

VII

PLV Proteins

Contents

| | | |
|-------|------------------------|------------|
| VII-1 | Introduction | 403 |
| VII-2 | Sequences | 404 |
| VII-3 | Alignments | 413 |
| | VII-3.1 Gag | 413 |
| | VII-3.2 Pol | 418 |
| | VII-3.3 Vif | 425 |
| | VII-3.4 Vpx | 428 |
| | VII-3.5 Vpr | 429 |
| | VII-3.6 Tat | 430 |
| | VII-3.7 Rev | 432 |
| | VII-3.8 Vpu | 434 |
| | VII-3.9 Env | 436 |
| | VII-3.10Nef | 443 |

VII-1 Introduction

The selection of Primate Lentivirus protein sequences for the following alignments was based on the sequences in the complete genome alignment as a starting alignment, and complete or nearly complete genes from other isolates were added if they increased the diversity of samples represented.

The annotation is mainly based on knowledge from HIV-1, and should therefore be taken with caution for HIV-2 and SIV sequences.

VII-2 Sequences

Sequences included in the PLV protein alignments.

| Name | Accession | Proteins | Author | Reference |
|------------------------|-----------|--|--------------------------|--|
| H1B.FR.83.HXB2 | K03455 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Wong-Staal, F. | <i>Nature</i> 313 (6000):277-284 (1985) |
| H1A1.UG.85.U455_U455A | M62320 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Oram, J.D. | <i>ARHR</i> 6 (9):1073-1078 (1990) |
| H1C.ET.86.ETH2220 | U46016 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Salminen, M.O. | <i>ARHR</i> 12 (14):1329-1339 (1996) |
| H1F1.BE.93.VI850 | AF077336 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Laukkanen, T. | <i>Virology</i> 269 (1):95-104 (2000) |
| H1O.BE.87.ANT70 | L20587 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Vanden Haesevelde, M. | <i>J Virol</i> 68 (3):1586-1596 (1994) |
| H1O.CM.91.MVP5180 | L20571 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Gurtler, L.G. | <i>J Virol</i> 68 (3):1581-1585 (1994) |
| H1O.CM.98.98CMA105 | AY169816 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.98.98CMABB197 | AY169811 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.SN.99.99SE_MP1299 | AJ302646 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Toure-Kane, C. | <i>ARHR</i> 17 (12):1211-1216 (2001) |
| H1O.US.99.99USTWLA | AY169814 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1N.CM.02.DJO0131 | AY532635 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Bodelle, P. | <i>ARHR</i> 20 (8):902-908 (2004) |
| H1N.CM.04.04CM_1015_04 | DQ017382 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Yamaguchi, J. | <i>ARHR</i> 22 (1):83-92 (2006) |

| Name | Accession | Proteins | Author | Reference |
|------------------------|-----------|--|----------------------------|--|
| H1N.CM.04.04CM_1131_03 | DQ017383 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Yamaguchi, J. | <i>ARHR</i> 22 (1):83-92 (2006) |
| H1N.CM.95.YBF30 | AJ006022 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Simon, F. | <i>Nat Med</i> 4 (9):1032-1037 (1998) |
| H1P.CM.06.U14788 | HQ179987 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Vallari, A. | <i>J Virol</i> 85 (3): 1403-7 (2011) |
| CPZ.CD.06.BF1167 | JQ866001 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Li, Y. | <i>J Virol</i> 86 (19):10776-10791 (2012) |
| CPZ.CD.90.ANT | U42720 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Vanden Haesevelde, M.M. | <i>Virology</i> 221 (2):346-350 (1996) |
| CPZ.CM.01.SIVcpzCAM13 | AY169968 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Nerrienet, E. | <i>J Virol</i> 79 (2):1312-9 (2005) |
| CPZ.CM.05.LB715 | KP861923 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Barbian, H.J. | <i>MBio</i> 2015 Apr 21;6(2) pii: e00296-15 doi: 10.1128/mBio00296-15 |
| CPZ.CM.05.SIVcpzDP943 | EF535993 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Van Heuverswyn, F. | <i>Virology</i> 368 (1):155-171 (2007) |
| CPZ.CM.05.SIVcpzEK505 | DQ373065 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Keele, B.F. | <i>Science</i> 313 (5786):523-526 (2006) |
| CPZ.CM.05.SIVcpzLB7 | DQ373064 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Keele, B.F. | <i>Science</i> 313 (5786):523-526 (2006) |
| CPZ.CM.05.SIVcpzMB66 | DQ373063 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Keele, B.F. | <i>Science</i> 313 (5786):523-526 (2006) |
| CPZ.CM.05.SIVcpzMB897 | EF535994 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Van Heuverswyn, F. | <i>Virology</i> 368 (1):155-171 (2007) |
| CPZ.GA.88.SIVcpzGAB2 | AF382828 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Bibollet-Ruche, F. | <i>ARHR</i> 20 (12):1377-1381 (2004) |

| Name | Accession | Proteins | Author | Reference |
|---------------------------|-----------|---|----------------|---|
| CPZ.TZ.00.TAN1 | AF447763 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Santiago, M.L. | <i>J Virol</i> 77 (3):2233-2242 (2003) |
| CPZ.TZ.06.SIVcpzTAN13 | JQ768416 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Takehisa, J. | Unpublished |
| CPZ.TZ.09.UG38 | JN091690 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Rudicell, R.S. | <i>J Virol</i> 85 (19): 9918-28 (2011) |
| CPZ.US.85.US_Marilyn | AF103818 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Gao, F. | <i>Nature</i> 397 (6718):436-441 (1999) |
| GOR.CM.04.SIVgorCP684con | FJ424871 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Takehisa, J. | <i>J Virol</i> 83 (4):1635-1648 (2009) |
| GOR.CM.07.SIVgor2139_287 | FJ424866 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Takehisa, J. | <i>J Virol</i> 83 (4):1635-1648 (2009) |
| GOR.CM.07.SIVgorCP2135con | FJ424863 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Takehisa, J. | <i>J Virol</i> 83 (4):1635-1648 (2009) |
| GOR.CM.12.SIVgor_BPID1 | KP004989 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | D'arc, M. | <i>PNAS USA</i> 112 (11); E1343-52 (2015) |
| GOR.CM.12.SIVgor_BQID2 | KP004991 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | D'arc, M. | <i>PNAS USA</i> 112 (11); E1343-52 (2015) |
| GOR.CM.13.SIVgor_BPID15 | KP004990 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | D'arc, M. | <i>PNAS USA</i> 112 (11); E1343-52 (2015) |
| MAC.US.x.239 | M33262 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Kestler, H. | <i>Science</i> 248 (4959):1109-1112 (1990) |
| H2A.DE.x.BEN | M30502 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Kirchhoff, F. | <i>Virology</i> 177 (1):305-311 (1990) |

| Name | Accession | Proteins | Author | Reference |
|----------------------|-----------|---|------------------|--|
| H2A.PT.x.ALI | AF082339 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Reeves, J.D. | <i>J Virol</i> 73 (9); 7795-804 (1999) |
| H2A.SN.86.ST_JSP4_27 | M31113 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Kumar, P. | <i>J Virol</i> 64 (2):890-901 (1990) |
| H2B.CI.x.EHO | U27200 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Rey-Cuille, M.A. | <i>Virology</i> 202 (1):471-476 (1994) |
| H2B.GH.86.D205_ALT | X61240 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Dietrich, U. | <i>Nature</i> 342 (6252):948-950 (1989) |
| H2G.CI.92.Abt96 | AF208027 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Brennan, C.A. | <i>ARHR</i> 13 (5):401-404 (1997) |
| H2U.CI.07.07IC_TNP3 | KC693505 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Ayouba, A. | <i>AIDS</i> 27 (15):2488-2491 (2013) |
| H2U.FR.96.12034 | AY530889 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Damond, F. | <i>ARHR</i> 20 (6):666-672 (2004) |
| ASC.UG.10.RT03 | KJ461716 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Lauck, M. | <i>Retrovirology</i> 11 (1):55 (2014) |
| ASC.UG.10.RT08 | KJ461715 | Gag, Pol, Vif, Tat, Rev, Env, Nef | Lauck, M. | <i>Retrovirology</i> 11 (1):55 (2014) |
| ASC.UG.10.RT11 | KJ461714 | Gag, Pol, Vif, Tat, Rev, Env, Nef | Lauck, M. | <i>Retrovirology</i> 11 (1):55 (2014) |
| COL.CM.x.CGU1 | AF301156 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Courgnaud, V. | <i>J Virol</i> 75 (2):857-866 (2001) |
| COL.UG.10.BWC01 | KF214240 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Lauck, M. | <i>Retrovirology</i> 10 (1):107 (2013) |

| Name | Accession | Proteins | Author | Reference |
|----------------------------|-----------|---|--------------------|---|
| COL.UG.10.BWC07 | KF214241 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Lauck, M. | <i>Retrovirology</i> 10 (1):107 (2013) |
| DEB.CM.04.SIVdeb04CMPF3061 | FJ919724 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Aghokeng, A.F. | <i>Infect Genet Evol</i> 10 (3); 386-96 (2010) |
| DEB.CM.99.CM40 | AY523865 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Bibollet-Ruche, F. | <i>J Virol</i> 78 (14):7748-7762 (2004) |
| DEB.CM.99.CM5 | AY523866 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Bibollet-Ruche, F. | <i>J Virol</i> 78 (14):7748-7762 (2004) |
| DRL.DE.11.D3 | KM378563 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Dietrich, U. | <i>ARHR</i> 31 (3); 353-7 (2015) |
| DRL.DE.11.D4 | KM378564 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Dietrich, U. | <i>ARHR</i> 31 (3); 353-7 (2015) |
| GRV.ET.x.GRI_677_gri_1 | M66437 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Fomsgaard, A. | <i>Virology</i> 182 (1):397-402 (1991) |
| GSN.CM.99.CN166 | AF468659 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Courgnaud, V. | <i>J Virol</i> 76 (16):8298-8309 (2002) |
| GSN.CM.99.CN71 | AF468658 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Courgnaud, V. | <i>J Virol</i> 76 (16):8298-8309 (2002) |
| LST.CD.88.SIVlhoest485 | AF188115 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Beer, B.E. | <i>J Virol</i> 74 (8):3892-3898 (2000) |
| LST.CD.88.SIVlhoest524 | AF188116 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Beer, B.E. | <i>J Virol</i> 74 (8):3892-3898 (2000) |
| LST.KE.x.lho7 | AF075269 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Hirsch, V.M. | <i>J Virol</i> 73 (2):1036-1045 (1999) |

| Name | Accession | Proteins | Author | Reference |
|--------------------------|-----------|---|----------------------|---|
| MAL.ZM.09.SIVagmMAL_ZMB | LC114462 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Carr, M. | <i>Arch Virol</i> 162 (2); 543-548 (2017) |
| MND-1.GA.x.MNDGB1 | M27470 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Tsujimoto, H. | <i>Nature</i> 341 (6242):539-541 (1989) |
| MND-2.CM.98.CM16 | AF367411 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Takehisa, J. | <i>ARHR</i> 17 (12):1143-1154 (2001) |
| MND-2.GA.x.M14 | AF328295 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Souquiere, S. | <i>J Virol</i> 75 (15):7086-7096 (2001) |
| MND-2.x.x.5440 | AY159322 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Hu, J. | <i>J Virol</i> 77 (8):4867-4880 (2003) |
| MNE.US.x.MNE027 | U79412 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Kimata, J.T. | <i>J Virol</i> 72 (1):245-256 (1998) |
| MON.CM.99.L1_99CML1 | AY340701 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Courgnaud, V. | <i>J Virol</i> 77 (23):12523-12534 (2003) |
| MUS-1.CM.01.CM1239 | EF070330 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Aghokeng, A.F. | <i>Virology</i> 360 (2):407-418 (2007) |
| MUS-2.CM.01.CM1246 | EF070329 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Aghokeng, A.F. | <i>Virology</i> 360 (2):407-418 (2007) |
| MUS-3.GA.09.09GabOI81 | KF304707 | Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env, Nef | Liegeois, F. | <i>Viruses</i> 6 (7):2880-2898 (2014) |
| OLC.CI.97.97CI12 | FM165200 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Liegeois, F. | <i>J Virol</i> 83 (1):428-439 (2009) |
| RCM.CM.00.SIVagi_00CM312 | HM803690 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Ahuka-Mundeke, S. | <i>J Gen Virol</i> 2010 Dec;91(Pt 12):2959-64 doi: 10.1099/vir0025767-0 Epub 2010 Aug 25 |

| Name | Accession | Proteins | Author | Reference |
|---------------------------|-----------|---|----------------------|---|
| RCM.CM.02.SIVrcm_02CM8081 | HM803689 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Ahuka-Mundeke, S. | <i>J Gen Virol</i> 2010 Dec;91(Pt 12):2959-64 doi: 101099/vir0025767-0 Epub 2010 Aug 25 |
| RCM.GA.x.SIVRCMGAB1 | AF382829 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Bibollet-Ruche, F. | <i>J Virol</i> 78 (14):7748-7762 (2004) |
| SAB.SN.x.SAB1 | U04005 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Jin, M.J. | <i>EMBO J</i> 13 (12):2935-2947 (1994) |
| SMM.CI.79.SIVsmCI2 | JX860430 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Fischer, W. | <i>J Virol</i> 86 (24):13217-13231 (2012) |
| SMM.SL.92.SIVsmSL92A | JX860432 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Fischer, W. | <i>J Virol</i> 86 (24):13217-13231 (2012) |
| SMM.SL.92.SL92B | AF334679 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Chen, Z. | <i>J Virol</i> 70 (6):3617-3627 (1996) |
| SMM.US.06.FTq | JX860414 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Fischer, W. | <i>J Virol</i> 86 (24):13217-13231 (2012) |
| SMM.US.11.SIVsmE660_FL8 | JQ864086 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Wu, F. | <i>J Virol</i> 86 (16); 8835-47 (2012) |
| SMM.US.86.CFU212 | JX860407 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Fischer, W. | <i>J Virol</i> 86 (24):13217-13231 (2012) |
| SMM.US.x.H9 | M80194 | Gag, Pol, Vif, Vpx, Vpr, Tat, Rev, Env, Nef | Courgnaud, V. | <i>J Virol</i> 66 (1):414-419 (1992) |
| SUN.GA.98.L14 | AF131870 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Beer, B.E. | <i>J Virol</i> 73 (9):7734-7744 (1999) |

| Name | Accession | Proteins | Author | Reference |
|---------------------------|-----------|---|----------------------|---|
| SYK.KE.x.KE51 | AY523867 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Bibollet-Ruche, F. | <i>J Virol</i> 78 (14):7748-7762 (2004) |
| SYK.KE.x.SYK173_COMGNM | L06042 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Hirsch, V.M. | <i>J Virol</i> 67 (3):1517-1528 (1993) |
| TAL.CM.00.266 | AY655744 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Liegeois, F. | <i>Virology</i> 349 (1):55-65 (2006) |
| TAL.CM.01.8023 | AM182197 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Liegeois, F. | <i>Virology</i> 349 (1):55-65 (2006) |
| TAN.UG.x.SIVagmTAN1 | U58991 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Soares, M.A. | <i>Virology</i> 228 (2):394-399 (1997) |
| VER.KE.x.AGM155 | M29975 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Johnson, P.R. | <i>J Virol</i> 64 (3):1086-1092 (1990) |
| VER.ZA.10.VSAC4004 | KR862336 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Bailey, A.L. | <i>J Virol</i> 2016 May 11 pii: JVI00573-16 |
| VER.ZA.10.VSAG1003 | KR862363 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Bailey, A.L. | <i>J Virol</i> 2016 May 11 pii: JVI00573-16 |
| VER.ZA.10.VSAM0022 | KR862356 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Bailey, A.L. | <i>J Virol</i> 2016 May 11 pii: JVI00573-16 |
| WRC.CI.98.98CI04 | AM713177 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Liegeois, F. | <i>J Virol</i> 83 (1):428-439 (2009) |
| WRC.GM.05.Pbt_05GM_X02 | AM937062 | Gag, Pol, Vif, Vpr, Tat, Rev, Env, Nef | Locatelli, S. | <i>Virology</i> 376 (1):90-100 (2008) |
| ASC.CD.13.SIVasc_13DRC_I8 | KY497574 | Tat, Rev | Ahuka-Mundeke, S. | <i>ARHR</i> 33 (7); 735-739 (2017) |
| H1B.US.90.WEAU160_GHOSH | U21135 | Vpu | Wei, X. | <i>Nature</i> 422 (6929):307-312 (2003) |
| H1D.CD.84.84ZR085 | U88822 | Vpu | Gao, F. | <i>J Virol</i> 72 (7):5680-5698 (1998) |
| H1G.SE.93.SE6165_G6165 | AF061642 | Vpu | Carr, J.K. | <i>Virology</i> 247 (1):22-31 (1998) |
| H1H.CF.90.056 | AF005496 | Vpu | Gao, F. | <i>J Virol</i> 72 (7):5680-5698 (1998) |
| H1J.SE.93.SE9280_7887 | AF082394 | Vpu | Laukkanen, T. | <i>ARHR</i> 15 (3):293-297 (1999) |
| H1K.CM.96.96CM_MP535 | AJ249239 | Vpu | Triques, K. | <i>ARHR</i> 16 (2):139-151 (2000) |
| H101_AE.TH.90.CM240 | U54771 | Vpu | Carr, J.K. | <i>J Virol</i> 70 (9):5935-5943 (1996) |

| Name | Accession | Proteins | Author | Reference |
|-----------------------------|-----------|----------|----------------------|--|
| H102_AG.NG.x.IBNG | L39106 | Vpu | Howard, T.M. | <i>ARHR</i> 10 (12):1755-1757 (1994) |
| H103_AB.RU.97.KAL153_2 | AF193276 | Vpu | Liitsola, K. | <i>ARHR</i> 16 (11):1047-1053 (2000) |
| H104_cpx.CY.94.94CY032_3 | AF049337 | Vpu | Gao, F. | <i>J Virol</i> 72 (12):10234-10241 (1998) |
| H1O.CM.94.BCF06 | AB485666 | Vpu | Takekawa, N. | Unpublished |
| H1O.CM.96.96CMA102 | AY169803 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.96.96CMABB009 | AY169806 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.96.96CMABB637 | AY169810 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.97.97CMABB447 | AY169813 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.97.97CMABB497 | AY169809 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.98.98CMA104 | AY169802 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.98.98CMABB141 | AY169807 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.98.98CMABB212 | AY169804 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.98.98CMU2901 | AY169812 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.98.98CMU5337 | AY169808 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.99.99CMU4122 | AY169815 | Vpu | Yamaguchi, J. | <i>ARHR</i> 19 (11):979-988 (2003) |
| H1O.CM.x.pCMO2_3 | AY618998 | Vpu | Tebit, D.M. | <i>Virology</i> 326 (2):329-339 (2004) |
| H1O.CM.x.pCMO2_5 | AY623602 | Vpu | Tebit, D.M. | <i>Virology</i> 326 (2):329-339 (2004) |
| H1O.FR.92.VAU | AF407418 | Vpu | Vartanian, J.P. | <i>J Gen Virol</i> 83 (Pt 4):801-805 (2002) |
| H1O.SN.99.99SE_MP1300 | AJ302647 | Vpu | Toure-Kane, C. | <i>ARHR</i> 17 (12):1211-1216 (2001) |
| H1O.US.x.I_2478B | AB485668 | Vpu | Takekawa, N. | Unpublished |
| H1N.CM.02.SJGddd | GQ324959 | Vpu | Vallari, A. | <i>ARHR</i> 26 (1):109-115 (2010) |
| H1N.CM.06.U14296 | GQ324962 | Vpu | Vallari, A. | <i>ARHR</i> 26 (1):109-115 (2010) |
| H1N.CM.06.U14842 | GQ324958 | Vpu | Vallari, A. | <i>ARHR</i> 26 (1):109-115 (2010) |
| H1N.CM.97.YBF106 | AJ271370 | Vpu | Roques, P. | <i>AIDS</i> 18 (10):1371-1381 (2004) |
| H1P.FR.06.RBF168 | GQ328744 | Vpu | Plantier, J.-C. | <i>Nat Med</i> 15 (8); 871-2 (2009) |
| CPZ.CM.05.LB715 | JX178450 | Vpu | Takehisa, J. | Unpublished |
| CPZ.CM.05.SIVcpzEK505.c2 | JN835460 | Vpu | Bibollet-Ruche, F. | <i>J Clin Invest</i> 122 (5):1644-1652 (2012) |
| CPZ.CM.05.SIVcpzMB897.c2 | JN835461 | Vpu | Bibollet-Ruche, F. | <i>J Clin Invest</i> 122 (5):1644-1652 (2012) |
| CPZ.CM.05.SIVcpzMT145.c2 | JN835462 | Vpu | Bibollet-Ruche, F. | <i>J Clin Invest</i> 122 (5):1644-1652 (2012) |
| CPZ.CM.05.SIVcpzMT145 | DQ373066 | Vpu | Keele, B.F. | <i>Science</i> 313 (5786):523-526 (2006) |
| CPZ.CM.98.CAM3 | AF115393 | Vpu | Corbet, S. | <i>J Virol</i> 74 (1):529-534 (2000) |
| CPZ.CM.98.CAM5 | AJ271369 | Vpu | Muller-Trutwin, M.C. | <i>J Med Primatol</i> 29 (3-4); 166-72 (2000) |
| CPZ.GA.88.GAB1 | X52154 | Vpu | Huet, T. | <i>Nature</i> 345 (6273):356-359 (1990) |
| CPZ.TZ.01.TAN2 | EF394357 | Vpu | Takehisa, J. | <i>J Virol</i> 81 (14):7463-7475 (2007) |
| CPZ.TZ.02.TAN3_1 | DQ374658 | Vpu | Takehisa, J. | <i>J Virol</i> 81 (14):7463-7475 (2007) |
| CPZ.TZ.06.TAN5 | JN091691 | Vpu | Rudicell, R.S. | <i>J Virol</i> 85 (19); 9918-28 (2011) |
| GOR.CM.07.SIVgorCP2139_1con | FJ424864 | Vpu | Takehisa, J. | <i>J Virol</i> 83 (4):1635-1648 (2009) |
| GOR.CM.07.SIVgorCP2139_2con | FJ424865 | Vpu | Takehisa, J. | <i>J Virol</i> 83 (4):1635-1648 (2009) |
| DEN.CD.x.CD1_CMO580407 | AJ580407 | Vpu | Dazza, M.C. | <i>J Virol</i> 79 (13):8560-8571 (2005) |
| MON.NG.x.NG1 | AJ549283 | Vpu | Barlow, K.L. | <i>J Virol</i> 77 (12):6879-6888 (2003) |
| MUS-1.CM.01.SIVmus_01CM1085 | AY340700 | Vpu | Cournaud, V. | <i>J Virol</i> 77 (23):12523-12534 (2003) |
| MUS-2.CM.01.CM2500 | EF070331 | Vpu | Aghokeng, A.F. | <i>Virology</i> 360 (2):407-418 (2007) |
| MUS-3.GA.11.11GabPts02 | KF304708 | Vpu | Liegeois, F. | <i>Viruses</i> 6 (7):2880-2898 (2014) |

Table with columns for membrane binding, Gag start, p17 start, and phosphorylation site. It contains multiple rows of amino acid sequences and their corresponding protein identifiers.

Table with columns for protein names, major homology region, p24 end_p2 start, p2 end_p7 start, and Zn motif. It lists various HIV-1 protein sequences and their conserved motifs across different regions.

| | Zn motif | p7 end_p1 start | p1 end_p6 start | Vpr binding | | | |
|---------------------------|------------------|-----------------|-----------------|---------------|-------------|-----------------|-------------|
| H1B.FR.83.HXB2 | GCWCKGKEGHQMKDCT | ERQANFLGK1WPS | YKGRPNFLQSRP | PTAPP | EESFRSGVETT | TPPOKQE | PIDKELYPLT |
| H1A1.UG.85.U455.U455A | | | | | A-I-GM-EKM | S-AKQE | LK-R-OT-VS |
| H1C.ET.86.ETH2220 | | RL | N | ESLRPEPTA | M-EVK-Q-E | PS-KQE | LK-R-ALTSLK |
| H1F1.BE.93.V1850 | R | GK | N | | A-GFRE-I | PA-KQE | QK-G |
| H1D.BE.87.AMT70 | | GN | Y-P | | M-EVK-Q-E | NOE-GGP-NE | FA |
| H10.CM.91.MVP5180 | | GN | Y-P | | M-EVKEQ-E | NOS-GDO-EE | FA |
| H10.CM.98.98CMA105 | | GN | Y-P | | M-EETK-Q-E | SOE-GSP-NE | FA |
| H10.CM.98.98CMA8197 | | GN | Y-P | | M-EMEK-Q-NR | EON-P-NE | FA |
| H10.SN.99.99SSE.MP1299 | | GN | Y-P | | M-TEEMK-Q-E | NOE-EDQ-NE | FA |
| H10.US.99.99STWLA | | GN | Y-P | | M-EVKEQ-E | NON-GECS-NE | FA |
| H1N.CM.02.GJ08131 | | EG | G-SP | | L-NYG-QE-RS | OGKERQE-NTEN | S-PLT |
| H1N.CM.04.04CM.1015.04 | | EG | G-SP | | L-YG-QE-GS | OGKEMQE-NOEK | TENS-PLT |
| H1N.CM.04.04CM.1131.03 | | EG | G-SP | | L-YG-QE-RS | ORKEMQE-NOEK | TENT-PLT |
| H1N.CM.95.YBF30 | | EG | S-SP | | L-YGFQE-KS | OGKEMQE-NOER | TENS-PLT |
| H1P.CM.96.U14788 | | KS | G | | M-EEEMTQ-N | KQXEEK-D-EKE | A |
| CPZ.CM.06.RT167 | | L | KS | | OTQEKLEM-R | SA-PME-V-NHW | NSKDFG |
| CPZ.CD.90.ANT | R-0 | L-N-PAT | NTGKV-PT-T | WMGC-V-KEEV-V | E | I-IYOEHEKR-QK | GLKGE |
| CPZ.CM.01.SIVcpzCAM13 | | G-V-S | RS-V-N-T | | A-YGOEKEK | NQEQE-G | NR-PLT |
| CPZ.CM.05.LB715 | | R-Q-KLS | LV-NS-P | EP | M-GFOQP | AOKMEVK-EEEOE | EKES-PLT |
| CPZ.CM.05.SIVcpzDP943 | | I-A-GD | G-SP-WSGGSK | G-K | I-YGOET-V | NKOGEKKE-MEVE | KEP-OLT |
| CPZ.CM.05.SIVcpzEK505 | | R-S | V-L-S | P-TTT-RK | I-YGOE-K | OGTER-E-KE-K | TESS-BLT |
| CPZ.CM.05.SIVcpzLB7 | | R-Q-NM | GK-V-L | SQ-P-EP | I-GFOPK | TOQOEDR-EEKTE | NK-PLT |
| CPZ.CM.05.SIVcpzMB66 | | RN-N | L-S | K-EP | I-GFRE-E | PAOKEEVK-SE-ETE | NK-OLT |
| CPZ.CM.05.SIVcpzMB897 | | Q-N-PD | MOR-P | K-EPP | I-GFODKI | DSLKOYVR-NEEOTE | NR-PLD |
| CPZ.GA.88.SIVcpzGAB2 | | ON | EG-V-GG | RS-P-RKV | P-T-GYAEQ-E | GKSQKE-EST | TEKES-YP |
| CPZ.TZ.09.TAN11 | R-0 | TRN | NTGV | LMGC-V-N | T | QOATFPVP | A-LEN |
| CPZ.TZ.06.SIVcpzTAN13 | R-0 | ARTNNTGV | RT | LMGC-V-N | N | QEGVQT-VP | PMKEGF |
| CPZ.TZ.09.UG38 | R-0 | GR | PNNSGV | RT | LMGC-V-N | T | TPVEP |
| CPZ.US.85.US.Marilyn | | AG | N | H-SPWSGGSK | EN-K | I-D-GYOE-V | QEKQK |
| GOR.CM.04.SIVgorCP684con | OD | RS | G | Y-P | GGE | YV-KQVQ | M-EEEK-E |
| GOR.CM.07.SIVgor2139.287 | OD | RS | G | Y-P | GGE | YV-KQVQ | M-EEKEM-R |
| GOR.CM.07.SIVgorCP2135con | OD | RS | G | Y-P | GGE | YV-KQVQ | M-EEK-E |
| GOR.CM.12.SIVgor.BPID1 | | K | SG | Y-P | GKG | YX-KQVQ | M-EEEMTQ-N |
| GOR.CM.12.SIVgor.BQID2 | | K | NG | Y-L | GGT | YV-KQAO | M-EVMK-K-E |
| GOR.CM.13.SIVgor.BPID15 | | K | IGK | Y-P | GKG | YV-KQVQ | M-EEEVQ-K-S |

| | | p8 end_p1 start | p1 end_p6 start | PTAP motif | | PSAP in HIV-2 B. U | |
|----------------------------|---------|-----------------|-----------------|----------------|----------------|--------------------|------------------|
| MAC.US.x.239 | | MD-V-AK-P | D-G | LGP-WG-K-R | PMAQVHQL | M | P-DPAVD |
| H2A.DE.x.BEN | | P-I-AN-P | G | LGP-RG-K-R | PVTOAPOGL | I | PADPAE |
| H2A.PT.x.ALI | | P-L-AN-P | G | LGP-WG-K-R | PVT-VPOGL | T | PAEPAAD |
| H2A.SN.86.ST.JSP4.27 | | A-I-AK-P | G | LGP-WG-K-R | PVAOIPQGL | T | PDPVED |
| H2B.CI.x.EH0 | | Q-I-SK-P | G | FGP-WG-K-R | P-VQAPQGL | V-S | PADPAE |
| H2B.GH.86.D295.ALT | | P-I-AK-P | G | LGP-WG-K-R | PMTQVPOGL | T-S | MNPAEGMTPGATPSAP |
| H2C.CI.92.Abt96 | | P-I-AK-P | G | FGP-WG-K-R | PMTQVPOGL | T-S | PMDPAVD |
| H2U.CI.07.071C.TMP3 | | P-I-AK-P | G | FGP-WG-K-R | PMTQIPQGL | T | PADPAVD |
| H2U.FR.96.12034 | | E-P-N-AS-P | | TKVG | LG-WG-K-R | PMAQVPOGL | T |
| ASC.UG.10.RT63 | K-F | R-MARO-RAS | EGAQGG | AFLDKHLGRPKG-R | PVVO | S | LTLQAD |
| ASC.UG.10.RT68 | K-F | R-MARO-RAS | EGAQGG | AFLDKHLGRPKG-R | PVVO | S | LTLQAD |
| ASC.UG.10.RT111 | K-F | R-MARO-RAS | EGAQGG | AFLDKHLGRPKG-R | PVVO | S | LTLQAD |
| COL.CM.x.CGU1 | RGAPRRP | VR-FT-NQ-MQR-P | NK | ATRGVELQ | ATF | A | KMSKD |
| COL.UG.10.BWC01 | FG | QL-IQR-Q | NK | AGRGIDLQ | AVF | A | MSKMQO |
| COL.UG.10.BW07 | FG | QL-IQR-Q | NK | AGRGIDLQ | AVF | A | MSKMQO |
| DEB.CM.04.SIVdeb04CMPF3061 | R-F | Q-T-MA-N-P | KAAPFSKN | KLP-WGRRA-F | IEQHQH | WIGIQY-EG | KMSQRK |
| DEB.CM.99.CM40 | K-F | Q-IA-N-G | QTPR | NTP-WASR-R | E | GIGKYI-EST | ADQVST |
| DEB.CM.99.CM5 | K-F | Q-IA-N-G | QVPR | NAP-WASR-R | E | GIGKYI-EST | ADQVST |
| DRL.DE.11.D3 | FR | ALD-MLRN | PK-V | NTP-WGS-K-R | P-ASL | T-S | LPGY |
| DRL.DE.11.D4 | FR | ALD-MLRN | PK-V | NTP-WGS-K-R | P-ASL | T-S | MPGF |
| GRV.ET.x.GRI.677.gri.1 | K-F | RA-IA-N-P | NG | YGH-WGAK-R | V-Y-GDTVG | LE | METAYDPA |
| GSN.CM.99.CN16 | K-F | RA-IA-N-P | NG | YGH-WGAK-R | V-Y-GDTVG | LE | METAYDPA |
| GSN.CM.99.CN71 | K-F | A-LARO-KTG | TPAKEL-D | LG-RGRR-A-R | PVTS | LL-S | AEEV |
| LST.CD.88.SIVLhoest485 | N | SKE-RFAQ-PKP | GKGV | YGP-WRN-P | PLMGGNAG | VV-S | M-RNS |
| LST.CD.88.SIVLhoest524 | N | AKD-R-AQ-PKP | GKGV | YGP-WRN-PS | PTMGGAAG | VV-S | M-KVPP |
| LST.KE.x.lho7 | N | SKE-RFAQ-P | KPKGV | YGP-WSSKP-R | YPLGGGAAG | RI-S | M-AP |
| MAL.ZM.99.SIVagMAL_ZMB | K-F | I-MA-K | G-V | YGR-WMGTK-R | AAT-GEPR | S | AYDPA |
| MND-1.GA.x.MNDGB1 | F | DL-I-RN-P | KMV | NTP-WGS-K-R | PAMPL | T-S | MPG |
| MND-2.CM.98.CM16 | F | DL-I-RN-P | KMV | NTP-WGS-K-R | PAMPL | T-S | MPG |
| MND-2.GA.x.M14 | F | DT-M-RN-P | KMV | NAP-WGS-K-R | PAMPL | T-S | MPG |
| MND-2.x.x.5440 | F | DP-L-RN-P | PKMV | NTP-WGS-K-R | PAMPL | T-S | MPG |
| MNE.US.x.MNE027 | QH | V-AK-P | G | FGP-WG-K-R | PMAQMPQGL | T | P-DPAVD |
| MON.CM.99.L1.99CML1 | K-YN | OP-LA-P | POPP | KONKGV | N-F | GP-KG-R | PLTS |
| MUS-1.CM.01.CM1239 | K-F | R-LA-O-RSE | GAKS | DPLGLQRR | R-P-TS | LL | EQ |
| MUS-2.CM.01.CM1246 | R-F | LARO-RD | TGCSAV | KDLGFRSRK | PVTS | LM | D |
| MUS-3.GA.09.09gab0181 | K-F | R-MA-O-N | TSTGRV | NP | FGPO-RG-R | PVTO | LV-G |
| RCM.CM.00.SIVag1.00CM312 | T-Y | E-D-IL-P | GG000Q5KV | KWR | GKGAEPLNMQORTE | MDIF | S-T |
| RCM.CM.02.SIVrcm.02CM8081 | D-D | L | K-V | SAP | WSR-K-R | PLTSL | T-S |
| RCM.GA.x.SIVRCMGAB1 | SPD | LL | K-I | RLP | WGO-K-R | PLTSL | T-S |
| SAB.SN.x.SAB1 | SPD | LL | K-V | FGP | WGR-K-R | PLTSL | T-S |
| SMM.CI.79.SIVsMC12 | S | V-AK-P | G | FGP-WG-K-R | PMTQLOOGL | T | MDBAVAPAMPDP |
| SMM.SI.92.SIVsLS192A | S | V-AK-P | G | FGP-WG-K-R | PMTQLOOGL | T | MDBAVAPAMPDP |
| SMM.SL.92.SL92B | E | RIOAN-P | NQK-G | LGP-WG-K-R | PMTQITSL | T-S | PDPAA-R |
| SMM.US.06.FTQ | M | I-AK-P | VG | FGP-WG-K-R | PMTQMPQGL | T | DLARD |
| SMM.US.11.SIVsME660.FL8 | G | T-V-AK-P | G | LGP-WG-K-R | PMAQMPQGL | I | P-DPAVD |
| SMM.US.86.CFU212 | M | V-AK-P | G | FGP-WG-K-R | PMAQVHQL | T | PADPAVD |
| SMM.US.x.H9 | M | V-AK-P | G | FGP-WG-K-R | PMAQMPQGL | T | P-DPAVD |
| SUN.GA.x.14 | PPGS | M-KOAO-P | QK-V | LGP-WG-K-R | PVM | S | LEDLT |
| SYK.KE.x.KE51 | K-FN | GT-IARO | ROPKRGQNPFP | G | WG-NRKP-A | PVM-E | T |
| SYK.KE.x.SYK173.COMGNM | K-FN | T-LARA | ROPKRGQNPFP | G | WGVSRP-A | PVRSE | S |
| TAL.CM.00.266 | PPGT | Y-P-IA-O-S | SA | S | FLGKIPVVRG-R | PVVO | S |
| TAL.CM.01.8023 | PPGS | F-QM-RAAQ-R | QP | A | FLGKIPVVRG-R | PVVO | S |
| TAN.UG.x.SIVtanTAN1 | K-L | P-LA-R | SG | RMP | TWGT-K-R | EGGAV | S |
| VER.KE.x.AGM155 | K-L | P-LA-R | G-V | YGR-WMGTK-R | PAATLGA | S | MPAHGFPTGSP |
| VER.ZA.10.VSAC4004 | K-F | T-MA-R | G-V | YGR-WMGSK-R | AATAGEEDR | TS | PP |
| VER.ZA.10.VSAG1003 | K-L | P-MA-R | G-V | YGR-WMGSK-R | AAT-GEEDR | T | PP |
| VER.ZA.10.VSAM0822 | R-L | P-MA-R | G-V | YGR-WAGSK-R | AAT-GEEDN | F | S |
| WRC.CI.99.99C104 | RGQPT | G-ET-VORN-P | KSR | T | NGR | GOMPSN | YPVQKMLNFGPATT |
| WRC.CM.05.Pdt.05GM.X02 | GPPT | G-T-IQ-N-P | ETR | V | NGR | GORPSN | YPVQKMLNFGV |

| | Vpr binding p6 end, Gag end | | |
|----------------------------|--------------------------------|--|--|
| H1B.FR.83.HXB2 |SLR.SLFGNDPSS..0* | | |
| H1A1.UG.85.U455_U455A |LK.....L.. | | |
| H1C.ET.86.ETH2220 |HLL..... | | |
| H1F1.BE.93.VI850 |SLK..... | | |
| H10.BE.87.ANT70 |K.....T.. | | |
| H10.CM.91.MVP5180 |K.....T.. | | |
| H10.CM.98.98CMA105 |K.....T.. | | |
| H10.CM.98.98CMAB197 |K.....T.. | | |
| H10.SN.99.99SE_MP1299 |K.....T.. | | |
| H10.US.99.99USTWLA |K.....XT.. | | |
| H1N.CM.02.DJ00131 |L..... | | |
| H1N.CM.04.04CM_1015_04 |L..... | | |
| H1N.CM.04.04CM_1131_03 |L..... | | |
| H1N.CM.95.YBF30 |K.....T.. | | |
| H1P.CM.06.U14788 |K.....I..S.. | | |
| CPZ.CM.06.SF1167 |K.....I..S.. | | |
| CPZ.CD.90.ANT |K.....K.. | | |
| CPZ.CM.01.SIVcpzCAM13 |K.....S.. | | |
| CPZ.CM.05.LB715 |K.....S..0.. | | |
| CPZ.CM.05.SIVcpzDP943 |K.....S.....E.. | | |
| CPZ.CM.05.SIVcpzEK505 |S.....L.. | | |
| CPZ.CM.05.SIVcpzLB7 |K.....L.. | | |
| CPZ.CM.05.SIVcpzMB66 |K.....L.. | | |
| CPZ.CM.05.SIVcpzMB897 |K.....L..Q0E.. | | |
| CPZ.GA.88.SIVcpzGAB2 |K.....S.. | | |
| CPZ.TZ.09.TAN1 |K.....T..S.. | | |
| CPZ.TZ.06.SIVcpzTAN13 |K.....T..S.. | | |
| CPZ.TZ.09.UG38 |K.....I..S.. | | |
| CPZ.US.85.US_Marilyn |K.....S.....E.. | | |
| GOR.CM.04.SIVgorCP684con |K.....T.. | | |
| GOR.CM.07.SIVgor2139_287 |K.....T.. | | |
| GOR.CM.07.SIVgorCP2135con |K.....T.. | | |
| GOR.CM.12.SIVgor_BP1D1 |K.....T.. | | |
| GOR.CM.12.SIVgor_BQ1D2 |K.....T.. | | |
| GOR.CM.13.SIVgor_BP1D5 |K.....T.. | | |
| | p6 end Gag end | | |
| MAC.US.x.239 |EDLLH-N....G.. | | |
| H2A.DE.x.BEN | LLHLEORETPHREETEDLLH-N...K.. | | |
| H2A.PT.x.ALI | LLHLEOGETPHKVEYTEDLLH-N...K.. | | |
| H2A.SN.86.ST_JSP4_27 | FLQLEKQETPCKRETEDLLH-N...K.. | | |
| H2B.CI.x.EH0 |EDLLH-N...E.. | | |
| H2B.GH.86.D205_ALT |EDLLH-N...E.. | | |
| H2G.CI.92.Abt96 |EXLLH-S...D.. | | |
| H2U.CI.07.07IC_TNP3 |EDLLH-N...D.. | | |
| H2U.FR.96.12034 |EDLLH-N...D.. | | |
| ASC.UG.10.RT03 |D..... | | |
| ASC.UG.10.RT08 |D..... | | |
| ASC.UG.10.RT11 |D..... | | |
| COL.CM.x.CGU1 |D..... | | |
| COL.UG.10.BWC01 |D.....0..S.. | | |
| COL.UG.10.BWC07 |K.....E..LL..GQM.. | | |
| DEB.CM.04.SIVdeb04CMPF3061 |DLSD..K.....E.. | | |
| DEB.CM.99.CM40 |DLMG..K.....E.. | | |
| DEB.CM.99.CM5 |DLMG..K.....E.. | | |
| DRL.DE.11.D3 |N.....S.....QL.. | | |
| DRL.DE.11.D4 |N.....S.....QL.. | | |
| GRV.ET.x.GRI_677_gri_1 |VED..V..S.....G.. | | |
| GSN.CM.99.CN166 |E..... | | |
| GSN.CM.99.CN71 |E..... | | |
| LST.CD.88.SIVlhoest485 |VDEPC..SFFF..PD.. | | |
| LST.CD.88.SIVlhoest524 |VDEPC..NFFF..PD.. | | |
| LST.KE.x.lho7 |VDEPC..N.....PD.. | | |
| MAL.ZM.09.SIVagmMAL_ZMB |WED..Y..N.....EG.. | | |
| MND-1.GA.x.MNDGB1 |EDFH..ST..QE.. | | |
| MND-2.CM.98.CM16 |N.....S.....T..QL.. | | |
| MND-2.GA.x.M14 |N.....S.....T..QL.. | | |
| MND-2.x.x.5440 |N.....S.....T..QL.. | | |
| MNE.US.x.MNE027 |EDLLH-N...E.. | | |
| MON.CM.99.L1_99CML1 |A.....E.. | | |
| MUS-1.CM.01.CM1239 |OD..LE..K.. | | |
| MUS-2.CM.01.CM1246 |D..Q..K.. | | |
| MUS-3.GA.09.09GAb0I81 |D..Q..E.. | | |
| OLC.CI.97.07C112 |E..... | | |
| RCM.CM.00.SIVag1_00CM312 |PLYPSLA..... | | |
| RCM.CM.02.SIVrcm_02CM8081 |S.....N..... | | |
| RCM.GA.x.SIVRCMGAB1 |A.....N.....S..... | | |
| SAB.SN.x.SAB1 |T.....N.....S.....QL.. | | |
| SMM.CI.79.SIVsmCI2 |S.....F..... | | |
| SMM.SL.92.SIVsmSL92A |EDLLHLN...D.. | | |
| SMM.SL.92.SL92B |EDLLH-N...E.. | | |
| SMM.US.06.FTq |EDLLH-N...E.. | | |
| SMM.US.11.SIVsmE660_FL8 |EDLLH-S...E.. | | |
| SMM.US.86.CFU212 |EDLLH-N...E.. | | |
| SMM.US.x.H9 |EDLLH-N...D.. | | |
| SUN.GA.98.L14 |EDLLH-N...E.. | | |
| SYK.KE.x.KE51 |CQEPC..N..M..PE..P.. | | |
| SYK.KE.x.SYK173_COMGNM |D..... | | |
| TAL.CM.00.266 |E..... | | |
| TAL.CM.01.8023 |D..... | | |
| TAN.UG.x.SIVagmTAN1 |LED..Y..S..... | | |
| VER.KE.x.AGM155 |WNEGY..N.....E.. | | |
| VER.ZA.10.VSAC4004 |WED..Y..S.....G.....L.. | | |
| VER.ZA.10.VSAG1003 |WED..Y..K.....G..... | | |
| VER.ZA.10.VSAM0022 |WED..Y.....G..... | | |
| WRC.CI.98.98C104 |GLPEWPSLK...E.. | | |
| WRC.GM.05.Pbt_05GM_X02 |VEYPSLK...EP..... | | |

500
494
505
499
499
499
500
498
499
513
512
512
513
493
529
523
507
501
507
512
504
500
510
508
525
525
515
506
490
490
493
499
491

511
522
522
522
520
522
510
510
517
520
520
483
492
494
545
536
533
522
522
514
513
511
513
512
516
518
503
518
519
519
507
511
506
517
514
481
518
524
519
555
519
507
508
514
508
509
508
522
562
555
528
532
524
521
525
523
524
514
503

| | | | | | | | | |
|---------------------------|---------------|----------|-------------------|---------|-------|--|-------------------------------|-----|
| H1B.FR.83.HXB2 | FFREDLAFLOGKA | REFSSEQT | RANSPTRRELOVWGRDN | NSPSEAG | ADROG | TVSFNFOVTLWQRPLVTIKIGGOLKEALLTDGADTVLEEMSL | PGRWKPMIGGIGGFKVROYDQIILEICGH | 125 |
| H1A1.UG.85.U455.U455A | N | Q | E | S | W | D | G | K |
| H1C.ET.86.ETH2220 | N | Q | E | S | W | D | G | K |
| H1F1.BE.93.V1850 | N | Q | E | S | W | D | G | K |
| H1E.87.ANT70 | N | Q | E | S | W | D | G | K |
| H10.CM.91.MVP5180 | N | Q | E | S | W | D | G | K |
| H10.CM.98.98CMA105 | N | Q | E | S | W | D | G | K |
| H10.CM.98.98CMA197 | N | Q | E | S | W | D | G | K |
| H10.SN.99.99SME1299 | N | Q | E | S | W | D | G | K |
| H10.SN.99.99STWLA | N | Q | E | S | W | D | G | K |
| H1N.CM.02.J008131 | N | Q | E | S | W | D | G | K |
| H1N.CM.04.04CM.1015.04 | N | Q | E | S | W | D | G | K |
| H1N.CM.04.04CM.1131.03 | N | Q | E | S | W | D | G | K |
| H1N.CM.05.YBF30 | N | Q | E | S | W | D | G | K |
| H1R.CM.06.U14788 | N | Q | E | S | W | D | G | K |
| CPZ.CM.06.P1167 | N | Q | E | S | W | D | G | K |
| CPZ.CM.09.ANT | N | Q | E | S | W | D | G | K |
| CPZ.CM.01.SIVcpzCAM13 | N | Q | E | S | W | D | G | K |
| CPZ.CM.05.LB715 | N | Q | E | S | W | D | G | K |
| CPZ.CM.05.SIVcpzDP943 | N | Q | E | S | W | D | G | K |
| CPZ.CM.05.SIVcpzEK505 | N | Q | E | S | W | D | G | K |
| CPZ.CM.05.SIVcpzLB7 | N | Q | E | S | W | D | G | K |
| CPZ.CM.05.SIVcpzMB66 | N | Q | E | S | W | D | G | K |
| CPZ.CM.05.SIVcpzMB897 | N | Q | E | S | W | D | G | K |
| CPZ.GA.88.SIVcpzGAB2 | N | Q | E | S | W | D | G | K |
| CPZ.TZ.09.UG38 | N | Q | E | S | W | D | G | K |
| CPZ.TZ.09.UG38 | N | Q | E | S | W | D | G | K |
| CPZ.US.85.US.Marilyn | N | Q | E | S | W | D | G | K |
| GOR.CM.04.SIVgorCP684con | N | Q | E | S | W | D | G | K |
| GOR.CM.07.SIVgorP2139.287 | N | Q | E | S | W | D | G | K |
| GOR.CM.07.SIVgorCP2135con | N | Q | E | S | W | D | G | K |
| GOR.CM.12.SIVgor.BPID1 | N | Q | E | S | W | D | G | K |
| GOR.CM.12.SIVgor.BPID2 | N | Q | E | S | W | D | G | K |
| GOR.CM.13.SIVgor.BPID15 | N | Q | E | S | W | D | G | K |

| | | | | | | | | |
|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| MAC.US.x.239 | MAC.US.x.239 | MAC.US.x.239 | MAC.US.x.239 | MAC.US.x.239 | MAC.US.x.239 | MAC.US.x.239 | MAC.US.x.239 | MAC.US.x.239 |
| H2A.DE.x.BEN | H2A.DE.x.BEN | H2A.DE.x.BEN | H2A.DE.x.BEN | H2A.DE.x.BEN | H2A.DE.x.BEN | H2A.DE.x.BEN | H2A.DE.x.BEN | H2A.DE.x.BEN |
| H2A.PT.x.ALI | H2A.PT.x.ALI | H2A.PT.x.ALI | H2A.PT.x.ALI | H2A.PT.x.ALI | H2A.PT.x.ALI | H2A.PT.x.ALI | H2A.PT.x.ALI | H2A.PT.x.ALI |
| H2A.SB.S6.S2.JSP4.27 | H2A.SB.S6.S2.JSP4.27 | H2A.SB.S6.S2.JSP4.27 | H2A.SB.S6.S2.JSP4.27 | H2A.SB.S6.S2.JSP4.27 | H2A.SB.S6.S2.JSP4.27 | H2A.SB.S6.S2.JSP4.27 | H2A.SB.S6.S2.JSP4.27 | H2A.SB.S6.S2.JSP4.27 |
| H2B.CI.x.EH0 | H2B.CI.x.EH0 | H2B.CI.x.EH0 | H2B.CI.x.EH0 | H2B.CI.x.EH0 | H2B.CI.x.EH0 | H2B.CI.x.EH0 | H2B.CI.x.EH0 | H2B.CI.x.EH0 |
| H2C.GH.86.D295.17 | H2C.GH.86.D295.17 | H2C.GH.86.D295.17 | H2C.GH.86.D295.17 | H2C.GH.86.D295.17 | H2C.GH.86.D295.17 | H2C.GH.86.D295.17 | H2C.GH.86.D295.17 | H2C.GH.86.D295.17 |
| H2C.CI.92.Abt1c | H2C.CI.92.Abt1c | H2C.CI.92.Abt1c | H2C.CI.92.Abt1c | H2C.CI.92.Abt1c | H2C.CI.92.Abt1c | H2C.CI.92.Abt1c | H2C.CI.92.Abt1c | H2C.CI.92.Abt1c |
| H2U.CI.07.071C.TMP3 | H2U.CI.07.071C.TMP3 | H2U.CI.07.071C.TMP3 | H2U.CI.07.071C.TMP3 | H2U.CI.07.071C.TMP3 | H2U.CI.07.071C.TMP3 | H2U.CI.07.071C.TMP3 | H2U.CI.07.071C.TMP3 | H2U.CI.07.071C.TMP3 |
| H2U.FR.96.12034 | H2U.FR.96.12034 | H2U.FR.96.12034 | H2U.FR.96.12034 | H2U.FR.96.12034 | H2U.FR.96.12034 | H2U.FR.96.12034 | H2U.FR.96.12034 | H2U.FR.96.12034 |
| ASC.UG.10.RT83 | ASC.UG.10.RT83 | ASC.UG.10.RT83 | ASC.UG.10.RT83 | ASC.UG.10.RT83 | ASC.UG.10.RT83 | ASC.UG.10.RT83 | ASC.UG.10.RT83 | ASC.UG.10.RT83 |
| ASC.UG.10.RT08 | ASC.UG.10.RT08 | ASC.UG.10.RT08 | ASC.UG.10.RT08 | ASC.UG.10.RT08 | ASC.UG.10.RT08 | ASC.UG.10.RT08 | ASC.UG.10.RT08 | ASC.UG.10.RT08 |
| ASC.UG.10.RT11 | ASC.UG.10.RT11 | ASC.UG.10.RT11 | ASC.UG.10.RT11 | ASC.UG.10.RT11 | ASC.UG.10.RT11 | ASC.UG.10.RT11 | ASC.UG.10.RT11 | ASC.UG.10.RT11 |
| COL.CM.x.CGU1 | COL.CM.x.CGU1 | COL.CM.x.CGU1 | COL.CM.x.CGU1 | COL.CM.x.CGU1 | COL.CM.x.CGU1 | COL.CM.x.CGU1 | COL.CM.x.CGU1 | COL.CM.x.CGU1 |
| COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 |
| COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 | COL.UG.10.BW07 |
| COL.UG.04.SIVdeb04CMPF3061 | COL.UG.04.SIVdeb04CMPF3061 | COL.UG.04.SIVdeb04CMPF3061 | COL.UG.04.SIVdeb04CMPF3061 | COL.UG.04.SIVdeb04CMPF3061 | COL.UG.04.SIVdeb04CMPF3061 | COL.UG.04.SIVdeb04CMPF3061 | COL.UG.04.SIVdeb04CMPF3061 | COL.UG.04.SIVdeb04CMPF3061 |
| DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 |
| DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 | DEB.CM.99.CM40 |
| DRL.DE.11.D3 | DRL.DE.11.D3 | DRL.DE.11.D3 | DRL.DE.11.D3 | DRL.DE.11.D3 | DRL.DE.11.D3 | DRL.DE.11.D3 | DRL.DE.11.D3 | DRL.DE.11.D3 |
| DRL.DE.11.D4 | DRL.DE.11.D4 | DRL.DE.11.D4 | DRL.DE.11.D4 | DRL.DE.11.D4 | DRL.DE.11.D4 | DRL.DE.11.D4 | DRL.DE.11.D4 | DRL.DE.11.D4 |
| GRV.ET.x.GRI.677.gri.1 | GRV.ET.x.GRI.677.gri.1 | GRV.ET.x.GRI.677.gri.1 | GRV.ET.x.GRI.677.gri.1 | GRV.ET.x.GRI.677.gri.1 | GRV.ET.x.GRI.677.gri.1 | GRV.ET.x.GRI.677.gri.1 | GRV.ET.x.GRI.677.gri.1 | GRV.ET.x.GRI.677.gri.1 |
| GSM.CM.99.CN16 | GSM.CM.99.CN16 | GSM.CM.99.CN16 | GSM.CM.99.CN16 | GSM.CM.99.CN16 | GSM.CM.99.CN16 | GSM.CM.99.CN16 | GSM.CM.99.CN16 | GSM.CM.99.CN16 |
| GSM.CM.99.CN71 | GSM.CM.99.CN71 | GSM.CM.99.CN71 | GSM.CM.99.CN71 | GSM.CM.99.CN71 | GSM.CM.99.CN71 | GSM.CM.99.CN71 | GSM.CM.99.CN71 | GSM.CM.99.CN71 |
| LST.CD.88.SIVLhoest485 | LST.CD.88.SIVLhoest485 | LST.CD.88.SIVLhoest485 | LST.CD.88.SIVLhoest485 | LST.CD.88.SIVLhoest485 | LST.CD.88.SIVLhoest485 | LST.CD.88.SIVLhoest485 | LST.CD.88.SIVLhoest485 | LST.CD.88.SIVLhoest485 |
| LST.CD.88.SIVLhoest524 | LST.CD.88.SIVLhoest524 | LST.CD.88.SIVLhoest524 | LST.CD.88.SIVLhoest524 | LST.CD.88.SIVLhoest524 | LST.CD.88.SIVLhoest524 | LST.CD.88.SIVLhoest524 | LST.CD.88.SIVLhoest524 | LST.CD.88.SIVLhoest524 |
| LST.KE.x.lho7 | LST.KE.x.lho7 | LST.KE.x.lho7 | LST.KE.x.lho7 | LST.KE.x.lho7 | LST.KE.x.lho7 | LST.KE.x.lho7 | LST.KE.x.lho7 | LST.KE.x.lho7 |
| MAL.ZM.99.SIVagmMAL_ZMB | MAL.ZM.99.SIVagmMAL_ZMB | MAL.ZM.99.SIVagmMAL_ZMB | MAL.ZM.99.SIVagmMAL_ZMB | MAL.ZM.99.SIVagmMAL_ZMB | MAL.ZM.99.SIVagmMAL_ZMB | MAL.ZM.99.SIVagmMAL_ZMB | MAL.ZM.99.SIVagmMAL_ZMB | MAL.ZM.99.SIVagmMAL_ZMB |
| MND.1.GA.x.MND6B1 | MND.1.GA.x.MND6B1 | MND.1.GA.x.MND6B1 | MND.1.GA.x.MND6B1 | MND.1.GA.x.MND6B1 | MND.1.GA.x.MND6B1 | MND.1.GA.x.MND6B1 | MND.1.GA.x.MND6B1 | MND.1.GA.x.MND6B1 |
| MND.2.CM.98.CM16 | MND.2.CM.98.CM16 | MND.2.CM.98.CM16 | MND.2.CM.98.CM16 | MND.2.CM.98.CM16 | MND.2.CM.98.CM16 | MND.2.CM.98.CM16 | MND.2.CM.98.CM16 | MND.2.CM.98.CM16 |
| MND.2.GA.x.M14 | MND.2.GA.x.M14 | MND.2.GA.x.M14 | MND.2.GA.x.M14 | MND.2.GA.x.M14 | MND.2.GA.x.M14 | MND.2.GA.x.M14 | MND.2.GA.x.M14 | MND.2.GA.x.M14 |
| MND.2.x.x.5440 | MND.2.x.x.5440 | MND.2.x.x.5440 | MND.2.x.x.5440 | MND.2.x.x.5440 | MND.2.x.x.5440 | MND.2.x.x.5440 | MND.2.x.x.5440 | MND.2.x.x.5440 |
| MNE.US.x.MNE027 | MNE.US.x.MNE027 | MNE.US.x.MNE027 | MNE.US.x.MNE027 | MNE.US.x.MNE027 | MNE.US.x.MNE027 | MNE.US.x.MNE027 | MNE.US.x.MNE027 | MNE.US.x.MNE027 |
| MON.CM.99.11.99CM11 | MON.CM.99.11.99CM11 | MON.CM.99.11.99CM11 | MON.CM.99.11.99CM11 | MON.CM.99.11.99CM11 | MON.CM.99.11.99CM11 | MON.CM.99.11.99CM11 | MON.CM.99.11.99CM11 | MON.CM.99.11.99CM11 |
| MUS.1.CM.01.CM1239 | MUS.1.CM.01.CM1239 | MUS.1.CM.01.CM1239 | MUS.1.CM.01.CM1239 | MUS.1.CM.01.CM1239 | MUS.1.CM.01.CM1239 | MUS.1.CM.01.CM1239 | MUS.1.CM.01.CM1239 | MUS.1.CM.01.CM1239 |
| MUS.2.CM.01.CM1246 | MUS.2.CM.01.CM1246 | MUS.2.CM.01.CM1246 | MUS.2.CM.01.CM1246 | MUS.2.CM.01.CM1246 | MUS.2.CM.01.CM1246 | MUS.2.CM.01.CM1246 | MUS.2.CM.01.CM1246 | MUS.2.CM.01.CM1246 |
| MUS.3.GA.09.09gab0181 | MUS.3.GA.09.09gab0181 | MUS.3.GA.09.09gab0181 | MUS.3.GA.09.09gab0181 | MUS.3.GA.09.09gab0181 | MUS.3.GA.09.09gab0181 | MUS.3.GA.09.09gab0181 | MUS.3.GA.09.09gab0181 | MUS.3.GA.09.09gab0181 |
| MUS.UG.x.H9 | MUS.UG.x.H9 | MUS.UG.x.H9 | MUS.UG.x.H9 | MUS.UG.x.H9 | MUS.UG.x.H9 | MUS.UG.x.H9 | MUS.UG.x.H9 | MUS.UG.x.H9 |
| RCM.CM.09.SIVag1.00CM312 | RCM.CM.09.SIVag1.00CM312 | RCM.CM.09.SIVag1.00CM312 | RCM.CM.09.SIVag1.00CM312 | RCM.CM.09.SIVag1.00CM312 | RCM.CM.09.SIVag1.00CM312 | RCM.CM.09.SIVag1.00CM312 | RCM.CM.09.SIVag1.00CM312 | RCM.CM.09.SIVag1.00CM312 |
| RCM.CM.02.SIVrcm.02CM8081 | RCM.CM.02.SIVrcm.02CM8081 | RCM.CM.02.SIVrcm.02CM8081 | RCM.CM.02.SIVrcm.02CM8081 | RCM.CM.02.SIVrcm.02CM8081 | RCM.CM.02.SIVrcm.02CM8081 | RCM.CM.02.SIVrcm.02CM8081 | RCM.CM.02.SIVrcm.02CM8081 | RCM.CM.02.SIVrcm.02CM8081 |
| RCM.GA.x.SIVRCMAGB1 | RCM.GA.x.SIVRCMAGB1 | RCM.GA.x.SIVRCMAGB1 | RCM.GA.x.SIVRCMAGB1 | RCM.GA.x.SIVRCMAGB1 | RCM.GA.x.SIVRCMAGB1 | RCM.GA.x.SIVRCMAGB1 | RCM.GA.x.SIVRCMAGB1 | RCM.GA.x.SIVRCMAGB1 |
| SAB.SN.x.SAB1 | SAB.SN.x.SAB1 | SAB.SN.x.SAB1 | SAB.SN.x.SAB1 | SAB.SN.x.SAB1 | SAB.SN.x.SAB1 | SAB.SN.x.SAB1 | SAB.SN.x.SAB1 | SAB.SN.x.SAB1 |
| SMM.CI.79.SIVsmC12 | SMM.CI.79.SIVsmC12 | SMM.CI.79.SIVsmC12 | SMM.CI.79.SIVsmC12 | SMM.CI.79.SIVsmC12 | SMM.CI.79.SIVsmC12 | SMM.CI.79.SIVsmC12 | SMM.CI.79.SIVsmC12 | SMM.CI.79.SIVsmC12 |
| SMM.CI.92.SIVsmL92A | SMM.CI.92.SIVsmL92A | SMM.CI.92.SIVsmL92A | SMM.CI.92.SIVsmL92A | SMM.CI.92.SIVsmL92A | SMM.CI.92.SIVsmL92A | SMM.CI.92.SIVsmL92A | SMM.CI.92.SIVsmL92A | SMM.CI.92.SIVsmL92A |
| SMM.SL.92.SL92B | SMM.SL.92.SL92B | SMM.SL.92.SL92B | SMM.SL.92.SL92B | SMM.SL.92.SL92B | SMM.SL.92.SL92B | SMM.SL.92.SL92B | SMM.SL.92.SL92B | SMM.SL.92.SL92B |
| SMM.US.06.FtQ | SMM.US.06.FtQ | SMM.US.06.FtQ | SMM.US.06.FtQ | SMM.US.06.FtQ | SMM.US.06.FtQ | SMM.US.06.FtQ | SMM.US.06.FtQ | SMM.US.06.FtQ |
| SMM.US.11.SIVsmE660.FL8 | SMM.US.11.SIVsmE660.FL8 | SMM.US.11.SIVsmE660.FL8 | SMM.US.11.SIVsmE660.FL8 | SMM.US.11.SIVsmE660.FL8 | SMM.US.11.SIVsmE660.FL8 | SMM.US.11.SIVsmE660.FL8 | SMM.US.11.SIVsmE660.FL8 | SMM.US.11.SIVsmE660.FL8 |
| SMM.US.86.CFU212 | SMM.US.86.CFU212 | SMM.US.86.CFU212 | SMM.US.86.CFU212 | SMM.US.86.CFU212 | SMM.US.86.CFU212 | SMM.US.86.CFU212 | SMM.US.86.CFU212 | SMM.US.86.CFU212 |
| SMM.US.x.H9 | SMM.US.x.H9 | SMM.US.x.H9 | SMM.US.x.H9 | SMM.US.x.H9 | SMM.US.x.H9 | SMM.US.x.H9 | SMM.US.x.H9 | SMM.US.x.H9 |
| SMM.US.98.114 | SMM.US.98.114 | SMM.US.98.114 | SMM.US.98.114 | SMM.US.98.114 | SMM.US.98.114 | SMM.US.98.114 | SMM.US.98.114 | SMM.US.98.114 |
| SYK.KE.x.KE51 | SYK.KE.x.KE51 | SYK.KE.x.KE51 | SYK.KE.x.KE51 | SYK.KE.x.KE51 | SYK.KE.x.KE51 | SYK.KE.x.KE51 | SYK.KE.x.KE51 | SYK.KE.x.KE51 |
| SYK.KE.x.SYK173.COMGNM | SYK.KE.x.SYK173.COMGNM | SYK.KE.x.SYK173.COMGNM | SYK.KE.x.SYK173.COMGNM | SYK.KE.x.SYK173.COMGNM | SYK.KE.x.SYK173.COMGNM | SYK.KE.x.SYK173.COMGNM | SYK.KE.x.SYK173.COMGNM | SYK.KE.x.SYK173.COMGNM |
| TAL.CM.00.266 | TAL.CM.00.266 | TAL.CM.00.266 | TAL.CM.00.266 | TAL.CM.00.266 | TAL.CM.00.266 | TAL.CM.00.266 | TAL.CM.00.266 | TAL.CM.00.266 |
| TAL.CM.01.8023 | TAL.CM.01.8023 | TAL.CM.01.8023 | TAL.CM.01.8023 | TAL.CM.01.8023 | TAL.CM.01.8023 | TAL.CM.01.8023 | TAL.CM.01.8023 | TAL.CM.01.8023 |
| TAN.UG.x.SIVtanTAN1 | TAN.UG.x.SIVtanTAN1 | TAN.UG.x.SIVtanTAN1 | TAN.UG.x.SIVtanTAN1 | TAN.UG.x.SIVtanTAN1 | TAN.UG.x.SIVtanTAN1 | TAN.UG.x.SIVtanTAN1 | TAN.UG.x.SIVtanTAN1 | TAN.UG.x.SIVtanTAN1 |
| VER.KE.x.AGM215 | VER.KE.x.AGM215 | VER.KE.x.AGM215 | VER.KE.x.AGM215 | VER.KE.x.AGM215 | VER.KE.x.AGM215 | VER.KE.x.AGM215 | VER.KE.x.AGM215 | VER.KE.x.AGM215 |
| VER.ZA.10.VSAC4004 | VER.ZA.10.VSAC4004 | VER.ZA.10.VSAC4004 | VER.ZA.10.VSAC4004 | VER.ZA.10.VSAC4004 | VER.ZA.10.VSAC4004 | VER.ZA.10.VSAC4004 | VER.ZA.10.VSAC4004 | VER.ZA.10.VSAC4004 |
| VER.ZA.10.VSAC1003 | VER.ZA.10.VSAC1003 | VER.ZA.10.VSAC1003 | VER.ZA.10.VSAC1003 | VER.ZA.10.VSAC1003 | VER.ZA.10.VSAC1003 | VER.ZA.10.VSAC1003 | VER.ZA.10.VSAC1003 | VER.ZA.10.VSAC1003 |
| VER.ZA.10.VSAC0022 | VER.ZA.10.VSAC0022 | VER.ZA.10.VSAC0022 | VER.ZA.10.VSAC0022 | VER.ZA.10.VSAC0022 | VER.ZA.10.VSAC0022 | VER.ZA.10.VSAC0022 | VER.ZA.10.VSAC0022 | VER.ZA.10.VSAC0022 |
| WRC.CI.99.99C104 | WRC.CI.99.99C104 | WRC.CI.99.99C104 | WRC.CI.99.99C104 | WRC.CI.99.99C104 | WRC.CI.99.99C104 | WRC.CI.99.99C104 | WRC.CI.99.99C104 | WRC.CI.99.99C104 |
| WRC.CM.05.Pdt.05GM.X02 | WRC.CM.05.P | | | | | | | |

Table with columns for protein names (e.g., H1B, H1A1, H1C, etc.), sequence alignments, and residue numbers (281-296). Headers include 'protease end_p66, p51 RT start', 'M41L', 'K70R D67N', and 'D110 catalytic site'.

Table with columns for protein names (e.g., MAC, H2A, H2A, etc.), sequence alignments, and residue numbers (293-296). Headers include 'protease end_p51 RT start' and 'D catalytic site'.

Table with columns for protein names (e.g., H1B, H1A1, H1C), amino acid sequences, and accession numbers. Includes a sub-section for p51 RT end_p15 RNase H start.

Table with columns for protein names (e.g., MAC, H2A, H2B), amino acid sequences, and accession numbers. Includes a sub-section for p51 RT end_p15 RNase H start.

H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1C.ET.86.ETH2220
H1F.BE.93.VI850
H10.BE.87.AMT70
H10.CM.91.MVP5180
H10.CM.98.98CMA105
H10.CM.98.98CMA8197
H10.CM.98.98CMA8197
H10.SN.99.99SSE.MP1299
H10.US.99.99STWLA
H1N.CM.02.J008131
H1N.CM.04.04CM.1015.04
H1N.CM.04.04CM.1131.03
H1N.CM.95.YBF30
H1P.CM.06.U14788
CPZ.CD.07.P1167
CPZ.CD.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.LB715
CPZ.CM.05.SIVcpzDP943
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMB897
CPZ.GA.88.SIVcpzGAB2
YRSY.TA.09.TAN1
CPZ.TZ.09.UG38
CPZ.US.85.US.Marilyn
GOR.CM.04.SIVgorCP684con
GOR.CM.07.SIVgor2139.287
GOR.CM.07.SIVgorCP2135con
GOR.CM.12.SIVgor.BPID1
GOR.CM.12.SIVgor.BPID2
GOR.CM.13.SIVgor.BPID5

RQKVVTLTDDTNNQKTELQAIYALQDSGLEVINIVTDSQYALGIIQAQDPOSESELVNIQIEQLIKKEKVLAWYPAHKIGGNEQDKLVSAQIR.
KVLFLDGIKQAQDEHEKYSNWRAMASDFNLPPVVAKEIVASCDKCLQK.
GEAMHGQVDCSPGIWQLDCTHLE
...
I-K EG -A -E-VL-KE-KQA -V -A-T-D-P-QK -E-MT-S

MAC.US.x.239
H2A.DE.x.BEN
H2A.PT.x.ALI
H2A.SN.86.ST.JSP4.27
H2B.CI.x.EH0
H2B.GH.86.D295.127
H2G.CI.92.Abt96
H2U.CI.07.07IC.TMP3
H2U.FR.96.12034
ASC.UG.10.RT83
ASC.UG.10.RT88
ASC.UG.10.RT11
COL.CM.x.CGU1
COL.UG.10.BWC07
COL.UG.10.BWC07
COL.UG.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DRL.DE.11.D3
DRL.DE.11.D4
GRV.ET.x.GRI.677.gri.1
GSN.CM.99.CN166
GNS.CM.99.CN71
LST.CD.88.SIVLhoest485
LST.CD.88.SIVLhoest524
LST.KE.x.Lho7
MAL.ZM.99.SIVagmAL.ZMB
MND-1.GA.x.MNDGB1
MND-2.CM.98.CM16
MND-2.GA.x.M14
MND-2.x.x.5440
MNE.US.x.MNE027
MON.CM.99.L1.99CML1
MUS-1.CM.01.CM1239
MUS-2.CM.01.CM1246
MUS-3.GA.09.09gab0181
RCM.CM.00.SIVag1.00CM312
RCM.CM.02.SIVrcm.02CM8081
RCM.GA.x.SIVRCMGAB1
SAB.SN.x.SAB1
SMM.CT.79.SIVsmCT2
SMM.SL.92.SL92A
SMM.SL.92.SL92B
SMM.US.06.FtQ
SMM.US.11.SIVsmE660.FL8
SMM.US.86.CFU212
SMM.US.x.H9
SUN.GR.99.L14
SVK.KE.x.KE51
SYK.KE.x.SYK173.COMGNM
TAL.CM.00.266
TAL.CM.01.8023
TAN.UG.x.SIVtanTAN1
VER.KE.x.AGM155
VER.ZA.10.VSAC4004
VER.ZA.10.VSAG1003
VER.ZA.10.VSAM0922
WRC.CI.99.99C104
WRC.GM.05.Pbt.05GM.X02

KD-KV-EO-QA-E-FLM-T- PKA-IV-VM-TGC-TE-R-EM-SEI-V-...
...
N-IQKI-N-Q-I-LRM-E-SDK-I-V-V-KDH-T-P-MQE-DL-KG-EI-TI

p66 RT, p15 RNase H end, p31 Integrase start

p15 RNase H end, p31 Integrase start

AKLLWKGEVAVIQDNS.DIKVVP RRKAKIIRDY GKQMAGDDCVASRODED.....
 M-G
 A-G
 E-I
 -Q-----KG-----E-----T-SM-G-T-SESVEOPSEIP
 -Q-----KG-----E-----T-SM-N-T-SESMEOPSEIP
 -Q-----KG-----E-----T-SM-G-T-SETVEOPSEIP
 -Q-----KG-----E-----X-SM-G-T-GESMEOPSEIP
 -Q-----KG-----H-----T-SM-G-T-SENVEOPSEIP
 -Q-----KG-E-----T-M-G-T-SENMEQXSEIS
G-----G-A-NQDMA
G-----G-NQDME
G-----G-NQDME
G-----G-NOEME
 -T-----G-----DT-R-SESLEQSG
 -R-----KEGE-E-----V-AGG-DT-R-SESLEQSG
 -Q-----K-QE-E-----KE-R-KIEDR-DL-G-N
 -T-----Q-G-E-----V-N-M-GG-SQGM
 -R-----K-KE-EV-----GSM-G-SQNL
L-QE-E-----E-AN-L-D-N
L-QE-E-----NONME
L-KE-----VE-N-L-D-S
 -S-----K-G-----KH-V-L-GG-NQNM
 -R-----KEGE-----AGGMD-N
 -R-----KEGE-----KE-AGS-DD-N-A
 -R-----KEGE-----V-K-SAGSMDD-N-AQNM
KE-E-EV-----K-A-SM-G-S
 -T-----KG-----DT-G-SESLEQSS
 -T-----KG-----DT-G-SESLEQSS
 -T-----KG-----DT-G-SESLEQSS
 -T-----KG-----DT-G-SESMEQSG
 -Q-----KG-----X-SM-XG-T-GESMEQSG
 -T-----KG-----N-DT-G-SESLEQSG

MAC.US.x.239 GE-----ILKVGIT-----K-----GGKEV-SSSHMEDTGEAREVA
 H2A.DE.x.BEN GE-----IVKVGIT-----GROEL-SSPHLEGAREDEGEMAC
 H2A.PT.x.ALI GE-----D-IVKVGIT-----GROEL-SGPHLEGAREDEGEMAC
 H2B.SI.x.EHO GE-----D-IVKVGIT-----K-----GROEM-SGSHLEGAREDEGEMAC
 H2B.CI.x.EHO GE-----I-KVGT-E-I-----N-GGKEL-CSADVEDTMOAREVAQSN
 H2B.GH.86.D295.ALT GE-----I-KVGT-E-----H-GGKGL-CSADMEDTMOAREVAQSD
 H2G.CI.92.Abt96 GE-----I-KVGT-E-----K-GGKEL-SSTNMDTROTGEVA
 H2U.CI.07.07IC.TNP3 GE-----I-KVGT-----I-----K-GGKEL-SGSHLEGARODGEMAOALGEIP
 H2U.FR.96.12034 -D-----L-KVGT-----I-----GGKEL-SSNLEGA-KVREMAPDQTP
 ASC.UG.10.RT03 -Q-----KTTTGELLT-K-----KP-ASQNVGNTEGMDSPRKEDTSLD
 ASC.UG.10.RT08 -Q-----KTTTGELLT-K-----KP-ANQWVGE-TEAMDSPRKEDTSLD
 ASC.UG.10.RT11 -Q-----KTTTGELLT-K-----KP-ASQVSG-KENMDSPRKEDTSLD
 COL.CM.x.CGU1 GE-V-----RNSQGLF-K-V-T-LQYGEDV-SENLL-NGQKEAETVKGMD
 COL.UG.10.BWC01 GO-V-----KNOEGSLF-K-V-TKVNRYGERMVSKALLPNEQ-EAGEVKEVD
 COL.UG.10.BWC07 GO-V-----RSTEGTLF-K-V-S-ITHTGTD-VVSEDLPLNGKEKPGETETLD
 DEB.CM.04.SIVdeb04CMPF3061 -T-----KLDTE-L-----V-K-DVGSKTDQTHE
 DEB.CM.99.CM40 -T-----CKTEVG-----V-KE-----D-DSKVNQTSNE
 DEB.CM.99.CM5 -T-----CKTEVG-----V-K-----GVDSKTDQTPSE
 DRL.DE.11.D3 -E-V-----KEGT-L-I-----K-----NVDSHNME
 DRL.DE.11.D4 -E-V-----KEGT-L-I-----K-----N-DSHTNME
 GRV.ET.x.GRI.677.gri.1 -R-I-----KEGE-----K-ERKTM-SEGSMEGVREANKQHEGSDLDQDE
 GSN.CM.99.CN166 -V-----V-TOAGE-IT-----KP-EAKTE-VGG-THTNN
 GSN.CM.99.CN71 -V-----V-TOAGE-IT-----KP-EAK-E-VGS-AHTSN-REEGRMAD
 LST.CD.88.SIVLhoest485 -QI-----VKEGD-T-FS-----LVK-EGPKDSGSLDNN
 LST.CD.88.SIVLhoest524 -Q-----VKEGE-N-FSI-----LVK-EGPKDSGSSMDN
 LST.KE.x.lho7 -Q-----VKEGE-N-FS-----LVK-EGPKDSGSLDNN
 MAL.ZM.99.SIVaomMAL_ZMB -Q-I-----I-KEADG-L-I-----K-QQKCVG-EDHVESRO
 MND-1.GA.x.MNDG81 GI-----LKYQE-E-I-----C-K-----EGKNSQ-NLESV
 MND-2.CM.98.CM16 -E-I-----KEGT-L-----K-----TVDS-PN-EA
 MND-2.GA.x.M14 -E-I-----KEGT-L-----K-----TVDSNPHEM-ETA
 MND-2.x.x.5440 -E-I-----KEGV-L-----K-----TVDSNPHEMGR
 MNE.US.x.MNE027 GE-----ILKVGIT-----K-GGKEV-SSSHMEDTGEAREVA
 MON.CM.99.L1.99CML1 -R-----KTEEG-LT-----KP-TTENVGGDTNOYNLRKODGLAN
 MUS-1.CM.01.CM1239 -A-----VKTEQGVIT-----KP-AKENVGSKSNTGDHRKEDGLDN
 MUS-2.CM.01.CM1246 -R-----L-KTKEGE-VT-----K-TRQNV-SEPD-VHVRKEDGLAD
 MUS-3.GA.09.09gab0I81 -T-----KTDLGEV-T-----KP-NA-KNVGSKSNTGDIREEDGLVDQGN
 OLC.CI.97.07C112 -E-I-----LLKEGD-KYFS-L-----V-KP-VDSSANV
 RCM.CM.00.SIVa01.00CM312 -T-I-----GO-L-----C-K-RKVDVSETSMEGRO-KD
 RCM.CM.02.SIVrcm.02CM8081 -I-----EETG-L-----K-RKVDSEANLAGRO-EN
 SAB.SN.x.SAB1 -I-----EQG-EL-TI-----K-----L-SOAPLENGRGTAGEVD
 SMM.CI.79.SIVsmC12 -C-N-----I-KVGT-----I-----K-GGKELGNSPYLENP-EDGKMAQPD
 SMM.SL.92.SIVsmS192A -E-I-----K-VG-E-----I-----K-GGKEL-S-SDMNETTQVRETAQLSEISKV
 SMM.SL.92.SL92B -E-----K-VGT-E-----I-----K-GROEMGSSASMEDO
 SMM.US.06.FTq GE-----I-KVGT-----K-AGGKEL-SGTHVEDTREAEMA
 SMM.US.11.SIVsmE660.FL8 GE-----ILKVGIT-E-----K-GGKEL-SGSHLEDTGEAREVA
 SMM.US.86.CFU212 GE-----ILKVGIT-----K-GGKEL-SGSHLEDTREAEMA
 SMM.US.x.H9 GE-----ILKVGIT-E-----K-GGKEL-SGSHLEDTGEAREVA
 SUN.GS.98.L14 -S-----KQGE-N-L-----LVK-GESSSVEN
 SYK.KE.x.KE51 -S-----TPDQOIVA-----SSD-ERVDSGTHLEITSKSN
 SYK.KE.x.SYK173.COMGNM -R-----V-TEEG-FA-----T-H-ER-DSGSH-ENDPKT
 TAL.CM.00.266 -H-I-----LRTDGEVIT-----KP-QALGNKIDLE-SKEQ-AEMGRDN
 TAL.CM.01.8023 -H-V-----LRTDGEVIT-----KP-QA-GNKTDLGSKEQ-AEMGRDN
 TAN.UG.x.SIVaomTAN1 -T-----KEGE-EL-----K-ERKTYG-KTNMGE
 VER.KE.x.AGM155 -R-I-----LKEGE-EL-----K-EPRKTLG-ETHLEGAGGSDHOMAG.DS
 VER.ZA.10.VSAC4004 -T-I-----K-QG-EL-----K-ERK-MG-ENHLEGLGTNN
 VER.ZA.10.VSAG1003 -T-I-----K-QE-EL-----K-##KRMG-ENHVEGIRRPDN
 VER.ZA.10.VSAM0022 -V-I-----L-VEPRF-G-IT-K-----K-GKVGNEHDHMEGIGRADN
 WRC.CI.98.98C104 -G-V-----L-V-TPEG-LLT-----KWNMGEQDRSSSNKD
 WRC.GM.05.Pbt_05GM_X02 -G-V-----L-V-TPEG-LLT-----KTWDGKM#

1003
1003
1003
1000
1010
1010
1010
1010
1010
1010
1010
1010
1013
1014
1014
1015
1008
1007
1004
1003
1013
1015
1014
1007
1011
1012
1011
1000
999
999
1006
1003
1003
1003
1003
1008
1007
1006

1020
1033
1033
1033
1036
1035
1016
1022
1034
1018
1021
1021
991
993
995
1011
1008
1008
1013
1013
1041
1016
1017
1020
1021
1021
1020
1010
1010
1011
1017
1012
1016
1018
1018
1012
1012
1017
1016
995
1018
1024
1019
1040
1028
1023
1006
1020
1016
1016
1017
1041
1037
1020
1021
1020
1032
1024
1024
1026
1022
1007

Table of HIV sequences with columns for protein name, accession number, and amino acid sequence. Includes sequences for Vif start, Vif, and Vif end.

Table of HIV sequences with columns for protein name, accession number, and amino acid sequence. Includes sequences for Vif start, Vif, and Vif end.

| | | |
|----------------------------|----------------------------|-----|
| | Vif end | |
| H1B.FR.83.HXB2 |* | 192 |
| H1A1.UG.85.U455_U455A |* | 193 |
| H1C.ET.86.ETH2220 |* | 193 |
| H1F1.BE.93.VI850 |* | 193 |
| H10.BE.87.ANT70 |* | 193 |
| H10.CM.91.MVP5180 |* | 193 |
| H10.CM.98.98CMA105 |* | 192 |
| H10.CM.98.98CMABB197 |* | 193 |
| H10.SN.99.99SE_MP1299 |* | 193 |
| H10.US.99.99USTWLA |* | 196 |
| H1N.CM.02.DJ08131 |* | 193 |
| H1N.CM.04.04CM_1015_04 |* | 193 |
| H1N.CM.04.04CM_1131_03 |* | 193 |
| H1N.CM.95.YBF30 |* | 193 |
| H1P.CM.06.U14788 |* | 196 |
| CPZ.CD.06.BF1167 |* | 197 |
| CPZ.CD.90.ANT |* | 197 |
| CPZ.CM.01.SIVcpzCAM13 |* | 194 |
| CPZ.CM.05.LB715 |* | 193 |
| CPZ.CM.05.SIVcpzDP943 |* | 193 |
| CPZ.CM.05.SIVcpzEK505 |* | 193 |
| CPZ.CM.05.SIVcpzLB7 |* | 193 |
| CPZ.CM.05.SIVcpzMB66 |* | 193 |
| CPZ.CM.05.SIVcpzMB897 |* | 193 |
| CPZ.GA.88.SIVcpzGAB2 |* | 193 |
| CPZ.TZ.00.TAN1 |* | 199 |
| CPZ.TZ.06.SIVcpzTAN13 |* | 199 |
| CPZ.TZ.09.UG38 |* | 199 |
| CPZ.US.85.US_Marilyn |* | 193 |
| GOR.CM.04.SIVgorCP684con |* | 193 |
| GOR.CM.07.SIVgor2139_287 |* | 193 |
| GOR.CM.07.SIVgorCP2135con |* | 193 |
| GOR.CM.12.SIVgor_BPID1 |* | 193 |
| GOR.CM.12.SIVgor_BQID2 |* | 199 |
| GOR.CM.13.SIVgor_BPID15 |* | 199 |
| | Vif end | |
| MAC.US.x.239 |* | 215 |
| H2A.DE.x.BEN |* | 216 |
| H2A.PT.x.ALI |* | 216 |
| H2A.SN.86.ST_JSP4_27 |* | 216 |
| H2B.CI.x.EH0 |* | 217 |
| H2B.GH.86.D205_ALT |* | 217 |
| SMM.CI.79.SIVsmCI2 |* | 217 |
| SMM.SL.92.SIVsmSL92A |* | 214 |
| SMM.SL.92.SL92B |* | 218 |
| SMM.US.06.Ftq |* | 215 |
| SMM.US.11.SIVsmE660_FL8 |* | 215 |
| SMM.US.86.CFU212 |* | 215 |
| SMM.US.x.H9 |* | 215 |
| H2G.CI.92.Abt96 |* | 216 |
| H2U.CI.07.07IC_TNP3 |* | 215 |
| H2U.FR.96.L2034 |* | 214 |
| SUN.GA.98.L14 |* | 242 |
| ASC.UG.10.RT11 |* | 242 |
| ASC.UG.10.RT03 |* | 242 |
| ASC.UG.10.RT08 |* | 241 |
| COL.CM.x.CGUI |* | 172 |
| COL.UG.10.BWC01 |* | 173 |
| COL.UG.10.BWC07 |* | 172 |
| DEB.CM.04.SIVdeb04CMPF3061 | GSLSNRLHKEGPENDLATSGRRLVLS | 251 |
| DEB.CM.99.CM40 | GCIGH | 231 |
| DEB.CM.99.CM5 | GGVSN | 231 |
| DRL.DE.11.D3 |* | 218 |
| DRL.DE.11.D4 |* | 218 |
| GRV.ET.x.GRI_677_gri_1 |* | 220 |
| GSN.CM.99.CN166 |* | 238 |
| GSN.CM.99.CN71 |* | 238 |
| LST.CD.88.SIVlhoest485 | YHSK | 228 |
| LST.CD.88.SIVlhoest524 | HFS | 227 |
| LST.KE.x.lho7 | HPSQ | 227 |
| MND-1.GA.x.MNDGB1 |* | 223 |
| MND-2.CM.98.CM16 |* | 213 |
| MND-2.GA.x.H14 |* | 213 |
| MND-2.x.x.5440 |* | 214 |
| MNE.US.x.MNE027 |* | 215 |
| MON.CM.99.L1_99CML1 |* | 256 |
| MUS-1.CM.01.CM1239 |* | 236 |
| MUS-2.CM.01.CM1246 |* | 239 |
| MUS-3.GA.09.09gab0I81 |* | 241 |
| OLC.CI.97.97CI12 |* | 204 |
| RCM.CM.00.SIVagi_00CM312 |* | 221 |
| RCM.CM.02.SIVrcm_02CM8081 |* | 221 |
| RCM.GA.x.SIVRCMGAB1 |* | 217 |
| SAB.SN.x.SAB1 |* | 235 |
| SYK.KE.x.KE51 |* | 234 |
| SYK.KE.x.SYK173_COMGNM |* | 222 |
| TAL.CM.00.266 |* | 228 |
| TAL.CM.01.8023 |* | 228 |
| TAN.UG.x.SIVagmTAN1 |* | 239 |
| VER.KE.x.AGM155 |* | 233 |
| VER.ZA.10.VSAC4004 |* | 228 |
| VER.ZA.10.VSAG1003 |* | 224 |
| MAL.ZM.09.SIVagmMAL_ZMB |* | 223 |
| VER.ZA.10.VSAM0022 |* | 223 |
| WRC.CI.98.98CI04 |* | 208 |
| WRC.GM.05.Pbt_05GM_X02 |* | 209 |

| | Vpx start | Vpx end |
|---------------------------|--|---------|
| MAC.US.x.239 | MSD..PRERIP..PGNSGEETIGEAFEWLNRTVEEINREAVNHLPRELIFQVWORSWEYWHDEQGMSPSYVKYRYLCLIQKALFMHCKKGCRLGEGHGAGGW.RPG..PPPPPPGLA* | 112 |
| H2A.DE.x.BEN | -T...-V...-E-I-AL...-R...-A-T...-M...-I-T-F-R-T-W-DM-RE-LEDQ...-V- | 114 |
| H2A.PT.x.ALI | -AN...-TV...-D...-AL...-R...-Q-T...-M...-MYT-FM-T-G...-P...-S...-V- | 113 |
| H2A.SN.86.ST.JSP4_27 | -AG...-T...-D...-A...-R...-I-T...-M...-M-I-S-R-T-G...-P...-S...-V- | 113 |
| H2B.CI.x.EH0 | ...-V...-V...-E-T-L-HL-V...-K-A-RE...-I-T...-M...-M-I-FA-G-R...-P...-S...-V- | 113 |
| H2B.GH.86.D295.ALT | ...-V...-D...-V...-A-E-IT-L-V...-A-RE...-I-T...-L-M-M-V-YT...-Q...-P...-S...-V- | 112 |
| H2G.CI.92.Abt96 | ...-V...-E-D...-V...-E-T-I...-R...-A-RE...-A-T...-L-M-M-V-YT...-Q...-P...-S...-V- | 112 |
| H2U.CI.07.071C.TNP3 | ...-D...-VE...-ES-TA-Q...-G...-T...-T...-Y...-R-Q-R...-P...-S...-V- | 112 |
| H2U.FR.96.12634 | -G...-E-V...-S-E-RD-A-Q...-Q-R...-E-T...-Q...-P...-S...-V- | 113 |
| MNE.US.x.MNE027 | ...-T...-T...-YS-AV-G-K...-R...-F...-QA-TQ...-L...-C-V...-T-R...-S...-A...-V- | 113 |
| SMM.CI.79.SIVsmCI2 | -GS.A...-D...-D...-E-V-T...-R...-AA-T...-VQ-M-M-Q-FR-T-R...-SQ...-T...-F- | 111 |
| SMM.SL.92.SIVsmSL92A | -T...-HN-AL-OT-Q...-R-C-V-Y-A-VQ-M-M-Q-FR-T-R...-SQ...-T...-F- | 113 |
| SMM.SL.92.SL92B | ...-E...-D...-A...-R...-R-A-T...-V...-R...-T...-T...-F- | 113 |
| SMM.US.06.FI0 | ...-E...-G...-A...-R...-M-E-T...-V...-R...-T...-T...-F- | 113 |
| SMM.US.11.SIVsmE660_FL8 | ...-L-V...-R-C-F...-M-M-R-N-R-E...-Q...-T...-F- | 113 |
| SMM.US.86.CFU212 | -X...-X...-D-H-X...-A...-R-X...-M-V-T-X...-GE...-T...-F- | 113 |
| SMM.US.x.H9 | -AEGQ-A-A-AE-VGA...-V-LE--Q-SLLR-Q-RL-FHP-FL-RL-NTCM-HY-ALQL-FI-S...-L-L-M--FQ0-S-VQGR-PPPLRPAGDRL...-V- | 116 |
| DRL.DE.11.D3 | -AEGQ-A-A-AE-VGA...-V-LE--Q-SLLR-Q-RL-FHP-FL-RL-NTCM-HY-ALQL-FI-S...-L-L-M--FQ0-S-VQGR-PPPLRPAGDRL...-V- | 116 |
| DRL.DE.11.D4 | -AEGQSV-A-AE-VGA...-V-LE--Q-SLLR-Q-QL-FHP-FL-RL-NTCM-HY-ALQL-FI-S...-L-LH-M-T-FQ0-S-VRGR-PPPLRPAGDRL...-V- | 117 |
| MND-2.CM.98.CM16 | -AE..GAPE...E...GA...-V-DLNT-E-SL-K-Q-RL-FHP-FL-RL-NACT-H--RHQR-L-A...-L-MN-M-T-MQ0-P-RS...-P...-MV- | 100 |
| MND-2.GA.x.M14 | -AE...-A-EA-QGA...-V-GLEQ-E-SL-QV...-QL-FHP-FL-RL-NTCV-H-RLRRTL-N-A...-L-M--M-V-MQ0-P-RS...-S...-MV- | 100 |
| MND-2.x.x.5440 | -AE...-A-EA-EGA...-V-GLEQ-ETSL-R...-RL-FHP-FL-RL-NTCV-H--RHQR-LD-A...-L-MH-MYT-MQ0-P-RN...-RP...-M- | 100 |
| RCM.CM.00.SIVagi_00CM312 | -AE...-A-EG-TGA...-VEFOPWLARMLY...-Q-RL-F-G-F-HL-RTCV...-RTLE-AG...-L-M--T-MRS-KLR...-DPPRO-ERV...-IL-MQ- | 106 |
| RCM.CM.02.SIVrcm_02CM8081 | -AE...-A-EV-TGA...-VEFOPWLARMLY...-Q-RL-FHP-F-RL-RTCV-H--L-R-LE-AG...-L-M--T-MRS-LR...-DPPRO-ERV...-IL-MQ- | 106 |
| RCM.GA.x.SIVRCMGAB1 | -AE...-A-EV-TGA...-AEFQPWL.RDML-KV-L-RL-FHP-F-RL-RTCV-H--VHQR-LE-AA...-L-M--I-QT-SQR...-PNPR-AV-ERI...-TIL-M- | 105 |

Table with columns: Vpr start, oligomerization, amphipathic alpha-helix, H(S/N)RIG motifs, Vpr end in HXB2, Vpr end. Rows list various HIV-1 sequences and their corresponding protein motifs.


```

Tat end
H1B.FR.83.HXB2 .....
H1A1.UG.85.U455_U455A .....
H1C.ET.86.ETH2220 .....
H1F1.BE.93.VI859 .....
H1O.BE.87.ANT70 ..... RTSEQ
H1O.CM.91.MVP5180 ..... RISGQ
H10.CM.98.98CMA105 ..... RTSGQ
H10.CM.98.98CMABB197 .....
H10.SN.99.99SE_MP1299 .....
H10.US.99.99USTWLA .....
H1N.CM.02.DJ08131 .....
H1N.CM.04.04CM_1015_04 .....
H1N.CM.04.04CM_1131_03 .....
H1N.CM.95.YBF30 .....
H1P.CM.06.U14788 ..... TT
CPZ.CD.06.BF1167 .....
CPZ.CD.90.ANT ..... WT
CPZ.CM.01.SIVcpzCAM13 .....
CPZ.CM.05.LB715 .....
CPZ.CM.05.SIVcpzDP943 .....
CPZ.CM.05.SIVcpzEK505 .....
CPZ.CM.05.SIVcpzLB7 .....
CPZ.CM.05.SIVcpzMB66 .....
CPZ.CM.05.SIVcpzMB897 .....
CPZ.GA.88.SIVcpzGAB2 .....
CPZ.TZ.00.TAN1 ..... RTSGTSSSGYTRPFKTSSSGSSGSACKH
CPZ.TZ.06.SIVcpzTAN13 ..... RTSGTSSSGSTRPVSTSSGCGWPKSP
CPZ.TZ.09.UG38 ..... RTSGPSSCGSTRPVKTSAGSSGR.S
CPZ.US.85.US_Marilyn .....
GOR.CM.04.SIVgorCP684con ..... TISEPWLCGSTGH
GOR.CM.07.SIVgor2139_287 .....
GOR.CM.07.SIVgorCP2135con .....
GOR.CM.12.SIVgor_BP1D1 .....
GOR.CM.12.SIVgor_BQ1D2 ..... TSGQ
GOR.CM.13.SIVgor_BP1D15 .....

Tat end
MAC.US.x.239 .....
H2A.DE.x.BEN .....
H2A.PT.x.ALI .....
H2A.SN.86.ST_JSP4_27 .....
H2B.CI.x.EH0 ..... RFAN
H2B.GH.86.D205_ALT .....
H2G.CI.92.Abt96 .....
H2U.CI.07.07IC_TNP3 .....
H2U.FR.96.12034 .....
ASC.CD.13.SIVasc_13DRC_I8 .....
ASC.UG.10.RT03 .....
ASC.UG.10.RT08 .....
ASC.UG.10.RT11 .....
COL.CM.x.CGU1 .....
COL.UG.10.BWC01 .....
COL.UG.10.BWC07 .....
DEB.CM.04.SIVdeb04CMPF3061 .....
DEB.CM.99.CM40 ..... RISNWC
DEB.CM.99.CM5 .....
DRL.DE.11.D3 ..... AKK.Q
DRL.DE.11.D4 ..... AKK.Q
GRV.ET.x.GRI_677_gri_1 .....
GSN.CM.99.CN166 .....
GSN.CM.99.CN71 ..... TT
LST.CD.88.SIVlhoest485 .....
LST.CD.88.SIVlhoest524 .....
LST.KE.x.lho7 .....
MAL.XM.2009.SIVagmMAL_ZMB ..... RGSIGARD
MND-1.GA.x.MNDGB1 .....
MND-2.CM.98.CM16 .....
MND-2.GA.x.M14 ..... GGQSK
MND-2.x.x.5440 ..... GSQSKGAG
MNE.US.x.MNE027 .....
MON.CM.99.L1_99CML1 .....
MUS-1.CM.01.CM1239 .....
MUS-2.CM.01.CM1246 ..... KTSTT
MUS-3.GA.09.09GaboI81 .....
OLC.CI.97.97C112 ..... H
RCM.CM.00.SIVagi_00CM312 .....
RCM.CM.02.SIVrcm_02CM8081 .....
RCM.GA.x.SIVRCMGAB1 .....
SAB.SN.x.SAB1 .....
SMM.CI.79.SIVsmCI2 .....
SMM.SL.92.SIVsmSL92A .....
SMM.SL.92.SL92B .....
SMM.US.06.FTq .....
SMM.US.11.SIVsmE660_FL8 .....
SMM.US.06.CFU212 .....
SMM.US.x.H9 .....
SUN.GA.98.L14 .....
SYK.KE.x.KE51 .....
SYK.KE.x.SYK173_COMGNM .....
TAL.CM.00.266 .....
TAL.CM.01.8023 ..... RTSTASSATVTRP
TAN.UG.x.SIVagmTAN1 .....
VER.KE.x.AGM155 ..... RGPVVGAGN
VER.ZA.10.VSAC4004 ..... RADTGSAD
VER.ZA.10.VSAG1003 ..... GGSAGTSD
VER.ZA.10.VSAM0022 ..... GRTVGP5D
WRC.CI.98.98CI04 .....
WRC.GM.05.Pbt_05GM_X02 .....

```

100
102
101
103
115
116
114
97
98
98
102
102
102
102
112
101
115
100
108
103
101
100
102
101
100
137
137
134
134
103
124
83
83
83
116
72

131
131
137
134
139
129
134
129
129
103
106
106
106
86
87
90
99
111
97
125
125
95
105
135
108
109
109
112
108
113
123
127
135
105
102
109
117
108
116
114
107
107
134
133
106
131
133
130
133
133
133
102

| | | |
|---------------------------|-----------|-----|
| | Rev end | |
| H1B.FR.83.HXB2 |* | 116 |
| H1A1.UG.85.U455_U455A |* | 124 |
| H1C.ET.86.ETH2220 |* | 108 |
| H1F1.BE.93.VI859 |* | 117 |
| H10.BE.87.ANT70 |* | 103 |
| H10.CM.91.MVP5180 |* | 104 |
| H10.CM.98.98CMA105 |* | 102 |
| H10.CM.98.98CMABB197 |* | 103 |
| H10.SN.99.99SE_MPI299 |* | 104 |
| H10.US.99.99USTWLA |* | 104 |
| H1N.CM.02.DJ08131 |* | 107 |
| H1N.CM.04.04CM_1015_04 |* | 105 |
| H1N.CM.04.04CM_1131_03 |* | 105 |
| H1N.CM.95.YBF30 |* | 105 |
| H1P.CM.06.06BWC97 |* | 103 |
| CPZ.CD.06.8F1167 | NSN.....* | 126 |
| CPZ.CD.90.ANT |* | 119 |
| CPZ.CM.01.SIVcpzCAM13 |* | 124 |
| CPZ.CM.05.LB715 |* | 111 |
| CPZ.CM.05.SIVcpzDP943 |* | 125 |
| CPZ.CM.05.SIVcpzEK505 |* | 104 |
| CPZ.CM.05.SIVcpzLB7 |* | 124 |
| CPZ.CM.05.SIVcpzMB66 |* | 124 |
| CPZ.CM.05.SIVcpzMB897 |* | 123 |
| CPZ.GA.88.SIVcpzGAB2 |* | 124 |
| CPZ.TZ.09.TAN1 |* | 106 |
| CPZ.TZ.06.SIVcpzTAN13 |* | 96 |
| CPZ.TZ.09.UG38 |* | 111 |
| CPZ.US.85.US_Marilyn |* | 106 |
| GOR.CM.04.SIVgorCP684con |* | 105 |
| GOR.CM.07.SIVgor2139_287 |* | 105 |
| GOR.CM.07.SIVgorCP2135con |* | 105 |
| GOR.CM.12.SIVgor_BPID1 |* | 105 |
| GOR.CM.12.SIVgor_BQID2 |* | 109 |
| GOR.CM.13.SIVgor_BPID15 |* | 105 |

| | | |
|----------------------------|--------------------------|-----|
| | Rev end | |
| MAC.US.x.239 |* | 108 |
| H2A.DE.x.BEN |* | 104 |
| H2A.PT.x.ALI |* | 108 |
| H2A.SN.86.ST_JSP4_27 |* | 108 |
| H2B.CI.x.EH0 | QTRG.....* | 151 |
| H2B.GH.86.0205_ALT | QTRT.....* | 154 |
| H2G.CI.92.Abt96 |* | 101 |
| H2U.CI.07.07IC_TNP3 |* | 103 |
| H2U.FR.96.12034 |* | 99 |
| ASC.CD.13.SIVasc_13DRC_I8 |* | 111 |
| ASC.UG.10.RT03 |* | 102 |
| ASC.UG.10.RT08 |* | 102 |
| ASC.UG.10.RT11 |* | 102 |
| COL.CM.x.CGU1 |* | 94 |
| COL.UG.10.BWC01 |* | 98 |
| COL.UG.10.BWC97 | EVATICDEC.....RA | 117 |
| DEB.CM.04.SIVdeb04CMPF3061 |* | 98 |
| DEB.CM.99.CM40 |* | 91 |
| DEB.CM.99.CM5 |* | 88 |
| DRL.DE.11.D3 |* | 87 |
| DRL.DE.11.D4 |* | 87 |
| GRV.ET.x.GRI_677_gri_1 | SKKSAARP.....* | 148 |
| GSN.CM.99.CN166 |* | 125 |
| GSN.CM.99.CN71 |* | 121 |
| LST.CD.88.SIVlhoest485 | TKEKMEIQIQREKWSQ.....* | 174 |
| LST.CD.88.SIVlhoest524 |* | 149 |
| LST.KE.x.Lho7 |* | 149 |
| MND-1.GA.x.MNDGB1 |* | 88 |
| MND-2.CM.98.CM16 |* | 91 |
| MND-2.GA.x.M14 |* | 89 |
| MND-2.x.x.5440 |* | 88 |
| MNE.US.x.MNE027 |* | 108 |
| MON.CM.99.L1_99CML1 |* | 113 |
| MUS-1.CM.01.CM1239 |* | 107 |
| MUS-2.CM.01.CM1246 |* | 100 |
| MUS-3.GA.09.09GabO181 |* | 114 |
| OLC.CI.97.07CI2 |* | 124 |
| RCM.CM.00.SIVag1_00CM312 |* | 108 |
| RCM.CM.02.SIVrcm_02CM8081 |* | 105 |
| RCM.GA.x.SIVRCMGAB1 |* | 94 |
| SAB.SN.x.SAB1 |* | 100 |
| SMM.CI.79.SIVsmCI2 |* | 104 |
| SMM.SL.92.SIVsmSL92A |* | 104 |
| SMM.SL.92.SL92B |* | 100 |
| SMM.US.06.FTq |* | 107 |
| SMM.US.11.SIVsmE660_FL8 |* | 101 |
| SMM.US.86.CFU212 |* | 106 |
| SMM.US.x.H9 |* | 101 |
| SUN.GA.98.L14 |* | 151 |
| SYK.KE.x.KE51 |* | 85 |
| SYK.KE.x.SYK173_COMGNM |* | 78 |
| TAL.CM.00.266 |* | 126 |
| TAL.CM.01.8023 |* | 112 |
| TAN.UG.x.SIVaemTAN1 |* | 104 |
| VER.KE.x.AGM155 |* | 90 |
| VER.ZA.10.VSAC4004 |* | 90 |
| MAL.ZM.2009.SIVagmMAL_ZMB |* | 87 |
| VER.ZA.10.VSAG1003 |* | 90 |
| VER.ZA.10.VSAM0022 |* | 90 |
| WRC.CI.98.98CI04 |* | 124 |
| WRC.GM.05.Pbt_05GM_X02 | KQQHQGNYC...QTLHKP.....* | 146 |

| | Vpu end | |
|-----------------------------|-----------------|-----|
| H1B.FR.83.HXB2 |* | 82 |
| H1A1.UG.85.U455_U455A |* | 30 |
| H1B.US.90.WEAU160_GHOSH |* | 82 |
| H1C.ET.86.ETH2220 |* | 82 |
| H1D.CD.84.84ZRO85 |* | 82 |
| H1F1.BE.93.VI7850 |* | 82 |
| H1G.SE.93.SE6165_G6165 |* | 82 |
| H1H.CF.90.056 |* | 81 |
| H1J.SE.93.SE9280_7887 |* | 82 |
| H1K.CM.96.96CM_MP35 |* | 81 |
| H101.AE.TH.90.CM240 |* | 81 |
| H102.AG.NG.x.IBNG |* | 82 |
| H103.AB.RU.97.KAL153_2 |* | 79 |
| H104.cpx.CY.94.94CY032_3 |* | 82 |
| H10.BE.87.ANT76 |* | 86 |
| H10.CM.91.MVP5180 |* | 86 |
| H10.CM.94.BCF06 |* | 86 |
| H10.CM.96.96CMA102 |* | 86 |
| H10.CM.96.96CMABB009 |* | 86 |
| H10.CM.96.96CMABB637 |* | 115 |
| H10.CM.97.97CMABB447 |* | 86 |
| H10.CM.97.97CMABB497 |* | 84 |
| H10.CM.98.98CMA104 |* | 86 |
| H10.CM.98.98CMA105 |* | 86 |
| H10.CM.98.98CMABB141 |* | 87 |
| H10.CM.98.98CMABB197 |* | 86 |
| H10.CM.98.98CMABB212 |* | 86 |
| H10.CM.98.98CMU2901 |* | 84 |
| H10.CM.98.98CMU5337 |* | 86 |
| H10.CM.99.99CMU4122 |* | 112 |
| H10.CM.x.pCM02_3 |* | 112 |
| H10.CM.x.pCM02_5 |* | 112 |
| H10.FR.92.VAU |* | 86 |
| H10.SN.99.99SE_MP1299 |* | 86 |
| H10.SN.99.99SE_MP1300 |* | 86 |
| H10.US.99.99USTWLA |* | 112 |
| H10.US.x.I.2478B |* | 87 |
| H1N.CM.02.DJ00131 |* | 75 |
| H1N.CM.02.SJGddd |* | 103 |
| H1N.CM.04.04CM_1015_04 |* | 101 |
| H1N.CM.04.04CM_1131_03 |* | 113 |
| H1N.CM.06.U14296 |* | 68 |
| H1N.CM.06.U14842 |* | 101 |
| H1N.CM.95.YBF30 |* | 110 |
| H1N.CM.97.YBF106 |* | 101 |
| H1P.CM.06.U14788 |* | 79 |
| H1P.FR.06.RBF168 |* | 79 |
| CPZ.CD.06.BF1167 |* | 82 |
| CPZ.CD.90.ANT |* | 135 |
| CPZ.CM.01.SIVcpzCAM13 |* | 108 |
| CPZ.CM.05.LB715 |* | 81 |
| CPZ.CM.05.LB715 |* | 81 |
| CPZ.CM.05.SIVcpzDP943 |* | 82 |
| CPZ.CM.05.SIVcpzEK505_c2 |* | 105 |
| CPZ.CM.05.SIVcpzEK505 |* | 105 |
| CPZ.CM.05.SIVcpzLB7 |* | 80 |
| CPZ.CM.05.SIVcpzMB66 |* | 87 |
| CPZ.CM.05.SIVcpzMB897_c2 |* | 83 |
| CPZ.CM.05.SIVcpzMB897 |* | 83 |
| CPZ.CM.05.SIVcpzMT145_c2 |* | 114 |
| CPZ.CM.05.SIVcpzMT145 |* | 114 |
| CPZ.CM.98.CAM3 |* | 86 |
| CPZ.CM.98.CAM5 |* | 115 |
| CPZ.GA.88.GAB1 |* | 112 |
| CPZ.GA.88.SIVcpzGAB2 |* | 123 |
| CPZ.TZ.00.TAN1 |* | 84 |
| CPZ.TZ.01.TAN2 |* | 84 |
| CPZ.TZ.02.TAN3_1 |* | 84 |
| CPZ.TZ.06.SIVcpzTAN13 |* | 84 |
| CPZ.TZ.06.TAN5 |* | 84 |
| CPZ.TZ.09.UG38 |* | 83 |
| CPZ.US.85.US_Marilyn |* | 85 |
| GOR.CM.04.SIVgorCP684con |* | 81 |
| GOR.CM.07.SIVgor2139_287 |* | 107 |
| GOR.CM.07.SIVgorCP2135con |* | 107 |
| GOR.CM.07.SIVgorCP2139_1con |* | 107 |
| GOR.CM.07.SIVgorCP2139_2con |* | 107 |
| GOR.CM.12.SIVgor_BPID1 |* | 79 |
| GOR.CM.12.SIVgor_BPID2 |* | 86 |
| GOR.CM.13.SIVgor_BPID15 |* | 80 |
| DEN.CD.x.CD1_CM0580407 |* | 65 |
| GSN.CM.99.CN166 |* | 107 |
| GSN.CM.99.CN71 |* | 77 |
| MON.CM.99.L1_99CML1 |* | 88 |
| MON.NG.x.NGI |* | 98 |
| MUS-1.CM.01.CM1239 |* | 76 |
| MUS-1.CM.01.SIVmus_01CM1085 |* | 77 |
| MUS-2.CM.01.CM1246 |* | 78 |
| MUS-2.CM.01.CM2500 |* | 78 |
| MUS-3.GA.09.09Gab0181 | SSA.....* | 140 |
| MUS-3.GA.11.11GabPts02 | SPFKYHRGI.....* | 148 |

Table with columns for glycosylation NCS, glycosylation NIS, glycosylation NDT, glycosylation NTS, glycosylation NKT, glycosylation NVS, glycosylation NGT, and glycosylation NGT. Rows list various HIV proteins and their amino acid sequences.

| | glycosylation NVS | | glycosylation NGS | | glycosylation NNT | | glycosylation NCT | | V3 tip | | glycosylation NKT | | glycosylation NKT | | CD4 binding | | | | |
|----------------------------|-------------------|-------------|-------------------|--------|-------------------|----------|-------------------|--------|---------|-------|-------------------|-------|-------------------|-----------|-------------|----------|---------|-------|-----|
| | | | | | | | | | | | | | | | | | | | |
| H1B.FR.83.HXB2 | VSTVQ | CHGIRPVVSTQ | LLQNSGLAEE | EEVIVR | SVNFT | DNAKTIIV | QVLTNSVEIN | TRPNNT | TRKRIRI | ORGP | GRAFV | TIGK | IGNMROAH | NIISRAKWN | NTLQIA | ASKLREOF | GNMKTII | FKO55 | 365 |
| H1A1.UG.85.U455 | U455A | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 363 |
| H1C.ET.86.ETH2220 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 360 |
| H1F.BE.93.V1850 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 359 |
| H10.BE.87.ANT70 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 359 |
| H10.CM.91.MVP5180 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 366 |
| H10.CM.98.98CMA105 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 366 |
| H10.CM.98.98CMA197 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 371 |
| H10.CM.98.98CMA8185 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 371 |
| H10.SN.99.99SE.MP1299 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 370 |
| H10.US.99.99I517WLA | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 365 |
| H1N.CM.02.DJ08131 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 370 |
| H1N.CM.04.04CM.1015 | 04 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 351 |
| H1N.CM.04.04CM.1131 | 03 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 366 |
| H1N.CM.95.YBF30 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 367 |
| H1P.CM.96.U14788 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 374 |
| CPZ.CM.06.BZ1167 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 378 |
| CPZ.CM.90.ANT | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 378 |
| CPZ.CM.01.SIVcpzCAM13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 368 |
| CPZ.CM.05.LB715 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 349 |
| CPZ.CM.05.SIVcpzDP943 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 390 |
| CPZ.CM.05.SIVcpzEK505 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 349 |
| CPZ.CM.05.SIVcpzLB7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 354 |
| CPZ.CM.05.SIVcpzMB66 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 387 |
| CPZ.CM.05.SIVcpzMB897 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 358 |
| CPZ.GA.88.SIVcpzGAB2 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 366 |
| CPZ.TZ.09.UT0127 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 379 |
| CPZ.TZ.06.SIVcpzTAN13 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 369 |
| CPZ.TZ.09.UG38 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 354 |
| CPZ.US.85.Marilyn | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 354 |
| GOR.CM.04.SIVgorCP684con | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 345 |
| GOR.CM.07.SIVgor2139 | 287 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 356 |
| GOR.CM.07.SIVgorCP2135con | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 349 |
| GOR.CM.12.SIVgor | BP1D1 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 383 |
| GOR.CM.12.SIVgor | BP1D2 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 380 |
| GOR.CM.13.SIVgor | BP1D15 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 380 |
| ASC.UG.10.RT63 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 370 |
| ASC.UG.10.RT68 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 380 |
| ASC.UG.10.CW1 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 373 |
| COL.CM.x.CW1 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 385 |
| COL.UG.10.BW01 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 372 |
| COL.UG.10.BW07 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 378 |
| DEB.UG.04.SIVdeb04CMPF3061 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 358 |
| DEB.UG.09.CM165 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 381 |
| DEB.UG.09.CM3 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 359 |
| DRL.DE.11.D5 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 420 |
| DRL.DE.11.D4 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 400 |
| GRV.ET.x.GRI.677.gri.1 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 374 |
| GSM.CM.09.CM165 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 381 |
| GSM.CM.09.CN71 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 386 |
| LST.CD.88.SIVlhoest485 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 405 |
| LST.CD.88.SIVlhoest524 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 403 |
| LST.KE.x.lho7 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 408 |
| L2.GA.x.MNDG81 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 401 |
| MND.2.CM.x.M16 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 395 |
| MND.2.CM.x.M14 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 395 |
| MND.2.x.x.5440 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 395 |
| MNE.US.x.MNE027 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 416 |
| MON.CM.91.CM193 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 367 |
| MON.CM.91.CM239 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 377 |
| MUS.2.CM.01.CM1246 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 385 |
| MUS.3.GA.09.09GaB0312 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 386 |
| OLC.CI.97.97C112 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 395 |
| RCM.CM.00.SIVag1.00CM312 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 378 |
| RCM.CM.02.SIVrcm.02CM8081 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 377 |
| RCM.GA.x.SIVRCMGAB1 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 369 |
| SAB.SN.x.SAB1 | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | 379 |

Table with columns for glycosylation NST, glycosylation NNT, CD4 binding, glycosylation NIT, glycosylation NES, and CD4 binding. Rows list protein sequences and their corresponding amino acid positions (e.g., H1B.FR.83.HXB2, H1A1.UG.85.U455.U455A).

Table with columns for protein names (e.g., H1B, H1A1, H1C), accession numbers, and amino acid sequences. The sequences are aligned across three domains: transmembrane domain, gp41 cytoplasmic tail start, and glycosylation NSG. The table lists various HIV-1 strains and their corresponding protein sequences.

H1B.FR.83.HXB2
H1A1.UG.85.U455.U455A
H1C.ET.86.ETH2220
H1F1.BE.93.V1850
H10.BE.87.ANT70
H10.CM.01.MVP5180
H10.CM.98.98CMA105
H10.CM.98.98CMA8197
H10.SN.99.99S5.MP1299
H10.US.99.99U5.TWLA
H10.CM.02.DJ08131
H1N.CM.04.04CM.1015.04
H1N.CM.04.04CM.1131.03
H1N.CM.95.YBF30
H1P.CM.06.U14788
CPZ.CO.06.BF1167
CPZ.CO.90.ANT
CPZ.CM.01.SIVcpzCAM13
CPZ.CM.05.LB715
CPZ.CM.05.SIVcpzDP943
CPZ.CM.05.SIVcpzEK505
CPZ.CM.05.SIVcpzLB7
CPZ.CM.05.SIVcpzMB66
CPZ.CM.05.SIVcpzMB897
CPZ.GA.88.SIVcpzGAB2
CPZ.TZ.09.TAN1
CPZ.TZ.06.SIVcpzTAN13
CPZ.TZ.09.UG38
CPZ.US.85.US.Marilyn
GOR.CM.04.SIVgorCP684con
GOR.CM.07.SIVgor2139.287
GOR.CM.07.SIVgorCP2135con
GOR.CM.12.SIVgor.BPID1
GOR.CM.12.SIVgor.BPID2
GOR.CM.13.SIVgor.BPID15
ASC.UG.10.RT63
F.UG.10.RT68
ASC.UG.10.RT11
COL.CM.x.CGU1
COL.UG.10.BWC01
COL.UG.10.BWC02
DEB.CM.04.SIVdeb04CMPF3061
DEB.CM.99.CM40
DEB.CM.99.CM5
DRL.DE.11.D3
DRL.DE.11.D4
GRV.ET.x.GRI.677.gri.1
GSM.CM.99.CM166
GSM.CM.99.CM71
LST.CO.88.SIVlhoest485
LST.CO.88.SIVlhoest524
LST.KE.x.lho7
MND.1.GA.x.MNDBG1
MND.2.CM.98.CM16
MND.2.GA.x.M14
MND.2.x.x.5440
MNE.US.x.MNE027
MON.CM.99.L1.99CML1
MUS.1.CM.01.CM239
MUS.2.CM.01.CM1246
MUS.3.GA.09.09GAB0181
OLC.CI.97.97C112
RCM.CM.00.SIVag1.00CM312
RCM.CM.02.SIVrcm.02CM0801
RCM.GA.x.SIVRCMGAB1
SAB.SN.x.SAB1
MAC.US.x.239
H2A.DE.x.BEN
H2A.PT.x.ALI
H2A.SN.86.ST.JSP4.27
H2B.CI.x.EH0
H2B.GH.86.D205.ALT
H2C.CI.92.AB996
H2U.CI.07.07IC.TNP3
H2U.FR.96.12034
SMM.CI.79.SIVsmC12
SMM.SL.92.SIVsmSL92A
SMM.SL.92.SIVsmSL92B
SMM.US.06.F10
SMM.US.11.SIVsmE660.FL8
SMM.US.86.CFU212
SMM.US.x.H9
TAN.UG.x.SIVagmTAN1
VER.KE.x.AGM155
VER.ZA.10.VSAC4004
MAL.ZM.2009.SIVagmMAL.ZMB
VER.ZA.10.VSAG1003
VER.ZA.10.VSAGM022
WRC.CI.98.98C104
WRC.GM.05.Pbt.05GM.X02

glycosylation NAT
.....WNLLQYSQELKNSAVSLLNATAIAVAEGTRDRIEVV...QGACRAIRHIPRRIROGLER..ILL.*
.....-L-GR--I-IT-D-V-V--GWI-----IG--TIG--LN-----A-A-Q-
.....GS-V-GL-K-IN-T--V-G--F-LI--RTW-FCN-----A-A-Q-
.....G-TR--T-F--V--I-L--R-G-VLN-----A-A-
KIINVC.RIC.AAVT-L-Q-T-DTL-V-NW-GI-AGI--RTIG-N-----S-
KTIEAC.RLC.GAVM--L-Q-TN-DTI-VS-NW-GI-LGL--RIGOGFL-----A--V-
KITSAC.RIC.AAVT-L-Q-T-DTV-V--NW-GI-SGI--AIG-G-N-T-----S-
KIINAC.RIC.GAFT--L-Q-T-D-I-V--NW-SI-SGI--RIG-G-N-----F-A-
KITSAC.RIC.IAVI--L-Q-T-DTI-V--NW-VTI-LGI--RIG-G-LN-----S-
RITEAC.RLC.AAVT-L-Q-T-DTV-V--GSW-NI-SG--RTG-LN-----A-X-
.....GIIA-GK-D-I-TI-V-----G-L-LA--RIG-G-L-----A-
.....-GI-A-GK-D-I-T-----G-L-LA--RIG-G-L-----T-V-
.....-GX-A-GK-D-I-XTX-----X-LA--RIG-G-L-V-----T-V-
.....-GI-A-GK-RD-I-T--V-----I-LA--RIG-G-L-----A-I-
.....NCLRDC.FAACG-T--Q-T-V-N-S-NW-Q-A-G--QIG-GFLN-----L-S-
RRLIHLTF.LH-ET-GILIRQHCIVCIDTL-EFTGWW-G-AL-RV-VDI-----T-I-A-N-
.....N-STL--TIIRTEIKNIDRL-W-G-K-SILLAL--TIV-I-EV-----I-A-N-
.....RGII--GR-RT-T-D-----G--I-RA-----IVF-I-GN-----A-
.....KGV--GR-V-I-DT-----I-A--RFG-G-LN-T-----A-
.....GVIA--GR-QI-T-DT-----I-L--RIG-G-L-----S-
.....GIIA--GR--IN-DT-V-----I-L--RIG-G-L-----F-A-
.....QGI--GK-T-T-DT-----I-I--RIG-G-L-----F-A-
.....RGI--G-T-I-DT-----I-IA--RFG-G-LN-----A-
.....KSI--G-T-I-D-----I-IA--RVG-G-L-----A-
.....KGTG--L-RT-I-DT--I--R-TI-A-TRIG-G-L-----A-E-
RLLIIVGVRQI.IEKSNTYAS.LQATYAL-NFTGWW-I-G-A-TRIG-G-L-----A-E-
LKWRLF.RDW-ANAYVIV-I-FIRN-DRL-NFT-WW-IL-GA-VNIF-G-N-T-----I-A-N-
.....WNMLTRDC.FH-ET-GILIRQHCIVCIDTL-EFTGWW-G-AL-RV-VDI-----T-I-A-N-
LKSROL...FDW-SNTYSI-RT-LIQADIRL-NFTGWW-L-AG-AFVAOG-N-----I-A-N-
.....GVIA-AR-V-T-DT-----I-LT--RRLFLG-I-----S-X-
VNNLTRDC.FAFIA-G--Q-I-DCV-VWT-DW-Q-AIA--RIG-G-LN-----S-
VNNLTRDC.FAFIA-G--Q-I-DCV-VWT-NW-Q-AIA--RIG-G-LN-----S-
VNNLTRDC.FAFIA-G--Q-I-DCV-VWT-DW-Q-AIA--RIG-G-LN-----S-
.....GLNCRVDC.LAVCG-V-OR-TN-DT-V--W-Q-LIG--RIG-G-LN-----S-
TFEACKTL.KATA--L-L-OR-XTN-DTV-V--NW-SI-LG--RFG-G-LN-----L-S-
GINCLRDC.LAVCG-A-QQ-T-DTV-VR-DW-Q-L-G--RIG-G-LN-----L-S-
AEATRRR...VAYIH-GL-F-AAVI-AMD-L-F-TWNW-EA-LAYL--RRLAOG-IA-----A-I-L-N-
EQRVL...VAYIH-GL-F-AAVI-AMD-L-F-TWNW-EA-LAYL--RRLAOG-IA-----A-I-L-N-
AEATRRR...IAYI-GL-F-AAVI-GWDLT-F-TWNW-EA-LVAL--RRLAOFETIA-----A-I-L-N-
POHPPTTSDK.RRS..A--T-RYL-LVPRGSLQIL-TLQ-WLRSR--ARGW-RAPEYL-GWIYDRPO--GPA-
SORPETPNTR.RTVF...T-W-L-LVPRCPDLSIP-TLS-WLRSR--TRRW-RASEYL-GWLFDRPE--SPA-
GEOPLPDGET.HQPRRMLP-A-WRL-LVPRSANL-IPOTLQ-WLRSR--TSG-G-RALESCG-WIR-DPO--GPA-
EQRVL...GY--GWK-F-DTI-WAGGELGOTFORWALVAL-L-GRS-E-LA--A--A-L-F-N-
ERFRVL...GYC--GWK-F-DTI-WAGGELGOWLQIGIAEALQGL--VR-GG-RV-A--A-L-P-N-
EHLRVV...GYC--GWK-F-DTI-WAGGELTWEWLOGIAOVA-QL--VWGG-NLLA-A--A-L-L-N-
.....QEAGRGIC.TLLREAAQRLLWRGGROFGLSAARAL--R-LAQGVAN-----A-I-L-N-
.....QEAGRSIC.ALLQEAAERLWRGGROGLSTARAL--RAFAQ-LAN-----A-V-L-N-
.....Y--GL-Q-OEA-TG-AQ-L-R-AR-AWG-LGAI--RS-Y-VINS--V--K-V-G-
EVAARV...IAYIS-GI--QAA-TGI-DSL-F-TWNW-EA-LHAC--RRWV-EFLA-----A-I-LFN-
EVAARA...AAYIS-GI--QAAVSGI-DSL-VFTWNV-EPLVOT--GRWV-EFLA-----A-I-L-N-
LFFHQHQLWKASSGMMV.ENAQKI-RWSW-KIRNRNRDIIAASHARDIRLG-KKRWR-FRG-SGF--PSETTETL-
LSYHQHQLWQTLRSFLV.ESAQKI-SWIR-KSRNRNG-LSKTNGTINIRL-NRR-W--FRF-SGL--SSEATETL-
LFFHQHQLWQTLRSRMMV.ENAQKIQSW-REKCRNRNG-LSDTRKNIQLGKKRWRML-FGG-SGL--SSEATETL-
VQEA...IRS.LGS-LYSFAASWRF-ETWGLQ-IAAA--GTC-W-L--SA-P-N-
.....KTAGGKIE-VLOGYAQGLWSRGHOW-LSTAACFR--AIA-G-IN-----A-V-L-N-
.....KAAGARICOVLOSFAORLWSGGHOL-LSVIRGAA--AIG--GN-----A-V-L-N-
.....KAAGAALYOVLOGFAORLWSRGYOL-LSG-RGAA--AFG-G-WN-----A-A-L-N-
.....LTY--GWSYFOEAVQVWRSATETL-GANGDLW-TL--GRVG-W-LA-----E-L-T-
.....RRL-LAY-GI--QAAVT--D-L-RFTIYW--ALLHAG--GRLV-WA-----A-I-F-N-
RRCSELLARL.LAYI--QAAATAA-DQL-F-TWNW-EPILOF--RRWV-EFLA-----A-I-LFN-
EALGRFIAH...A-GI-K-QAAVSGA-D-L-F-TWNW-EPILOFS--KIW-KFLA-----A-I-L-N-
ARLREVIRRI.IAY--GL--QEA-TGF-D-L-F-TWNW-EA-LQ-L--RRVA-EFLA-----A-I-LFN-
.....WT-QKSLQHWVRTLKQKRO...HGINLQOKK--RRGSPLOKAEQLGREGAE.TTGV-
.....AKV-DGW-F-AF-W-GEMA-ENAYYTRGLQSLA--ROLAGWPATCG-----L-N-
.....AKI-DGW--EFSGW-AEMAQOAYVWRGLCVTA--DIA-WPTVC-----F-L-N-
.....AKI-GW-F-EFSSWFAEMALONAYYTRGLCA-A--RDFAGWPAMVC-----L-CN-
QDIQQRTRQ...TAHLEYGWQELKAAAASL-VAV-QAA-SASDS-RHACRSIV-VIAH--M-E--WFN-
.....Env end
.....gp41 end

MAC.US.x.239
H2A.DE.x.BEN
H2A.PT.x.ALI
H2A.SN.86.ST.JSP4.27
H2B.CI.x.EH0
H2B.GH.86.D205.ALT
H2C.CI.92.AB996
H2U.CI.07.07IC.TNP3
H2U.FR.96.12034
SMM.CI.79.SIVsmC12
SMM.SL.92.SIVsmSL92A
SMM.SL.92.SIVsmSL92B
SMM.US.06.F10
SMM.US.11.SIVsmE660.FL8
SMM.US.86.CFU212
SMM.US.x.H9
TAN.UG.x.SIVagmTAN1
VER.KE.x.AGM155
VER.ZA.10.VSAC4004
MAL.ZM.2009.SIVagmMAL.ZMB
VER.ZA.10.VSAG1003
VER.ZA.10.VSAGM022
WRC.CI.98.98C104
WRC.GM.05.Pbt.05GM.X02

.....LTY--GWSYFHEAVQVWRSATETL-GANGDLW-TL--RRGG-W-LA-----L-T-
.....AAQ--GCEWIOEAFQAFART-RETL-GANGWLW-AA--RRIG-G-LAV-----A-L-A-
.....AAY--GCEWIOEAFRAFARIARETLTNTWRDLWGA--WVG-R-LAV-----A-I-A-
.....TAY--GGEWIOEAFRAFAR--GETLTNAWRGFWGTL--GGIG-G-LAV-----A-I-A-
.....PAY-R-GISWFOEAIQAAAR-AGETL-SAARTSWG-L-RR-AGE-IA-----A-L-A-
.....VAY--GISWFOEALRAAAR--GETL-SAGETLW-AL--RRWV-EFLA-----L-T-
.....AAYFS-GFRW-QEACTLRAT-GETLSTWLRKML--RRVG-G-LA-----L-T-
QYGCQWF...QEA--AGWKAXRETLA-AWRDLWETLGRVGRGILAIPE-----F-L-A-
.....IAY--GW-YF-EAFQAFK-ARETLRTRGRELW-TL--GRVG-W-LA-----F-L-A-
.....IAY--GW-YFOEACQEAAR-ARETL-SAGRDLW-TL--GRIG-GLLAA-----L-AF-
.....VAYIR-GW-YFIE-LOEAL-AROTLVSTGRITQTL--GHIG-G-AA-----L-T-N-
G...YIR-GWYFIE-LOEAL-AROTLVSTGRITQTL--GRVG-G-LA-----L-M-N-
.....LTY--GWSWFOEA-QAAFK-ARETL-SAWRDLW-IL--RRVGGWMLA-----L-T-
.....IAY--GWRYFOEAVQAWKFARETL-SAWRDLW-TL--GRVG-G-LA-----V-L-T-
.....LTYIG-GCLWF-EALQAAGR-AGOTL-SAWRDLW-TL--GRVG-G-LN-----L-T-
.....LTY--GWRY-QEA-QXWKFRETL-SAWRDLW-TL--GRVG-G-LA-X-----L-T-
TSFHGQRLWQTLRGMGL-HLIQATSRIRNACRSTERYSSQKASRIFSL-RKRPKWRT-GSRIPSETTETL-
.....FIS-GVT--EA-I-GREV-HQNVAIQAL-LAYA--PRVAENVAAL--L--Y-N-
.....FIS-GFN--IA-A-GREIRDW-AIWOAIYAAT--RRVVE-VAAL--L--I-Y-N-
TAAAAATLQY...GFS-IOAAWRDGTI-L-RFTWNV-NG-LAAA--RW-GE-AA-T--Y-CFT-
TAAV...ATI--GFS-IOAAWGDVLT-L-RFTWRW-SG-LAAA--RW-VQ-AA-T--Y-C-A-
TLLLQLLRQE.CQYI--GW-QF-EG-ARSE-F-L-S-ASASRSLWNAC--RS-Y--LEH--M-E--WFN-
.....Y--GLG--AA-QEAV-HL-SFARNAAHWLAC--RS-Y--INS--V--E-V-N-
.....Y-H-GLR--OTA-QE-AQV-V-S-AONAAHOIWLAC--RSIY-TVINS--V-A-Q--N-
.....Y--FGF--AA-QEAAR-V-SFAONAAHOIWLAC--RS-Y--INS--V-EC--N-
.....Y--GLG--OTA-QEAARTV-RLAQNLAGHOIWLTC--RS-Y--VVDV--V-EV--N-
.....GFE--AA-QEAAR-V-GVAQNAHQIWLII--RS-Y--INS--V-EI--N-
I...SQAHACTRKLTKSRKETEK-TSEISSLKSR--SS-VNKL-T-LCSSLNKSRRR--N-
L...CGAAYAAALQEKIAARRQKKEK-IYTPGSRET...SRSSIRGITA-LCTNLNPCRGGQ--

856
856
852
853
860
873
868
874
879
871
840
833
853
885
859
865
913
849
846
846
878
842
857
871
871
865
837
859
855
850
890
888
893
887
847
871
882
838
843
835
847
862
851
900
884
855
882
887
915
917
913
879
867
868
891
883
870
892
883
893
844
855
851
868

880
861
863
860
853
860
860
882
868
887
886
887
882
888
888
886
886
841
833
873
891
880
871
878
877
869
873
885
877

Table of HIV sequence alignments for PLV proteins. Columns include protein names (e.g., H1B, H1C, H1A1), amino acid sequences, and various annotations like 'myristoylation', 'Nef start', 'acidic cluster', 'phosphorylation', and 'poly-P helix'. The table is organized into sections for different proteins and mutations.

| | HXB2 premature Nef end | | normal Nef end |
|-----------------------------|---------------------------------------|----------------------------------|------------------------|
| H1B.FR.83.HXB2-LAI-IIIB-BRU | GLEGLIHSQRORODILDWIYHTQGYFPD>QNYTPGPG |VRYPLTFGWGKYLVPVPEKIDEEAN | .KGENTSLLHPVSLGHMDPERE |
| H1C.ET.86.ETH2220 | ---Y-KK-E---V-N-F-W | ---F---D-SEV-I-E-NC-A-E-ED | -K-K-H-RR-M---Y-D |
| H1F1.BE.93.V1850 | ---Y-KK-G-T---V-N-F-W | ---F---D-EFV-K-E-NC-M-A-E-ED | -R-K-S-LR-I-R-FYOD |
| H1A1.UG.85.U455-U455A | -D-Y-KK-E-V---F-W | ---F---D-AEV-T-G-N-ICO-V-E-K | -M-K-T-LK-R-Y---FY-D |
| H1N.CM.02.0300131 | ---V-W-RK-E---M-IL-WH | ---I-V-F-LSAEAV-E-D-NA-ICO-V-DHK | -V-S-RR-I---DFY |
| H1N.CM.05.YBF30 | ---V-W-RK-E---N-I-W | ---I-V-F-LSAEAV-E-D-NA-ICO-V-DHK | -V-S-RR-I---DFY |
| H1N.CM.04.04CM-1015-04 | ---V-W-RK-E---V-F-W | ---I-V-F-LSAEAV-E-D-NA-ICO-V-EHK | -V-S-RR-L-DFY |
| H1N.CM.04.04CM-1131-03 | ---V-W-RK-E---V-F-W | ---I-V-F-LSAEAV-E-D-NA-ICO-V-EHK | -V-S-RR-L-DFY |
| CPZ.CM.05.SIVcpzEK505 | ---V-W-TK-E---V-I-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.CM.05.SIVcpzLB7 | ---V-VY-K-E---GV-I-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.CM.01.SIVcpzCM13 | ---V-VY-K-E---V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.GA.88.SIVcpzGAB2 | ---V-Y-K-E---V-WN | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.CM.05.SIVcpzDP943 | ---V-Y-R-E---V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.US.85.US-Marilyn | ---V-Y-R-E---V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.CM.05.LB715 | ---V-Y-R-E---V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.CM.05.SIVcpzMB66 | ---V-Y-KK-E---V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.CM.05.SIVcpzMB897 | ---Y-E-E---V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| H10.BE.87.ANT70 | ---Y-HK-AE-V-N-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| H10.CM.98.98CMA105 | ---Y-HK-AE-VH-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| H10.CM.98.98CMAAB197 | -D-Y-HK-AE-V-N-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| H10.CM.99.995E-MP1299 | -D-Y-HK-AE-V-N-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| H10.CM.91.MVP5180 | -D-Y-HK-AE-V-N-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| H10.US.99.99USTWLA | -D-Y-HK-AE-V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| GOR.CM.12.SIVgor-BQID2 | -D-Y-HK-AE-V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| H1P.CM.06.U14788 | -D-VY-PE-AE-V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| GOR.CM.12.SIVgor-BPID1 | -D-VY-PE-AE-V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| GOR.CM.13.SIVgor-BPID15 | -D-VY-PE-AE-V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| GOR.CM.04.SIVgorCP684con | -D-VF-A-S-V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| GOR.CM.07.SIVgorCP2135con | -D-VF-A-S-E-V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| GOR.CM.07.SIVgor2139-287 | -D-VF-A-S-V-F-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.CO.06.BF167 | -D-YW-RK-AA-T-MFN-I-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.TZ.08.TAN1 | -D-PF-TK-AA-M-M-N-I-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.TZ.09.UG38 | -D-PF-TK-AA-M-M-N-I-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.TZ.06.SIVcpzTAN13 | -D-X-Y-T-AA-T-M-N-V-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| CPZ.CO.90.ANT | -D-X-XXLK-AA-M-MFN-I-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| RCM.CM.00.SIVragi_09CM312 | -D-IYV-I-HR-IYLENEH-II-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| RCM.GA.X.SIVRCM1 | -D-IYV-I-HR-IYLENEH-II-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |
| RCM.CM.02.SIVrcm-02CM8081 | -D-IYV-I-HR-IYLENEH-II-W | ---I-V-F-LTSEEV-Q-D-NV-MCO-E-DK | -V-S-RV-R---FYO |

max HIV-1 similarity

| | max HIV-1 similarity | | normal Nef end |
|------------------------------|---|--------------------------------|--|
| MAC.US.X.239 | ---IYY-A-HR---IYLEKEE-II-W-D-S |I-KM-LW---NVSDEA.QE | .D-EHY-M-AQTSQW-.PWGE-A-K-K-PT-YTYE-VYRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| SHM.CI.79.SIVsmCI2 | ---IYY-E-HK---IYLENEE-II-W-S |I-KM-LW---NVSDEA.QE | .D-RHC-V-AQISL-.PWGE-A-K-K-PT-YTYE-VYRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| SHM.CI.92.B192B | -D-IYY-E-HK---IYLENEE-II-W-S |I-KM-LW---NVSDEA.QE | .D-RHC-V-AQISL-.PWGE-A-K-K-PT-YTYE-VYRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| MNE.US.X.MNE627 | ---IYY-E-HK---IYLEKEE-IV-W-D-S |P-K-LW---NVSDEA.QE | .E-ENY-AQTSQW-.PWGE-V-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| H2U.CI.07.07IC-TNP3 | ---IYYN-HR---MYLENEE-IVA-W-E |F-K-LW---HVSDEA.S | .D-THC-V-AQTSQW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| SHM.US.06.FTQ | ---IYYN-HR---IYLEKEE-II-W-S |I-KM-LW---NVSDEA.QE | .D-THC-V-AQTSQW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| SHM.US.11.SIVsmE660-FL8 | ---IYYNEK-HR---MYMEKEE-II-W-S |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| SHM.US.X.H9.1 | ---IYYN-HR---IYLENEE-II-W-S |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| SHM.US.86.CFU212 | -D-IYY-E-HR---MYLEKEE-II-W-S |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| H2U.FR.96.12034 | -D-IYY-E-HR---IYLENEE-II-W-S |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| SHM.SL.92.SIVsmSL92A | -D-IYY-E-HK---SYL-NEE-II-W-R |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| H2A.DE.X.BEN | -D-IYY-E-HR---IYLEKEE-II-W-H |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| H2A.CI.86.VD J5P4-27 | -D-IYY-E-HR---IYLEKEE-II-W-H |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| H2B.CI.X.ALI | -D-IYY-E-HR---IYLENEE-IVS-W-H |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| H2B.CI.X.EH0 | -D-IYY-E-HK---TYLENEE-IVS-W-H |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| H2B.GH.86.D205-ALT | ---IYY-E-HR---TYFENE-IVS-W-H |I-MY-LW---DVSDEA.QE | .D-THC-M-AQT-QW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| H2G.CI.92.ABT96 | ---IYY-E-HR---IYMEKE-II-W-A |I-KX-LW---TVSDEA.QE | .D-THC-V-AQTSQW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| ASC.UG.10.RT03 | ---FY-PE-HKK---YA-NW-LV-GW-Q |P-F-TC-LW---DVSDEA.QE | .D-THC-V-AQTSQW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| ASC.UG.10.RT01 | ---FY-PE-HKK---YA-NW-LV-GW-Q |P-F-TC-LW---DVSDEA.QE | .D-THC-V-AQTSQW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| ASC.UG.10.RT08 | ---FY-PE-HOK---YA-VW-LT-GW-Q |P-F-TC-LW---DVSDEA.QE | .D-THC-V-AQTSQW-.PWGE-A-K-K-PT-YTYE-YTRY-E-GSKSGLSEEEVRRRLTARGLL..NMADKKKTR- |
| GSN.CM.99.CN166 | ---MFYCED-HOK-EQYA-LEW-LV-GWLSF |I-TIP-F-IC-R-ATTDES-PG | .DDEYL-T-AYOGRS-EDHKEF-VFS-C-K-IKSGTLDLOQEERKRR..LTANRFL |
| GSN.CM.99.CN71 | ---AMFYCED-HOK-ESYC-YEW-IV-GWLOW |I-TMP-F-WC-R-ATTDES-PG | .DDEYL-T-AYOGRS-EDHKEF-VFS-C-K-IKSGTLDLOQEERKRR..LTANRFL |
| MON.CM.99.L1-99CML1 | ---GMFYCED-HRKIEQYA-LEW-LV-GWLE |I-TMP-F-WC-R-ATTDES-PG | .DDEYL-T-AYOGRS-EDHKEF-VFS-C-K-IKSGTLDLOQEERKRR..LTANRFL |
| MUS-1.CM.01.CM1239 | ---MFYCEE-HOK-FYTY-VOW-IV-GWLOF |I-TLP-F-LWC-RE-ATTDES-PG | .DDEYL-T-AYOGRS-EDHKEF-VFS-C-K-IKSGTLDLOQEERKRR..LTANRFL |
| MUS-2.CM.01.CM1246 | ---MFYCID-LOK-ETYA-VKSKIV-RSLI-D |I-TLP-F-LWC-RE-ATTDES-PG | .DDEYL-T-AYOGRS-EDHKEF-VFS-C-K-IKSGTLDLOQEERKRR..LTANRFL |
| MUS-3.GA.09.SIVmus-09Gab0181 | ---DRMFYCYD-HOK-ETYA-MEW-IV-GWLR |I-TMP-F-LWC-RE-ATTDES-PG | .DDEYL-T-AYOGRS-EDHKEF-VFS-C-K-IKSGTLDLOQEERKRR..LTANRFL |
| SYK.KE.X.K51 | -D-MFFC-K-HE-0-YLKNH-VID-DIT-S |I-LWE-A-N-TEGYLVDE | .EDTLMHHPAAGVGAS-EDPHRN-M-N-NPH-YTPGWEMARQLEERTGKR |
| SYK.KE.X.SYK173-COMGNM | -D-MNVCYCEK-DE-H-YLQNEH-ID-RI-S |I-T-LWE-N-TEGGL-YE | .EHTLL-HPASGQSSSMG-PHVE-LOPPPGTPGWEMARQLEERTGKPKPEQASLAKNIS |
| TAL.CM.00.266 | -K-IFYNK-HA-N-YAHNEW-IL-W-S-E |I-K-ILW-C-THEDDT | .EDGLH-P-AYDGOA-EDPWGES-V-Y-EK-YTPG-KMAEYDRLERKRMILLAPPQTASS |
| TAL.CM.01.8823 | -K-IFYNK-HA-N-YAHNEW-IL-W-S-E |I-K-ILW-C-THEDDT | .EDGLH-P-AYDGOA-EDPWGES-V-Y-EK-YTPG-KMAEYDRLERKRMILLAPPQTASS |
| DEB.CM.04.SIVdeb04CMPF3061 | ---DIFYA-A-HA-E-HAONEW-II-GWLO-E |I-KY-FLF---IADPPEE | .D-RNI-DAHOGQ-EDPKYR-V-K-Q-YCYK-GHEAHTKETHRR-CMFP-KRK |
| DEB.CM.99.CM40 | ---DMVFE-A-HA-E-HAONEW-II-GWLO-K |I-KY-FLF---IADPPEE | .D-RNI-DAHOGM-EDPKYR-V-K-Q-YCYK-GHEAHR-EHRR-CMFP-KRK |
| DEB.CM.99.CM5 | ---DIFYA-A-HA-E-HAONEW-II-GWLO-E |I-KY-FLF---IADPPEE | .D-RNI-DAHOGQ-EDPKYR-V-K-Q-YCYK-GHEAHR-EHRR-CMFP-KRK |
| SAB.SN.X.SAB1 | ---IYY-E-KK---YALNEW-IDW-G-A |I-KY-FLF---IADPPEE | .D-RNI-DAHOGQ-EDPKYR-V-K-Q-YCYK-GHEAHR-EHRR-CMFP-KRK |
| GRV.ET.X.GRI-677-gr1-1 | -D-IYY-E-EK-N-YALNEW-IDW-A-S |I-RV-F-F-DLHEEA-R | .C-RHC-V-AQMGEDP-GIDHGE-V-K-PM-VEFK-GI-FSDMHATVGKGN |
| MAL.ZM.09.SIVagMAL-ZMB | -D-IYY-E-DK-N-YALNEW-IDW-A-S |I-RC-F-F-DLHEEA-R | .C-RHC-V-AQMGEDP-GIDHGE-V-K-PM-VEFK-GI-FSDMHATVGKGN |
| VER.KE.X.AGM155 | -D-IYY-E-DK-N-YALNEW-IDW-A-S |I-RC-F-F-DLHEEA-R | .C-RHC-V-AQMGEDP-GIDHGE-V-K-PM-VEFK-GI-FSDMHATVGKGN |
| VER.ZA.10.VSAG1003 | -S-IYY-E-NK-T-YALNEW-IDW-A-S |I-R-F-F-DLHEEA-R | .C-RHC-V-AQMGEDP-GIDHGE-V-K-PM-VEFK-GI-FSDMHATVGKGN |
| VER.ZA.10.VSAG0022 | -S-IYY-E-DK-T-YALNEW-IDW-A-S |I-R-F-F-DLHEEA-R | .C-RHC-V-AQMGEDP-GIDHGE-V-K-PM-VEFK-GI-FSDMHATVGKGN |
| VER.ZA.10.VSAG4004 | -D-IYY-E-DK-T-YALNEW-IDW-A-S |I-R-F-F-DLHEEA-R | .C-RHC-V-AQMGEDP-GIDHGE-V-K-PM-VEFK-GI-FSDMHATVGKGN |
| TAN.UG.X.SIVagntAN1 | ---IYW-PK-EQ-N-YALNEW-IDW-A-S |I-K-R-F-F-DVSDEA.QE | .D-RHC-AQIOWES-.PWKET-V-K-NPL-VQYNPD-SFKDMHGLV |
| COL.CM.X.CGU1 | -I-RTPE-D-LIEQYA-IEW-CLKWLE-EDEL-EDGALKED |K-VA-LW---YI-OLGEYAYS | .YDL-S-TS |
| COL.UG.10.BWC01 | -I-RTPE-D-LIEQYA-IEW-VLKWLE-EDDETENMOIREE |K-VA-LW---YI-OLGEYAYS | .YDL-S-TS |
| COL.UG.10.BWC07 | -I-RTPE-DRL-ERYC-IEW-VLS-WLE-EDETENBDGTPKEE |L-AC-LW---LLDSIGWYAC | .YDM-NRPVT |
| OLC.CI.97.97C112 | ---MNY-P-ENLIV-YALNEW-IIGNF-ATEVEEFGT |I-K-K-P-LWD-R-LPETLEE | .RG |
| WRC-F.98.98C104 | ---YTFY-P-E-VN-YAFNEW-LTEGW-A-E |I-K-K-W---HSESNOSDOGF-KNL | .GAOSES-PWGEH-V-K-P-CDF-SR-PGHIAOAGDAIETLHKELKRKH |
| WRC.GM.05.Pbt-0185GM-X02 | ---DIFW-P-EE-VN-YA-IEW-IVEGW-A-E |I-K-K-W---HSESNOSDOGF-KNL | .GAOSES-PWGEH-V-K-P-CDF-SR-PGHIAOAGDAIETLHKELKRKH |
| DRL.DE.11.D3 | -K-IWW-L-ES---YAONEW-FIKGW-S-K |I-KV-FLF---DV-D-DLV | .QPCNR-NSSOMGPI-S-GER-M-Q-AS-YTFO-CIH---E-GHVTISRHDKE..EPDCCKRRKWO |
| DRL.DE.11.D4 | -K-IWW-HK-ET---YAONEW-FIKGW-S-R |I-KV-FLF---V-E-DVLD | .QPCNR-NSSOMGPI-PNGER-M-Q-PT-YSFK-CIH---E-GHARSIEH.IAE..EPACCKRRKWO |
| MND-1.GA.X.MNDGB1 | -G-IYW-R-EE---YAONEW-FE-GW-Q-T |I-K-FLF-E-SRAI-PLEE | .AA-NH-SSO-GIOE#-P-GER-M-SGTL-I-PMTLQH |
| MND-2.CM.98.CM16 | -D-IYW-K-EE---YAONEW-FI-TW-S-D |I-K-RFLF-C-AVPP-DOE | .N-CNK-QS0-GIOE#-P-GER-M-K-G-YTFY-PTIR-EY-CVTSLSIYAYAKEHPEDCKRRKWO |
| MND-2.X.X.5440 | -IYW-K-EE---YAONEW-LI-NW-K |I-K-FLF---AVSP-PDEE | .D-CNR-NSSO-GIOE#-P-GER-M-K-G-YTFY-PTIR-EY-CVTSLSIYAYAKEHPEDCKRRKWO |
| MND-2.GA.X.M14 | -IYW-R-EE---YAONEW-FI-W-E-S |I-KR-FLF-I-QVPP-DOE | .Q-CNR-NSSO-GIOE#-PWGER-M-K-PE-YTFY-PTK-KEYRHVTSLSIYAY..EKSDCKRRKWO |
| SUN.GA.98.L14 | ---WY-RT-EE---YAONEW-FITG-D-K |I-KA-LW-A-TI-EDRDP | .HPCOA-SSOQGVNE-PWGER-I-T-PT-YDFR-IQK---E-HVTSLVG |
| LST.KE.X.LH07 | ---YW---I-YCENEW-LIG-FM-D |T---LWO-E-AC-EYKDP | .D-TQC-SSO-GVLE-PWGER-I-H-NPM-VDFI-LRKO-AKTO-NTAFAPD |
| LST.CO.98.SIVlhoest485 | ---YW---I-YCENEW-LIG-FM-D |T---LWO-E-AC-EYKDP | .D-TQC-SSO-GVLE-PWGER-I-H-NPM-VDFI-LRKO-AKTO-NTAFAPD |
| LST.CO.98.SIVlhoest524 | ---YW---I-YCENEW-LIG-FM-D |T---LWO-E-AC-EYKDP | .D-TQC-SSO-GVLE-PWGER-I-H-NPM-VDFI-LRKO-AKTO-NTAFAPD |