R-----
H2A.GM.CAM2CG                                K-L-Q-----Q-----ET--RCNE--G-L-H-----N-----R-----A-----A--K--NA-R-----I-----A-----P-----P-----
H2A.GM.FG                                    R-R-Q-----R-----A-A--HWDEL--R-----DKG-R-----Q-----T-----DS-R-----L-----Q--F-----P-----F-----
H2A.GM.MDS                                   E-R-----R-----QTP-EAS--HWDKL--L--G-----R-----R-----A-----A-----G-----E-----R-----R-----L-----
H2A.SN.HIV2ST                                H-R-PR-Q-----A-A--YWNb--RF--D-----D-----R-----Q-----KD--A--R--N--R-----K-----P-----HR--I-----L--TR--
H2A.SN.ROD                                   H-R-PR-Q-----A-A--YWNb--RF--D-----D-----R-----Q-----KD--A--R--N--R-----K-----P-----HR--I-----L--TR--

CONSENSUS_H2B                                <-| env cds                                ITR -->
H2B.CI.EHO                                   MGASGSKrSkqQ?GLRErLLR??e?PyG?lSgerregSrgsGgSKdLNspSCsGma?gAE.....GGGqQD?DesD.EdnE.?vgyVrPp??vPLRpMTYkLAVdMshfIkeKgg
H2B.CI.UC1                                   p-----ARRG-R-ES--O-R-LQY-----G-----D--KTL-----K--S--D-----R-----R-----G-----F-----L-----E-----
H2B.GH.D205                                  R-----R-----TQ-E--K--EGQ-K-----R-----R-----PR-----T-D-----NR-----S-----M--I-----
-----R-ER-Q--K--VP-R-R-----E-----E-----R-----R-----A-----A-----D--.VGAX--xI-----Q-----

H2D.LR.F0784                                  H2D.LR.F0784                                  xEKGG

H2G.CI.ABT96                                  MGASGKKRAPQQLGLRKRLLQARGEYgkLWEGLLEEGYSQYREEsgkGLSSLSCEGQYTOGQFMNTPWRNPATERAKLAYR:QNDDVDSDDND.LVGVFVYP.KVPLRVMSYKLAIDMSHFIKEKGG
MAC.US.MAC239                                MGGAISmrRSkp?gDLRQrLLRARGEYgRLLIgeVEdgsQsIGLgldKGLSSLSCEGQKYNQgqYMNTPWRNPAEErEKLaYRKQNmDDIDEEDDD.LVGvsVrp.kVPLRtMsYKLAIDMSHFIKEKGG
MAC.US.MAC251-1A11                           T-----R--S-----RST-----E-----E-----E-----E-----E-----S-----$-----
MAC.US.MM142                                  KK-----PR-----N-----FKG-----E-----K-----P-----I-----IP-EA.R-----A-----T-----
MAC.US.MM251                                  A-----K-----A-----K-----R-----R-----K-----K-----K-----K-----A-----T-----
MAC.US.MMP11                                  -----T-----K-----K-----G-----R-----R-----G-----R-----K-----K-----K-----A-----T-----

MNE.US.MNE                                   MGATSKRRSKSPGDLRQQLRARGEYGRLEWELLEDGYSQSLGSDKSSLLSCEGQKYNQGFNTPWKNPAGEREKLAYRKQNIIDIDEEDND.LVGVFVYP.RVPLRIISYKLAIDMSHFIKEKGG

CONSENSUS_SMM                                <-| env cds                                ITR -->
SMM.-.ENFW25                                 mgg?skkq?rYgGnlrERLLqARGeTYGrLwEglF?gYsQSlgasKglsslLscEbpQkY?eQ?MNTPWRNPAAErAKL?YRQNMDDVD?eddd.LvgG?V7P.rVpLR?MtykLAIIDMSHFIKEKGG
SMM.-.NEFW61                                  .....xRRRSKSA-D-Q-R-----LGEV-D-----GLD-----G--NO--Y-----G--NO--Y-----E--E--A--K-----E-----V--M-----T--S-----
SMM.-.SMMPB7E                                 .....x--D--Q-R-----x-LGEV-D-----GLD-----G--NO--Y-----G--NO--Y-----E--E--A--K-----E-----F--M-----T--S-----V-----
SMM.US.MM32H                                  AI-RRRSKSA-D-Q-R-----LGEV-D-----GLD-----G--NO--Y-----G--NO--Y-----E--E--A--K-----E-----V--M-----T--S-----V-----
SMM.US.SIU04982                               VT--RxA--x-----R-----R-----E-----E-----E-----E-----E-----E-----E-----E-----E-----E-----E-----E-----
SMM.US.SIVSMMH9                              VT--RR--Y-----Y-----GE--OD-----OD-----C--F-----G-----G-----NE--C--S-----C--S-----V--I-----
SMM.US.SMMF236                              AI--YK-----Y-----Y-----GE--OD-----OD-----C--F-----G-----G-----SA--C--S-----C--S-----V--I-----
SMM.US.SMPEB14-15                            VT--RR--Y-----Y-----GE--OD-----OD-----D--L--R--GLD--DWNLH--S--G--G--S--F-----DN--E-----A--H-----A--S-----
SMM.US.SIVSMM670                            --AAG--RSKQD-G-----D--L--R--GLD--DWNLH--S--G--G--S--F-----DN--E-----A--H-----A--S-----

STM.US.STM                                   MGASGKKQRKHGELRELLRARGEYgkLLEGLGEGSPGQASDKGLNshSCEPQRVNEGQFMNTPWKNPAAESAKLEYRQNMDDVDEEDND.LVGVAVHP.RVPLREMTYKLAIDLSHFIKSKGG
    
```

