

V

HIV-2/SIVsmm proteins

| | |
|--|-----|
| Introduction | 641 |
| Table of HIV-2/SIVsmm protein alignments | 642 |
| Gag | 644 |
| Pol | 649 |
| Vif | 658 |
| Vpx | 660 |
| Vpr | 661 |
| Tat | 662 |
| Rev | 663 |
| Env | 664 |
| Nef | 672 |

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, SIVcpzs are the most related to HIV-1, so they are 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, SIVl'hoest, 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.

| Name | Accession | Region | Author | Reference |
|-----------------|-----------|--|-------------------------|---|
| H2A.-.CBL21 | U05350 | ENV | Breuer, J | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.-.CBL22 | U05351 | ENV | Breuer, J | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.-.CBL23 | U05352 | ENV | Breuer, J | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.-.HIV2JAU1 | L28935 | VIF | Jacinto, A | Unpublished (1994) |
| H2A.CI.UC2 | U38293 | ENV GAG NEF POL REV TAT VIF VPX | Barnett, SW | <i>Virology</i> 222 (1):257–261 (1996) |
| H2A.DE.BEN | M30502 | ENV GAG NEF POL REV TAT VIF VPR VPX | Kirchhoff, F | <i>Virology</i> 177 (1):305–311 (1990) |
| H2A.DE.PEI2 | U22047 | ENV GAG NEF POL REV TAT VIF VPR VPX | Talbott, R | <i>Proc Natl Acad Sci USA</i> 90 (9):4226–4230 (1993) |
| H2A.FR.96151 | AF170034 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2A.FR.96199 | AF170043 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2A.FR.96201 | AF170035 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2A.FR.96202 | AF170040 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2A.FR.96205 | AF170031 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2A.FR.96323 | AF170032 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2A.FR.96329 | AF170045 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2A.GH.GH1 | M30895 | ENV GAG NEF POL REV TAT VIF VPR VPX | Hasegawa, A | <i>ARHR</i> 5 :593–604 (1989) |
| H2A.GM.CBL24 | U05353 | ENV | Breuer, J | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.GM.D194 | J04542 | ENV GAG NEF POL REV TAT VIF VPR VPX | Kuehnel, H | <i>Proc Natl Acad Sci USA</i> 86 (7):2383–2387 (1989) |
| H2A.GM.ISY | J04498 | ENV GAG NEF POL REV TAT VIF VPR VPX | Franchini, G | <i>Proc Natl Acad Sci USA</i> 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 | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.GW.CAM2CG | D00835 | ENV GAG NEF POL REV TAT VIF VPR VPX | Tristem, M | <i>J Gen Virol</i> 72 :721–724 (1991) |
| H2A.GW.CAM3 | U05355 | ENV | Breuer, J | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.GW.CAM4 | U05356 | ENV | Breuer, J | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.GW.CAM5 | U05357 | ENV | Breuer, J | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.GW.CAM6 | U05358 | ENV | Breuer, J | <i>J Gen Virol</i> 76 (Pt 2):333–45 (1995) |
| H2A.GW.FG | J03654 | ENV GAG NEF POL REV TAT VIF VPR VPX | Zagury, JF | <i>Proc Natl Acad Sci USA</i> 85 (16):5941–5945 (1988) |
| H2A.GW.HIV-2ALI | L25445 | ENV | Costa Taveira, N | <i>ARHR</i> 10 :223–224 (1994) |
| H2A.GW.MDS | Z48731 | ENV GAG NEF POL REV TAT VIF VPR VPX | Becker, M | Unpublished (1995) |
| H2A.SN.A1871 | U81840 | VPR | Stivahtis, GL | <i>J Virol</i> 71 (6):4331–4338 (1997) |
| H2A.SN.A2057 | U81842 | VPR | Stivahtis, GL | <i>J Virol</i> 71 (6):4331–4338 (1997) |
| H2A.SN.A2057 | U81849 | VPX | Stivahtis, GL | <i>J Virol</i> 71 (6):4331–4338 (1997) |
| H2A.SN.A640 | U81838 | VPR | Stivahtis, GL | <i>J Virol</i> 71 (6):4331–4338 (1997) |
| H2A.SN.A640 | U81845 | VPX | Stivahtis, GL | <i>J Virol</i> 71 (6):4331–4338 (1997) |
| H2A.SN.H2A7924A | U81836 | VPR | Stivahtis, GL | <i>J Virol</i> 71 (6):4331–4338 (1997) |
| H2A.SN.HIV2ST | M31113 | ENV GAG NEF POL REV TAT VIF VPR VPX | Kumar, P | <i>J Virol</i> 64 (2):890–901 (1990) |
| H2A.SN.ROD | M15390 | ENV GAG NEF POL REV TAT VIF VPR VPX | Clavel, F | <i>Nature</i> 324 (6098):691–695 (1986) |
| H2B.-.GH-2 | D10458 | ENV | Kawamura, M | <i>Virology</i> 188 :850–853 (1992) |
| H2B.CI.EHO | U27200 | ENV GAG NEF POL REV TAT VIF VPR VPX | Rey-Cuille, MA | <i>Virology</i> 202 (1):471–476 (1994) |
| H2B.CI.UC1 | L07625 | ENV GAG NEF POL REV TAT VIF VPR VPX | Castro, BA | <i>Virology</i> 178 (2):527–34 (1990) |
| H2B.FR.96200 | AF170057 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2B.FR.96307 | AF170059 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2B.FR.96339 | AF170053 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |

| | | | | |
|--------------------|----------|--|----------------|--|
| H2B.FR.97227 | AF170052 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2B.FR.97244 | AF170056 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2B.FR.97378 | AF170055 | ENV | Damond, F | <i>Virology</i> 280 (1):19–30 (2001) |
| H2B.GH.D205 | X61240 | ENV GAG NEF POL REV TAT VIF VPR VPX | Kreutz, R | <i>ARHR</i> 8 (9):1619–1629 (1992) |
| H2C.LR.22381G | L33077 | GAG | Gao, F | <i>J Virol</i> 68 :7433–7447 (1994) |
| H2C.LR.2238POLB7 | M87138 | POL | Gao, F | <i>Nature</i> 358 :495–499 (1992) |
| H2C.LR.HIV22238E | M87118 | ENV | Gao, F | <i>Nature</i> 358 :495–499 (1992) |
| H2D.LR.FO784 | L33083 | GAG | Gao, F | <i>J Virol</i> 68 :7433–7447 (1994) |
| H2D.LR.FO784 | M87069 | ENV | Gao, F | <i>Nature</i> 358 :495–499 (1992) |
| H2D.LR.FO784 | M87115 | NEF | Gao, F | <i>Nature</i> 358 :495–499 (1992) |
| H2D.LR.FO784PA | M87110 | POL | Gao, F | <i>Nature</i> 358 :495–499 (1992) |
| H2D.LR.FORTC2 | M87111 | POL | Gao, F | <i>Nature</i> 358 :495–499 (1992) |
| H2E.SL.PA | L33087 | GAG | Gao, F | <i>J Virol</i> 68 :7433–7447 (1994) |
| H2E.SL.PA | L33093 | POL | Gao, F | <i>J Virol</i> 68 :7433–7447 (1994) |
| H2G.CI.ABT96 | AF208027 | ENV GAG NEF POL REV TAT VIF VPR VPX | Brennan, CA | <i>ARHR</i> 13 (5):401–404 (1997) |
| MAC.-.MM132ZL | L22809 | ENV | Kodama, T | <i>J Virol</i> 67 :6522–6534 (1993) |
| MAC.US.MM142 | M16403 | ENV GAG NEF POL REV TAT VIF VPR VPX | Chakrabarti, L | <i>Nature</i> 328 (6130):543–547 (1987) |
| MAC.US.MM239 | M33262 | ENV GAG NEF POL REV TAT VIF VPR VPX | Kestler, H | <i>Science</i> 248 :1109–1112 (1990) |
| MAC.US.MM251 | M19499 | ENV GAG NEF POL REV TAT VIF VPR VPX | Franchini, G | <i>Nature</i> 328 (6130):539–543 (1987) |
| MNE.-.MNE170 | AF361745 | POL | Diamond, TL | Unpublished |
| MNE.-.SIU06277 | U06277 | ENV | Chackerian, B | <i>J Virol</i> 68 (6):4080–4085 (1994) |
| MNE.-.SIU06352 | U06352 | ENV | Chackerian, B | <i>J Virol</i> 68 (6):4080–4085 (1994) |
| MNE.-.SIU06377 | U06377 | ENV | Chackerian, B | <i>J Virol</i> 68 (6):4080–4085 (1994) |
| MNE.US.MNE | M32741 | ENV GAG NEF POL REV TAT VIF VPR VPX | Benveniste, RE | Unpublished (1990) |
| SMM.-.ENVW25 | X86727 | NEF | Whatmore, AM | <i>J Virol</i> 69 (8):5117–5123 (1995) |
| SMM.-.NEFW61 | X90853 | NEF | Whatmore, AM | <i>J Virol</i> 69 (8):5117–23 (1995) |
| SMM.-.P209C15 | L20009 | ENV | Hynes, NA | <i>ARHR</i> 9 :803–806 (1993) |
| SMM.-.SMP209 | L20008 | ENV | Hynes, NA | <i>ARHR</i> 9 :803–806 (1993) |
| SMM.LR.SMMLIB1 | M62651 | POL | Marx, PA | <i>J Virol</i> 65 :4480–4485 (1991) |
| SMM.SL.SIU17646 | U17646 | GAG | Chen, Z | <i>J Med Primatol</i> 24 (3):108–115 (1995) |
| SMM.SL.SIVsmSL92f | U48815 | GAG | Chen, Z | <i>J Virol</i> 70 (6):3617–3627 (1996) |
| SMM.SL.SMMSL92B | AF334679 | ENV GAG NEF POL REV TAT VIF VPR VPX | Chen, Z | <i>J Virol</i> 70 (6):3617–3627 (1996) |
| SMM.US.MM32H | D01065 | ENV GAG NEF POL REV TAT VIF VPR VPX | Rud, EW | <i>J Gen Virol</i> 75 (Pt 3):529–543 (1994) |
| SMM.US.SIU04982 | U04982 | NEF | Hirsch, VM | <i>J Virol</i> 68 :2649–2661 (1994) |
| SMM.US.SIVSMMB670 | M90048 | ENV NEF | Mullins, JI | Unpublished (1992) |
| SMM.US.SIVSMMH9 | M80194 | GAG NEF VPR | Courgnaud, V | <i>J Virol</i> 66 :414–419 (1992) |
| SMM.US.SMM17EC1 | AY033233 | ENV | Anderson, MG | <i>Virology</i> 195 (2):616–626 (1993) |
| SMM.US.SMMF236 | X14307 | ENV GAG NEF POL REV TAT VIF VPR VPX | Hirsch, VM | Unpublished |
| SMM.US.SMMM7 | M27256 | POL | Li, Y | Unpublished (1989) |
| SMM.US.SMMPBJ14-15 | L03295 | ENV NEF POL VIF VPR VPX | Dewhurst, S | <i>Nature</i> 345 :636–640 (1990) |
| SMM.US.SMMPBJA | M31325 | GAG REV TAT | Dewhurst, S | <i>Nature</i> 345 :636–640 (1990) |
| STM.US.STM | M83293 | ENV GAG NEF POL REV TAT VIF VPR VPX | Novembre, FJ | <i>Virology</i> 186 (2):783–787 (1992) |
| STM.US.STMAK2 | X60667 | GAG | Khan, AS | <i>J Virol</i> 65 (12):7061–7065 (1991) |

HIV-2/SIVsmm proteins

| | Gag p24 Capsid \ / x peptide / p9 NC | |
|-------------------|---|-----|
| MAC.US.MM239 | PIPVGNIYRRWTLGLQKCVRMVNPNTIILDYKQGFEPFQSVVDFYKSLRAEQTDAAVKVNWMTQTLIIQNANPDCKLVKGLGVNPTLLEEMLTACQGVGGQKARLMAEALKEALAPVIFPFAAAQQR | 385 |
| CONSENSUS_H2A | -V-----I-----I-----P-----m-----nt-a----- | 382 |
| H2A.CI.UC2 | -V-----I-----V-----A-P-----M-----T-P----- | 385 |
| H2A.DE.BEN | -V-----I-----K-----I-----M-----MG-S----- | 385 |
| H2A.DE.PEI2 | -V-----I-----I-----P-----M-----A----- | 385 |
| H2A.GH.GH1 | -V-----I-----I-----P-----M-----T-P----- | 386 |
| H2A.GM.D194 | -V-----I-----S-----P-----M-----S-----T-A----- | 385 |
| H2A.GM.ISY | -V-----I-----S-----P-----M-----V-S-----MR-A----- | 384 |
| H2A.GM.ALI | -V-----I-----I-----P-----M-----MT-A----- | 385 |
| H2A.GM.CAM2CG | -V-----I-----I-----P-----I-----V-----MG-P----- | 385 |
| H2A.GM.FG | -V-----I-----IN-----P-----M-----T-----VM-A----- | 383 |
| H2A.GM.MDS | -V-----I-----I-----P-----M-----M-----MT-A----- | 385 |
| H2A.SN.HIV2ST | -V-----I-----S-----P-----I-----M-----M-----A----- | 385 |
| H2A.SN.ROD | -V-----I-----P-----P-----V-----M-----VIG-A----- | 385 |
| H2B.CI.EHO | -V-----I-----I-----P-----M-----I-----T-STN-----P----- | 380 |
| H2B.CI.UC1 | -V-----I-----I-----P-----M-----I-----T-A----- | 380 |
| H2B.GH.D205 | -V-----I-----I-----P-----M-----I-----T-A-----V----- | 380 |
| H2C.IR.22381G | -V-----I-----I-----P-----M-----I-----T-A-----V----- | 261 |
| H2D.IR.FO784 | -----I-----T----- | 261 |
| H2E.SL.PA | -----D-----S----- | 261 |
| H2G.CI.ABT96 | -----X-----X-----X-----M-----N-TAL-----K----- | 386 |
| MAC.US.MM142 | -----I-----I-----P-----P-----T-----A-----K----- | 385 |
| MAC.US.MM251 | -----I-----I-----P-----P-----K-----K----- | 385 |
| MNE.US.MNE | -----S-----P-----P-----M-----G-L-----K----- | 385 |
| CONSENSUS_SMM | -----I-----P-----P-----P-----d-t-t?-l-----v-k----- | 382 |
| SMM.SL.SIU17646 | -V-----I-----V-----I-----P-----X----- | 276 |
| SMM.SL.SIVsmSL92f | -V-----I-----IS-----K-----V-----I-----G-----X----- | 273 |
| SMM.SL.SMMSL92B | -V-----D-----P-----P-----M-----M-D-----TGSLVAAOFRGAA | 386 |
| SMM.US.MM32H | -----X-----I-----S-----X-----X-----D-----TOG-L-X-V--K----- | 385 |
| SMM.US.SIVSMH9 | -----X-----I-----S-----P-----X-----M-----R-DL-----V--K----- | 386 |
| SMM.US.SMMP236 | -----I-----I-----P-----P-----I-----D-----TOG-L-----V--K----- | 386 |
| SMM.US.SMMPB5A | -----V-----V-----I-----A-P-----M-----FO-G-L-----Q----- | 386 |
| STM.US.STMAK2 | -----V-----V-----I-----T-----A-PS-----R-----VFQ-D-L-----Q----- | 387 |

| | | | | |
|-----------------|-----------------------------|------|------------|-----|
| MAC.US.MM239 | VTEDLLHL..... | p6 \ | NSLFGGDQ\$ | 510 |
| CONSENSUS_H2A | -----EggetPhret?TEDLLHL | | -----k | 517 |
| H2A.CI.UC2 | -----EQGETPRREA.TEDLLHL | | -----K | 521 |
| H2A.DE.BEN | -----EQGETPRREA.TEDLLHL | | -----K | 521 |
| H2A.DE.PEI2 | -----EQGETPRREA.TEDLLHL | | -----K | 521 |
| H2A.GH.GH1 | -----EQGETPRREA.TEDLLHL | | -----K | 521 |
| H2A.GH.D194 | -----EQGETPRREA.TEDLLHL | | -----K | 521 |
| H2A.GH.ISY | -----EQGETPRREA.TEDLLHL | | -----N | 521 |
| H2A.GW.ALI | -----EQGETPRREA.TEDLLHL | | -----K | 521 |
| H2A.GW.CAM2CG | -----EQGETPRREA.TEDLLHL | | -----K | 521 |
| H2A.GW.FG | -----RFQAEPTCRET.TEDLLHL | | -----K | 519 |
| H2A.GW.MDS | -----RFQAEPTCRET.TEDLLHL | | -----T | 521 |
| H2A.SN.HIV2ST | -----F-Q-EKQETPCRET.TEDLLHL | | -----K | 521 |
| H2A.SN.ROD | -----EQGETPRREPPTEDLLHL | | -----K | 522 |
| H2B.CI.EHO | -----EQGETPRREPPTEDLLHL | | -----E | 519 |
| H2B.CI.UC1 | -----EQGETPRREPPTEDLLHL | | -----E | 521 |
| H2B.GH.D205 | -----EQGETPRREPPTEDLLHL | | -----E | 521 |
| H2G.CI.ABT96 | --XX-----S-----D | | | 509 |
| MAC.US.MM142 | ----- | | | 506 |
| MAC.US.MM251 | ----- | | | 506 |
| MNE.US.MNE | ----- | | | 506 |
| CONSENSUS_SMM | ----- | | | 498 |
| SMM.SI.SMMSL92B | ----- | | -----e | 507 |
| SMM.US.MM32H | ----- | | -----E | 510 |
| SMM.US.SIVSMMH9 | ----- | | -----E | 507 |
| SMM.US.SMMPZ36 | ----- | | -----E | 507 |
| SMM.US.SMMPBJA | ----- | | -----E | 507 |
| STM.US.STMAK2 | -----V----- | | | 510 |
| STM.US.STM | ----- | | -----E | 507 |

Pol p15 RNase H \ Pol p31 Integrase
MAC.US.NM239 VLFLEK.IEPAQEEHDKYHSNVKELVFKFGLPRIVARQIVDTCDKCHKGEAIGHQANSDLGTWQMDCTHLEKGLIIVAVHVASGFIEAEVIFQETGRQTALFLLKLAGRWPITHLHTDNGANFASQEVK
CONSENSUS_H2A
H2A.CI.UC2
H2A.DE.BEN
H2A.DE.PE12
H2A.GH.GH1
H2A.GM.D194
H2A.GM.ISY
H2A.GM.ALI
H2A.GM.CAM2CG
H2A.GM.FG
H2A.GM.MDS
H2A.SN.HIV2ST
H2A.SN.ROD
H2B.CI.EHO
H2B.CI.UC1
H2B.GH.D205
H2C.LR.2238POLB7
H2D.LR.FO784PA
H2D.LR.FORTC2
H2G.CI.ABT96
MAC.US.NM142
MAC.US.NM251
MNE.US.MNE
CONSENSUS_SMM
SMM.LR.SMMLIB1
SMM.SJ.SMMSLJ92B
SMM.US.NM32H
SMM.US.SMMF236
SMM.US.SMMW7
SMM.US.SMMPBJ14.15
STM.US.STM

894
872
889
889
889
889
907
869
866
868
869
889
889
870
889
890
890
186
186
658
949
890
890
890
811
281
871
894
854
282
888
875

.....e.....sH--I-qt--N--aq-Q--V--AE--S-----S-----T-----
.....E--I--IH--I--OL--N--AQ-Q--V--AEI--V--S-----S-----P--T-----
.....E--II--TH--I--LL--NS-AQ-Q--V--AEI--V--Y-----S-----P--T-----
.....E--E--SH--K--I--N--AQ-Q--VDAE--S-----S-----T-----
.....R--E--M--TH--I--OL--N--AQ-Q--V--AEI--V--S-----S-----T-----
.....E--I--TH--I--OL--N--AQ-Q--V--AEI--V--S-----S-----P--T-----
.....R--E--AH--NL--N--AQ-Q--V--AE--S-----S-----T-----
.....E--T--SH--NL--N--AQ-Q--V--AE--S-----S-----S--V--T-----
.....E--T--CH--DI--QL--N--AQ-Q--V--AEV--S-----S-----T-----
.....E--I--SH--I--KL--N--AHVQ--V--AE--S-----S-----T-----
.....E--E--SH--QL--N--AQ-Q--V--A-----S-----T-----
.....E--E--SH--KL--N--TQ-Q--V--AE--S-----S-----T-----
.....E--E--SH--I--NL--NS-AQ-Q--V--AE--S-----S-----T-----
I-----E--N--H--I--QL--NS--Q-----V--E-----S-----T--D-----
.....E--G--H--QL--K--NS--Q-----V--V--AE--S-----T--D-----
.....E--G--H--I--QL--K--NS--Q-----V--V--A-----T--PS-----
.....-ERF--A--TH--I--NL--K--N--HV--Q-----V--A-----A-----T-----
.....D--E--I--CH--I--KL--I-----R-----R-----V--AE--S-----T-----
.....D--E--I--KL--I-----R-----R-----V--AE--VT-----X-----T-----
---X---EXP--TH--I--QL--K--NSYHS--Q-----V--V-----V-----T-----
.....S--I-----L--K-----V-----V-----S-----T-----
.....L--L--K-----V-----V-----S-----T-----
.....L--L--K-----V-----V-----S-----T-----
.....?--e-----I--Y--I--L--K-----V--AE-----S-----T-----
.....S--D--AQ--YNI--QL--K--NA--N--Q-----T--AEV-----R-----I--S--K-----T-----
---K--X---E-----L--K-----V--AE-----S-----T-----
.....E-----R--L--K-----Q-----R--V--TE--RI--I-----V--V-----S-----T-----
.....E-----L--K-----V--AE-----S-----T-----
.....E-----I--L--K-----V--AE-----S-----V-----T-----

| | | |
|--------------------|---|------|
| MAC.US.MM239 | MYAWWAGIEHTFVPYFQSQGVVEAMNHHKKNQIDRIRIQANSVETIVLMAVHCMNFKRGGIGDMTPAERLINWITTEQEIQQQSKNSKFKNFRVYREGRDQLWKGPGELLWKGEGAVLLKYGTDI | 1024 |
| CONSENSUS_H2A | -----s-----T-----L-a-l-----f-----d-----V----- | 1002 |
| H2A.CI.UC2 | -----S-----T-----L-R-N-K-Q-----V-----A----- | 1019 |
| H2A.DE.BEN | -----S-----T-----L-R-N-K-Q-----V-----V----- | 1019 |
| H2A.DE.PE12 | -----T-----S-----LHA-L-----F-----D-----V----- | 1019 |
| H2A.GH.GH1 | -----I-----S-----L-R-N-K-Q-----D-----V-----A----- | 1019 |
| H2A.GM.D194 | -----I-----S-----L-R-N-K-Q-----D-----V-----A----- | 1037 |
| H2A.GM.ISY | -----I-----S-----L-R-N-K-Q-----D-----V-----A----- | 989 |
| H2A.GM.ALI | -----V-----V-----L-A-L-----P-----Q-----V----- | 996 |
| H2A.GM.CAM2CG | -----S-----S-----L-T-L-----P-----D-----V----- | 998 |
| H2A.GM.FG | -----S-----S-----L-A-L-----F-----D-----V----- | 999 |
| H2A.GM.MDS | -----I-----S-----L-A-L-----F-----D-----V-----E----- | 1019 |
| H2A.SN.HIV2ST | -----I-----S-----L-A-L-----F-----D-----V----- | 1019 |
| H2A.SN.ROD | -----I-----S-----V-A-----LQ-----D-----V-----A----- | 1000 |
| H2B.CI.EHO | -A-I-Q-----E-----D-V-I-V-----T-----L-T-L-Q-----D-----I-----E----- | 1019 |
| H2B.CI.UC1 | -A-I-Q-----D-V-I-V-----A-----F-A-L-Q-Q-----L-I-----E----- | 1020 |
| H2B.GH.D205 | -----V-----Q-----L-D-V-I-V-----T-----F-A-L-Q-Q-----I-----E----- | 237 |
| H2C.LR.2238POLB7 | -----V-----V-----TM-----X----- | 237 |
| H2D.LR.FO784PA | -----K-----K-----D-----I-----T-----I-----E----- | 979 |
| H2G.CI.ABT96 | -----S-----K-----K-----D-----I-----T-----I-----E----- | 1020 |
| MAC.US.MM142 | -----V-----QA-----S-----K-----K-----D-----I-----T-----I-----E----- | 1020 |
| MAC.US.MM251 | -----V-----Q-----K-----K-----D-----I-----T-----I-----E----- | 1020 |
| MNE.US.MNE | -----M-----L-----V-----V----- | 1020 |
| CONSENSUS_SMM | -----q-----L-----K-----D-----I-----T-----L-----V-----e----- | 940 |
| SMM.LR.SMML1B1 | -----N-Q-----T-----K-----D-----I-----T-----L-----V-----X----- | 397 |
| SMM.SI.SMMS192B | -----L-V-QS-----DL-----KN-----K-----E-----L-----A-----L-----T-----YLN-Q-----Q-----A-----VI-----E----- | 1001 |
| SMM.US.RM32H | -----I-----I-----I-----V----- | 1024 |
| SMM.US.SMMF236 | -----Q-----Q-----K-----R-----T-----E-----I-----F-----E----- | 984 |
| SMM.US.SMMW7 | -----S-----D-----K-----R-----T-----E-----I-----F-----E----- | 402 |
| SMM.US.SMMPBJ14.15 | -----Q-----Q-----T-----T-----IVS-----V-----V----- | 1018 |
| STM.US.STM | -----Q-----Q-----T-----T-----D-----T-----V-----K-----L-----V-----V----- | 1005 |

Vpx_CDS
MAC.US.MM239
CONSENSUS_H2A
H2A.DE.BEN
H2A.DE.PE12
H2A.GH.GH1
H2A.GM.D194
H2A.GM.ISY
H2A.GM.ALI
H2A.GM.CAM2CG
H2A.GM.FG
H2A.GM.MDS
H2A.SN.7924A
H2A.SN.A1871
H2A.SN.A2057
H2A.SN.A640
H2A.SN.HIV2ST
H2A.SN.ROD
H2B.CI.BHO
H2B.CI.UC1
H2B.GH.D205
H2G.CI.ABT96
MAC.US.MM142
MAC.US.MM251
MNE.US.MNE
CONSENSUS_SMM
SMM.SI.SMMSL92B
SMM.US.MM32H
SMM.US.SIVSMMH9
SMM.US.SMMF236
SMM.US.SMMPBJ14.15
STM.US.STM

M...EERPPENEGQPEWVVELELKEEALKHFDPRLLITALGNHIIYNRHGDTLEGAGELIRIQRALFMHFRGGCIHSRIGQPGGNGFLSALPPSRSMI.
-LEAPT-f--dgt-p--g--I-I-r-i--I-Y-t--I-V-t--a-g--t-x-----tp-g-q
-TEAPT-F--DGT-R-DLGS-D-I-T-R-I--R--Y-H--R-KT--V--A-NR-----
-TEAPA-F--D-T-P-G-G--IGI-R-R--T--Y-CA--S-R-NV--V--A-KI--TR-ET-F--TP-G-Q-\$
-TEAPT-F--DGT-R-LGGD-IRI-G-I--I--Y-HS--P--R--V-L-A-NR--S-TRRRT-PP-A-TP-G-
-TEAPT-F--DGT-R-LGST-I-T-K-I--C-I--Y--R--V--V-I-A-DR--K-TRREA-CP-A-TP-G-H
-TEAPA-F--DGT-P--G--I-R-I--R--M-I--GY-T--R-F--NA--T--A-G--R-R-----TP-N-Q
-TEAPT-F--AGM--HQGAR--I-R-I--R--I-RDI--G-A--R-R--V--T--A-N--TR-----TP-R-\$
-TEAPT-L--DGT-P--G--I-I-R-IE--R--I-R-IT--R--RS--R--T--A-G--TR-----TP-R-\$
-TEAPT-L--DRT-P--G--I-I-R-IT--R--I-R-IO--V--R-ID--KY-T--R--XK--V--T--A-G--TR-----TP-Q\$
-AEAPT-L--GT-P--G--I-I-R-IR--G--I-T-R-IR--K--V--T--X--G--TR-----TP-G-Q\$
-AEAPA-F--\$DGT-P--GE--I-I-R-V--N-S-I--Y--T--G--R--VS--V--T--A-G--TR-----TV-TP-G-\$
-AEAPT-F--DGT-P--G--IOI-R-IG--I-T-R-I-\$--I-T-R-I--R--V--T--A-G--TR-----TP-D-Y\$
-TEAPT-S--DRT-P--G--I-I-R-I--G--I-I-R-I--R--KV--T--A-G--TR-----TP-G-R
-AEAPT-L--VDGT-L--G--I-I-R-I--D--I-Q--N--F--N--KL--L--Q--\$
-AEAVP-I--DKN--D--I-Q--D--E--ED-M--I-Q--R--F--S--A--K--L--L--A-Q--\$
-AEAPAP-T--S--E--IG-I--I-Q--R--N--E--IG-I--I-Q--R--F--S--A--K--L--L--A-Q--\$
-AEAPAP-I--N--E--IG-I--I-Q--R--N--E--IG-I--I-Q--R--F--S--A--K--L--L--A-Q--\$
-A...I--D-A-----I-----I-----Y--D-----K-----L--D-RC--GNA--T-----GVF\$
.....K-----I-----S-----N-----T-----
.....D-----K-----T-----R-----
.....t-----d-a-----y-d-----i-----r-----s-----t-----gV-
.....HA--D-TMP-----IRD-----YV-DTY--I-----I-K--K--L--H--T-----
.....T--D-A--SIVSMMH9-----X--NX-----X--Y--D-----G--I-----R-----S--X-----T--X-GV-\$
.....A--D-A-----I-----N-----I-----D-----I-----S--A-----SR-----V-\$
.....T--D-A-----I-----N-----I-----D-----I-----R-----S-----T--X-GV-\$
.....TH--D-----I--Q--R-----S--Y--D-----K-----AT--T-GV-

HIV-2/SIVsmm proteins

| | | | | | |
|-----------------|--|----------------------|-----------------------|------------------------------------|-----|
| MAC.US.MM239 | Vpr_CDS \ | / Rev_CDS | exon \ | / exon | 130 |
| CONSENSUS_H2A | MEIPLREQNSLESNRRSSCISADASTPESANLGEIEILSQLRPLEACVNTCYCKKCCYHCQFCFLKKGKIGICYSRKRRTPKKAKANTSSASN.KPISNRRTRHCQPEKAKKETEKAVATAPGLGR | | | | |
| H2A.CI.UC2 | ---KaP-s---Y--P--T--g-va?Q-l-kg | ---w--Rkgr | t-?hs--D--S--t--gns | ---kQ-k-l--tt-e-dc | 126 |
| H2A.DE.BEN | ---KAP-S--K-Y--P--T--WEVAAQ-L-KQ | ---W-VKGG | ---RT-THPP-TPD | ---S-IO-GDSR-T-KQ--P-TP--TS | 130 |
| H2A.DE.PE12 | ---SKAP-S--M-C--P--T--R-VTAQ-L-KQ | ---R-S | ---R-T-TPSP-PD | ---S-T-GDS--T-EO-K-S-AT-V-TC--Q | 130 |
| H2A.GH.GH1 | ---H-KAP-S--Y--P--T--QVTAQ-L-KQ | ---W-RKGR | ---HS--D--S--T | ---GNS--T-KOTK-P-T-LE--R--Q\$ | 130 |
| H2A.GH.DL94 | ---K-P-S--G-Y--P--T--R-VTAQ-R-KQ | ---W-AKQR | ---R-T-THS--D--S--T | ---GDS--T-EO-K-T-TMV-TCS--H | 130 |
| H2A.GW.ISY | ---K-PGS--MPY--P--T--Q-VAVQ-L-KQ | ---DRKGR | ---S-T-THPPP-D--S--T | ---GDS--T-KQ-K-P-TI-VS-C--H | 130 |
| H2A.GW.CAM2CG | ---KAP-S--C--P--RT-QQ | ---DRKGR | ---S-T-HS-P-D--S--T | ---GNS--T-KQ-K-L-LTILEAD--SHIY | 134 |
| H2A.GW.FG | ---I-KAP--Y--P--HT-Q-VA-Q-L-RQ | ---W-DRKGR | ---HS--D--S--T | ---NS--BQ-K-L-TTLG-DC-P--SHIY | 130 |
| H2A.GW.MDS | ---KAP-G--G-Y--P--T--Q-AAQGLVSP-D--Y | ---W-RKGR | ---T-HP--D--S--T | ---NS--A-KQ-K-L-AT-E-D-- | 130 |
| H2A.SN.HIV2ST | ---KAP-S--K-C--PF-RT-Q-VA-Q-L-RQ | ---R-KGR | ---T-THP-PTPD | ---S-T-GNS--T-KQ-K-L-AT-E-DL-P--\$ | 130 |
| H2B.CI.EHO | ---I-K--S--N--SGH--ST--GV-N-QGLD-R | ---R...K-SS-R--TT | ---P-NRSL-A | ---GDS--T-KQ-KE--TTR--DL-P--SNTS | 133 |
| H2B.CI.UC1 | ---I-Q--S--K--S--P--ST--PVVN-QG-D-Q | ---W-DH...K-SS-R-VTA | ---D,ESL-AN-GDS | ---T-KQ-TK--TKGL-DL-P-- | 128 |
| H2B.GH.D205 | ---S-----K-CR-L--TF-ETVDA-GLEGTQA--Y | ---S-K | ---L--N--V--P-R | ---T--SF--ES--A--NX--K-K\$ | 112 |
| H2G.CI.ABT96 | ---K-----G--T--A-P-L--E | ---S--K-HR | ---P-I-L-K | ---A | 130 |
| MAC.US.MM142 | ---K-----r-h-s-v--? | ---E | ---D | ---K-EQ--A | 130 |
| MAC.US.MM251 | ---DA-I..... | ---T-AKP...-VK | ---PFDT--OS--SG-N | ---K-EQ--K-TE-D-DC--\$ | 117 |
| MNE.US.MNE | ---K-----CR-H-S--V-VP | ---H...-H | ---T-PPP--SL-T-A-NR | ---K-E--TE--ADL-- | 104 |
| CONSENSUS_SMM | ---K-----R-H-S--V-D--S-E | ---Q...-Q | ---T--D--SL-R-A-N | ---K-E--AE--DL--\$ | 131 |
| SMM.SI.SMMSL92B | ---K-----S--R--S-P--T--V-A--GL--QE--W | ---E-C-K-F | ---L--VT--NRSL-T-A-NS | ---K-EQ-KE--TE-EST--K | 128 |
| SMM.US.MM32H | | | | | 128 |
| SMM.US.SMMPZ36 | | | | | 128 |
| SMM.US.SMMPBJA | | | | | 129 |
| STM.US.STM | | | | | 129 |
| H2A.GW.ALI | IS\$ | | | | 136 |
| H2B.CI.EHO | TSRFAN\$ | | | | 139 |

Table with columns for protein identifiers (e.g., MAC.US.MM239, CONSENSUS_H2A), alignment characters (e.g., PNASITPKWNETWQEWKRVDFLEENITALLLEA...), and residue numbers (739, 707, 726, etc.).

| | | |
|--------------------|-------------|-----|
| MAC.US.MM239 | LNWADKKETR | 262 |
| CONSENSUS_H2A | PFS | 252 |
| H2A.CI.UC2 | PYSE\$ | 257 |
| H2A.DE.BEN | PYSE | 257 |
| H2A.DE.PEI2 | PFS\$ | 253 |
| H2A.GH.GH1 | PYS | 256 |
| H2A.GM.D194 | PYSE | 257 |
| H2A.GM.ISY | PFS | 261 |
| H2A.GW.ALI | PFSKMRN | 261 |
| H2A.GW.CAM2CG | PFS | 260 |
| H2A.GW.FG | PFS | 180 |
| H2A.GW.MDS | PFN\$ | 255 |
| H2A.SN.HIV2ST | PFS | 255 |
| H2A.SN.ROD | PFS | 256 |
| H2B.CI.EHO | PTE\$ | 238 |
| H2B.GH.DZ05 | PTD | 241 |
| H2D.IR.FO784 | YKT-----G\$ | 137 |
| H2G.CI.APT96 | PIE\$ | 256 |
| MAC.US.MM142 | -E---R---S | 263 |
| MNE.US.MNE | -K---R--S | 263 |
| CONSENSUS_SMM | -k-----s | 254 |
| SMM.-.ENW25 | -----s | 258 |
| SMM.-.NEFW61 | -----s | 251 |
| SMM.SL.SMMSL92B | P | 255 |
| SMM.US.MM32H | -----s | 263 |
| SMM.US.SIVSMMB670 | SK-----S\$ | 263 |
| SMM.US.SIVSMMH9 | -K---K-S\$ | 261 |
| SMM.US.SM62 | YK---R-S\$ | 263 |
| SMM.US.SMMF236 | XK-----S\$ | 263 |
| SMM.US.SMMPBJ14.15 | -K---K-S\$ | 261 |
| STM.US.STM | -K-----S | 264 |