

# VI

## HIV-2/SIV Proteins

### Contents

---

|        |                              |            |
|--------|------------------------------|------------|
| VI-1   | Introduction . . . . .       | <b>369</b> |
| VI-2   | Annotated features . . . . . | <b>370</b> |
| VI-3   | Sequences . . . . .          | <b>371</b> |
| VI-4   | Alignments . . . . .         | <b>374</b> |
| VI-4.1 | Gag . . . . .                | 374        |
| VI-4.2 | Pol . . . . .                | 378        |
| VI-4.3 | Vif . . . . .                | 385        |
| VI-4.4 | Vpx . . . . .                | 387        |
| VI-4.5 | Vpr . . . . .                | 388        |
| VI-4.6 | Tat . . . . .                | 389        |
| VI-4.7 | Rev . . . . .                | 390        |
| VI-4.8 | Env . . . . .                | 392        |
| VI-4.9 | Nef . . . . .                | 398        |

---

### 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 recombinant 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, Stumptailed 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        | 374  |
| p15                         | Gag     | 135      | 374  |
| p27                         | Gag     | 136      | 374  |
| p27                         | Gag     | 364      | 376  |
| p2                          | Gag     | 365      | 376  |
| p2                          | Gag     | 381      | 376  |
| p8                          | Gag     | 382      | 376  |
| p8                          | Gag     | 433      | 376  |
| p1                          | Gag     | 434      | 376  |
| p1                          | Gag     | 447      | 376  |
| p6                          | Gag     | 448      | 376  |
| PTAP motif                  | Gag     | 458-461  | 376  |
| PSAP in HIV-2 B, U          | Gag     | 476-479  | 376  |
| p6 end                      | Gag     | 510      | 377  |
| Gag end                     | Gag     | 510      | 377  |
| Pol p15 start (-1 from Gag) | Pol     | 1        | 378  |
| p15                         | Pol     | 67       | 378  |
| protease                    | Pol     | 68       | 378  |
| protease                    | Pol     | 166      | 379  |
| p51 RT start                | Pol     | 167      | 379  |
| D catalytic site            | Pol     | 276      | 379  |
| DD catalytic site           | Pol     | 351      | 380  |
| p51 RT end                  | Pol     | 605      | 381  |
| p15 RNase H start           | Pol     | 606      | 381  |
| p15 RNase H end             | Pol     | 725      | 382  |
| p31 Integrase start         | Pol     | 726      | 382  |
| p31 Integrase end           | Pol     | 1018     | 384  |
| Pol end                     | Pol     | 1018     | 384  |
| Vif start                   | Vif     | 1        | 385  |
| Vif end                     | Vif     | 214      | 386  |
| Vpx start                   | Vpx     | 1        | 387  |
| Vpx end                     | Vpx     | 112      | 387  |
| Vpr start                   | Vpr     | 1        | 388  |
| Vpr end                     | Vpr     | 101      | 388  |
| Tat start                   | Tat     | 1        | 389  |
| exon 1 end                  | Tat     | 99       | 389  |
| exon 2 start                | Tat     | 100      | 389  |
| Tat end                     | Tat     | 130      | 389  |
| Rev start                   | Rev     | 1        | 390  |
| exon 1 end                  | Rev     | 24       | 390  |
| exon 2 start                | Rev     | 25       | 390  |
| Rev end                     | Rev     | 107      | 390  |
| Env start                   | Env     | 1        | 392  |
| signal peptide end          | Env     | 22       | 392  |
| gp120 start                 | Env     | 23       | 392  |
| V3                          | Env     | 311-344  | 393  |
| gp120 end                   | Env     | 525      | 395  |
| gp41 start                  | Env     | 526      | 395  |

| Feature                  | Protein | Location | Page |
|--------------------------|---------|----------|------|
| Env end                  | Env     | 879      | 397  |
| Nef start                | Nef     | 1        | 398  |
| R17Y mutation            | Nef     | 17       | 398  |
| max HIV-1 similarity     | Nef     | 153-182  | 398  |
| premature stop in Mac239 | Nef     | 93       | 398  |
| normal Nef end           | Nef     | 262      | 399  |

### 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> <b>248</b> (4959):1109-1112 (1990)             |
| A.CI.88.UC2               | U38293    | All                                    | Barnett, S.W.    | <i>Virology</i> <b>222</b> (1):257-61 (1996)                  |
| A.DE.x.BEN                | M30502    | All                                    | Kirchhoff, F.    | <i>Virology</i> <b>177</b> (1):305-311 (1990)                 |
| A.DE.x.PEI2               | U22047    | All                                    | Kraus, G.        | <i>ARHR</i> <b>14</b> (1):65-77 (1998)                        |
| A.GH.x.GH1                | M30895    | All                                    | Hasegawa, A.     | <i>ARHR</i> <b>5</b> (6):593-604 (1989)                       |
| A.GM.87.D194              | J04542    | All                                    | Kuehnel, H.      | <i>Proc Natl Acad Sci USA</i> <b>86</b> (7):2383-2387 (1989)  |
| A.GM.x.ISY                | J04498    | All                                    | Franchini, G.    | <i>Proc Natl Acad Sci USA</i> <b>86</b> (7):2433-2437 (1989)  |
| A.GM.x.MCN13              | AY509259  | All                                    | Schmitz, C.      | <i>J Virol</i> <b>78</b> (4):2006-2016 (2004)                 |
| A.GM.x.MCR35              | AY509260  | All                                    | Schmitz, C.      | <i>J Virol</i> <b>78</b> (4):2006-2016 (2004)                 |
| A.GW.86.FG                | J03654    | Gag, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Zagury, J.F.     | <i>Proc Natl Acad Sci USA</i> <b>85</b> (16):5941-5945 (1988) |
| A.GW.x.ALI                | AF082339  | Gag, Pol, Vif, Vpx, Tat, Rev, Env, Nef | Azevedo-Pereira  | Unpublished   |
| A.GW.x.MDS                | Z48731    | All                                    | Becker           | Unpublished   |
| A.IN.95.CRIK_147          | DQ307022  | All                                    | Santhosh, C.V.   | <i>ARHR</i> <b>24</b> (10):1315-1317 (2008)                   |
| A.IN.x.NNVA               | EU980602  | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env | Desai            | Unpublished   |
| A.SN.85.ROD               | M15390    | All                                    | Clavel, F.       | <i>Nature</i> <b>324</b> (6098):691-695 (1986)                |
| B.CI.88.UC1               | L07625    | All                                    | Barnett, S.W.    | <i>J Virol</i> <b>67</b> (2):1006-1014 (1993)                 |
| B.CI.09.pPRD320-20HIV2_56 | AB485670  | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env | Takekawa         | Unpublished   |
| B.CI.x.EHO                | U27200    | All                                    | Rey-Cuille, M.A. | <i>Virology</i> <b>202</b> (1):471-476 (1994)                 |
| B.GH.86.D205              | X61240    | All                                    | Dietrich, U.     | <i>Nature</i> <b>342</b> (6252):948-950 (1989)                |
| B.JP.01.KR020             | AB100245  | All                                    | Kusagawa, S.     | <i>ARHR</i> <b>19</b> (11):1045-1049 (2003)                   |
| G.CI.92.ABT96             | AF208027  | All                                    | Brennan, C.A.    | <i>ARHR</i> <b>13</b> (5):401-404 (1997)                      |
| AB.CI.90.7312A            | L36874    | All                                    | Gao              | Unpublished   |

| Name                  | Accession | Proteins                               | Author          | Reference  |
|-----------------------|-----------|--|-----------------|--|
| AB.CM.03.03CM_510_03  | EU028345  | Gag, Vif, Vpx, Tat, Rev, Env, Nef      | Yamaguchi, J.   | <i>ARHR</i> <b>24</b> (1):86-91 (2008)   |
| U.FR.96.12034         | AY530889  | All                                    | Damond, F.      | <i>ARHR</i> <b>20</b> (6):666-672 (2004)   |
| MAC.US.x.251_1A11     | M76764    | All                                    | Marthas, M.L.   | <i>J Med Primatol</i> <b>18</b> (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); <i>VACCINES 92: MODERN APPROACHES TO NEW VACCINES INCLUDING PREVENTION OF AIDS: 229-235</i> ; Cold Spring Harbor Laboratory Press, New York (1992) |
| MAC.US.x.251_BK28     | M19499    | Gag, Pol, Vif, Vpx, Vpr, Rev, Env, Nef | Hirsch, V.      | <i>Cell</i> <b>49</b> (3):307-319 (1987)   |
| MAC.US.x.MM142        | Y00277    | All                                    | Chakrabarti, L. | <i>Nature</i> <b>328</b> (6130):543-547 (1987)   |
| MAC.US.x.17EC1        | AY033233  | All                                    | Anderson, M.G.  | <i>Virology</i> <b>195</b> (2):616-626 (1993)  |
| MAC.US.x.17EFR        | AY033146  | All                                    | Flaherty, M.T.  | <i>J Virol</i> <b>71</b> (8):5790-5798 (1997)  |
| MAC.US.x.SMM142B      | BD131285  | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev      | Alizon, M.      | Patent: JP 2002030099-A 2 29-JAN-2002; INSTITUT PASTEUR  |
| MAC.US.x.r80025       | AY576480  | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Nef | O'Connor, D.H.  | <i>J Virol</i> <b>78</b> (24):14012-14022 (2004)   |
| MAC.US.x.r90131       | AY576481  | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Nef | O'Connor, D.H.  | <i>J Virol</i> <b>78</b> (24):14012-14022 (2004)   |
| MNE.US.82.MNE_8       | M32741    | All                                    | Kimata, J.T.    | <i>J Virol</i> <b>72</b> (1):245-256 (1998)  |
| MNE.US.x.MNE027       | U79412    | All                                    | Kimata, J.T.    | <i>J Virol</i> <b>72</b> (1):245-256 (1998)  |
| SMM.SL.92.SL92B       | AF334679  | All                                    | Chen, Z.        | <i>J Virol</i> <b>70</b> (6):3617-3627 (1996)  |
| SMM.US.x.F236_H4      | X14307    | All                                    | Hirsch, V.M.    | <i>Nature</i> <b>339</b> (6223): 389-92 (1989)   |
| SMM.US.x.PBJ14_15     | L03295    | All                                    | Dewhurst, S.    | <i>Nature</i> <b>345</b> (6276):636-640 (1990)   |
| SMM.US.x.PBJA         | M31325    | All                                    | Dewhurst, S.    | <i>Nature</i> <b>345</b> (6276):636-640 (1990)   |
| SMM.US.x.PBJ_6P6      | L09212    | All                                    | Novembre, F.J.  | <i>J Virol</i> <b>67</b> (5):2466-2474 (1993)  |
| SMM.US.x.PGM53        | AF077017  | All                                    | Novembre, F.J.  | <i>J Virol</i> <b>72</b> (11):8841-8851 (1998)   |
| SMM.US.x.SME543       | U72748    | All                                    | Hirsch, V.      | <i>J Virol</i> <b>71</b> (2):1608-1620 (1997)  |
| STM.US.x.STM          | M83293    | All                                    | Novembre, F.J.  | <i>Virology</i> <b>186</b> (2):783-787 (1992)  |
| MAC.US.x.97074        | AY599198  | Pol, Vpr, Rev                          | O'Connor, D.H.  | <i>J Virol</i> <b>78</b> (24):14012-14022 (2004)   |
| MAC.US.x.MAC239_87082 | AY600249  | Pol, Vpr, Rev                          | O'Connor, D.H.  | <i>J Virol</i> <b>78</b> (24):14012-14022 (2004)   |
| SMM.US.x.H9           | M80194    | Pol                                    | Courgnaud, V.   | <i>J Virol</i> <b>66</b> (1):414-419 (1992)  |
| SMM.US.x.PBJ_143      | M80193    | Pol                                    | Courgnaud, V.   | <i>J Virol</i> <b>66</b> (1):414-419 (1992)  |
| A.GW.87.CAM2CG        | D00835    | Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Tristem, M.     | <i>J Gen Virol</i> <b>72</b> (PT 3):721-724 (1991)   |

| Name           | Accession | Proteins                      | Author         | Reference  |
|----------------|-----------|-------------------------------|----------------|--|
| MAC.US.x.1937  | AY611495  | Pol, Vpx,<br>Vpr, Nef         | O'Connor, D.H. | <i>J Virol</i> <b>78</b> (24):14012-14022 (2004) |
| MAC.US.x.2065  | AY611493  | Pol, Vpx,<br>Vpr, Rev,<br>Nef | O'Connor, D.H. | <i>J Virol</i> <b>78</b> (24):14012-14022 (2004) |
| MAC.US.x.80035 | AY611486  | Pol, Vpr,<br>Rev              | O'Connor, D.H. | <i>J Virol</i> <b>78</b> (24):14012-14022 (2004) |

|                           | Gag p15 start   | p15_p27                        |     |
|---------------------------|---|--------------------------------|-----|
| MAC_US.x.239              | MGVNRNSVLSGKKADELEKIRLRLPVGKKKYMLKHVVAAANELDRFGLAESLLENKEGCKQKILSVLAPLVPTGSENLSKLYNTVCVIWCIHAEKVKHTEEAQIVQRHLVVTETGTTETMPKTSRPTAPSSGRGGNYPVQQ | IGGNYVHLPLSPRTLNAWVKLIEEKKFGAE | 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--S--V-----V-----V-----                          |                                | 169 |
| A.DE.x.PE12               | --A--S--R--V-----G--R--I-----GK-----S-----IT--D-----F-----L-----D--G-----A--ADK--S--A--P-----A--S--V-----V-----V-----                         |                                | 169 |
| A.GH.x.GH1                | --A--R-----V--S--R--I-----K-----S-----T--D-----F-----L-----D-----KL--GA--A--K--S-----P--R--F--TG--I--V-----V--D-----V-----                    |                                | 170 |
| A.GM.87.D194              | --A--R-----V--R--R-----I-----S-----K--E-----F-----L-----D-----KLA--A--A--K--NI-----P--K--F--A--I--V-----V-----V-----                          |                                | 169 |
| A.GM.x.ISY                | --AK--R-----T--G--R--I-----T-----S-----I--F-----F--T--L-----D--G--R--G-----A--E--A--K--NI-----P--K--F--V--T--V-----V-----V-----               |                                | 169 |
| A.GM.x.MCN13              | --A--K-----R-----G--R--I-----S-----R--T--D-----F-----L-----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-----L-----D--G-----AA--A--K--N-----P--K--F--V--T--V-----V-----V-----                            |                                | 169 |
| A.GW.86.FG                | --A--R-----R-----G--K--I-----S-----T--D-----F-----V--V-----D--G-----I--AA--IE--A--K--S-----P--EQ--F--VA--T--V--D-----V-----                   |                                | 167 |
| A.GW.x.ALI                | --A--S--R-----R-----V--G--R--I-----S-----R--K--E-----F-----F-----D--G-----KFA--A--A--K--N-----P--G--F--H--T--V-----V-----                     |                                | 169 |
| A.GW.x.MDS                | --AKS--R-----R-----G--R--I-----K--T-----S-----T--D-----F-----D--KL--A--A--K--N-----P--EG--F--R--T--V-----V-----V-----                         |                                | 169 |
| A.IN.x.CRTK_147           | --AK--R-----R-----G--RV-----K--T-----S-----T--D-----F-----D--KL--A--A--K--N-----P--EG--F--R--T--V-----V-----V-----                            |                                | 169 |
| A.IN.x.NNVA               | --A--R-----R-----G--R--I-----K-----S-----T--D--M-----F-----D--G--R--A--A--K--S-----EK--H--V--T--I-----V-----V-----                            |                                | 169 |
| A.SN.85.ROD               | --A--S-----T--V--G--R--C--II--V-----S--H--T-----F--Y--L--D--K--A--AAD--K--A--P-----A--M-----V-----V-----                                      |                                | 164 |
| B.CI.88.UC1               | --A--G-----T--V--G--R--C--II--V-----S--Q--H--T-----M-----F--Y--L--D--K--A--S--ADK--K--A--P-----A--V-----V--D-----V-----                       |                                | 164 |
| B.CI.09.pPRD320-20HIV2_56 | --A--G-----T--V--G--R--I--V--E--Y--T-----R--GS--R--RK--G-----F--L--D--K--A--AAD--K--AM--K--SK--T--RL--A-----A--S-----V-----V-----             |                                | 164 |
| B.CI.x.EH0                | --A--G-----T--V--G--R--I--V--E--Y-----S--H--T--E-----F--I--F--L--D--K--A--AAD--K--A--NK--P-----LA--T-----V-----V-----                         |                                | 164 |
| B.GH.86.D205              | --A--G-----T--V--G--C--I--V--Y-----S--H--T--E-----F--Y--L--V--D--D--K--A--S--AAN--K--A--K--P-----VA--T-----V-----V-----                       |                                | 164 |
| B.JP.01.KR020             | --A--S-----V--G-----S-----L--Q-----V-----KV--A--I--P-----V-----V-----V-----   |                                | 169 |
| G.CI.92.ABT96             | --A--G-----T--V--G--R--I--V-----N--S-----T--M-----F--L--K--D--KLA--S--D--KT--TADK--A--T-----S--VA--V-----V--L-----V-----                      |                                | 164 |
| AB.CI.90.7312A            | --A--S-----T--V--G--R--I--V-----T--S-----K-----F--L--K--A--A--DA--K--TA--NK--P-----A--V-----V-----V-----                                      |                                | 163 |
| AB.CM.03.03CM_510_03      | --A--#-----V--V-----A-----L-----A--E--A--K--A--X--P--GR-----A--V-----V-----   |                                | 168 |
| U.FR.96.12034             | --A-----G-----A--I-----V-----V-----   |                                | 169 |
| MAC_US.x.251.1A11         | --A-----G-----A-----V-----V-----  |                                | 169 |
| MAC_US.x.251.32H.PJ5      | --A-----G-----A-----F-----T-----V-----  |                                | 169 |
| MAC_US.x.251.BK28         | --A-----G-----M--A-----F-----T-----V-----   |                                | 169 |
| MAC_US.x.MM142            | --A-----G-----M--A-----F-----T-----V-----   |                                | 169 |
| MAC_US.x.17EC1            | --A-----G-----M--A-----F-----T-----V-----   |                                | 169 |
| MAC_US.x.17EFR            | --A-----G-----M--A-----F-----T-----V-----   |                                | 169 |
| MAC_US.x.SMM142B          | --A-----G-----M--A-----F-----T-----V-----   |                                | 169 |
| MAC_US.x.r00025           | --A-----X-----F-----T-----V-----  |                                | 169 |
| MAC_US.x.r00131           | --A-----X-----F-----T-----V-----  |                                | 169 |
| MNE_US.82.MNE.8           | --A-----G-----A-----V--T-----V-----   |                                | 169 |
| MNE_US.x.MNE027           | --A-----G-----D-----A-----KR-----T-----V-----   |                                | 169 |
| SMM_SL.92.SL92B           | --A--G-----V--G--R-----II--R-----S--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-----D-----ADR--A--P-----V-----V-----   |                                | 169 |
| SMM_US.x.PB114.15         | --A-----V-----G--R--K--I-----L-----ADR--A--P--K-----T-----V-----  |                                | 169 |
| SMM_US.x.PB1A             | --A-----V-----G--R--Q--I-----L-----ADR--A--P--K-----T-----V-----  |                                | 169 |
| SMM_US.x.PBJ.6P6          | --A-----V-----G--R--K--I-----L-----ADR--A--P--K-----T-----V-----  |                                | 169 |
| SMM_US.x.PGMS3            | --A-----V-----G--R--K--I-----D--I--A-----L-----ADR--A--P--K-----T-----V-----  |                                | 169 |
| SMM_US.x.SME543           | --A-----V-----G--R--K--I-----T-----ADR--A--P-----V-----V-----   |                                | 169 |
| STM_US.x.STM              | --A--S-----V--G-----S-----IT--E-----F-----V--K-----ANK--A--P-----V-----V-----   |                                | 169 |

|                           |   |    |    |   |     |
|---------------------------|---|----|----|---|-----|
| MAC.US.x.239              | VVPGFQALSEGCTPYDINQMLNCVGDHQAMQIIRDIINEEAADWDLQHPQ  | AP | .. | QQGQLREPSGSDIAGTTSSVDEIQWYRQNPVPGNIYRRWQLGLQKCVRMYPNTILDVKQGPKEPFQSYVDRFYKSLRAEQTDAAVKNNMTQTLIIQNANPCKLVL | 335 |
| A.CI.88.UC2               | -----Q-----E-----A-I-G-L-PA-D-R-----T-----P-V-----I-----K-----V-----S-----A-P-----V-----                  |    |    |   | 336 |
| A.DE.x.BENJ               | -----S-I-G-L-PA-D-R-----T-----P-V-----I-----K-----I-----S-----A-P-----V-----                              |    |    |   | 336 |
| A.DE.x.PE12               | -----V-I-G-L-PA-D-R-----T-E-----F-A-----I-----K-----I-----S-----A-P-----V-----                            |    |    |   | 336 |
| A.GH.x.GH1                | -----D-----A-I-G-L-PA-D-R-----T-E-----P-V-----I-----I-----S-----A-P-----V-----                            |    |    |   | 337 |
| A.GM.87.D194              | -----A-I-G-L-PA-D-R-----T-----P-V-----I-----I-----S-----A-P-----V-----                                    |    |    |   | 336 |
| A.GM.x.ISY                | -----V-I-G-L-PA-D-R-----T-E-----E-----E-----V-----I-----S-----A-P-----V-----S-----                        |    |    |   | 335 |
| A.GM.x.MCN13              | -----V-----V-----I-G-L-PA-D-R-----T-----F-P-V-----I-----I-----S-----A-P-----V-----                        |    |    |   | 336 |
| A.GM.x.MCR35              | -----V-----V-----I-G-L-PA-D-R-----T-----F-P-V-----I-----I-----S-----A-P-----V-----                        |    |    |   | 336 |
| A.GW.86.FG                | -----VA-----I-G-L-PA-D-R-----T-E-----F-P-V-----I-----I-----IN-----A-P-----V-----                          |    |    |   | 334 |
| A.GW.x.ALI                | -----VA-----I-G-L-PA-D-R-----T-E-----F-PR-V-----I-----I-----I-----A-P-----V-----                          |    |    |   | 336 |
| A.GW.x.MDS                | -----A-----I-G-L-PA-D-R-----T-----F-A-V-----I-----I-----I-----A-P-----V-----                              |    |    |   | 336 |
| A.IN.95.CR1K.147          | -R-----A-----I-G-L-PA-D-R-----T-E-----F-A-V-----I-----I-----I-----A-P-----V-----                          |    |    |   | 336 |
| A.IN.x.NNVA               | -----V-----V-----I-G-L-PA-D-R-----T-E-----E-----F-A-V-----I-----I-----I-----A-P-----V-----II-----         |    |    |   | 336 |
| A.SN.85.ROD               | -----E-----V-----I-G-L-PA-D-R-----T-E-----F-P-V-----I-----I-----I-----A-P-----V-----II-----               |    |    |   | 336 |
| B.CI.88.UC1               | -----O-----I-G-L-PA-D-R-----T-E-----A-V-----I-----I-----S-----A-P-----V-----                              |    |    |   | 331 |
| B.CI.89.pPRD320-20HIV2_56 | -----O-----I-G-L-PA-D-R-----T-E-----A-V-----I-----I-----S-----A-P-----V-----                              |    |    |   | 331 |
| B.CI.x.EH0                | -----O-----S-----G-M-PA-D-R-----T-E-----P-V-----I-----I-----I-----A-P-----V-----                          |    |    |   | 331 |
| B.GH.86.D205              | -----O-----S-----G-M-PA-D-R-----T-E-----A-V-----I-----I-----I-----A-P-----V-----                          |    |    |   | 331 |
| B.JP.01.KR020             | -----E-----D-----O-----S-----G-M-PA-D-R-----T-E-----P-----I-----I-----I-----A-P-----V-----                |    |    |   | 331 |
| G.CI.92.ABT96             | -----L-----E-----E-----X-----Q-G-P-PA-X-T-----TIE-----TH-----X-----X-----X-----X-----P-----V-----         |    |    |   | 336 |
| AB.CI.90.7312A            | -----E-----V-----E-----D-----O-----V-----G-L-PA-D-R-----T-E-----G-S-----I-----I-----I-----A-T-RA-E-V----- |    |    |   | 331 |
| AB.CM.03.03CM.510_03      | -----E-----D-----O-----S-----G-L-PA-D-R-----T-E-----G-----I-----I-----I-----P-S-E-V-----                  |    |    |   | 330 |
| U.FR.96.12034             | -----E-----E-----D-----T-----NQ-G-P-PA-----T-E-----P-----I-----I-----I-----P-----V-----                   |    |    |   | 335 |
| MAC.US.x.251.1A11         | -----   |    |    |   | 335 |
| MAC.US.x.251.32H.PJ5      | -----   |    |    |   | 335 |
| MAC.US.x.251.BK28         | -----   |    |    |   | 335 |
| MAC.US.x.MM142            | -S-----L-----Q-----T-E-----P-----   |    |    |   | 335 |
| MAC.US.x.17EC1            | -----   |    |    |   | 335 |
| MAC.US.x.17EFR            | -----   |    |    |   | 335 |
| MAC.US.x.SMM142B          | -S-----L-----Q-----T-E-----P-----   |    |    |   | 335 |
| MAC.US.x.r80025           | -----   |    |    |   | 335 |
| MAC.US.x.r90131           | -----   |    |    |   | 335 |
| MNE.US.82.MNE.8           | -----E-----E-----Q-----T-----S-----P-----   |    |    |   | 335 |
| MNE.US.x.MNE027           | -I-----E-----E-----Q-----T-----K-----S-----IR-----P-----  |    |    |   | 335 |
| SMM.SL.92.SL92B           | -----E-----E-----RG-QQPAQPA-G-----TPS-----E-----A-----V-----D-----P-----                                  |    |    |   | 336 |
| SMM.US.x.F236.H4          | -----G-L-PA-D-R-----T-----P-----P-----P-----  |    |    |   | 336 |
| SMM.US.x.PB114.15         | -----G-I-PP-R-----T-----S-----P-----  |    |    |   | 336 |
| SMM.US.x.PB1A             | -----G-I-PP-R-----T-----P-----  |    |    |   | 336 |
| SMM.US.x.PB1.6P6          | -----G-I-PP-R-----T-----S-----P-----S-----  |    |    |   | 336 |
| SMM.US.x.PGM53            | -----G-I-PA-D-R-----T-E-----P-----  |    |    |   | 336 |
| SMM.US.x.SME543           | -----E-----E-----G-L-PA-D-R-----T-E-----P-----  |    |    |   | 336 |
| STM.US.x.STM              | -----L-----E-----KE-----M-----PG-L-PA-----PE-----V-----T-----A-PS-----R-----                              |    |    |   | 337 |

|                           | p27_p2  | p2_p8   | p8_p1                                   | p1_p6                    | PTAP motif               | PSAP in HIV-2 B, U |              |                                      |                                       |  |   |        |     |
|---------------------------|---|---|---|--------------------------|--------------------------|--------------------|--------------|--------------------------------------|---------------------------------------|--|---|--------|-----|
| MAC.US.x.239              | KGLGVNPTLEEMLTACQGVGGPQKARLMAEALKEALPVPIPFAAAQQR... | .GPRK.PIKWCNCKEGHSARQCRAPRRQGCWKCGKMDHVMKCPDRQAGFLGLGPWGGKPRNFMAQVHQLMPTAPPEDPAVDLLKNYMLGKQREKORE.....S |   |                          |                          |                    | 487          |                                      |                                       |  |   |        |     |
| A.CI.88.UC2               | -----M-----   | -----T P-----   | -----R.T.R-----                         | -----K-----              | PG.I.N-----              | M.SR-----          | V.AP.I-----  | A.-A.-EK.L.Q.RK--QRERPYKEVTE.DLLHLEQ | 498                                   |  |   |        |     |
| A.DE.x.BEN                | -----M-----   | -----MG.S-----  | -----A.RY-----                          | -----K-----              | PG.I.N-----              | R-----             | VT.AP.I----- | A.-A.-ER.-Q.RK--QRERPYKEVTE.DLLHLEQ  | 498                                   |  |   |        |     |
| A.DE.x.PE12               | -----M-----   | -----I-----   | -----A-----                             | -----R.T-----            | D-----                   | SG.-N.E-----       | I-----       | VTR.P.T-----                         | A.-A.-EK.L.Q.RK-K.QKMRPYKEVTE.DLLHLEQ | 498                                      |   |        |     |
| A.GH.x.GH1                | -----M-----   | -----S-----   | -----T.A-----                           | -----R.A.R-----          | -----K-----              | SG.I.N-----        | M.-R.Q-----  | A.-AP.I-----                         | I-----                                | EK.-Q.RK--QRERPYKEVTE.DLLHLEQ            | 498                                     |        |     |
| A.GM.87.D194              | -----M-----   | -----I-----   | -----T.A-----                           | -----K.R.A-----          | -----K-----              | SG.I.N-----        | V.P.S-----   | T-----                               | M-----                                | EK.-Q.RK--QRERPYKEVTE.DLLHLEQ            | 497                                     |        |     |
| A.GM.x.ISY                | -----M-----   | -----I-----   | -----T.A-----                           | -----K.R.T-----          | -----K-----              | PG.I.N-----        | V.M.R-----   | V.-P.T-----                          | V-----                                | E.-Q.-R--QRERPYKEVTE.DLLHLEQ             | 498                                     |        |     |
| A.GM.x.MCN13              | -----M-----   | -----I-----   | -----T.A-----                           | -----K.R.T-----          | -----K-----              | PG.I.N-----        | V.M.R-----   | V.-P.T-----                          | V-----                                | EK.-Q.-R--QRERPYKEVTE.DLLHLEQ            | 498                                     |        |     |
| A.GM.x.MCR35              | -----M-----   | -----I-----   | -----T.A-----                           | -----K.R.T-----          | -----K-----              | PG.I.N-----        | V.M.R-----   | V.-P.T-----                          | V-----                                | EK.-Q.-R--QRERPYKEVTE.DLLHLEQ            | 498                                     |        |     |
| A.GW.86.FG                | -----M-----   | -----T-----   | -----V.M.A-----                         | -----T.F-----            | -----W.S-----            | SG.-N.E-----       | I.H.A-----   | A.-P.T-----                          | L-----                                | EK.-Q.-R--QRERPYKEVTE.DLLRFEQ            | 496                                     |        |     |
| A.GW.x.ALI                | -----M-----   | -----I-----   | -----T.A-----                           | -----R.T-----            | -----K-----              | PG.L.N-----        | E-----       | VTR.P.T-----                         | AE.-A.-EQ.-Q.RK--QRERPYKEVTE.DLLHLEQ  | 498                                      |   |        |     |
| A.GW.x.MDS                | -----M-----   | -----I-----   | -----T.A-----                           | -----R.T-----            | -----K-----              | PG.L.N-----        | E-----       | VTR.P.T-----                         | AE.-A.-EQ.-Q.RK--QRERPYKEVTE.DLLHLEQ  | 498                                      |   |        |     |
| A.IN.95.CR1K_147          | -----M-----   | -----I-----   | -----T.A-----                           | -----R.T-----            | -----K-----              | PG.L.N-----        | E-----       | VTR.P.T-----                         | AE.-A.-EQ.-Q.RK--QRERPYKEVTE.DLLHLEQ  | 498                                      |   |        |     |
| A.IN.x.NNVA               | -----M-----   | -----G-----   | -----MG.T-----                          | -----V-----              | -----T.R-----            | -----K-----        | PG.I.N-----  | M.-R-----                            | V.-IP.T-----                          | M-----                                   | EK.-Q.RK--QRERPYKEVTE.DLLHLEQ           | 498    |     |
| A.SN.85.R0D               | -----M-----   | -----I-----   | -----VIG.A-----                         | -----A.F-----            | -----K-----              | PG.I.TN-----       | -----        | V.-P.T-----                          | V-----                                | EK.-Q.-R--QRERPYKEVTE.DLLHLEQ            | 498                                     |        |     |
| B.CI.88.UC1               | -----M-----   | -----I-----   | -----T.A-----                           | -----K..AGKRGIVT-----    | -----V.-T.K-----         | QG.I.S-----        | E-----       | T.-P.-VT.S-----                      | M-----                                | EGMTPRGATPSAPPADPAV.MLKSYMKGROORE-       | 498                                     |        |     |
| B.CI.09.pPRD320-20HIV2_56 | -----I-----   | -----I-----   | -----T.A-----                           | -----K..AGKR.T.T-----    | -----G.-T.K-----         | PG.-S.Q-----       | -----        | V-----                               | F-----                                | P.RVT.S-----                             | M.-TQ-MTPQGATPSAPPADPAV.MLKNYMOLGKQORE- | 497    |     |
| B.CI.x.EH0                | -----M-----   | -----I-----   | -----T.STN-----                         | -----P..AGKR.TVT-----    | -----A.-T.K-----         | QOG.I.S-----       | -----        | F-----                               | V.-AP.IV.S-----                       | MN.-FGMTPQGATPSAPPADPAE.MLKNYMOLGKQKQEN  | 496                                     |        |     |
| B.GH.86.D205              | -----M-----   | -----I-----   | -----T.A-----                           | -----V.-K..AGKRGIVT----- | -----Q.-T-----           | TG.I.S-----        | -----        | F-----                               | T.-P.-VT.S-----                       | MN.-EGMTPRGATPSAPPADPAV.MLKSYMOMGRQORE-  | 498                                     |        |     |
| B.JP.01.KR020             | -----M-----   | -----I-----   | -----T.P.L-----                         | -----V..AGKRGIVT-----    | -----I.-T-----           | QG.I.S-----        | -----        | F-----                               | V.-AP.TL.S-----                       | MN.-ENMTPQGAMPSAPPADPAV.MLKDYMLGRKQKQGG  | 498                                     |        |     |
| G.CI.92.ABT96             | -----M-----   | -----X-----   | -----N.TAL-----                         | -----K..T.GKRST-----     | -----TV-----             | PG.I-----          | -----        | F-----                               | T.-P.-T.S-----                        | M-----                                   | X.-RK.K-----                            | Q      | 486 |
| AB.CI.90.7312A            | -----PH-----  | -----I-----   | -----T.A-----                           | -----K..AG.RCTVT-----    | -----G.-T.K-----         | PG.N.N-----        | -----        | E.-V-----                            | T.-P.-EIV.S-----                      | MNT.-EGKTHOGATPSAPPADPAV.MLKSYMOLGKQOREK | 498                                     |        |     |
| AB.CM.03.03CM_510_03      | -----M-----   | -----I-----   | -----K-----                             | -----T.A-----            | -----F.-K..TGKRGIVT----- | -----G.-T.-K-----  | -----        | QG.I.S-----                          | E.-V-----                             | T.-P.-VT.S-----                          | MN.-EXMTPQGVTPSAPPADPAV.MLKSYMOMGRQORE- | 497    |     |
| U.FR.96.12034             | -----M-----   | -----I-----   | -----Q.M-----                           | -----V.-G.-R.-VR-----    | -----T.-K-----           | E.-PG.-N.-S-----   | ETKV-----    | E-----                               | P.-T-----                             | AE.-TPTAPPADPAVDLLKSYMQQGKK...QKEN       | 498                                     |        |     |
| MAC.US.x.251.1A11         | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 487    |     |
| MAC.US.x.251.32H.PJ5      | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 487    |     |
| MAC.US.x.251.BK28         | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 483    |     |
| MAC.US.x.MM142            | -----T-----   | -----A-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 483    |     |
| MAC.US.x.17EC1            | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 487    |     |
| MAC.US.x.17EFR            | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 487    |     |
| MAC.US.x.SMM142B          | -----T-----   | -----A-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 483    |     |
| MAC.US.x.r80025           | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 487    |     |
| MAC.US.x.r90131           | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 487    |     |
| MNE.US.82.MNE.8           | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 483    |     |
| MNE.US.x.MNE027           | -----M-----   | -----I-----   | -----K-----                             | -----T-----              | -----K-----              | -----              | -----        | -----                                | -----                                 | -----                                    | -----                                   | 483    |     |
| SMM.SL.92.SL92B           | -----M-----   | -----I-----   | -----M.D.-TGSLVA-QFRGAAGQGN-PI-R-F----- | -----T-----              | -----K-----              | EEGRIQ.N-----      | NQK-----     | -----                                | TTS.-T.S-----                         | D.-ARIV.E-LEKAQREKT.....R                | 484                                     |        |     |
| SMM.US.x.F236.H4          | -----M-----   | -----I-----   | -----R.DQL-----                         | -----V.-K.-Q.-I-----     | -----K-----              | TG-----            | -----        | -----                                | MP.I-----                             | -----                                    | KM.RK-----                              | N      | 484 |
| SMM.US.x.PB114.15         | -----I-----   | -----D-----   | -----TOG.L-----                         | -----V.-K.-Q.-I-----     | -----K-----              | AG-----            | -----        | -----                                | MP.T-----                             | -----                                    | KM.RR-----                              | N      | 484 |
| SMM.US.x.PB3A             | -----I-----   | -----D-----   | -----TOG.L-----                         | -----V.-K.-Q.-I-----     | -----K-----              | AG-----            | -----        | -----                                | MP.T-----                             | -----                                    | KM.RR-----                              | N      | 484 |
| SMM.US.x.PBJ.6P6          | -----I-----   | -----D-----   | -----TOG.L-----                         | -----V.-K.-Q.-I-----     | -----K-----              | AG-----            | -----        | -----                                | MP.T-----                             | -----                                    | KM.RR-----                              | N      | 484 |
| SMM.US.x.PGMS3            | -----M-----   | -----I-----   | -----R.DQL-----                         | -----V.-K.-Q.-R.T-----   | -----K-----              | TG-----            | -----        | -----                                | MP.T-----                             | -----                                    | R.-KM.RK-----                           | N      | 484 |
| SMM.US.x.SME543           | -----M-----   | -----I-----   | -----R.DQL-----                         | -----V.-K.-Q.-R.T-----   | -----K-----              | TG-----            | -----        | -----                                | MP.T-----                             | -----                                    | KM.RK-----                              | N      | 484 |
| STM.US.x.STM              | -----M-----   | -----I-----   | -----VFQ.D.L-----                       | -----Q.G.R.TV-----       | -----A.-T.K.-KG-----     | QOG.Q-----         | E.-V-----    | F-----                               | H-----                                | IP.-T-----                               | A.-RS-----                              | K----- | 484 |

```

p6 end
Gag end
MAC.US.x.239      REKPYK.EVTEDLLHLSLFGGQ*
A.CI.88.UC2      G-T-RR--A-----K---
A.DE.x.BEN       -T-HR--E-----K---
A.DE.x.PE12      G-T-H--A-----K---
A.GH.x.GH1       GKA-HR--A-----K---
A.GM.87.D194     G-T-HR.GA-----K---
A.GM.x.ISY       G-T-HR--T-----N---
A.GM.x.MCR13     G-A-CR--T-----K---
A.GM.x.MCR35     G-A--R--T-----*K--
A.GW.86.FG       A-T-CR--T-----K---
A.GW.x.ALI       G-T-H-- -------K---
A.GW.x.MDS       GKT-C--T-----T---
A.IN.95.CR1K_147 GGA-CG--M-----K---
A.IN.x.NNVA      G-A-CG--T-----K-L-
A.SN.85.ROD      G-T--REPP-----K---
B.CI.88.UC1      --R-----E---
B.CI.09.pPRD320-20HIV2_56 --R-----E---
B.CI.x.EH0       --R-----E---
B.GH.86.D205    --R-----E---
B.JP.01.KR020   --R-----E---
G.CI.92.ABT96   --N-----XX---S---D---
AB.CI.90.7312A  QGR-----E---
AB.CM.03.03CM_510_03 --R-----E---
U.FR.96.12034   --R-----N---
MAC.US.x.251.1A11 -----M-----
MAC.US.x.251.32H.PJ5 -----
MAC.US.x.251.BK28 -----
MAC.US.x.MM142  -G-----
MAC.US.x.17EC1  -----
MAC.US.x.17EFR  -----
MAC.US.x.SMM142B -G-----
MAC.US.x.r80025 -----
MAC.US.x.r90131 -----
MNE.US.82.MNE.8 KR-----E---
MNE.US.x.MNE027 KR-----A-----E---
SMM.SL.92.SL92B --SR-----E---
SMM.US.x.F236.H4 --R-----E---
SMM.US.x.PB114_15 --R-----E---
SMM.US.x.PB1A   --R-----E---
SMM.US.x.PB1_6P6 --R-----E---
SMM.US.x.PGM53  --R-----E---
SMM.US.x.SME543 --R-----E---
STM.US.x.STM    -KT-----E---

```

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

|                           | Pol p15 start (-1 from Gag)   | p15 protease   |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
|---------------------------|---|--|-------------------|------|----|-----|--------|-------|--------------|----------------|---------------------|----------------------|------------|------------|----------|----------|-------|----------|-----|-----|----|-----|----|----|------|------|-----|-----|---|---|-----|-----|-----|
| MAC.US.x.239              | FFRPWSMGKEAPQFPHGSSASGADANCSPRGSPSCGSAKELHAVGQAAERKAER.....KOREALQ          | GGDRGFAAPQFSLWRRPVVTAHIEGQPVVEVLDTGADDSIVTGIELGPHYTPKIVGGIGGFINTKEYKNIVEIYVGLGKRIGKTIIMTGDTPINIFGR | 155               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| A.CI.88.UC2               | ---DGLT---L-R-P-S---T-ST-SRSGS-PVR-IF-A-EK-GAEGETIOGGDG-RLTAPRAG-DTS-R---   | L---K---Y-D---A---DN---V---K-N-VRA   | 168               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| A.DE.x.BEN                | ---VGPT---S-L-RDP-P---T-STSGRS-S-TVG-IV-AREK-GAEGETIOGGDG-GLAAPRAE-DTS-R--- | L---K---Y-D---A---DN---K-N-VRA   | 168               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| A.DE.x.PE12               | ---D-P---S-L-RDP-PA---T-ST-SR-SRP-R-VL-AREE-AENETIOGGDR-GLTAPRTR-DTT-R---   | L---K---YV---A---SN-S---K-N-KV-A   | 168               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| A.GH.x.GH1                | ---DG---L-R-P-S---T-ST-SRS-S-IGKIV-A-ER-GAEGETIOGGDG-RLTAPRAGKSTS-R---      | L---K---Y-V---A-Q-DN-V---I-I-K-N-VRA   | 168               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| A.GM.87.D194              | ---DGPT-A---L-R-P-S---T-ST-NRS-S-PVG-IY-AREK-AEGETIOGGDG-GLTAPRAG-D-P-R---  | L-T---K---F-D---A---DN---K-N-VRA   | 168               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| MAC.US.x.97074            | X   |  | 155               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| MAC.US.x.MAC239.87082     | PL  | RT   | 155               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| MAC.US.x.MM142            | PL  | RT   | 151               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| MAC.US.x.SMM142B          |   |  | 155               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| MAC.US.x.r80025           |   |  | 155               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| MAC.US.x.r90131           |   |  | 155               |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| MNE.US.82.MNE.8           | L-P   | PN   | K                 | 151  |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| MNE.US.x.MNE027           | L-P   | PN   | K                 | 151  |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.SL.92.SL92B           | P   | S-L-ADNI-I-TVST  | SSPDSOGVSRESTEGKD | EEQ  | G  | R   | R-L    | L     | A            | G              | R                   | RT-K                 | V-RE-L     | 148        |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.US.x.F236.H4          | A-P   | PD   | T                 | S    | TE | EG  | K      | G     | E            | T              | NG                  | I                    | Y          | E          | I        | A        | N     | D-K-K    | V   | 151 |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.US.x.H9               | A-P   | PD   | T                 | X    | S  | TE  | EG     | K     | G            | E              | T                   | NG                   | I          | Y          | E        | A        | N     | D-K-K    | V   | 151 |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.US.x.PB114.15         | A-P   | PD   | T                 | S    | TE | EG  | K      | G     | E            | T              | NG                  | I                    | Y          | E          | A        | N        | D-K-K | V        | 151 |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.US.x.PBJA             | A-P   | PD   | T                 | S    | TE | EG  | K      | G     | E            | T              | NG                  | I                    | Y          | E          | A        | N        | D-K-K | V        | 151 |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.US.x.PBJ.143          | A-P   | PD   | T                 | X    | S  | TE  | EG     | K     | G            | E              | T                   | NG                   | I          | Y          | E        | A        | N     | D-K-K    | V   | 151 |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.US.x.PBJ.6P6          | A-P   | PD   | T                 | S    | TE | EG  | K      | G     | E            | T              | NG                  | I                    | Y          | E          | A        | N        | D-K-K | V        | 151 |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.US.x.PGMS3            | A-P   | PD   | T                 | S    | TE | EG  | K      | G     | E            | T              | G                   | A                    | Y          | E          | A        | N        | D-K-K | V        | 151 |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| SMM.US.x.SMS543           | V-P   | PD   | T                 | S    | TE | EG  | K      | G     | E            | T              | G                   | A                    | Y          | E          | A        | N        | D-K-K | V        | 151 |     |    |     |    |    |      |      |     |     |   |   |     |     |     |
| STM.US.x.STM              | L-P   | L  | PNT               | PS   | R  | EK  | A      | E     | E            | EDT            |                     |                      |            |            |          |          | LQ    | V        | F   | N   | K  | 151 |    |    |      |      |     |     |   |   |     |     |     |
| A.GM.x.ISY                | A-T   | L  | PKFA              | NT   | ST | N   | S      | PTG   | V            | AREKT          | AETKTIORS           | DR                   | GLAASRAR   | DTT        | RD       | L        | LK    | YV       | D   | A   | SN | S   | D  | R  | N    | KVRA | 168 |     |   |   |     |     |     |
| A.GM.x.MCN13              | DGPP  | L  | R                 | PA   | NT | ST  | SRS    | S     | PTG          | Y              | ARKKTK              | TERETIORS            | DR         | GLTAPRAG   | GTM      | RD       | L     | LK       | YV  | A   | SN | S   | K  | N  | KVRA | 168  |     |     |   |   |     |     |     |
| A.GM.x.MCR35              | DGPP  | L  | R                 | PA   | NT | ST  | SRS    | S     | PTG          | IY             | ARKKTK              | TERETIORS            | DR         | GLTAPRAG   | STI      | RD       | L     | LK       | YV  | A   | SN | S   | K  | N  | KVRA | 168  |     |     |   |   |     |     |     |
| A.GW.87.CAM2CG            | D-PL  | L  | R                 | P    | ST | NT  | ST     | I     | S            | S              | TG                  | IY                   | AREK       | GAETETIORS | DR       | GLTAPRTR | GPM   | DN       | L   | K   | A  | SN  | S  | N  | KVRA | 168  |     |     |   |   |     |     |     |
| A.GW.x.ALI                | A-P   | S  | L                 | RNP  | SA | INT | ST     | SRA   | S            | P              | GAVY                | A                    | EK         | K          | AEREA    | IORS     | DR    | GLTAPRAG | DTT | R   | L  | K   | Y  | ED | K    | N    | V-A | 168 |   |   |     |     |     |
| A.GW.x.MDS                | D-PL  | L  | R                 | PGSA | NT | ST  | SRS    | S     | PTG          | IY             | AREK                | GAERETIORS           | DR         | GLAAPRAG   | KDTM     | DN       | L     | N        | A   | SN  | S  | N   | N  | V  | KVRA | 168  |     |     |   |   |     |     |     |
| A.IN.95.CRTK.147          | DKPP  | S  | L                 | R    | P  | ST  | NT     | STSN  | S            | PTG            | IY                  | AREK                 | GAERETIORS | DR         | GPTAPRAG | GT       | R     | N        | L   | K   | Y  | A   | SN | S  | N    | KVRA | 168 |     |   |   |     |     |     |
| A.IN.x.NNVA               | DGPP  | L  | R                 | PNPA | NT | ST  | N      | S     | S            | PIG            | IY                  | AREK                 | GAERETIORS | DR         | GPTAPRAG | ST       | R     | N        | RL  | K   | Y  | A   | SN | S  | N    | KVRA | 168 |     |   |   |     |     |     |
| A.SN.85.ROD               | TGPL  | L  | R                 | P    | SA | T   | ST     | S     | S            | S              | TG                  | IY                   | AREKT      | AERETIORS  | DR       | GLTAPRAG | GDTI  | GATN     | L   | K   | Y  | A   | SN | S  | N    | KVRA | 169 |     |   |   |     |     |     |
| B.CI.88.UC1               | VRTL  | S  | L                 | DP   | S  | TI  | T      | D     | R            | HDTSGGDTIC     | PC                  | SSSGDAEKLHEDGETA     | AREP       | T          |          |          |       |          |     |     |    |     |    |    |      |      | VRA | 169 |   |   |     |     |     |
| B.CI.09.pPRD320-20HIV2.56 | IRTL  | S  | L                 | P    | EG | TI  | T      | D     | NT           | HDTSGGDSICTPC  | SSSGDAEKLHAAREAA    | AREP                 | T          |            |          |          |       |          |     |     |    |     |    |    |      |      | VRA | 169 |   |   |     |     |     |
| B.CI.x.EH0                | VRPL  | S  | R                 | PGTP | DS | I   | A      | DE    | I            | RHDTSGGDSICTPC | SSRGDAEKLHATREAA    | E                    | E          |            |          |          |       |          |     |     |    |     |    |    |      |      | VRA | 168 |   |   |     |     |     |
| B.GH.86.D205              | VRTL  | S  | L                 | DP   | S  | TI  | T      | DE    | R            | HDTSGGDTIC     | PC                  | SSSGDAEKLHADGETTEREP | T          |            |          |          |       |          |     |     |    |     |    |    |      |      | VRA | 169 |   |   |     |     |     |
| B.JP.01.KR020             | I-PL  | S  | L                 | CSP  | TP | DT  | I      | TTNE  | REHDTSGGNAIC | PC             | SSSGDVEGLHAAREKAERG | G                    | T          |            |          |          |       |          |     |     |    |     |    |    |      |      | VRA | 169 |   |   |     |     |     |
| G.CI.92.ABT96             | V-TL  | S  | L                 | DP   | P  | S   | SISTTD | SRPTE | X            | EE             | KG                  |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      | VRA | 151 |   |   |     |     |     |
| AB.CI.90.7312A            | A-TL  | S  | L                 | P    | DS | I   | A      | DEH   | R            | QDTSGGDTIC     | PC                  | SSSGDAEKLHATREAA     | E          | AREAG      |          |          |       |          |     |     |    |     |    |    |      |      | VRA | 169 |   |   |     |     |     |
| U.FR.96.12034             | A-PL  | K  | S                 | L    | P  | PA  | T      | P     | SPSSRT       | S              | DPDPSSSGSSGS        | KELHATGQE            | TKGE       | TI         | G        | K        | I     | Y        | K   | N   | A  | K   | E  | Y  | T    | E    | E   | KFH | N | K | T   | R   | 165 |
| MAC.US.x.17EC1            |   |  |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   |     | 155 |     |
| MAC.US.x.17EFR            |   |  |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   | 155 |     |     |
| MAC.US.x.1937             | X   | X  |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   | 155 |     |     |
| MAC.US.x.2065             |   |  |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   | 155 |     |     |
| MAC.US.x.251.1A11         |   |  |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   | 155 |     |     |
| MAC.US.x.251.32H.PJ5      |   |  |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   | 155 |     |     |
| MAC.US.x.251.BK28         |   |  |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   | 155 |     |     |
| MAC.US.x.80035            |   |  |                   |      |    |     |        |       |              |                |                     |                      |            |            |          |          |       |          |     |     |    |     |    |    |      |      |     |     |   |   | 151 |     |     |

|                           | protease_p51 RT start |               | D catalytic site |                          |               |              |              |            |             |               |              |            |            |         |     |
|---------------------------|-----------------------|---------------|------------------|--------------------------|---------------|--------------|--------------|------------|-------------|---------------|--------------|------------|------------|---------|-----|
| MAC.US.x.239              | LLLTALGMSLNFP         | IAKVEPVKVALK  | PGKIDGPKLQW      | PLSKKEIKVALREICEK.MEKDGL | EEAAPTNPYNTPT | PTFAIKKKDKNK | WRMLIDFRELNR | VTDFTEVQLG | IPHPAGLAKRR | ITVLDIGDAYFSI | PLDDEEFROYTA | FTLPSVNNAE | PGKRIYIKVL | PQGWKGS | 324 |
| A.CI.88.UC2               | -I-----               | L-V-RI-I-IT   | R-----           | T-VE-K-----              | E-----        | E-----       | E-----       | I-----     | K-----      | V-----        | Y-D-P-----   | A-V-----   |            |         | 337 |
| A.DE.x.BEN                | -I-----               | L-V-I-I-T     | R-----           | T-E-K-----               | E-----        | E-----       | E-----       | I-----     | K-SI-V----- | H-D-----      | A-M-----     |            |            |         | 337 |
| A.DE.x.PE12               | -I-----               | L-V-D-I-I     | VR-----          | T-E-K-----               | RE-----       | E-----       | E-----       | K-E-I----- | KR-----     | H-D-----      | T-----       |            |            |         | 337 |
| A.GH.x.GH1                | -I-----               | L-V-I-I-T     | R-R-----         | T-E-K-----               | E-----        | E-----       | E-----       | I-----     | K-----      | V-----        | H-D-----     |            |            |         | 337 |
| A.GM.87.D194              | -I-AT-----            | L-V-LD-I-T    | R-----           | T-E-K-----               | RE-----       | E-----       | E-----       | I-----     | K-----      | V-----        | H-D-----     |            |            |         | 337 |
| MAC.US.x.97074            | -----                 | L-----        |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.MAC239_87082     | -----                 | X-----        |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.MM142            | -----                 | L-----        |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.SMM142B          | -----                 | L-----        | SP-----          |                          |               |              |              |            |             |               |              |            |            |         | 320 |
| MAC.US.x.r80025           | -----                 |               | E-----           |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.r90131           | -----                 |               |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MNE.US.82.MNE_8           | -----                 | E-----        | T-----           | R-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| MNE.US.x.MNE027           | -----                 |               | T-----           | R-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| SMM.SL.92.SL92B           | -I-A-V-Y-V-YT         | -K-E-M-R----- | Q-T-----         | E-R-----                 | N-D-----      | KM-E-----    | KEME-----    | V-P-----   | I-Q-----    |               |              |            |            |         | 317 |
| SMM.US.x.F236_H4          | -M-----               | L-X-I-T-----  | E-R-----         | I-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| SMM.US.x.H9               | -M-----               | L-X-I-T-----  | E-R-----         | I-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| SMM.US.x.PB114_15         | -M-----               | L-I-T-----    | R-----           | I-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| SMM.US.x.PBJA             | -M-----               | L-I-T-----    | R-----           | I-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| SMM.US.x.PBJ_143          | -M-----               | L-X-I-T-----  | R-----           | I-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| SMM.US.x.PBJ_6P6          | -M-----               | L-X-I-T-----  | R-----           | I-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| SMM.US.x.PG53             | -M-----               | L-V-I-T-----  | R-----           | I-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| SMM.US.x.SME543           | -M-----               | L-V-I-T-----  | E-R-----         | I-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| STM.US.x.STM              | -----                 | V-----        | T-----           | I-----                   | E-K-----      |              |              |            |             |               |              |            |            |         | 320 |
| A.GM.x.ISY                | -I-----               | L-V-I-T-----  | QR-----          | TR-E-K-----              | RE-----       | E-----       | E-----       | K-----     | KR-----     | V-----        | Y-D-----     |            |            |         | 337 |
| A.GM.x.MCN13              | -I-----               | L-V-I-I-IM    | R-----           | T-E-K-----               | RE-----       | E-----       | E-----       | K-----     | K-----      | I-----        | H-D-----     | I-----     |            |         | 337 |
| A.GM.x.MCR35              | -I-----               | L-V-I-I-IM    | R-----           | T-E-K-----               | RE-----       | E-----       | E-----       | K-----     | K-----      | I-----        | H-D-----     | I-----     |            |         | 337 |
| A.GW.87.CAM2CG            | -I-----               | L-V-I-I-IM    | R-----           | T-E-K-----               | E-----        |              |              |            |             |               |              |            |            |         | 337 |
| A.GW.x.ALI                | -I-----               | L-V-I-IE-R    | R-----           | T-E-K-----               | T-RE-----     |              |              |            |             |               |              |            |            |         | 337 |
| A.GW.x.MDS                | -I-----               | L-V-I-I-IM    | R-----           | T-E-K-----               | RE-----       | E-----       | E-----       | K-----     | K-----      | I-----        | H-D-----     |            |            |         | 337 |
| A.IN.95.CR1K_147          | -I-V-----             | L-V-I-I-IM    | R-----           | T-E-K-----               | E-----        |              |              |            |             |               |              |            |            |         | 337 |
| A.IN.x.NNVA               | -I-----               | L-V-I-I-K     | R-----           | T-E-K-----               | RE-----       | E-----       | E-----       | K-E-I----- | KR-----     | V-----        | H-D-----     |            |            |         | 337 |
| A.SN.85.ROD               | -I-----               | L-V-I-I-IM    | R-----           | T-E-K-----               | E-----        |              |              |            |             |               |              |            |            |         | 338 |
| B.CI.88.UC1               | -I-NT-----            | T-V-I-K-----  | IR-----          | L-K-----                 | E-----        |              |              |            |             |               |              |            |            |         | 338 |
| B.CI.09.pPRD320-20HIV2_56 | -I-NT-----            | T-V-I-K-----  | IR-----          | L-K-----                 | E-----        |              |              |            |             |               |              |            |            |         | 338 |
| B.CI.x.EH0                | -I-NS-----            | T-V-RI-Q-E    | IR-----          | L-K-----                 | E-----        | S-----       |              |            |             |               |              |            |            |         | 337 |
| B.GH.86.D205              | -I-NT-----            | T-V-I-K-----  | IR-----          | L-K-----                 | E-----        |              |              |            |             |               |              |            |            |         | 337 |
| B.JP.01.KR020             | -I-NT-----            | T-V-I-K-----  | IR-----          | L-K-----                 | E-----        | #-----       |              |            |             |               |              |            |            |         | 338 |
| G.CI.92.ABT96             | -I-VK-----            | T-V-I-Q-----  | M-IR-----        | L-Q-----                 | D-QE-----     | T-----       |              |            |             |               |              |            |            |         | 320 |
| AB.CI.90.7312A            | -I-NT-----            | T-V-I-Q-----  | IR-----          | L-Q-----                 | E-S-----      |              |              |            |             |               |              |            |            |         | 338 |
| U.FR.96.12034             | -----                 | E-G-EAL-IGG   | IR-----          | LE-K-----                | E-----        |              |              |            |             |               |              |            |            |         | 334 |
| MAC.US.x.17EC1            | -----                 |               |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.17EFR            | -----                 |               |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.1937             | -----                 |               | N-----           |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.2065             | -----                 |               |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.251_1A11         | -----                 | L-----        | V-----           |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.251_32H_PJ5      | -----                 | L-----        |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |
| MAC.US.x.251_BK28         | -----                 | L-----        | T-----           | V-----                   |               |              |              |            |             |               |              |            |            |         | 320 |
| MAC.US.x.80035            | -----                 |               |                  |                          |               |              |              |            |             |               |              |            |            |         | 324 |



MAC.US.x.239
A.CI.88.UC2
A.DE.x.BEN
A.DE.x.PE12
A.GH.x.GH1
A.GM.87.D194
MAC.US.x.97074
MAC.US.x.MAC239.87082
MAC.US.x.MM142
MAC.US.x.SMM142B
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.SMS543
STM.US.x.STM
A.GM.x.ISY
A.GM.x.MCN13
A.GM.x.MCR35
A.GM.87.CAM2CG
A.GW.x.ALI
A.GW.x.MDS
A.IN.95.CR1K.147
A.IN.x.NNVA
A.SN.85.ROD
B.CI.88.UC1
B.CI.89.pPRD320-20HIV2.56
B.CI.x.EH0
B.GH.86.D205
B.JP.01.KR020
G.CI.92.ABT96
AB.CI.90.7312A
U.FR.96.12034
MAC.US.x.17EC1
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

Table with columns for sequence identifiers and amino acid sequences. The header is 'p51 RT end\_p15 RNase H start'. The table contains multiple rows of sequences, each corresponding to an identifier from the left column. The sequences are aligned, showing conserved and variable regions. Some positions are marked with 'X' or 'N' to indicate specific amino acids or gaps.



MAC.US.x.239  
 A.CI.88.UC2  
 A.DE.x.BEII  
 A.DE.x.PE12  
 A.GH.x.GH1  
 A.GM.87.D194  
 MAC.US.x.97074  
 MAC.US.x.MAC239\_87082  
 MAC.US.x.MM142  
 MAC.US.x.SMM142B  
 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.x.STM  
 A.GM.x.ISY  
 A.GM.x.MCR13  
 A.GM.x.MCR35  
 A.GW.87.CAM2CG  
 A.GW.x.ALI  
 A.GW.x.MDS  
 A.IN.95.CR1K\_147  
 A.IN.x.NNVA  
 A.SN.85.ROD  
 B.CI.88.UC1  
 B.CI.89.pPRD320-20HIV2\_56  
 B.CI.x.EH0  
 B.GH.86.D205  
 B.JP.01.KR020  
 G.CI.92.ABT96  
 AB.CI.90.7312A  
 U.FR.96.12034  
 MAC.US.x.17EC1  
 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

|   |      |
|---|------|
| WPITHLHTDNGANFASQEVKMWAWWAGIEHTFGVPYNPQSOGVVEAMNHHLKNQIDRIREQANSVETIVLMAVHCMNFKRRGGIGDMTPAERLINMITTEQEIQFQSKNSKFKNFRVYREGRDQLWKGPGELLWKGEGAVILKVGTDIKVVPRRKAKIKDYGGGKEV | 1003 |
| -----P-T-----V-OS-----S-----T-----L-R-N-K-Q-----V-A-----R-RO-L  | 1016 |
| -----P-T-----V-OS-----S-----T-----L-R-N-Q-----V-A-----R-RO-L  | 1016 |
| -----T-----T-OS-----TM-----LHA-L-L-Q-F-----D-V-I-----R-RR-L   | 1016 |
| -----S-T-----I-OS-----S-----T-----L-R-N-Q-----D-V-A-I-----R-RO-L  | 1016 |
| -----P-T-----I-OS-----S-----T-----L-R-N-K-Q-----D-V-A-----R-RO-L  | 1016 |
| -----X-----   | 1003 |
| -----   | 1003 |
| -----   | 999  |
| -----   | M    |
| -----   | M    |
| -----   | 999  |
| -----   | 1003 |
| -----   | 1003 |
| -----T-----M-----L-----V-----   | 999  |
| -----T-----M-----L-----V-----   | 999  |
| --K-----T-----L-V-OS-----DL-KN-K-E-L-A-----IV-L-T-YLN-Q-Q-----A-----VI-E-----RQ-M   | 996  |
| -----T-----V-OS-----T-----I-----V-----  | L    |
| -----T-----X-0-X-----X-----T-----I-----I-----V-----   | L    |
| -----T-----0-----IVS-----I-----V-----   | L    |
| -----T-----0-----I-----V-----   | L    |
| -----T-----X-0-X-----X-----T-----I-----I-----V-----   | L    |
| -----T-----0-----IVS-----I-----V-----   | L    |
| -----T-----0-----I-----V-----   | L    |
| -----T-----0-----I-----V-----R-----   | L    |
| -----T-----0-----I-----V-----   | L    |
| -----V-----T-----0-----T-----D-T-V-----K-L-----V-----V-----   | 999  |
| -----T-----V-OS-----TM-----V-V-----L-A-L-F-N-Q-----D-V-I-----R-PRO-M  | 1016 |
| -----T-----V-OS-----T-----T-----LHA-L-F-----D-V-----R-RO-L  | 1016 |
| -----T-----V-OS-----T-----T-----LHA-L-F-----D-V-----R-RO-L  | 1016 |
| -----S-V-----T-----V-OS-----S-----T-----S-----L-A-L-P-F-----D-V-II-----R-RO-L   | 1015 |
| -----S-V-----T-----V-OS-----S-----D-----S-----L-A-L-F-----D-V-I-----R-RO-L  | 1016 |
| -----T-----I-OS-----S-----D-----S-----L-A-L-F-----D-V-I-----R-RODM  | 1016 |
| -----T-----I-OS-----S-----T-----S-----L-A-L-F-----D-V-II-----R-RO-M   | 1016 |
| -----T-----V-OS-----S-----A-----L-A-L-F-----D-V-II-----R-RO-  | 1016 |
| -----T-----I-OS-----S-----T-----I-----S-----L-A-L-D-F-----LV-II-----R-RO-M  | 1017 |
| -----T-D-A-I-0-----D-V-I-V-A-----V-----F-A-L-Q-Q-----LI-E-I-----RH-   | L    |
| -----T-D-A-I-0-----E-----D-V-I-V-A-----V-----F-A-L-Q-Q-----LV-E-I-----RH-   | L    |
| -----T-D-A-I-0-----E-----D-V-I-V-A-----IV-----D-----I-E-I-----RH-   | L    |
| -----T-D-A-I-V-0-----L-D-V-I-V-A-----V-----F-A-L-Q-Q-----I-E-I-----RH-  | GL   |
| -----T-D-A-I-V-0-----IV-A-----L-T-KL-Q-----F-----I-E-I-----RH-  | L    |
| -----T-----V-0A-----S-K-K-D-----I-----I-----I-E-----999   |      |
| -----T-D-----V-0-----E-TM-A-----V-----F-A-L-Q-Q-----I-E-I-----RH-   | L    |
| -----T-----V-0-----E-TM-A-----I-S-----L-----AD-----LI-----I-----R-----L   | 1012 |
| -----   | 1003 |
| -----   | 1003 |
| -----X-----X-----X-----V-----   | 1003 |
| -----   | 1003 |
| -----S-----T-----   | M    |
| -----   | 1003 |
| -----V-----   | 999  |
| -----   | 1003 |

|                           | p31 Integrase end<br>Pol end\ |      |
|---------------------------|-------------------------------|------|
| MAC.US.x.239              | DSSSHMEDTGEAR.EVA*            | 1019 |
| A.CI.88.UC2               | ---P-L-GAR-DG.----            | 1033 |
| A.DE.x.BEN                | ---P-L-GAR-DG.-M-C            | 1033 |
| A.DE.x.PE12               | ----L-G-R-DG.----             | 1033 |
| A.GH.x.GH1                | ---HLEGARE-DG.----            | 1033 |
| A.GM.87.D194              | ----L-GAR-DG.----             | 1033 |
| MAC.US.x.97074            | -----                         | 1020 |
| MAC.US.x.MAC239_87082     | -----                         | 1020 |
| MAC.US.x.MM142            | -----                         | 1016 |
| MAC.US.x.SMM142B          | -----                         | 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           | G--AS---Q*KNSGAL--            | 1013 |
| SMM.US.x.F236_H4          | --G--L-----                   | 1016 |
| SMM.US.x.H9               | --G--L-----                   | 1016 |
| SMM.US.x.PBJ14_15         | --G--L-----                   | 1016 |
| SMM.US.x.PBJA             | --G--L-----                   | 1016 |
| SMM.US.x.PBJ_143          | --G--L-----                   | 1016 |
| SMM.US.x.PBJ_6P6          | --G--L-----                   | 1016 |
| SMM.US.x.PGH53            | --GP-L-----                   | 1016 |
| SMM.US.x.SME543           | --G--L-----                   | 1016 |
| STM.US.x.STM              | --G--L---R--G----             | 1016 |
| A.GM.x.ISY                | --G--L-GAR-DG.-M--            | 1033 |
| A.GM.x.MCN13              | --G--L-GAR-DG.-M--            | 1033 |
| A.GM.x.MCR35              | --G--L-GAR-DG.-M--            | 1033 |
| A.GW.87.CAM2CG            | ----L-GAR-NG-----             | 1032 |
| A.GW.x.ALI                | --GP-L-GAR-DG.----            | 1033 |
| A.GW.x.MDS                | --GP-L-G-R-DG.----            | 1033 |
| A.IN.95.CR1K_147          | --GP-L-G-R-DG.----            | 1033 |
| A.IN.x.NNVA               | --GP-L-G-R-DG.----            | 1033 |
| A.SN.85.ROD               | --G--L-GAR-DG.-M--            | 1034 |
| B.CI.88.UC1               | -CGTD---RQ---M-Q              | 1034 |
| B.CI.09.pPRD320-20HIV2_56 | -CGADV---RQ---M-Q             | 1034 |
| B.CI.x.EH0                | -C-ADV---MQ---M-Q             | 1033 |
| B.GH.86.D205              | -C-AD---RQ---M-Q              | 1033 |
| B.JP.01.KR020             | -C-AN---Q---M-Q               | 1034 |
| G.CI.92.ABT96             | ---TN---RQTG.----             | 1016 |
| AB.CI.90.7312A            | -C-TDV---RQ---M-Q             | 1034 |
| U.FR.96.12034             | ---NL-GAEKV-.M-L              | 1029 |
| 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         | -----                         | 1020 |
| MAC.US.x.251_32H_PJ5      | -----                         | 1020 |
| MAC.US.x.251_BK28         | -----                         | 1016 |
| MAC.US.x.80035            | -----                         | 1020 |

Vif start

```

MAC.US.x.239
A.CI.88.UC2
A.DE.x.BEN
A.DE.x.PE12
A.GH.x.GH1
A.GM.87.D194
A.GM.x.ISY
A.GM.x.MCN13
A.GM.x.MCR35
A.GW.86.FG
A.GW.87.CAM2CG
A.GW.x.ALI
A.GW.x.MDS
A.IN.95.CR1K.147
A.IN.x.NNVA
A.SN.85.R0D
B.CI.88.UC1
B.CI.89.pPRD320-20HIV2.56
B.CI.x.EH0
B.GH.86.D205
B.JP.01.KR020
G.CI.92.ABT96
AB.CI.90.7312A
AB.CM.03.03CM_510_03
U.FR.96.12034
MAC.US.x.17EC1
MAC.US.x.251.1A11
MAC.US.x.251.32H.PJ5
MAC.US.x.251.BK28
MAC.US.x.17EPR
MAC.US.x.MN142
MAC.US.x.SMM142B
MAC.US.x.r00025
MAC.US.x.r00131
MNE.US.82.MNE.8
MNE.US.x.MNE027
SMM.SL.92.SL92B
SMM.US.x.F236.H4
SMM.US.x.PB114.15
SMM.US.x.PBJA
SMM.US.x.PBJ.6P6
SMM.US.x.PGM53
SMM.US.x.SME543
STM.US.x.STM
MEEKKRWIAPVPTWRIPERLE.RWHSLIKLYKLYKTKDLQKVCVPHFKVGVWAWMTCSRVIPLQEGSHLEVOGVYHWTPEKGLWSTYAVRITWYKSNFMTDVTPNYADILLHSTYFPCFTAGEVRRAIRGEQLLSCCRFPRAHKYQVPSLQYLAKVVS..DVRSOGENPT 167
---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---Q0.ND.P---RKG.A 168
---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---Q0.ND.P---RKGTA 168
---GE.V---V.G.M.K.V---HR---EG---H---GN.I.A.N---S---TER---DC.S.I---S---K---NY.Q---RSK.L---F---V---Q0.NG.P---KNST. 168
---G.N.VPTWRVPGR.M---V---R---R---EE.R---H---KGE.I.A.N---SHS.L---TER---D---I---S---K---NY.Q---V---V---Q0.ND.P---RKGTA 168
---G.N.V---V.G.M---V.H---R---EE.R---H---EGE.I.A.N---SHS.L---TEK---DC.S.I---S---K---NY.Q---A---V---Q0.NG.P---RKGAA 168
---DQG---V---V.G.M.K---R---ED---R---H---KGN.I.A.N---S.S.M---TER---DC.S.I---S---K.S---NY.Q---SK---F---V---Q0.NDKP.RD.T 168
---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---Q0.ND.P---RNGTP 168
---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---Q0.ND.P---RNGTP 168
---G.V---I---V.G.M---V---R---E---H---K.N.I.A.N---SHS.L---TEK---DC.T.I---S---K---KY---RS---F---V---Q0.ND---NSA. 168
---G.S---V---V.G.M.K.V---G---EQ---H---RGD.R.I.A.N---S---M---TEK---DC.T.I---S---K---KY---RS---F---V---Q0.ND.P---RDRT. 168
---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---S---F---V---Q0.NGKP.RNST. 168
---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---Q0.NG.P---RDST. 168
---D---V---KV.G.M.K.V---E---H---KDN.I.I.A.N---E---S---TE---DCT.S.I---A---K.I---NY.Q---RT---T---F---V---Q0.ND---RDGT. 168
---G.V---V.G.M.K.V---R---E---H---NNN.N.I.A.N---S---TEK---D---S.I.G---A---K.V---NY.Q---T---T---F---V---Q0.ND---RDST. 168
---D---V---V.G.M.K.V---E---H---KGN.I.A.N---S.S---TEK---DC.V.I---K---NY---RA---F---V---Q0.ND.P---RDST. 168
---G.N.V---G---K.V---HR---E---Q.S---H---K.EAY---N---R.F.S---L---KRS.Y---DV.Q---GS.S---N---K.I---Y.NY.S---EG---F---R---IQEGKD.G---SA. 169
---N.V---G---V---HR---E---Q.S---H---K.AY---I---N---R.L.S---L---RES.Y---DV.Q---GS.S---N---K.I---Y.NY.S---DR---F---R---QEGKNG---SA. 169
---N---C---R---O.S---H---K.A---N---R.F.S---L---ERS.Y---DV.R---GS.SS---N---K.I---H.NY.S---TG---F---R---QEGKNG---ST. 169
---D.V---G---R.GE---O.S---H---I---NK.AW---N---R.F.S---L---ER---Y---DV.Q---GS.S.S---N---K.I---Y.NY.S---EG---F---R---QEGKNG---SA. 169
---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---K.I---Y.NY.S---RG---F---R---QEGNGP---SAA. 169
---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
---G.N.V---K.G---K.V---HR.GE---O.S---H---KG.AC---N---R.F.S---L---EES.Y---DV.Q---G.S---N---K.I---Y.NY.S---EG---F---R---QEGKNG---SA. 169
---D.V---K.G.A---R---GE---O.S---H.X---DKXAX.X---N---R.L---L---XER---Y---DV.Q---GS.S---N---K.I---Y.NY.S---EG---F---R---IOEGKNG---K.T. 169
---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
MAC.US.x.251.1A11
MAC.US.x.251.32H.PJ5
MAC.US.x.251.BK28
MAC.US.x.17EPR
MAC.US.x.MN142
MAC.US.x.SMM142B
MAC.US.x.r00025
MAC.US.x.r00131
MNE.US.82.MNE.8
MNE.US.x.MNE027
SMM.SL.92.SL92B
SMM.US.x.F236.H4
SMM.US.x.PB114.15
SMM.US.x.PBJA
SMM.US.x.PBJ.6P6
SMM.US.x.PGM53
SMM.US.x.SME543
STM.US.x.STM
---G---V---V---SR.IVE---C---FH---RE.E.A---H---Y.A---E---V---N---S---EK---DV.Q.A---A.HA.Q---V---Y.GYAV---HSS.Q---L---LQ.ND.PK.K--- 169
---N.V---G---K---H---N---MA---H---RDET---N.A---S---R---D---T---SE---K---K---K---N---T---H---D--- 167
---N.V---G---K---H---N---A---H---KDEA---N---R---D---T---G---SE---K---K---K---N---T---H--- 167
---N.V---G---K---H---N---A---H---KDEA---N---R---D---T---G---SE---K---K---K---N---T---H--- 167
---N.V---K.G---K---FR---A---H---RDE---N.A---H---RK---D---T---G---SE---Q---K---K---K---N---T---H---R.D--- 167
---N.V---K.G---K---FR---A---H---RDKT---N---H---R---D---T---G---SE---Q---K---K---K---N---T---H---R.D--- 167
---V---G---H---N---E.S.A---H---GEA---N---E---TR---S---DC.Q---G---K---TK---N---E---H---R---TA 167

```

|                           | Vif end   |     |
|---------------------------|---|-----|
| MAC.US.x.239              | WKQWRDRNRRLMAKQNSRGDKQRGGKPPTKGANFPLAKVLGILA*   |     |
| A.CI.88.UC2               | R----HW---V-R-DY-SLE-----APR-H--V---E---        | 214 |
| A.DE.x.BEN                | R----HW---V-REDH-SL--G-SE-SAPR-H--V---E---      | 216 |
| A.DE.x.PE12               | 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.GM.87.D194              | R----HW---V-R-DY-SL--G-SE-SAPR-H--V---E---      | 216 |
| A.GM.x.ISY                | R----NY---L-R-DG-SH---SE--AQ--Y--V---E---       | 216 |
| A.GM.x.MC113              | R----Y---QL--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                | R-R-G-Y---R-D--Y---SES-PTR-H--E-E-E---          | 216 |
| A.GW.87.CAM2CG            | R----Y---L-R-D--SY---SES-AP--Y--V---E---        | 216 |
| A.GW.x.ALI                | R----Y---V-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.95.CR1K 147          | R-R--Y---L-R-DG-SH---ES-AQ-TH---E-----          | 216 |
| A.IN.x.NNVA               | R-R--Y---L-REDG-SH---SES-AQ--H--V---E---        | 216 |
| A.SN.85.R0D               | 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.09.pPRD320-20HIV2_56 | R-R-G-S---I--RK-TSRTQ-GSSQSFAPRTH-----E-----    | 217 |
| B.CI.x.EH0                | R-R--NS---I--RD-I-TSQ-SSSQSLAQ-TY-----E-----    | 217 |
| B.GH.86.D205              | R-R--NS--SI-L-RK--NNRAQ-GS-Q-FAPRTH-----E-----  | 217 |
| B.JP.01.KR020             | 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.CI.90.7312A            | R-R--S---I--RK-NSRTQ-GSSQ-FAPRTH-----E-----     | 217 |
| AB.CM.03.03CM_510_03      | RE-R--N---SI-X-RK--NNRAQ-GSSQ-LAPRTH-----E----- | 217 |
| U.FR.96.12034             | R-FR--NC---FQL-RK-CERYQ-GSSATSSQ-TY-----A----   | 214 |
| MAC.US.x.17EC1            | -----S-----S-----D-----                         | 215 |
| MAC.US.x.251_1A11         | -----S-----S-----D-----                         | 215 |
| MAC.US.x.251_32H_PJ5      | -----S-----S-----D-----                         | 215 |
| MAC.US.x.251_BK28         | -----S-----S-----D-----                         | 215 |
| MAC.US.x.17EPR            | -----S-V-----E-----                             | 215 |
| MAC.US.x.MM142            | -----S-V-----E-----                             | 215 |
| MAC.US.x.SMM142B          | -----S-----S-----D-----                         | 215 |
| MAC.US.x.r80025           | -----S-----S-----VD-----                        | 215 |
| MAC.US.x.r90131           | -----S-----S-----D-----                         | 215 |
| MNE.US.82.MNE_8           | -----S-----S-----D-----                         | 215 |
| MNE.US.x.MNE027           | -----S-----S-----D-----                         | 215 |
| SMM.SL.92.SL92B           | R----N-----VT-Y-G-PESNSSTT--RV-----E-----GR     | 217 |
| SMM.US.x.F236_H4          | -----N-K-----RN--GSS-S-AE-----E-----            | 215 |
| SMM.US.x.PB114_15         | -----N-----L-R-----RN--GSSEFAE-T-----V-         | 215 |
| SMM.US.x.PBJA             | -----N-----L-R-----RN--GSSEFAE-T-----E-----     | 215 |
| SMM.US.x.PBJ_6P6          | -----N-----L-R-----RN--GSSEFAE-T-----E-----     | 215 |
| SMM.US.x.PGM53            | -----N-----I-----RN--GSSESLAE-----E-----        | 215 |
| SMM.US.x.SME543           | -----GN-----RN--GSS-S-AE-----E-----             | 215 |
| STM.US.x.STM              | R----G--GSI-V-T--G--H-P--S--S-E-TD-----E-----   | 215 |

|                           | Vpx start   | Vpx end |
|---------------------------|---|---------|
| MAC.US.x.239              | MSDPREIRPPGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVWQRSWEYWHDEQGMSPSYVKRYRLCLIQKALFMHCKKGCRCLEGEHGAGGWR.PGPPPPPPPLA*                             | 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-----E-----I-----AL-----R-----A-----T-----M-----I-----T-----F-----R-----T-----W-----DM-----RE-----LEDQ-----V-----                     | 114     |
| A.DE.x.PE12               | -A-----K-----TV-----G-----V-----A-----E-----M-----A-----R-----DL-----Q-----T-----R-----M-----Y-----M-----I-----V-----T-----G-----P-----V----- | 112     |
| A.GH.x.GH1                | -T-----V-----D-----I-----AL-----R-----D-----T-----M-----V-----I-----F-----R-----T-----G-----P-----S-----V-----                                | 113     |
| A.GM.87.D194              | -A-----V-----D-----I-----AL-----A-----D-----T-----T-----IM-----V-----Y-----I-----F-----T-----R-----P-----V-----                               | 113     |
| A.GM.x.ISY                | -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-----D-----T-----M-----T-----R-----T-----G-----P-----V-----  | 113     |
| A.GM.x.MCR35              | -T-----TV-----E-----N-----D-----I-----A-----R-----R-----Q-----T-----M-----T-----R-----T-----G-----P-----V-----                                | 113     |
| A.GW.86.FG                | -T-----TV-----E-----A-----D-----A-----H-----R-----T-----T-----M-----MY-----A-----RDGT-----G-----M-----QK-----GD-----Q-----V-----              | 113     |
| A.GW.87.CAM2CG            | -T-----TV-----E-----A-----D-----A-----R-----R-----Q-----T-----M-----V-----I-----F-----R-----T-----G-----P-----V-----                          | 113     |
| A.GW.x.ALI                | -AN-----TV-----D-----AL-----R-----Q-----T-----M-----MY-----T-----FM-----T-----G-----P-----S-----V-----  | 113     |
| A.GW.x.HDS                | -A-----TV-----EQ-----A-----D-----A-----R-----R-----T-----T-----M-----MY-----I-----GR-----T-----R-----P-----V-----                             | 113     |
| A.IN.95.CR1K.147          | -T-----TV-----A-----D-----A-----G-----A-----T-----M-----I-----I-----VRR-----T-----G-----P-----S-----V-----                                    | 113     |
| A.IN.x.NNVA               | -T-----TV-----A-----E-----A-----G-----A-----T-----M-----M-----I-----ARR-----T-----G-----G-----S-----L-----WV-----                             | 113     |
| A.SN.85.R0D               | -T-----TV-----A-----E-----A-----R-----R-----E-----T-----I-----VY-----VR-----T-----R-----P-----V-----  | 113     |
| AB.CI.90.7312A            | -V-----D-----V-----A-----E-----IV-----L-----A-----R-----D-----I-----T-----L-----M-----V-----FA-----T-----RG-----P-----R-----                  | 112     |
| AB.CM.03.03CM_510_03      | -V-----D-----X-----E-----IT-----L-----V-----C-----A-----R-----S-----T-----L-----M-----M-----V-----YT-----T-----RQ-----X-----S-----            | 112     |
| B.CI.88.UC1               | -V-----D-----D-----E-----IT-----L-----V-----C-----A-----RE-----S-----T-----L-----M-----M-----V-----YT-----Q-----P-----S-----                  | 112     |
| B.CI.09.pPRD320-20HIV2_56 | -V-----D-----V-----A-----E-----LA-----L-----K-----K-----A-----R-----I-----T-----L-----M-----MY-----IA-----T-----RR-----P-----Q-----           | 112     |
| B.CI.x.EH0                | -V-----D-----V-----E-----T-----L-----HL-----V-----K-----A-----RE-----I-----T-----M-----M-----I-----FA-----G-----R-----P-----S-----            | 112     |
| B.GH.86.D205              | -V-----D-----V-----A-----E-----IT-----L-----V-----A-----RE-----I-----T-----L-----M-----M-----V-----YT-----Q-----P-----S-----                  | 112     |
| B.JP.01.KR020             | -V-----D-----V-----E-----IT-----L-----V-----A-----E-----TI-----T-----M-----Y-----LAR-----A-----R-----P-----S-----                             | 112     |
| G.CI.92.ABT96             | -X-----E-----D-----E-----XT-----I-----R-----A-----T-----L-----M-----M-----V-----T-----QK-----P-----Q-----                                     | 112     |
| U.FR.96.12034             | -G-----E-----V-----S-----E-----RD-----A-----Q-----Q-----R-----E-----T-----L-----M-----M-----V-----Q-----P-----S-----I-----                    | 113     |
| MAC.US.x.17EC1            | .....X.....   | 113     |
| MAC.US.x.17EFR            | .....X.....   | 113     |
| MAC.US.x.1937             | .....X.....   | 113     |
| MAC.US.x.2065             | .....X.....   | 113     |
| MAC.US.x.251.1A11         | .....V-----Q-----   | 113     |
| MAC.US.x.251.32H.PJ5      | .....V-----Q-----   | 113     |
| MAC.US.x.251.BK28         | .....V-----Q-----M-----   | 113     |
| MAC.US.x.MM142            | .....V-----Q-----T-----   | 113     |
| MAC.US.x.r80025           | .....V-----Q-----T-----   | 113     |
| MAC.US.x.r90131           | .....V-----Q-----T-----   | 113     |
| MAC.US.x.5M0142B          | .....V-----Q-----T-----   | 113     |
| MNE.US.02.MNE.8           | .....V-----Q-----   | 113     |
| MNE.US.x.MNE027           | .....V-----Q-----   | 113     |
| SMM.SL.92.SL92B           | -T-----K-----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-----D-----D-----A-----R-----Y-----M-----E-----T-----V-----M-----E-----T-----E-----T-----   | 113     |
| SMM.US.x.PB14.15          | -----D-----D-----A-----R-----Y-----M-----V-----T-----M-----V-----T-----GE-----T-----  | 113     |
| SMM.US.x.PB3A             | -----D-----D-----A-----R-----M-----V-----T-----M-----V-----T-----GE-----T-----  | 113     |
| SMM.US.x.PBJ.6P6          | -----D-----D-----A-----R-----M-----V-----T-----M-----V-----T-----GE-----T-----  | 113     |
| SMM.US.x.PG053            | -----E-----G-----A-----R-----M-----E-----T-----V-----R-----E-----T-----S-----   | 113     |
| SMM.US.x.SME543           | -----E-----G-----A-----R-----M-----E-----T-----V-----R-----E-----T-----S-----   | 113     |
| STM.US.x.STM              | -----A-----E-----H-----D-----R-----PG-----S-----R-----  | 113     |

|                           | Vpr start  | Vpr end |     |
|---------------------------|--|---------|-----|
| MAC.US.x.239              | M...EERPPENEGPQREPWDEWVVEELKEEALKHFDPRLLTALGNHIYNRHGDTLEGAGELIRLQRALFMHFRGGCIHSRIGQPGGGNPLSAIPPSRSM* |         | 101 |
| A.CI.88.UC2               | .....#GRS-MIV-I---*R-----I-D-Y-----N-----R-K-----S-L-V-GSR---*TRRR-CP-T-TP-G-H                       |         | 82  |
| A.DE.x.BEN                | -TEAPT-F---DGT-R-DLGS-D-I-T-R-I---R-----I---Y-H-----R-KT---V--A-NR---*TRRR-CP-A-TP-A-H               |         | 105 |
| A.DE.x.PE12               | -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                  |         | 106 |
| A.GH.x.GH1                | -TEAPT-F---DGT-R-LGGD-IRI-G-I-----I---Y-HS-----P-R-----V-L-A-NR---S-TRRR-FP-A-TP-G-Y                 |         | 106 |
| A.GM.87.D194              | -TEAPT-F---DGT-R-LGST-I-T-K-I-----C-I---Y-----R-V---V-I-A-DR--K-TRRR-CP-A-TP-G-H                     |         | 106 |
| A.GM.x.ISY                | -TEAPA-F---DGT-P-G---I-I-R-I-----Y-Y-T-----R-V---T-A-G---R-R-----TP-N-Q                              |         | 106 |
| A.GM.x.MCN13              | -TEAPT-F---DGT-P-G---II-I-RKI-K-----Y-HT-----R-NV---A-RL---T-RT-TP-TSTP-T-Q                          |         | 106 |
| A.GM.x.MCR35              | -TEAPT-F---DGT-P-G---II-I-RKI-K-----Y-HT-----R-NV---A-RL---T-RT-TP-TSTP-T-Q                          |         | 106 |
| A.GW.86.FG                | -TEAPT-L---DRT-P-G-A-I-I-R-IE---R-----RY-T-----R---A-A-G---TR---TP-G-H                               |         | 106 |
| A.GW.87.CAM2CG            | -TEAPT-L---DGT-P-G---I-I-RDI-----G-A-----R-R-V---T-A-N---TR---TP-R                                   |         | 105 |
| A.GW.x.MDS                | -TEAPT-L---DGT-P-G---I-I-R-IT-----RS-----R-T-----R-V---T-A-G---TR---TP-Q                             |         | 106 |
| A.IN.95.CRIK_147          | -TEAPT-L---DGT-P-G---I-I-R-I-----R-----RY-D-Y-----R-V---T-A-G---R---T-TP-R                           |         | 106 |
| A.IN.x.NNVA               | -TEAPT-L---DRT-P-G---I-I-R-I-----R-----RY-T-Y---P-R-V---T-A-G---R---TP-R                             |         | 106 |
| A.SN.85.ROD               | -AEAPT-L---VDGT-L-G---II-I-R-I-----I---KY-T-----R-KV---T-A-G---TR---TP-N-Q                           |         | 106 |
| B.CI.88.UC1               | -AEAAP-T---S-----E-ED-M-I-Q---R-----F-S---A-K---L-A-Q  |         | 102 |
| B.CI.09.pPRD320-20HIV2_56 | -AEAAP-T---D-----E-ED-I-Q---R-----F-S---A-K---L-Q  |         | 102 |
| B.CI.x.EH0                | -EAAP-I---DKN-----EQ---I-Q---R-----F-S---N-KL---L-Q  |         | 102 |
| B.GH.86.D205              | -AEAAP-I---N-----E-IG-I-I-Q---R-----F-S---A-K---L-A-Q  |         | 102 |
| B.JP.01.KR020             | -TEAAP-T---D-----E-ED---Q---R-----F-G---IA---L-Q   |         | 102 |
| AB.CI.90.7312A            | -AEAAS-T---D-----E-ED---I-Q---R-----F-S---A-K---L-Q  |         | 106 |
| G.CI.92.ABT96             | -AEEI-T---D-A-----I-----I-----Y-D-----K---L-D-RC-GNA-T---GVF   |         | 103 |
| U.FR.96.12034             | .....A-I---DGA---G---R-I---VR-----Y-----K---I-L---R-SSR---TTV  |         | 98  |
| MAC.US.x.17EC1            | .....  |         | 102 |
| MAC.US.x.17EFR            | .....X-X-----T-----X   |         | 102 |
| MAC.US.x.1937             | .....  |         | 102 |
| MAC.US.x.2065             | .....  |         | 102 |
| MAC.US.x.251.1A11         | .....*---G-----G-----R   |         | 101 |
| MAC.US.x.251.32H.PJ5      | .....N-----T---GV  |         | 102 |
| MAC.US.x.251.BK28         | .....N-----T   |         | 98  |
| MAC.US.x.80035            | .....  |         | 102 |
| MAC.US.x.97074            | .....  |         | 102 |
| MAC.US.x.MAC239.87082     | .....  |         | 102 |
| MAC.US.x.MM142            | .....K-----I---S-S-----T   |         | 102 |
| MAC.US.x.SM1142B          | .....K-----I---S-S-----T   |         | 102 |
| MAC.US.x.r80025           | .....  |         | 102 |
| MAC.US.x.r90131           | .....  |         | 102 |
| MNE.US.82.MNE.8           | .....D-----Y-D-----K-----T---S---T---R   |         | 102 |
| MNE.US.x.MNE027           | .....D-----Y-D-----K-----T---S---T   |         | 102 |
| SMM.SL.92.SL92B           | .....H-A-D-TNP-----IRD-----YV-DTY---I---I-K-K-L-H-T---R-P---GS-SA-DV                                 |         | 101 |
| SMM.US.x.F236.H4          | .....A---D-A-----Y-D-----I---S-A---S-----GV  |         | 90  |
| SMM.US.x.PB114.15         | .....T---D-A-----I---N-----Y-D-----I---R---S-----T---GV  |         | 102 |
| SMM.US.x.PBJA             | .....T---D-A-----I---N-----Y-D-----I---R---S-----T---GV  |         | 102 |
| SMM.US.x.PBJ.6P6          | .....T---D-A-----I---N-----Y-D-----I---R---S-----T---GV  |         | 102 |
| SMM.US.x.PGM53            | .....T---D-A-----K---V-----Y-D-----I---A---SR---T---GV   |         | 102 |
| SMM.US.x.SME543           | .....A---D-A-----I---N-----Y-D-----I---S-A---SR---T---A  |         | 102 |
| STM.US.x.STM              | .....TH---D-----I-Q---R-----S---Y-D-----K-----R-----AT---T-GV  |         | 102 |

|                           | Tat start   | exon 1 | end | exon 2 start | Tat end |     |
|---------------------------|---|--------|-----|--------------|---------|-----|
| MAC.US.x.239              | METPLRQENSLLESSNRSSCISEADASTPESANLGEETLSQLYRPLEACYNCTCYCKCCYHCQFCFLKGLGICYEQSRKRRRTPKKAKANTSSASMKPISNRTRHCOPEKAKKETVEKAVATAPGLGR* |        |     |              |         | 130 |
| A.CI.88.UC2               | -----KAP-S-K-Y-P-P-T-MEVAQA-L-KQ-L-A-H-----P-T-P-----SF-L-S-----W-VKGG-----RT-THPP-TPD-S-ID-GDSR-T-KQ--P-TP--TS----               |        |     |              |         | 131 |
| A.DE.x.BEN                | -----KAP-S-KPY-P--T-R-VTAQ-L-KQ-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----              |        |     |              |         | 131 |
| A.DE.x.PE12               | -----SKAP-S-M-C-P--T-Q-VKSO-L-KQ-RL-----Q-----N-P-----L-----L-----RKGR-----HS-----D-S-T-GNS-T-KQTK-P-T-LE-R-----Q----             |        |     |              |         | 131 |
| A.GH.x.GH1                | -----H-KAP-S--Y-P--T-OGVTAQ-L-KQ-L-H-----T-S-----Q-SF-L-----W-ARKSR-----R-T-THS-----D-S-T-GDS-T-EQ-K-T-TTMV-TCS----               |        |     |              |         | 131 |
| A.GM.87.D194              | -----K-P-S--Y-P--T-R-VTAQ-R-KQ-L-A-H-----T-S-----Q-SF-L-----W-ARQGR-----R-T-THPPP-D-S-T-GDS-T-KQ-K-P-TT-VS-C--H----               |        |     |              |         | 131 |
| A.GM.x.ISY                | -----KAP-S-G-Y-P-RT-Q-VA-Q-L-Q-----T-N-K-F-G-F-L-N-----W-DRKGR-----S-T-HS-P-----D-S-T-GNS-T-KQ-K-LGTTLEAD-----                    |        |     |              |         | 131 |
| A.GM.x.MC13               | -----KAP-S-M-Y-P--T-R-VGSO-L-KQ-L-H-----P-N-K-----G-F-L-N-----W-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-----W-DRKGR-----T-HS-----D-S-T-GNS-----KQ-K-L-TTLE-R-----               |        |     |              |         | 131 |
| A.GW.86.FG                | -----KAP-S--C-G-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-GQ-T-Q-L-K-----Q-----E-D-S-----R-----L-----W-DRKGR-----HS-----D-S-T-NS-----A-KQ-K-L-AT-E-D-----               |        |     |              |         | 131 |
| A.GW.x.ALI                | -----K-PGS-MPY-P--T-Q-VAQ-L-KQ-----T-N-----E-----L-N-----W-DRKGR-----S-I-HS-----D-S-T-NS-----EKO-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-----THP-----D-S-T-GNS-T-KQ-K-L-AT-E-DL-P-----                   |        |     |              |         | 131 |
| A.IN.95.CR1K 147          | -----KAP-S--C-PFLHT-Q-VA-QG--Q-----AS-----L-----L-Q-----W-RKGR-----T-THPTLT-D-S-T-WNS-----E-K-L-TT-V-DLDP-----                    |        |     |              |         | 130 |
| A.IN.x.NNVA               | -----KVP-S--C-PF-HT-Q-VA-Q-VQE-----FQ-----D-S-----F-EL-Q-----W-DRKGR-----R-T-HP-TPD-S--T-NS-T-EQ-K-L-TT-D-DL-P-----               |        |     |              |         | 131 |
| A.SN.85.R0D               | -----KAP-S-K-C-PF-RT-Q-VA-Q-L-RQ-----T-N-S-----R-----M-N-----W-RKGR-----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-DH-----SS-R-VTA--DESL-AN-GDS-T-KQ-TK--TKGL-DL-P-----                         |        |     |              |         | 129 |
| B.CI.89.pPRD320-20HIV2 56 | -----I-SQ-----RF-----ST-GEG--QGLD-Q-----N-----F-----L-----W-DH-----SS-R-VTA--PDESL-TS-GDS-T-KQ-TK--EKGI-DL-P-----                 |        |     |              |         | 129 |
| B.CI.x.EH0                | -----I-K-S-N-SGH-ST-GV-N-QGLD-R-----K-S-----S-----L-----W-DRKGR-----R-----P-ESL-A-GDS-T-KQ-KE--TTR-D-P-S-----                     |        |     |              |         | 129 |
| B.GH.86.D205              | -----I-Q--S-K--P-ST-PVVN-QGLD-Q-----K-D-----L-----L-----W-DR-----SA-R-TTAP-PD-SL-A-GDS-T-KQ-KE--TTGT-D-P-----                     |        |     |              |         | 129 |
| B.JP.81.KR020             | -----GI-Q-----F-S--ST-EG-N-RGLD-Q-----R-K-----L-----L-----W-DH-----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-KK*-K-TK--DL-----                         |        |     |              |         | 129 |
| AB.CI.90.7312A            | -----I-QG-----RF-S--ST--VVN-QGLD-Q-----H-R-N-----L-----L-----W-DR-----SS-RN-TA--D-SV-T-NS-T-KQ-KE--TTG-DL-P-S-----                |        |     |              |         | 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-DH-----R-TTA-P-PD-S--T-DS-A-KQ-KE--TTGT-DL-P-----           |        |     |              |         | 127 |
| U.FR.96.12034             | -----I-K--S-K-CK-Q-ST-G-VG-QG-AA-----S-K-----L-----L-----P-R--S--V--YS-AP--SL-G-NS-KEK*-KAL-T--DL-----                            |        |     |              |         | 128 |
| MAC.US.x.251 1A11         | -----SGH--T--A-----L-----   |        |     |              |         | 131 |
| MAC.US.x.251 32H_PJ5      | -----T-----   |        |     |              |         | 131 |
| MAC.US.x.MM142            | -----Y--A-AI-----S--K-HR-----E-P-I-L-K-----A-----   |        |     |              |         | 131 |
| SMM.US.x.F236 H4          | -----K-----CR-H-S-V-VP-----K--R--H-----H-R-----T--PFP--SL-T-A-NR-K-E-----TE--ADL-----   |        |     |              |         | 117 |
| MAC.US.x.SMM142B          | -----Y--A-AI-----S--K-HR-----E-P-I-L-K-----A-----   |        |     |              |         | 131 |
| MAC.US.x.r80025           | -----P-----P-----   |        |     |              |         | 131 |
| MAC.US.x.r90131           | -----   |        |     |              |         | 131 |
| MNE.US.82.MNE 8           | -----K--G--T--A-P-L-----E-----E-----R-----D--K--K-EQ--A-----  |        |     |              |         | 131 |
| MNE.US.x.MNE027           | -----K--G--T--A-T-L-----E-----E-----R-----V-----D--K--K-EQ--T--V-----   |        |     |              |         | 131 |
| MAC.US.x.17EC1            | -----   |        |     |              |         | 131 |
| MAC.US.x.17EFR            | -----   |        |     |              |         | 131 |
| SMM.SL.92.SL92B           | -----TGIQ*KG-GK*-RYYKRHC-YTLGM-PIQGLDSQ-DQ-PWD-----T--S-F--F--L--R-----T-AKP-R--VK-----PFDT-QS-SG-N--K-EQ--K--TE-D-DC-----        |        |     |              |         | 127 |
| SMM.US.x.PBJ14_15         | -----K--S--R-H-S-V-D--S-E-----K--R--H-----O-R-----T--D-SL-R-A-N--K-E-----AE--DL-----  |        |     |              |         | 129 |
| SMM.US.x.PBJA             | -----K--S--R-H-S-V-D--S-E-----K--R--H-----O-R-----T--D-SL-R-A-N--K-E-----AE--DL-----  |        |     |              |         | 129 |
| SMM.US.x.PBJ 6P6          | -----K--S--R-H-S-V-D--S-E-----K--R--H-----O-R-----T--D-SL-R-A-N--K-E-----AE--DL-----  |        |     |              |         | 129 |
| SMM.US.x.PGM53            | -----K--G--YT-E-VL-----KVE--C-----S-----L-----F-P-ESL-T-A-NS-K-E-----TE--DL-----  |        |     |              |         | 131 |
| SMM.US.x.SME543           | -----K--YR-H-S-V-VP-----E-----P--K--R--H-----HR-----T-T-PLP--SL-T-NR-K-E--K-TE--ADL-----  |        |     |              |         | 129 |
| STM.US.x.STM              | -----K--S--R--S-P--T--V-A--GL--QE--W-----E-C-K-F-----L-VT-----T-R-R--VK-----TYPI--SL-T-A-NS-K-EQ-KE--TE-EST--K-----               |        |     |              |         | 129 |

|                           | Rev start   | exon 1 end | exon 2 start | Rev end |     |
|---------------------------|---|------------|--------------|---------|-----|
| MAC.US.x.239              | MSIHREEEELRKLRLIHLHQINPYTPGTANORRQRKRWRRRRQQLLALADRIYSFDPPTDTPLDLAIQ0LONLAIESIPDPPTNTPEALCDPTEDSRSPQD   |            |              |         | 107 |
| A.CI.88.UC2               | -CEKAD--ORE--R--YR--G--N-R--R--Q-LR--KLHTA--A-SS--W--H--G-T-REL--DL-S--DSNQLAET   |            |              |         | 103 |
| A.DE.x.BEN                | -ERAD--G-QGK-LR--0--S--N-R--R--Q-LR-V--NKLCAV--S--R--H--R-T-OEL--DL-S--NSNOGLAET  |            |              |         | 103 |
| A.DE.x.PE12               | -NGRAD--G-ORQK--R--0-L--R--N-R--RKQH--R-V--NS--T--A-S--R--R--G-T-OEL--L--SSESTNNQGLAETYNLSLPAIWRVDPDRSAPGPKDYERDSCEVERLTVGGNGTDROGNTCS5KQDQAGGRTCPVPR |            |              |         | 170 |
| A.GH.x.GH1                | -HEKADG--QE--R--H--S--N-R--Q--LR-V--KL-T--S--R--D--R-T-HEL--DL-S--NSNOGLAET   |            |              |         | 103 |
| A.GM.87.D194              | -RORAD--G-DEK--R--I--H-Q--S--N-R--R--Q-FR-V--TKLHTI--S--R--G--T-OEL--DL-S--NSNOGLAET  |            |              |         | 103 |
| A.GM.x.ISY                | -TERAD--GV-RK--R--0-L--R--N-R--E--K-I--T--A-P--QT--G--T-OEL--DL-S--LAETQGSLLPAVWRVDPDRSVPGPREGYKRDSYERGEELVGGSGTNRKGDTRSSTKDQAGSRNCPVPR               |            |              |         | 159 |
| A.GM.x.MCN13              | -TGRAD--RK--R--0-L--S--N-R--KQ--R-V--K-T--A-S--GRT--H--E-T-ODL--SF--SSEGTDSNO-LAKN  |            |              |         | 107 |
| A.GM.x.MCR35              | -TGRAD--RK--R--0-L--S--N-R--KQ--R-V--K-T--A-S--RT--H--E-T-ODL--SF--SSEGTDSNO-LAKN   |            |              |         | 107 |
| A.GW.86.FG                | -TERAD--G-ORR--R--0--S--N-R--KQ--R-I--S-T--A-S--R--H--G-T-ODL--L--SPESTNSNO-LAEA  |            |              |         | 107 |
| A.GW.87.CAM2CG            | -TERAD--G-ORR--R--0--S--N-R--KQ--R-I--S-T--A-P--RT--H--E-T-ODL--HP--SQ--LAEA  |            |              |         | 100 |
| A.GW.x.ALI                | -TERAG--D-ORR--R--0--S--N-R--Q--G-IV--FT--ASS--R-V-H--G-T-ODL--DL--SSESADNNQGLAET   |            |              |         | 107 |
| A.GW.x.MDS                | -NERAD--G-ORR--R--0--R--N-R--Q--R-I--S-T--A-S--QTV--G--T-ODL--HL--SQ--PAAT  |            |              |         | 100 |
| A.IN.95.CR1K 147          | -NERAD--G-ORR--R--0--S--N-R--Q--L-I-T--S-T--ANS-V-QT--G--T-OEL--HP--SQ--LAET  |            |              |         | 100 |
| A.IN.x.NNVA               | -TERAG--GPERK--R--0--R--N-R--Q--I--S-T--A-S-V-QT--G--T-OEL--HP--SQ--LAEA  |            |              |         | 100 |
| A.IN.85.ROD               | -NERAD--G-ORR--R--0--S--N-R--KQ--R-I--S-T--A-S--QT--H--G-T-OEL--HL--SQ--LAET  |            |              |         | 100 |
| B.CI.88.UC1               | -TTR-KD..Q-G-L--QT--S--N-R--K-GL-I--R-LS-S-EE--V-R-E-TV-DL-N-S--TA-AFTCIPPVWDQLVPRSNPSSNEGCRDSCCEHRKSPMESSQKDSGNSHRDPQEDOTRT                          |            |              |         | 150 |
| B.CI.89.pPRD320-20HIV2 56 | -TTR-KD..Q-G-L--QA--S--N-R--K-GL-I--R-S-S-EE--V-R-E-T-DL-N-S--TT-AFACIPIVWDQLAPRNPSSNEGCRDSCERKSPMGSSQKNGGNGHRGPOENQTRT                               |            |              |         | 150 |
| B.CI.x.EHO                | -NAR--D..Q-G-L--0--S--N-R--K-GL-I--HPL-S-EG--R--I-KDL-N-S--TA-ASTCIPIVWDQLVPRSNPSSQGGCRDSCERGEDLVGSPQESGRDHCTOEDOTRG                                  |            |              |         | 150 |
| B.GH.86.D205              | -TAR.GD..Q-E-L--0--S--N-R--K-GL-I--H-SIA-AEE-N--R--TV-DL-N-P-LNOSP--TT-APGCVPIVWDQLVPRSNPSSQGGCRDSCCEGRDLMGSSQESGNSHRDPQENQTRT                        |            |              |         | 153 |
| B.JP.01.KR020             | -TTR-KD..Q-G-LR--0--S--N-R--GL-I--0-HPL-S-EE-N--K-TV-DL-T-SI--TA-TSTCIPIVWDQLVPRSNPSSDEGCRDSCYKHKRGMGGSQKNSGNRDPQEDOTRTRPLVR                          |            |              |         | 157 |
| AB.CI.90.7312A            | -TAR-KD..Q-GI-L--0--S--N-R--K-GL-I--HPL-NS-AEE--RR--TV-DL-N-SS--TT-APVCVPIVWDQLVPRSNPSSQGGYGRDSCCEHREDLMGSSQEDGEGNHRSPQKNTGA                          |            |              |         | 150 |
| AB.CM.03.03CM_510_03      | -TTR--E--E-L--H--N--R--K-GL-I--R-SAA-AEE-N--R--TV-DL-N--TT-APVCVPIVWDQLVPRSNPSSSEGYERDSCCEGRDLMGSSQGGSDPRDPQKNTRT                                     |            |              |         | 148 |
| G.CI.92.ABT96             | --L-E--F--0--X--N-R-K-KQ-L-I--X-ARE--S-T-QDX-S-AV-KN-Q-PSN  |            |              |         | 100 |
| U.FR.96.12034             | --P-EG..VQR--N-R--Q-L-I--K-T--S-V--G--QDL-E-AVV-H--Q-TSAP   |            |              |         | 98  |
| MAC.US.x.17EC1            | -N-----R-----T-----I-----X-----   |            |              |         | 107 |
| MAC.US.x.17EFR            | -N-----R-----T-----I-----X-----   |            |              |         | 107 |
| MAC.US.x.2065             | -S-----R-----X-----I-----K-----N-----   |            |              |         | 107 |
| MAC.US.x.251_1A11         | -S-----R-----X-----I-----K-----N-----   |            |              |         | 107 |
| MAC.US.x.251_32H_PJ5      | -S-----S-----R-----X-----I-----K-----N-----   |            |              |         | 107 |
| MAC.US.x.251_BK28         | -S-----S-----R-----X-----I-----K-----N-----   |            |              |         | 107 |
| MAC.US.x.80035            | -S-----R-----X-----I-----K-----N-----   |            |              |         | 107 |
| MAC.US.x.97074            | -S-----R-----X-----I-----K-----N-----   |            |              |         | 107 |
| MAC.US.x.MAC239_87082     | -RS-TG--R-----S-S--KR--Q--I-----N--A  |            |              |         | 107 |
| MAC.US.x.MM142            | -RS-TG--R-----S-S--KR--Q--I-----I--LRRIR--A   |            |              |         | 107 |
| MAC.US.x.SM142B           | -S-----R-----X-----I-----K-----N-----   |            |              |         | 107 |
| MAC.US.x.r90131           | -S-AE--R-----S--N-R--Q--F--V--N--V--I--I-H--NP  |            |              |         | 107 |
| MNE.US.82.MNE 8           | -S-AE--R-----S--N-R--Q--F--V--N--V--I--I-H--NP  |            |              |         | 107 |
| MNE.US.x.MNE027           | -S-AE--R-----S--N-R--Q--F--V--N--V--I--I-H--NP  |            |              |         | 107 |
| SMM.SL.92.SL92B           | -Q-P-E--R-----I--A--S--N-R--K--I-IV--A-DI--FDS-S-Q-L-E-TV-T-R-QSA   |            |              |         | 99  |
| SMM.US.x.F236_H4          | -ST-E--R-----F--0--R--R--Q--I--V--G--EL--SA-P-N-VAKSP   |            |              |         | 100 |
| SMM.US.x.PBJ14_15         | T-SNGE--R-----S--D--R--R--Q--I--V--R--EL-N-ASA-P-K-AA-SP  |            |              |         | 100 |
| SMM.US.x.PBJA             | -SN-E--R-----D--R--R--Q--I--V--R--EL-N-ASA-P-K-AA-SP  |            |              |         | 100 |
| SMM.US.x.PBJ_6P6          | -SN-E--R-----D--R--R--Q--I--V--R--EL-N-ASA-P-K-AA-SP  |            |              |         | 100 |
| SMM.US.x.PGM53            | --SSAG--R-----F--0--R--R--Q--I--A-S--V--D-SL-N--SV--NP--R   |            |              |         | 104 |
| SMM.US.x.SME543           | -S-TE--R-----F--0--R--R--Q--I--N--DL--SA-T-K-AAKS   |            |              |         | 100 |
| STM.US.x.STM              | -DQ-E--R-----QF--0--N--R--Q--N--N--AS--V--G-S-QDL--P-L-KD-Q-TA-N  |            |              |         | 99  |



signal peptide end\_gp120 start

| Env start                 | signal peptide end_gp120 start   | 160 |
|---------------------------|--|-----|
| MAC.US.x.239              | MGCLGNQLLAIALLLSVYGIYCT.LYVTVFYGVPAMRNATIPFLFCATKNRDTWGTQCLPNDNGDYSEVALNVTFESFDANNVTVEQAIEDVHQLFETSICKPCVKLSPLCITMRCNKSETDRWGLTKSITTTASTTSTTASA.....KVDVNNETSSSCIAQDNCT.GL | 160 |
| A.CI.88.UC2               | EPGR--AV--T-ACL--K-Q-----V--S-----I-----D-Q-IP--A--D-----R-----T--VA-N--PVTGNNTNA-AKP-AARP-NP.....SYLIII--S-T-VGA-----   | 157 |
| A.DE.x.BEN                | EPGR--FVW--T-ACLV--S-Q-----I--K-S-----I-----D-Q-IP--A--D-----V--H-----T--VA-N--SRVQGNITTPNPRTSSTTSRPP.....SAASTI---N-ENNT-A  | 158 |
| A.GH.x.PE12               | DSR--IV--T-ACL--A-Q-----I--K-S-----R-----I-----D-Q-IP--A--D-----V--N-----V--T--VQ-E--STS-ESSNS-SEGS-VP.....EIL--T--TNNS-S-D  | 148 |
| A.GM.87.D194              | GKSL-CV-S-A-A-LV---Q-----V--S-----I--K--D-Q-IT--A--D-----V--S-----T--VA-S--STTNNITTTGTTGM.....SEI--T--PSYS-----  | 145 |
| A.GM.x.ISY                | EPGR--V--T-ACL--K-Q-----I--S-----I-----D-Q-IP--A--D-----R-----T--VA-N--ITSGTTATPSPP.....NITIID-N-T-GDN-----  | 146 |
| A.GM.x.MCN13              | SGKI--V-F--T-ACL---K-----V-K-S-----I-----D-Q-IP--A--D-I--V--N-----T--V--N--A-TESAVAT-SPSGP.....I-D-DP-QLN--S--   | 146 |
| A.GM.x.MCR35              | MGR--V--T-TCL---N-----I--S-----I-----D-Q-IT--A--D-----N-----T--VA--NTDARNITPTTASPRIT.....IKP-TEIS-N--RAN--S--  | 153 |
| A.GW.86.FG                | MGR--V--A-TCL---N-----I--S-----I-----D-Q-IT--A--D-----N-----T--VA--NTDARNITPTTASPRIT.....IKP-TEIS-N-T-RAN--S--   | 153 |
| A.GW.87.CAM2CG            | KGSK---V-A-A-L-H-K-QF---I--S-----I-----D-Q-IT--A--D-----V--N-----T--VA-N--TRNM-TWT-R-DTQ.....NITIIDNT-HARA-----  | 146 |
| A.GW.x.ALI                | ERGR--A-ACL--R-Q0---K-S-----I-----D-Q-IP--A--D-I--V--N-----T--VA-K--I-TS-TMIRTTTPS.....EAPISDN-P-RTN--S--  | 152 |
| A.GW.x.MDS                | MSSR---VT--A-ACLV--K-Q---K-S-----I-----D-Q-IP--A--D-----V--R-----T--A-K-SNIS-ESTTTPS-PGS.....TLKPLI--SDP-KA--PR--  | 151 |
| A.IN.95.CR1K.147          | TRMH---T-ACL--K-P00---I--S-----I-----D-E-I--A--D-----V--S-----T--VA-N--SRLSDSASTRITTTNA--A--TIR-TTIAP.....RNITIS-N-P-RA--S--   | 165 |
| A.IN.x.NNVA               | AHE-T---F-T-ACL--K-Q---I--K-S-----R-----I-----D-Q-IV--A--D-----N-----I--T--V---T-TTTAPT-TSAGSTA-P.....PI---NT-MHAN--S--  | 155 |
| A.SN.85.ROD               | TRERA--VV--T-TCL--K-Q---I--K-S-----I-----D-Q-IP--A--D-----V--N-----T--VA-E-KNR--GNATT-APT-PEINET.....TEIS-NT-VRG--S--  | 154 |
| B.CI.88.UC1               | M---M---A-ACLV---Q-----T-K-----R-----I-----D-Q-IT--A--D-----H-----T--VA-K-SST-SSTGNN-T-KS-STT-TP-D.....QEQEIS-DTP-ARA--S--   | 155 |
| B.CI.89.pPRD320-20HIV2.56 | AHTS-H-F-LL-I---FLGHKKN---I---V---T---V---T-ISV-I-A-----VD-S-----T--VA--NTG-NITTKPIITPI-TTKP.....SENLL-D-P-KN-T-P-I  | 156 |
| B.CI.x.EH0                | AHINSH--SL-I---CM-K-Q---I---V---A---V---S---T-ISV-I-A-----VD-N-----L--T--VA-N--STNN-RTNNT-A-T--GN--P.....IV--AIP-VKAN--S--I  | 153 |
| B.GH.86.D205              | AHVN-Y-VTL-I-I---YMGK.NF---I--K-S-----R-----V---D--T-IQ-I-A-----D--D--TK--S-----T--V--K--TWSSASKETTSSASLRS.....QTLI--D-K-ON-S-A-I  | 154 |
| B.JP.81.KR020             | AYFSSR-P-L-IGIS-FV-K-Q---I--K-V-I---I---V---D--D-I-I-A-----D--Q--VD--R-----T--VA-N-S-T-NPGKSSSTI--KPT--R.....GLKTI---DP-KN-S--   | 156 |
| G.CI.92.ABT96             | AHINKY-FA-L-I-I-FMGE.DF---I--K-V---I---V---D--D-I-I-A-----D--VA-N-----T--V--N-SNLN-TANT-ANNI-A.....SKNIS-L-GND--RD--I  | 152 |
| AB.CI.90.7312A            | AY---TX---T-F---K---I---SV-----D--L---A--D-----V--N-----T--T--A--T--NK---GTT-V-PA-VP--KM.....VTAEL-S-Q-LMY---I   | 160 |
| AB.CM.03.03CM_510_03      | GK-L-FV-S-A-A-L---K---V---S-----I-----D-Q-IP--A--D-----V--S-----T--VA-S--STTATTPPSTNN--TTEPT.....GGPEI--FP-MRT--   | 155 |
| U.FR.96.12034             | AYSSSC-P-TL---IC-FTXK.Q-X---I--X-XX-X---I-----D-Q-IT--A--X-----R-----T-XXVA-N-TTT-ATPQNT-TRN.....TTIEDNDP--SRN--   | 146 |
| MAC.US.x.251.1A11         | A---L-C-NACV---Q-----I---N-----V---S-----T--A--S-T-N---GTAA--T-PAMQES-K.....VI---EP--RNNS-S--  | 157 |
| MAC.US.x.251.32H.PJ5      | -----L-----L-----E-----T-----S-----TAAPKAMSE.....IN-I-----VH-----  | 159 |
| MAC.US.x.251.BK28         | -----L-----L-----E-----K-----L-----P-AP-A-----S-----ITTAAP-S-PVSE.....I-----N-----   | 162 |
| MAC.US.x.17EC1            | -----V-----M-----RV-----A-----   | 160 |
| MAC.US.x.17EFR            | -----V-----M-----RV-----A-----   | 160 |
| MAC.US.x.MN142            | -----I-Q-----D-----E-----K-----S-----T-TAK.....SVETR-I---P-VVH-----  | 163 |
| MNE.US.82.MNE.8           | -----F-A---I-Q-----V-R-----D--L--I---E-----H-----T--K---K---S---P-AIP.....KAEATKV---N-P-NH-----  | 160 |
| MNE.US.x.MNE027           | -----F-A---I-Q-----V-R-----D--L--I---E-----H-----T--K---K---S---AP-TK-TTI.....KEIEV---N-T-VNR-----   | 161 |
| SMM.SL.92.S192B           | A-P-LH--D-F--L-TW-A-Q---I--I---Q---V-----D--L--I---A--D-----N-----T--A-K--N---RAA--S-P-T-SPLTAASP-SGEET--D-M-TKNN--S--I  | 165 |
| SMM.US.x.F236.H4          | -----L-V--LE-C-V-Q---I--K---Q---D--L--I---A--D-----N-----T--A--T---GNAG--TTAIT.....TPSVAENVI--SNP--KNNS-A--  | 162 |
| SMM.US.x.PB314.15         | -----L-AS--V-Q---I--V-----D--L--I---A--D-----N-----T-----GTPAP-TTQ-T-QASTTPTSPITAKV--DSDP--KIN--   | 167 |
| SMM.US.x.PB3A             | -----L-AS--V-Q---I--V-----D--L--I---A--D-----N-----T-----GTPAP-TTQ-T-QASTTPTSPITAKV--DSDP--KIN--   | 167 |
| SMM.US.x.PB3.6P6          | -----L-AS--V-Q---I--V-----D--L--I---A--D-----N-----T-----GTPAP-TTQ-T-QASTTPTSPITAKV--DSDP--KIN--   | 167 |
| SMM.US.x.PGMS3            | -----A-----D-----L--I--R-----N-----T-----NEIP-TT-T-K-SPKAEATKAVI--SDP--SNN-----  | 165 |
| SMM.US.x.SME543           | -----L-V--LE-C-V-Q---K---R-----D--L-V-I-A-----N-----T--A--T---GRAE--TTAK--S---TTTTVTPKVI--GD---KNNS-A--  | 164 |
| STM.US.x.STM              | A-P-----ACLT---Q-----L-I---A--D-----N-----T--A--N--K---GKTV--VTP-AAA-A.....T-PEL-----VSNN-----   | 160 |

MAC.US.x.239  
A.CI.88.UC2  
A.DE.x.BEN  
A.DE.x.PE12  
A.GH.x.GH1  
A.GM.87.D194  
A.GM.x.ISY  
A.GM.x.MCN13  
A.GM.x.MCR35  
A.GW.86.FG  
A.GW.87.CAM2CG  
A.GW.x.ALI  
A.GW.x.MDS  
A.IN.95.CR1K 147  
A.IN.x.NNVA  
A.SN.85.ROD  
B.CI.88.UC1  
B.CI.09.pPRD320-20HIV2 56  
B.CI.x.EH0  
B.GH.86.D205  
B.JP.01.KR020  
G.CI.92.ABT96  
AB.CI.90.7312A  
AB.CM.03.03CM\_510\_03  
U.FR.96.12034  
MAC.US.x.251.1A11  
MAC.US.x.251.32H.PJ5  
MAC.US.x.251.BK28  
MAC.US.x.17EC1  
MAC.US.x.17EFR  
MAC.US.x.MN142  
MNE.US.02.MNE.8  
MNE.US.x.MNE027  
SMM.SL.92.SL92B  
SMM.US.x.F236.H4  
SMM.US.x.PB314.15  
SMM.US.x.PB3A  
SMM.US.x.PB3.6P6  
SMM.US.x.PG53  
SMM.US.x.SME543  
STM.US.x.STM

```

EQEQMTSCKFNMTGLKRDKKKEYMETWYSADLVCEQGN. TGNESRCYMNHCNTSVIOESC DKHYWDAIRFYCAAPPYALLRCNDTNYSGFMPCSKVWVSSCTRMMETQTSTWFGFNGTRAENRTYIYWHGRDNRTIISLNKYVNLTKCRPGRGKTVLPVTIMSLV 329
GD-G-VN---EQ-I-G-TD--D-V--DST-K--TNT--R---K---SMK---T---A-N-P-AA---H---H-K---V-I-L---HR 326
GY-E-MO-E--K-EQ--RR-KD--LE-V--DNITTA--GT--R---I-K---M---F---E-T-AA---R---K---I-L--- 323
GS-EVVD-R---QL-PQO-S--K-V--DTT-G--TSRK---T---M---LC---E-P-AAT---TH---H-K---S---I-LR-R- 315
GK-EIVN-Q-Y---E---0---K-V--SN-T.KDGKN---T---K---E---P-A-T---SIH-K---V-I-L--- 314
GK-EVVE-E---EQ-RR-DA--R-V--DKT-G--TGT--R---K---MK---F---E---AA---K---H-K---V-I-L---RR 312
RF-D-VE-Q---EL--0-S--K-V--SD-S.TDRK---T---M---FV---E-N-A-T---P-L---IL---E---V-I-L---RR 314
GE-EVVN-Q---E---0-S--K-V--GNGT--TDT---T---M---F---A-N-AAT---F---IH-K---V-I-L---R 319
GK-EVVN-Q---E---0-A--K-V--GNGT--T-T---T---M---F---A-N-AAT---F---IH-K---V-I-L---R 319
KE-E--D-Q-S---E--R-Q-T-A--K-V--DNNTS--SQ-K---T---M---F---A-N-AAT---F---K---NF---H-K---I-F--FK 313
E-KIVK-H---E--R-Q--S-V--DNSTDQ-T-TT---T---M---F-I--K---A-N-A-T---K---H---S-Y---V-I-L---OR 322
GD-E-VN-R---Q--P-Q--K-V--PF-T--T-OT---T---T---M---D-I--A-N-AAT---K---H---H-K---V-I-L---I 320
E-E-VK-Q---QI---Q--S-V--CEM--T-Q--S---K---K---F---Q-N-A-T---K---AS---SLH-K---V-I-L--- 334
GE-EVVN-E---V--P-T--R-VD--PDST--T-SRK---T---TM---F---A-N-A-T---L--SKSD---SIH-K---T-I-L---YK 323
RE-EIVI-Q---I--P-R--S-V--PDSTSTRNNSSSGO-MR-WYMRD---DM---E-N-T-AA---S---H---SIY-K---V-I-H-P--- 324
GE-ET-N-Q---E---Q--K-V--TN-S--T-OTQ---T---A-N-A-T---S---H---SIY-K---I-KOIML--H- 323
GL-NTVD-Y---R--E-Q-KD--EK--E-MGNS...TSTI--RT---SL---I---M--SK---T---M--SK---H---I-I---N 322
GL-DVVN-T---RQ-ER-Q--D-RR--E-G--T--RT---SL---I---M--SK---R---H---H-K---M-V-IRTV--L 315
CL-E--D-Q-K---ES-Q-KD--KO---K-T--RS--K-IKT--I---SL---F--K--N---LY---M--SK---S---H-K---M-V-IRTV--L 321
GE-EIMQ-N-S---R--EL-Q-KD--E--E-MNTR...KYT---IRT--TI---SL---FF--N--A---S---EK---T---SIH-K---V-IRTV--L 323
F-N-VT-R---V-EP-T--D--AE-VQ-NGS---I---HIKT--I---QL---F-FM-S--N---N---KS--V--K--IH---V---I 314
QS-S-VG---QR---Q--QX--S-K-SE--R---R---F---A-N-T---IX---I--- 329
GE-E-VD-Q---E--T-Q-S--K-V--SN-A.SDGRD---T---T---F---N---M--SK---IH-K---V-I-L--- 324
N-E-VT-N-S---EQ--R--Q-V--DKN-SSETLNKT--R---K---TM---T---F---Q-N-AA--I---H-K---V-I-L-P--- 316
PLV-R---Q-K--H-I---NT--SES-K---N---N---T--RG---S---Y--- 325
G---S---I---Q--- 328
G---T---T---S--D---I---T--- 328
T---T---T---S--D---K---T--- 331
N---N---K---C--D--C---N---R---H---I---A--- 329
P---S---C--D--C---N---R---H---I---A--- 332
P---RR---S--E---SK---N---SK---N--- 329
P---R---S--ED---K---N---SK---N--- 330
P-G-Q---Q-RQ---R---G--ES---Y---A-N---GS---I--- 332
P-G---N---R--I--SA--ES-K-H---Y---S--L-A-N---KS---R---E--- 330
P-V---R---NS--ETD-K---S---A-N-T---S---R--- 335
P-V---R---NS--ETD-K---S---A-N-T---S---R--- 335
P-V---R---NS--ETD-K---S---A-N-T---S---R--- 335
P-V-R---K---R-I--STT-SR--K---A-N---S---S---IR--- 334
P-G---I---R-I--PA--S-K-Q---S---A---NS--- 332
E-SLVG---R---S--I--NVT--E---R---L---T-A-N---S--- 328

```





|                           |   |     |
|---------------------------|---|-----|
| MAC.US.x.239              | TALLEEAQIQEKMYELQKLSNWDVFGNMFDLASWIKYIQYGVYIVVGVILLRIVIVIVQMLAKLRQGYRPFVSSPPSYFQQTHTIQDDPALPTREGKRDGEGGGGNSWPWQIEYIHFLIRQLIRLLTWLFSNCRITLLSRVYQILOPIQLRLSATLQRIREVLRTEL | 821 |
| A.CI.88.UC2               | S05-----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-ANSPTRRL-S-N-...-TA-DW-LKAA  | 807 |
| A.DE.x.BEN                | S05-----IL-----T-V-----H-I-A-A-V-L-SRF-K-----G-L-I-HK-RGO-AN-ET-E-V-GDS-YDL-P-N-VQ-HL-T-IG-YNI-D-KNSPTRRL-S-S-...-TA-DW-LKAA  | 802 |
| A.DE.x.PE12               | SRS-Q-----VII-I-A-----V-L-SR-K-----G-I-I-HK-WEO-D-ETDE-A-NSI-D-----P-A-----G-Y-V-KD-SFPT-L-F-S-ORA-TT-DW-LTIA   | 799 |
| A.GH.x.GH1                | S05-Q-----T-----V-A-----L-SR-K-----G-L-I-HT-RGO-AN-ET-E-DRDD-YDL*-P-N-----HL-T-G-YKI-D-TNSPTRRL-S-N-...-TA-DW-LKAA  | 793 |
| A.GM.87.D194              | S05-Q-----I-G-A-----L-SR-K-----G-L-I-HT-RGO-AN-ET-E-A-DDS-FGL-PLN-Q-HL-T-G-YNS-G-KNSPTRRL-S-S-...-TA-DW-LKAA  | 793 |
| A.GM.x.ISY                | SES-Q-----M-----I-VA-----V-SR-K-----G-I-I-HK-WEO-D-ET-E-V-NDV-SR-P-----L-----R-YNS-D-L-L-----L-DW-LKAA  | 788 |
| A.GM.x.MCN13              | S00-Q-----T-----II-IV-----SR-K-----G-L-I-HK-WEO-A-ET-E-V-NNV-D-----P-R-----H-----AG-YNI-N-ISLT-R-VF-S-ORA-TA-DW-DAA   | 801 |
| A.GM.x.MCR35              | S00-Q-----S-----T-----II-IV-----SR-K-----G-L-I-HK-WEO-A-ET-E-V-NNV-D-L*-P-R-----H-----G-YNI-N-ISLT-R-AF-S-ORA-TA-DW-DAA   | 800 |
| A.GW.86.FG                | S05-Q-----T-L-FT-VR-----V-IVA-----SR-K-----G-I-I-HK-QEO-A-ET-E-V-SN-DR-----P-A-----L-----G-YNI-D-ISP-F-S-ORA-TA-DW-LKAA   | 798 |
| A.GW.87.CAM2CG            | S05-L-----N-----T-----S-----I-V-----V-SR-K-----S-G-I-I-HK-EO-AS-ET-E-V-GN-DR-----G-YNI-N-IF-T-----N-----DW-PKVA   | 801 |
| A.GW.x.ALI                | SEQ-R-----T-----T-L-TA-V-----I-IVA-----V-SR-K-----G-I-I-HK-QEO-----GET-E-V-DNV-DRL-P-A-L-----HL-A-IG-Y-I-D-ISP-FRS-ORA-TT-DW-LKAA                                       | 804 |
| A.GW.x.MDS                | S0K-----Q-----I-----T-----I-V-----V-SR-K-----G-I-I-HK-QEO-D-ET-E-V-GN-DR-----P-A-----L-----R-YNS-D-IFLT-F-N-----DQ-LRTA   | 811 |
| A.IN.95.CR1K.147          | S05-Q-----S-----T-----C-----I-V-----V-SR-K-----G-I-I-HK-LEO-A-J-E-V-DN-YR*-P-A-----HL-----R-YNS-KD-SFLT-----N-----DW-LKTA   | 801 |
| A.IN.x.NNVA               | S05-Q-----I-----T-----I-A-V-----V-L-CSF-K-----G-I-I-HK-QEO-D-ET-E-V-DN-YR-----P-A-----LL-----R-YNS-KD-SFLT-----N-----DW-LRTA  | 799 |
| A.SN.85.R0D               | SKS-Q-----I-----T-V-----L-I-A-A-----V-SR-K-----G-I-I-HK-RGO-AN-ET-E-SN-DRY-----P-A-----R-Y-I-D-SFLT-L-Y-N-----DW-LRTA   | 800 |
| B.CI.88.UC1               | -----R-----FT-MA-RL-L-V-A-L-V-----M-----R-K-----T-IP-RKHRGO-AN-ET-DE-NE-AYR-----A-----RN-I-YNG-N-LKTS-----A-----QP-LS-A   | 799 |
| B.CI.09.pPRD320-20HIV2.56 | -Q-----H-----S-----FT-MA-RL-L-V-L-V-----VI-----R-K-----T-IP-HKHRGO-AN-ET-DE-RE-DYR-----A-----RN-I-YNG-N-LTS-----A-----QP-LS-A   | 789 |
| B.CI.x.EH0                | -K-----Q-I-S-----FT-MA-RL-L-V-I-IV-----A-I-----R-K-----T-IP-RK-RGO-AN-ET-EG-NNE-DYR-----P-----RD-I-Y-G-KTF-T-V-----QP-LPA   | 794 |
| B.GH.86.D205              | -----I-----T-----HL-L-A-LV-----V-----R-K-----T-IP-RK-RGO-AN-ET-EG-ND-DYR-----L-----RN-I-YNG-LKTF-HO-STN-----QP-LPA  | 801 |
| B.JP.01.KR020             | -K-V-----E-K-Q-S-----FT-A-I-L-VI-LVV-----L-R-K-----V-IP-HKGOGR-AS-ET-EG-DDDYR-----R-----RN-I-YNS-S-LKIF-P-LVF-----QP-LP-A   | 787 |
| G.CI.92.ABT96             | -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-LF-I-R-----A-H-L-AA   | 826 |
| AB.CI.90.7312A            | SES-Q-----V-----I-VA-----VI-----V-IGR-R-----G-----IR-HK-QEO-AN-ET-EG-ND-YR-----RN-I-YDG-LKTF-T-----A-----QP-LLFA  | 799 |
| AB.CM.03.03CM_510_03      | S0I-Q-----S-----T-----R-----A-VV-X-A-----XSR-K-----SG-V-I-HT-QGO-S-ET-EG-ND-DYR-----V-----L-----RN-I-YNG-LKTS-----P-LPA   | 788 |
| U.FR.96.12034             | -Q-A-----N-I-----T-V-VYL-L-V-A-I-I-----V-L-GS-K-----H-VE-IP-R-QEO-KGEI-EGA-DS-YR-----R-----H-----Y-S-D-IC-S-LF-SI-----R-HL-IA   | 809 |
| MAC.US.x.251.1A11         | -----R-----G-----A-----R-----819  | 819 |
| MAC.US.x.251.32H.PJ5      | -----G-----A-----A-----R-----820  | 820 |
| MAC.US.x.251.BK28         | -----I-V-----*-----T-----G-----A-----R-V-----822  | 822 |
| MAC.US.x.17EC1            | -----N-----V-----*-----T-----G-----A-----R-V-----821  | 821 |
| MAC.US.x.17EFR            | -----N-----T-----I-I-----V-----R-----*-----T-----K-KG-GS-----A-----F-----R-----L-----820  | 820 |
| MAC.US.x.MN142            | -----R-----R-----R-Q-----K-KG-GS-----A-----F-F-T-----V-----822  | 822 |
| MNE.US.82.MNE.8           | -----R-----R-----R-Q-----K-KG-GS-----A-----F-F-T-----V-----824  | 824 |
| MNE.US.x.MNE027           | -----R-----R-----R-Q-----K-KG-GS-----A-----F-F-T-----V-----824  | 824 |
| SMM.SL.92.S192B           | -QM-RL-----I-----T-VR-----FL-I-IV-----V-SR-----H-I-R-QE-AK-T-GE-N-DYR-----I-YN-LTR-A-A-N-QLC-----EIS-P-LV-R-AG  | 828 |
| SMM.US.x.F236.H4          | -O-----I-----T-R-----L-L-G-----V-R-----A-V-IP-HKGOEP-K-E-G-DR-SR-----S-DW-L-S-----V-S-T-----V-I-IA  | 827 |
| SMM.US.x.PBJ14.15         | -O5-----T-----I-----T-----L-L-G-----V-R-----A-V-IP-TGOE-K-E-G-DR-SR-----S-DW-L-NC-T-V-S-R-----A-I-VOIA  | 831 |
| SMM.US.x.PBJA             | -O5-----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-----A-I-VOIA   | 831 |
| SMM.US.x.PBJ.6P6          | -O5-----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-----A-I-VOIA   | 831 |
| SMM.US.x.PGMS3            | -QA-----I-----T-----L-L-VG-----V-R-----P-A-V-IP-HKGOEP-K-E-G-DR-SR-----N-----Y-L-I-VF-----V-I-S   | 827 |
| SMM.US.x.SME543           | -Q-----I-----T-----L-L-VG-----V-R-----P-A-V-IP-HKGOEP-K-E-G-DR-SR-----N-----Y-L-I-VF-----V-I-S  | 827 |
| STM.US.x.STM              | -Q-----V-----T-VR-----L-I-LVM-VA-----M-L-R-K-----CR-IP-HKGOEP-K-E-G-DR-IN-----T-----V-----YN-F-AC-----I-T-H-TF-I-RI-----V-LGAA  | 824 |



|                      | Nef start                      | R17Y mutation                      | premature stop in Mac239                | max HIV-1 similarity            |   |                                 |                        |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
|----------------------|--------------------------------|------------------------------------|---|---------------------------------|---|---------------------------------|------------------------|-------|-------|-------|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|-----|
| MAC.US.x.239         | MGGAISMRSPSGDLRQLRLRARGETYGRLL | GEVEDGYSQSPGLDKGLSSLS.C            | EGOK..YNOGQYMNTPWRNPAEEREKLAYRKONMDDIDE | DDD.LVGVSVR.PKVPLRTMSYKLAIDMSHF | IKEKGGLEGIYYSARRHRILDIYLEKEEGIIPDWQDYTS | SGPGI                           | 164                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.CI.88.UC2          | --ASG-KKL-KQ-QG-E              | -----GG-V-QCSASGGES--Q-SGREOK-P    | ---Q-R-D                                | ---T-M-GQ-ES-KQ                 | ---V-SD-N-Q                             | ---T-SR--P-T--V                 | ---MF-RD--L--V--N-H--V | 165   |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.DE.x.BEN           | --ASG-KKL-KH-RG-E              | -----DG-KORDASGGE--FOESGREON-P     | ---Q-Q-E-S                              | ---T-Q-DL-Q                     | ---V-SD--I-P-T-R-R-E-T                  | ---Q-MF-R--N-H--V               | 165                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.DE.x.PE12          | --ASG-KKC-SLQG-E               | -----C-GQWDSGAGE-L-FQE-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--V     | 165                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GH.x.GH1           | --ASG-KKH-KH-OR-E              | -----H-GG-VQOCNASGGE--QE-SG-QK-P   | ---Q-R-DF                               | ---T-I-GQ-KL-KQ                 | ---SS--P-T-R-A-T                        | ---V-KR--D-MF-RD--L--N-H--V     | 164                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GM.87.D194         | --ASG-KK-EH-QG-E               | -----GG-VKQRNASGGES--QE-SGREOK-P   | ---Q-Q-EF                               | ---T-AIGQ-NS-KQ                 | ---V-SD--P-M-R--E-T                     | ---F-RE--LF--N-H--T             | 165                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GM.x.ISY           | --ASG-KK--RG-QE                | -----GAC-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--N-H-L-V                 | 165                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GM.x.MCN13         | --ASG-KE--QG-E                 | -----C-GHCSGSGGE-L-QEESGREOR-P     | ---R-Q-DF                               | ---T-K-G-EL-KQ                  | ---R-V-SD--T-R--ELTHR--V                | ---M-E--A--N-H--                | 163                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GM.x.MCR35         | --ASG-KE--QG-E                 | -----C-GHCSGSGGE-L-QEESGREOR-P     | ---R-Q-DF                               | ---T-K-G-EL-KQ                  | ---R-V-SD--K-T-R--E-THR                 | ---M-E--K--N-H--                | 164                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GW.86.FG           | --ASG-KK-K-LQG-QE-Q            | -----C-GRCN-SGG-L-HE-SGREON-P      | ---R-Q-DFV                              | ---T-A--EL-KQ                   | ---V-LD--Q-PP-T-R--P-TF                 | ---V--LF-Q--L-D                 | 148                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GW.87.CAM2CG       | --ASG-KK--LQG-QE               | -----AG-C-ECYNAL-GESLR-QE-S-REON   | ---R-Q-DF                               | ---A-A-GK-N-Q                   | ---SD--PAT-R--T--V                      | ---LF-E--L-D--A--N--V           | 165                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GW.x.ALI           | --ASG-KK--G-LQG-E              | -----QTP--C-GQCSGSGG--Q-SGR-QKLP   | ---R-Q-DF                               | ---T--EL-KQ                     | ---V-LD-S--Y-T-RT--A-T--V               | ---L-R--MF-E--N-H--T            | 165                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.GW.x.MDS           | --ASG-KK--RG-QE                | -----AGAC-GHMD-LGGE--R-QE-S--QR-P  | ---Q-Q-DF                               | ---T-T-G-DS-Q                   | ---V-SD--Y-T-RT--A-T--V                 | ---MF-E--N-H--T                 | 163                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.IN.95.CR1K.147     | --ASG-KKH-K--RG-E              | -----QA-CVQGWK-LGGE--QE-S--QR-P    | ---Q-Q-DF                               | ---T-T-KG-DL-Q                  | ---V-SD--T-R--T-RM--L                   | ---MF-E-L-#                     | 139                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| A.SN.85.ROD          | --ASG-KKH--PRG-QE              | -----AGAC-GYWN-SGGE--RFQE-S-REQK-P | ---RQ-Q-DF                              | ---KD-A--NL-Q                   | ---V-SD--QR-T--P-THR                    | ---L-TR--MF-E-K-N--A--N-H--V    | 165                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| AB.CI.90.7312A       | --S-G-KKQ-KQQRG-E              | -----Q-GP--S-GRRGES--EES-R-N-P     | .....                                   | TA-GA-GGG--Q-DIEE               | DDE-N-E-Y--Q--P-T                       | -----F-E--T-F-N-R--VGG-N-Y      | 147                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| AB.CM.03.03CM_510_03 | --S-G-KK-EQQRG-E               | -----V-ERP-K-S-GRRERS-R--EES-D-N-P | .....                                   | NALGA-GGG--Q-A-ES-D-E           | ---AY--SR--P-T-T                        | -----L-D--E--T-F-N--VSG-N-H     | 146                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| B.CI.88.UC1          | --S-G-KKQ-KQQRG-E              | -----TQE-P-K-SEGQRKQS-R--S-D-N-P   | ---R.....                               | NA-RA-GGG--Q-DT                 | ---SDE-N-E-Y--NR-S-T-M                  | -----E--T-N--VSG-N-Y            | 147                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| B.CI.x.EH0           | --S-G-KKQ-KQPG-E               | -----RGPR-ESS--RQERSL-Y--S-N-P     | ---D.....                               | KTLGA-GGG--DS-E                 | DDE-N-E-R--G--P-TF--V--L-E              | ---F-E-K-T-N--VSG-N-H--V        | 147                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| B.GH.86.D205         | --S-G-KK-ERQOG-EK              | -----VPERP--S--RREOS-R--ES-D-N-P   | .....                                   | NARGA-GGG--Q-DA-E               | SDE-N-E-AICTK-I--P-T                    | ---V--Q--M-E--T-F-N--VSG-N-H--V | 148                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| B.JP.01.KR020        | --S-G-KKQ-KQRRG-E              | -----QTQEGP--SE-QRGQS-R--S-DSN-P   | ---R.....                               | SA-GAOGGG--Q-GIE                | ---SDE-N-E-Y--NR--P-T                   | ---W--E--L-T-N--VSG-N-H--V      | 147                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| G.CI.92.ABT96        | --S-G-KK-APQQLG-K-Q            | ---P-K-WEGL-E--YREESG              | ---Q-T-F                                | ---T-A--X-N-V-SD-N              | ---P-Y--V                               | -----E--M-Q--N-A--              | 165                    |       |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| U.FR.96.12034        | --S-KKH-KR-K-E-A               | ---N--W-GL-E--FELYQ-SG-D--RF       | ---A-GQG-SE-EF                          | ---T-TGKD-Q-KQ                  | ---V-E-N                                | ---A-R-A--T--V                  | ---D-F-Q--N--N--T      | 167   |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.17EC1       | .....                          | S                                  | .....                                   | K                               | .....                                   | I                               | .....                  | 165   |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.17EFR       | .....                          | G                                  | .....                                   | EX                              | .....                                   | T                               | .....                  | 165   |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.1937        | XX                             | X-X                                | X                                       | .....                           | .....                                   | .....                           | .....                  | 165   |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.2065        | .....                          | .....                              | .....                                   | .....                           | .....                                   | .....                           | .....                  | 165   |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.251.1A11    | T                              | ST                                 | E                                       | S                               | L                                       | E                               | .....                  | 165   |       |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.251.32H.PJ5 | R                              | KSA                                | K                                       | S                               | L                                       | G                               | R                      | ..... | 165   |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.251.BK28    | K                              | A                                  | K                                       | N                               | FKG                                     | S                               | L                      | G     | R     | ..... | 165 |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.MN142       | KK                             | K-PR                               | N                                       | FKG                             | S                                       | L                               | G                      | R     | ..... | 164   |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.r80025      | .....                          | .....                              | .....                                   | .....                           | .....                                   | .....                           | .....                  | ..... | 165   |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MAC.US.x.r90131      | .....                          | .....                              | .....                                   | .....                           | .....                                   | .....                           | .....                  | ..... | 165   |       |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MNE.US.82.MNE.8      | T                              | K                                  | KSP                                     | WE                              | L                                       | S                               | S                      | L     | ..... | 165   |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| MNE.US.x.MNE027      | T                              | KK                                 | KSP                                     | WE                              | L                                       | S                               | S                      | L     | ..... | 165   |     |       |       |       |       |       |       |       |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| SMM.SL.92.SL92B      | IT                             | SKPA-KH                            | KK                                      | S                               | DS                                      | WD                              | GL                     | E     | SQ    | Q     | S   | RA    | N     | CL    | TRD   | Y     | ..... | 166   |       |       |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| SMM.US.x.F236.H4     | KK                             | QYKRG                              | N                                       | E                               | Q                                       | WE                              | GL                     | E     | AS    | G     | Q   | V     | DE    | E     | I     | H     | R     | A     | T     | ..... | 165 |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| SMM.US.x.PB114.15    | VT                             | KKQR                               | RG                                      | N                               | YE                                      | Q                               | D                      | WE    | GL    | GE    | QD  | ASG   | ..... | ..... | ..... | ..... | ..... | ..... | ..... | 163   |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| SMM.US.x.PB3A        | VT                             | KKQR                               | RG                                      | N                               | YE                                      | Q                               | WE                     | GL    | GE    | QD    | ASG | ..... | ..... | ..... | ..... | ..... | ..... | ..... | ..... | 163   |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| SMM.US.x.PBJ.6P6     | A                              | G                                  | KKQ                                     | QR                              | G                                       | GEK                             | Q                      | H     | K     | WE    | GL  | R     | E     | GR    | DN    | LH    | F     | ..... | ..... | 165   |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| SMM.US.x.PGM53       | KK                             | QH                                 | RG                                      | N                               | E                                       | WE                              | GL                     | E     | AS    | G     | Q   | V     | DE    | E     | H     | A     | T     | ..... | ..... | 165   |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |     |
| STM.US.x.STM         | ASG                            | KKQR                               | KQH                                     | E                               | E                                       | K                               | E                      | GL    | GE    | SGP   | Q   | AS    | N     | H     | P     | R     | E     | F     | K     | A     | SA  | E | Q | V | E | N | A | H | R | E | T | L | S | E | M | V | N | A | 165 |

|                      | max HIV-1 similarity  | normal Nef end |     |
|----------------------|---|----------------|-----|
| MAC.US.x.239         | RYPKTFGWLKLVPNVNSDEAQ...EDEEHYLMHPAQTQSDWDDPWGEVLWKFDPDLTAYTEAYVRYPEEFGSKSGLSEEEVRRRLTARGLNMDADKKETR* |                | 262 |
| A.CI.88.UC2          | ---MC-----M-Q-E---D-TNC-----RY-IH-T-V-R-NSM-E-K-FTL-----H---P-K-WKAK-K---IPYSE-----                   |                | 258 |
| A.DE.x.BEN           | ---MY-----S-EL-Q-E---ANC-V-----RH-EH-T-V-Q-SM--N-K-FTL-----H---P-K-WKAK-K---IPYSE-----                |                | 258 |
| A.DE.x.PE12          | ---F-----D-PQG---E-D-C-L---GS--H-T-M-R--R--E-T-FN-----Y---P--WKAK-K---IPFS-----                       |                | 254 |
| A.GH.x.GH1           | ---MC-----D-Q-E---D-TN-T---RH-EH-T-L-R---D-K-FILH-----H---P-K-WKAK-K---IPYS-----                      |                | 256 |
| A.GM.87.D194         | ---MY-----DI-Q-E---V-TNC-V---RY-EH-T-V-R--M--S-K-FILH-----H---P-K-WKAK-K---IPYSE-----                 |                | 258 |
| A.GM.x.ISY           | ---MF-----T-PQ-GE---DT-TLC--S--V-RF--H-T-V---M--HE-TTFIL-----H---ME-DDWKAK-K---IPFS-----              |                | 257 |
| A.GM.x.MCN13         | ---MF-----D-TROEE---D-GT-C-L---RF--H-T-I---#---HE-K-FIL-----H---P--DWKAK-K---IPFS-----                |                | 257 |
| A.GM.x.MCR35         | ---MF-----D-TROEE---D-GT-C-L---RF--H-T-I---#---HE-K-FIL-----H---P--DWKAK-K---IPFS-----                |                | 214 |
| A.GW.86.FG           | ---MF-----DT-Q-GEDTETDT-T-C-L-----RH-MM-T-V--SM-LK-FT-----H---P-D-WKAK-K---IPFS-----                  |                | 181 |
| A.GW.87.CAM2CG       | ---MF-----DT-Q-GEDTETDT-T-C-L-----RH-MM-T-V--SM-LK-FT-----H---P-D-WKAK-K---IPFS-----                  |                | 261 |
| A.GW.x.ALI           | ---MF-----D-PO-GE---DT-T-C-L-V---RH-TH-T-V-R--K--HD-K-FILH-----Y---P-D-WKAK-K---IPFS-----K            |                | 258 |
| A.GW.x.MDS           | ---C-----DLPDEN---T-T-C-V---I-KF-QH-T-V-R-SL--E-T-FKL-----HQ--P--WKA-K---IPFN-----                    |                | 256 |
| A.IN.95.CR1K.147     | ---MF-----D-PQ-GE---DT-T-C-V---KF--H-T-V-E--L--S--FI-----H---P--WKA-K---IPFS-----                     |                | 172 |
| A.SN.85.ROD          | ---MF-----D-PQ-GE---DT-T-C-V---KF--H-T-V-E--L--S--FI-----H---P--WKA-K---IPFS-----                     |                | 257 |
| AB.CI.90.7312A       | ---H-----E-AAVTR--E-E-T-C-V---AA--H-T-V-Q-SL--SD-FN-F---YQ--P-K-WKA-K---IPTE-----                     |                | 240 |
| AB.CM.03.03CM_510_03 | ---Y-----ETPT--R--EGE-T-C---S--YH-T-V--L--HD-V-FN-F---YQ--P-K-WKAK-K---IPTD-----                      |                | 239 |
| B.CI.88.UC1          | ---R-----DIPE-ER---GA-TSC-V---I-S--IH-T-R--L--HD-V-FN---YQ--P-K-+KA-K---IPTD-----                     |                | 238 |
| B.CI.x.EH0           | ---F-----I-MIA-PE---DE-T-C-V---A--HE-T-V-Q-SL--D-V-FS-F---YQ--MP-K-WKAK-R---IPTE-----                 |                | 239 |
| B.GH.86.D205         | ---Y-----E-PAATR--E-E-T-C---I-S--IH-T-I-Q-SL--D-V-FN-F---YQ--P--WKA-K---IPTD-----                     |                | 241 |
| B.JP.01.KR020        | ---Y-----E-PAATR--E-E-T-C---I-S--IH-T-I-Q-SL--D-V-FN-F---YQ--P--WKA-K---IPTD-----                     |                | 239 |
| G.CI.92.ABT96        | ---X-----IPE-DK--A-TSC---A-P--VH-T-I-Q-SL--D-I-FN---YQ--P-K-WKV-K---IPTD-----                         |                | 257 |
| U.FR.96.12034        | ---W-MC-----E-DAN-K-----RL-VGS---CEE-H-A-V--SS--S-Q-FIKC-----K-----PVK-----NC                         |                | 260 |
| MAC.US.x.17EC1       | -----F-----   |                | 264 |
| MAC.US.x.17EFR       | ---H-----D-----C-I-----I-----P-----A-----   |                | 264 |
| MAC.US.x.1937        | -----C-X-----   |                | 264 |
| MAC.US.x.2065        | -----C-X-----   |                | 264 |
| MAC.US.x.251.1A11    | -----R--Q--K--P-----A--#--P-----R-----  |                | 264 |
| MAC.US.x.251.32H.PJ5 | -----R--Q--K--P-----A--#--P-----R-----  |                | 264 |
| MAC.US.x.251.BK28    | ---L-M---I-----V-----I-----K-K-A---E--R--S-----   |                | 263 |
| MAC.US.x.MN142       | ---L-M---I-----X-----I-----K-K-A---E--R--S-----   |                | 263 |
| MAC.US.x.r80025      | -----X-----   |                | 264 |
| MAC.US.x.r90131      | -----X-----   |                | 264 |
| MNE.US.82.MNE.8      | -----G-N-L-----V-----I-----K--R--S-----   |                | 264 |
| MNE.US.x.MNE027      | -----G-N-L-----V-----I-----K--R--S-----   |                | 264 |
| SMM.SL.92.SL92B      | ---F---Q---D---N---N-T-C-V---Y-HS--K--N-H---F-H---W---T---E---ANKP...KPQ-KMAD                         |                | 261 |
| SMM.US.x.F236.H4     | ---MHY---D---T-C-V---Y---E---S-K-FIK-----K---K#-----S-----  |                | 263 |
| SMM.US.x.PB114.15    | ---MF-----T-----K--N-K-F-EH---Q---K-Q---K---K-S-----  |                | 262 |
| SMM.US.x.PB3A        | ---MF-----T-----K--N-K-F-EH---Q---K-Q---K---K-S-----  |                | 262 |
| SMM.US.x.PBJ.6P6     | ---MF-----T-----K--N-K-F-EH---Q---K-Q---K---K-S-----  |                | 262 |
| SMM.US.x.PGMS3       | ---MF-----SA-----T-V---I-----SQ--R--FI-H---Q---K---K---K---S-----                                     |                | 264 |
| SMM.US.x.SME543      | ---MY---D---T-C-V---H-----E--S-K-FIK-----K-----IYK---R--S-----  |                | 264 |
| STM.US.x.STM         | ---Q-----DM-N---ED-GT--V---H-----V---L-H---F-H---PK--E---K-----S-----                                 |                | 264 |