HMMER Sequences in the Gag Alignment

A_U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
B_HXB2R	HIVHXB2R	K03455	Starcick,B.	Science 227, 538 (1985)
C_UG268	HIVUG268	L11799	Louwagie,J.J.	AIDS 7, 769 (1993)
D_ELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
F_BZ163B	HIV1BZ163B	S0585	Louwagie,J.J.	ARHR 10, 561 (1994)
G_LBV217	HIVLBV217	L11778	Louwagie,J.J.	AIDS 7, 769 (1993)
H_VI557	HIVVI557	L11793	Louwagie,J.J.	AIDS 7, 769 (1993)
O_ANT70C	HIVANT70C	L20587	Vanden Haesevelde,M.	JVI 68, 1586 (1994)
CPZANT	SIVCPZANT	U42720	Vanden Haesevelde,M.	Virology 221, 346 (1996)
CPZGAB	SIVCPZGAB	X52154	Huet,T.	Nature 345, 356 (1990)
A_ROD	HIV2ROD	M15390	Clavel,F.	Nature 324, 691 (1986)
B_EHOA	HIV2EHOA	U27200	Rey-Cuille,M.A.	Virology 202, 471 (1994)
SD_MM251	SIVMM251	M19499	Franchini,G.	Nature 328, 539 (1994)
STM_STM	SIVSTM	M83293	Novembre,F.J.	Virology 186, 783 (1992)
VER_AGM3	SIVAGM3	M30931	Baier,M.	Virology 176, 216 (1990)
GRI_AGM677	SIVAGM677	M66437	Fomsgaard,A.	Virology 182, 397 (1991)
TAN_AGM17	SIVAGM17	L19250	Hirsch	Virology 197, 426 (1993)
SAB_SAB1C	SIVSAB1C	U04005	Jin,M.J.	EMBO J. 13, 2935 (1994)
SYK_SYK	SIVSYK	L06042	Hirsch,V.M.	JVI 67, 1517 (1993)

HIV1 GAG

The following alignment was generated using the HMMER program. For simplicity, only representative types and subtypes are shown. An ordinary consensus sequence (lowercase signifies majority, uppercase signifies 50% or greater) was created from these sequences using MASE; this is not a "most likely sequence" based on a HMMER model (Part II). Annotation is based on HIV-1s, therefore the user shouod be cautious about its applicability to other PI sequences. Cleavage sites are designated by '\/'.

```
GAG_CONSENSUS ATGGGtGcgagagcgTCa???gtatTaagagGgagaaaattaGAt??aTggGAaaaaaTtaggtTacggC     65
A_U455        -------------------...--------C----A----------TC------G------C-----A---     67
B_HXB2R       -------------------...--------C---G--G--------CG----------C-----A---     67
C_UG268       -------------------...-----------CG-----------ACC-------------A----A---     67
D_ELI         -------------------...--------C---G-----------AA-------------C---------     67
F_BZ163B      -------------------...--------C---G-----C-----GC-------------C-----A---     67
G_LBV217      -------------------...--------C---G-----------GA----------------C-GA---     67
H_VI557       -------------------...--------C---G-----------GC------G------C----GA---     67
O_ANT70C      -----------T-----T...--G--G-C---A--C-----G---GC-------C-----------AA--     67
CPZGAB        -------------------....--TC---C----G-----------CGC---------G--C--C-TA---     67
CPZANT        -----A---G-G-----T...--T--G--G--AGAG--GC-----AC--------GT--C---C-T----     67
A_ROD         -----C------AAC--C....--C--G-------A----GC----GA--TA----G---C----------     67
B_EHOA        -----C-------GC--C....--C--GTC-----AG--GAC----GA--TA-----GG------------     67
SD_MM251      -----C------AAC--C....--C--GTC-----AG---GC----GA--TA------------C----A-     67
STM_STM       -----C------AGC--C....--C--GTC-----A----GC----GA--TA------G------------     67
VER_AGM3      -----G---GCTA-C---...-C-C---ATA-----C-------CAA--TT--GC-T--AC-AC-T--C-     67
GRI_AGM677    -----C-G-G-TCAC---...-C-C-GTC--------GCC-C--CACG-TC--G--G------C----T-     67
TAN_AGM17     -----A--AG--CAC--G...-C-C-GTC-A----G--T--G--CAC--TT--------AC-C-----C-     67
SAB_SAB1C     -----------TAAC---...-C-----T--T--G---C-----GC--TC---TC-G-C----------     67
SYK_SYK       -----A--AGCG-G---GGCAA------C---T---G-GC----CAG--AT--G------C--C-C--A-     70

                          <- INS-1 in HIV-1s
GAG_CONSENSUS C?gg?ggaAagAAaaa?TAta?a?T?aaaCAt?TagTaTgGgcAagcaaggAgcTgga?agaTttgca?t    126
A_U455        -A--G-----C-----A----G-C-G------T---------------G---------A-A---CA--C-    137
B_HXB2R       -A--G-----------A----A-T-A------A---------------G------A--AC----C---G-    137
C_UG268       -A--G---------TGC----TGA-G-----CC----C----------G--------GA---------C-    137
D_ELI         -A--A-----------A----G-C-A------A---------------G------A--AC---A----C-    137
F_BZ163B      -G--G-----------A----GCT-G--G---C---------------G------A--AC--------A-    137
G_LBV217      -A--G-----------A----G-A-G------T---------------G---------G---------C-    137
H_VI557       -A--G-----------A----G-C-G------T---------------G---------A---------C-    137
O_ANT70C      -A--ATCT--A--G--A----G-C-A-----CT---------------GA--------A-----C---TG    137
CPZGAB        -C--G----GA----GA----TGA-G------T---------------G---------A-----C---TG    137
CPZANT        -C--T--C--------G--C-TGA-A------C-G--T--------ATC-------C-GC-T-----GC-    137
A_ROD         -C--C-----------G--C-GGC-A------A-T--G------GCG--TA-AT----C-----C-G-T-    137
B_EHOA        -C--C----------GA--C-TGC-A------A-----------GTG--T--AT-A--G-------G-T-    137
C_2238                                                  %%-GCG--C--AT----C---------T-     27
D_FO784                                                 %%-GCA--T--AT----C-------G-T-     27
SD_MM251      -C--C-----------G--C-TGT-G--G---G-----------GCA--T--AT-A--T-------G-T-    137
STM_STM       -C--C--G--------G--C-TGT-G------G----G------GCA--C--AT----C-------G-T-    137
VER_AGM3      -GACC-----------G--CCA-A-T------T--A--------G-------AA----GC-C--C-GCC-    137
GRI_AGM677    -GAAC--G--A--G--G--CCA-A-T------T--A--------G-A--A--AA----AC------GGT-    137
TAN_AGM17     -GAAC-----------G---CA-C-A------T--A---A----G--------A----GC-C----GCC-    137
SAB_SAB1C     -CAAC--------G--G----AGT-A-G----C-------------------A--A--TC-----AGCC-    137
SYK_SYK       -GAAA--G-----GCGC--CCT-G-TCG----T----T-------AG------T----C-----C-GCC-    140

GAG_CONSENSUS ??a?gc??gcct?tTaGaaacaaaaGAaGG?tgtcaaaAaat??t???agt?tTa?a?cCatt?gt?ccg    179
A_U455        TA-CC-TG----T---------GC------A-----GC----AC-GGG-CAA---C-A---GCTC-C-A-    207
B_HXB2R       TA-TC-TG----G---------TC------C---AG-C----AC-GGG-CAGC--C-A----CCC-T-A-    207
C_UG268       TA-CT-TG----T---------TC------C---A--C----AA-GAA-CAGC--C-A---GCTC-T-A-    207
D_ELI         TA-TC-TG----T---------TC------C---A--C----AA-AGGGCAGC--C-A---GCTA-T-A-    207
F_BZ163B      TA-TC-TG----TC--------TC------C-----------AA-AGG-CAG---C-AT-G-CCC-T-A-    207
G_LBV217      TA-CC-TG----C----------C------T------C----AC-GCA-CAG--GC-A----CTC-C-A-    207
H_VI557       CA-CC-CGA---T-----C---GC------C----T-C--T-AA-AGA-CAGC--C-A---GCTC-TAA-    207
O_ANT70C      TA-TC-TGAG--AC-------TGC---G--TAA-G-G--GC-GC-ACAGCAG---G-G---GCTC-CAA-    207
CPZGAB        TG-CC-CG-G--AA-G----GT--G-----A---ACT---T-GT-ACA-CAA---G-G---GCTC-CAAA    207
CPZANT        CAGCT-CTC---TC--------TC------T---G----GGCTA-CCATCAA--GAGC--T-CCA-AGAA    207
A_ROD         AGCA-AGA----G--G--GT-------G--T-----------TC-TAC---T---G-T---A-G--A---    207
B_EHOA        GGCA-AGA---GG--G-G-T----------A--C-GG-----TCGGAA---C---GGA-----A--A--A    207
C_2238        AGCA-AAA----G--G--G--C--------A--C--------TT-AAC---G---G-A-----A--T---     97
D_FO784       AGCA-AAA----G--G----AT--------G-----------TC-CTCG--C---GCT-----A--A---     97
SD_MM251      AGCA-AAA----G--G--G-AC--------A-----------AC-TTCG--C---GCT-----A--G--A    207
STM_STM       GGCA-AAA----G--G--GT----------G-----------TA-AACG--C---G-A-----A--G--A    207
VER_AGM3      CC-T-AGA-AT-AC-----T--G-------A---A-G--G--CA-AGA---AC-CT-C--GC-A-AA--A    207
GRI_AGM677    AC-T-AGAAA--T-----------------C-----------CA-AGA---T---ACC--G--G-AA---    207
TAN_AGM17     CC-T-AGAAA--T--------------G--A-----------CA-AGA---A---GCC--T--A-AA--A    207
SAB_SAB1C     TTCA--AAAT--C-----G-----G-----GGTCGTC-----TC-TAGT--AC-CCTT--CC-A--T--T    207
SYK_SYK       TTCG-ACCAA--CA-G--GT-T--------C---G-------TC-CAC---CC--CTT-----A-AAG-T    210
```

```
GAG_CONSENSUS AcagG?tCagaaga??Taaaa?c?cT?Tttaatac?gt?tg?gtc?Tctggtgt?TaCAc??agaa?a?a    234
A_U455        -----AA-------AC-T-G-T-AT-A-A------A--AGCA---C---AT---G----TCA-AGGATAG    277
B_HXB2R       -----A--------AC-T-G-T-AT-A-A------A--AGCAAC-C---AT---G-G--TCA-AGGATAG    277
C_UG268       -----GA----T--AC-T-G-T-AT-A-A---C--A--AGCAACTC---AT---G----TAA--GGATAG    277
D_ELI         -----AA-------AC-T-G-T-AT-A-A------A--AGCAAC-C---AT---G----TAA--G-ATAG    277
F_BZ163B      -----A--------GC-T-G-T-AT-A-A------A--AGCA---C---AT-T-G----TCA-AG-GTAG    277
G_LBV217      -----AA----G--GC-T---T-AT-A------G-A--AGCA--TC--------A----TCA-AGGATAG    277
H_VI557       -----AA-------AC-TC--T-AT-A------CTAT-AGCAAC-C---AT--CG----TCA-AG-ATAG    277
O_ANT70C      -----G-----CAGCC-GC-GT-A--C-GG--CG-AA-AGTA--GC--------G-T---AACAG-T-T-    277
CPZGAB        -----C-------GAC-GCGGT-CT-G-----C--TC-GGCA--AC-G-----CA----TAGT--CATC-    277
CPZANT        -T-A-A--CCCT--AA---T-T-TT-G-----C--CA-T--T--TC-G-----CG----TAA--GGG-A-    277
A_ROD         -----T------A-TT-----AGT--T--------T--C--C---A-T-----CA-----GC----G-G-    277
B_EHOA        -----T------A-TT-----AGT--T-A------C--C--C---A---TT--CC-G---GC----G-G-    277
C_2238        -----T------A-TT----GAGC--T-A------TACC--C---A-T------T-----GC----C-G-    167
D_FO784       -----T------A-TT-----AGC--T-AC-----T--C--C---A--------T-----GC----G-G-    167
SD_MM251      -----C------A-TT-----AGC--T-A------T--C--C---A------CA-T---GC----G-G-    277
STM_STM       -----T------A-TT----GAGT--T--------T--C--C---A-T------A-C---GC----G-G-    277
VER_AGM3      -----G--G--G-GCT-----AGT--G-----CCTT--G--C--AT-G-TT--CG-----AA---TA-GG    277
GRI_AGM677    -----C--C--G-GGC-----G-T--G------TTGTGC--C---A-T-----CA-T---GC----C-G-    277
TAN_AGM17     -----A--G----GCT-G---AGT--G------TTGTGC--C---A-T------A-T---GC----C-G-    277
SAB_SAB1C     -----G-----GA-TT---TTG-T--C------TTATGC--C--GT-AGCT---A-T---GC----ATA-    277
SYK_SYK       -AT--G------A-TT-----AGC--T---GGC-TAA-T-CG---G-A---GCCG-T--TGCCA--A-GG    280

                          INS-1 ->
GAG_CONSENSUS aagT?aaagAcAC?gAggaagc??taga?aaa?????????????????????tagag????????????ga    268
A_U455        -T--A--------CA-------TT--A-T---..................A-----............--    317
B_HXB2R       -GA-A--------CA-------TT----C--G..................A-----............--    317
C_UG268       ----ACG------CA-------CT----C--G..................A-----............--    317
D_ELI         -T--A--------CA-------TT----A--G..................A-G---............--    317
F_BZ163B      -G--A--G-----CA-------TT----C--G..................C-----............--    317
G_LBV217      GG--A--------CA-A-----TC----GG--...................G-G--A............A-    317
H_VI557       -T--A--------CA-------TT----G--G..................T--A--............--    317
O_ANT70C      --A-TGG---T--GC--C-G--AA--C-A--G..................T--A--............--    317
CPZGAB        CT--AG-------AC--A----TC----AC-GCTAAAG......CGGCATCAT-GA............--    329
CPZANT        -GA-A--------A--AC----CG-TA-A-C-............GTGAAAA-GA-A............-T    323
A_ROD         ----G-----T--T--A-G---AAA-C-A-T-GTGCGG......AGACATC---T............-C    329
B_EHOA        ----G-----T--T--------AAA-.....G..................A---CACAGCGA...CAT..    321
C_2238        ----G---C-T--A------CGAAAC--GGT-GTAGAG......AGACATC---C............-C    219
D_FO784       ----G--------T----G---AAA-C-G-T-GTGCAG......AGACATC---T............-T    219
SD_MM251      ----G---C----T--------AAA-C-G-T-GTGCAG......AGACACC---T............-T    329
STM_STM       ----G---C-T--T--------AAAGC-AGTGGTAAAG......AGACATC-T-T............-T    329
VER_AGM3      ----G--------A--A-----AG---CA-T-GTAAGACAATGCTGCCATC---T............--    335
GRI_AGM677    ----G--------A--------TG---TA-C-...................G-TA--CAACACTACCAT..    327
TAN_AGM17     ----G--------A--------TG---TA-T-GTAAAACAACACTGCCATC---T............--    335
SAB_SAB1C     ----G-----T--A--------AAA--CG---GTAAAA......GAGGAAG-GCCA............-C    329
SYK_SYK       ----AG-------T---C----AAA-C-G---...................G--A-A............--    320

GAG_CONSENSUS a?????????????????????????????????????????????????????????????????????    269
A_U455        -.................................................................    318
B_HXB2R       -.................................................................    318
C_UG268       -.................................................................    318
D_ELI         -.................................................................    318
F_BZ163B      -.................................................................    318
G_LBV217      G................................................................    318
H_VI557       G................................................................    318
O_ANT70C      -.................................................................    318
CPZGAB        -.................................................................    330
CPZANT        -.................................................................    324
A_ROD         -.................................................................    330
B_EHOA        .................................................................    321
C_2238        -.................................................................    220
D_FO784       G................................................................    220
SD_MM251      G................................................................    330
STM_STM       -.................................................................    330
VER_AGM3      GAAAGAAAGA.......................................................    345
GRI_AGM677    .................................................................    327
TAN_AGM17     -.................................................................    336
SAB_SAB1C     -.................................................................    330
SYK_SYK       -GCATGTAATTGGAAAGATGACCCACCAGCGACATCTGGTGGACAAAGTGAAAATAGCAGTCAAAACATG    390
```

HIV1 GAG

```
GAG_CONSENSUS ??????gaa??aaa?aa?a??gag???a?a??????????????????????????????a??ca??    286
A_U455        ......ATGCA---T--G-ACA--CAA-GG.................................-CA--GC    349
B_HXB2R       .......--GCA---C--A-GTA--AAA-A-................................GCA--GC    349
C_UG268       .........---CA---C--G-TTC--CAA-A-..............................-CA....    345
D_ELI         .......--GCA---C--A-GTA--AAA-AG................................GCA--GC    349
F_BZ163B      .........---CA---C--A-GTC--CAA-AG..............................-CAA-GC    349
G_LBV217      ......AGGCA---G--C-GTC--CAA-A-................................-TA--GC    349
H_VI557       ......-C-CA---C--A-GCC--AAC-GG.................................-CA--GC    349
O_ANT70C      ......-T-ATGGGG-GC-GGA--...TCT................................GCGG-CG    346
CPZGAB        ......C--CAG-GC--A-CT--A...-GTAACTCAGGAAGCCGT......GAAGGGGGAGCC-GT--AG    385
CPZANT        ......ATGCAG-CAC-AGCA--A...-C-GGA............................AGT-GC--AA    358
A_ROD         .......---AC-GGA-CTGCA---...-A-................................-TG-CAA    358
B_EHOA        ......CT-GC-GCGG-C-CA--A...-A-................................-TG-CAG    349
C_2238        .........---ACT--A--TGCA--A...-A-..............................-TG-CAG    248
D_FO784       .........---AC-GGA-CTGCA--C...-A-..............................-TG-CAT    248
SD_MM251      .........---AC-GGA-CAGCA--A...-CT..............................-TG-CAA    358
STM_STM       .......--GACTGGA-CTGCAA-C...-A-................................-TG-CTG    358
VER_AGM3      AATGCA---AG---T-CA-CA---ACATCT................................-GTGGAC    382
GRI_AGM677    ......CT-GTGG-C--A-AT---AAAGC-................................GCTA-AA    358
TAN_AGM17     ......A--GA---A-CTGCA-CA...-CGCCATCTGGTGGC............CAAAAGCAA-ATT-CA    385
SAB_SAB1C     .........---ATG-CAG-A-GT-CC...-C-GCGACATCTAGTGGCCAAACTAAGGAACTGCAGGCAA-AA    391
SYK_SYK       GCTAGT--GAC-TCT-GTGGCC-AAAGGT-................................GTA--GC    427

GAG_CONSENSUS a?gca??????a??g???c?acagga?????????????c?????????????????????????????    300
A_U455        -G---.......GCA-CTAAC------.............................................    369
B_HXB2R       -A---.......GCA-CTGAC------CAC.........AGC..............................    375
C_UG268       ..-A-......-CA-CTGAC-A----...............................................    363
D_ELI         -A---.......GCA-CTGAC------AAC.........AAC..............................    375
F_BZ163B      -A--G......GCA-CTGAC-A---G..............................................    369
G_LBV217      -G---............GAT-A----GAC.........AAC..............................    369
H_VI557       -G---......-CA-GTGAT-A----AAT.........GGC...............................    375
O_ANT70C      CC--T......-AG-AAGAC---A-C............G-A..............................    369
CPZGAB        GC--T......-GT-CCT-TG-T--C..............................................    405
CPZANT        CC---......-GCAGAGGC-TGCTTCTG.........CGGCTGCTCCTGTTAAACAAACAGTGGTGTCA    413
A_ROD         GCA--......-GTAGAC-A----C-CCA.........T-T...............................    384
B_EHOA        CTATG......-GTAAAC-A-GTAA-CCA.........A-T...............................    375
C_2238        CCA--......-GTAGAC-A----CTCCA.........T-T...............................    274
D_FO784       CAA-G......-GTAGAC-A----C-CCA.........C-T...............................    274
SD_MM251      -AA--......-GTAGAC-A----C-CCA.........T-T...............................    384
STM_STM       CCA--......-GTAGAC-A----C-CCA.........C-T...............................    384
VER_AGM3      -AAAG......-AAAATGAC-AG---...GTAACAGTGC-A...............................    414
GRI_AGM677    -GAA-......-AT-AGA-A----CGCCA.........C-T...............................    384
TAN_AGM17     -CA--......GCT-CGC-AC-T--C..............................................    405
SAB_SAB1C     -GAA-......-AT-AGC-A----TGACA.........C-T...............................    417
SYK_SYK       -G-A-AAACAA-AG-CAG-A---CCGCCA.........C-T...............................    459

                                          p17 \/ p24
GAG_CONSENSUS ??????????ag????g???g?cgAaatTacCC??t?gt?ca?aatgt??a?gg?ca?t?ggt?CAccag    340
A_U455        ...........--C......A-T-A---------CA-A--G--A----CAC-A--G--ACCA--A------    423
B_HXB2R       ..........-ATCAG-TCA-C-A---------TA-A--G--G--CA-CC-G--G--AAT---A--T---    435
C_UG268       ............AAG-TCA-T-A------T--TA-A--G--G---C-CC-A--G--AAT---A------    420
D_ELI         ...........--CCAG-TCA-C-A------T--TA-A--G--G--CC-AC-G--G--AAT---A--T---    435
F_BZ163B      ................-TCA-T-A---------TA-A--A--G---C-TC-G--A--AAT---A--T---    423
G_LBV217      ...........--CCAA-TCA-C-A------T--TA-A--G--G----CAC-A--G--AAT---A------    429
H_VI557       ..........-ACAAGATCA-T-A------T--TA-A--A--A----CCC-A--A--GCC---A------    435
O_ANT70C      ...........--GCAG-CGG-T-A---------TA-A--ATCA----CGC-G--A--AAT---A--T---    429
CPZGAB        ................ATTA-TG----------CC-C--A--A----CTC-A--G--GAT---G--T---    459
CPZANT        GCGACATCTT--TGGC-AAG-GA----------CA-CA-AGTGG---CAGGA--AATAGCAAGG--T---    483
A_ROD         ...........--CGAGAAGG-AG----------A...--G--AC----AGGC--CA-C-ACACC--TATA    441
B_EHOA        ...........--C......C-C-T-GC---T--A...--G--GC-AA-AGCT--CA-T-ATTCC----T-    426
C_2238        ...........--TGGC-GCA-AG----------A...--G--GC-A--AGCA--CA-T-AT--C-----T    331
D_FO784       ...........--TGACAGAG-AA----------A...--G--GC-A--CGGC--AA-T-AT--C----TA    331
SD_MM251      ...........--CGGCAGAG-AG----------A...--A--AC-AA-AGGT--TA-C-AT--C----T-    441
STM_STM       ...........--TGGCAGAG-AG----------A...--A--GC-A--AGGC--CA-T-AT--C----TA    441
VER_AGM3      ..........CCTGGT-GCA-T-A-----T---A...-CA--AC-ACAGGGAAATGCA-G-A-A--TGT-    471
GRI_AGM677    ..........G-TGGC-AATCAA----------A...--AGTA---CAGA-TAATGCC-G---A------    441
TAN_AGM17     ............GGTAATCATG----------T...--AGTAC-ACAGA-TAAC--G-G---G--TACA    459
SAB_SAB1C     ............--TGGT-GCTCG-----------CA-T--C...-G---TA-TAAT--A-G---C------    474
SYK_SYK       ............--A.........G----C--T--AC-GC-CAGA---CCTC-AAAT--G-G-A-T--TACA    510
```

```
GAG_CONSENSUS cc??Tatcacc?aGaAC?tTaAAtgCaTGGGTaaAagtagTagAagaaAAga??Ttcag?gCagaagTag    403
A_U455        G-CT-------T--G--C--G-----------G--------------C---GCT-----CC--------A    493
B_HXB2R       G-CA-------T-----T-----------------------------G---GCT-----CC-------GA    505
C_UG268       G-CC-------T-----T--G------------------A----G------GCT-----CC----G---A    490
D_ELI         G-CA-------T-----T--G--C---------------A-----------GCT-----CC--------A    505
F_BZ163B      T-CA-------T-----T-----------------G--GA-------G---GCT--T--CC--------A    493
G_LBV217      --CA-------T--G--T-----------------------------------GCC-----TC--------A    499
H_VI557       G-CA-------T--G--C-----------------------------G---GCT--T--CC--------A    505
O_ANT70C      G-CA-C--C--C--G--T-----------------G-C------------AGCC--T-ACC-T---A-CA    499
CPZGAB        G-CA-C-----C-----C--G-----------------G-----G------GCA--T--CC-T--G---A    529
CPZANT        --AC-GA----A-----C--------C---------GTGT--------G--A-AT----ATC--------CA    553
A_ROD         --GC-GAGT--CC----CC-------C---------T-------G-----A-AG---G-G----------    511
B_EHOA        --GC--AGC--CC----C--------C---------T----------G----AG---G-G----------    496
C_2238        --TC--AGT--T-----AC-G--------------GT---------------AG--TG-G-----G----    401
D_FO784       --GC-GAGC--C-----A--------------T--GT----G--G--C--A-AA---G-G-----G----    401
SD_MM251      --AT--AGC--G-----A--------C---------T-GA----G-------AA--TG-A----------    511
STM_STM       --AT--AGC--A-----C--G--------------GT----G--G-------AA--TG-T--------G-    511
VER_AGM3      --CT-G-----AC-C--C--------G----------C------G--G--A-AA---G-A------A---    541
GRI_AGM677    --TT-G--T--GC-C--G--------G-----C---TGC--G--G-----A-GG-GGG-A----------    511
TAN_AGM17     --AC-G-----AC-C--CC-G--------------GAC---T--------A-GG--TG-A------A---    529
SAB_SAB1C     --AT-G--C--A--G--AC-T-----T-----------GA-T--G-------AG-----T-----G----    544
SYK_SYK       GGAG-CC--GTT--G--TC----AA-------GG-G-CT--GA-CTCC--A-AA--TGAT---TCCA---    580

GAG_CONSENSUS TaCCcatgTTt???GCacTatCaGAaGGagccac?cc?tAtGAt?T?AAt???aTGcTaAAt??agTgGG    461
A_U455        -----------TCA---T-------G--------C--AC-A---T-A---ATG-----G---GT------    563
B_HXB2R       -----------TCA---T----------------C--AC-A---T-A--CACC--------CAC------    575
C_UG268       -----------ACA---T----------------C--AC-A---T-A--CACC---------AC------    560
D_ELI         -----------TCA---T----------------C--AC-A---T-A--CACC--------CAC------    575
F_BZ163B      -----------TCA---T----------G-----T--AC-A---T-A---ACC---T-----AC------    563
G_LBV217      -----------TCA---T----------------T--AC-A---C-A---ACC---------AC------    569
H_VI557       -----------TCA---T----------------C--AC-A---T-A---GCC---T-----AT------    575
O_ANT70C      -T--T-----CATG---T-G-----G-----A-TTT-C------A-T---ACT---------GCCA-A--    569
CPZGAB        ----A------TCA--T--------G------TTA--TC-G---G-T--CACC---T-----GC---A--    599
CPZANT        -C--T------TCT--TT----------G--A--T--TC-----T-A--CACC-----T---GC---T--    623
A_ROD         -G--AGGA---CAG-----C--------CTG---G--C------A-C--CCAA-----T---TGT-----    581
B_EHOA        ----AGGA---CAG--------------GTG---C--T------A-T---CAG---T-----TGT--A--    566
C_2238        ----AGGA---CAG-----GA-------GTG---A--C------A-T---CAAT--------TGT--A--    471
D_FO784       -G--AGG----CAG--------G-----CTG---T--G-----CA-C---CAG---------TGT--A--    471
SD_MM251      -G--AGGA---CAG-----G--------CTG---C--C-----CA-T---CAG---T-----TGT-----    581
STM_STM       -G--AGGA---CAG--------------CTGT--C--C-----CA-C---CAAC--T-----TGT-----    581
VER_AGM3      -G--------CCAG--TT----------GTG---A--C-----CA-C---CAA-----T---GTCC----    611
GRI_AGM677    -C--------CCAA-----C-----G---TGTCTCT-C------G-A---CAG-----C---GT-A-A--    581
TAN_AGM17     -G---------CAG--T--C--------CTG-CTTAGC-----CA-T---CAA---T-----GTCA-A--    599
SAB_SAB1C     -----------AGC--C---G----------T-TC--T------A-C---CAA-----C---GC---A--    614
SYK_SYK       -G--AT-A---CAG--T---A--------TTT-TT--T------C-C---GGC-----T---GCT--A--    650

GAG_CONSENSUS ?GaaCAtCAaGcaGCcaTgCAaaT??TaAaaGA?at?ATcAATGAgGAaGc?Gcaga?TgGGAtagg??a    522
A_U455        G-G---C--G-----T--------GT-------T-CC--------------T-----G-----C---TT-    633
B_HXB2R       G-G---------------------GT-------G-CC--------------T-----A--------AGTG    645
C_UG268       G-G---------------------GT-------C-CC-----------G--T-----A---------TT-    630
D_ELI         G-G---------------------GC-------G-CC--------A-----T-----A---------TT-    645
F_BZ163B      G-G---------------------GC-------C-CC--------------T-----A-----C--ATT-    633
G_LBV217      G-G------------T--------GC----G--T-CT--------------T-----G---------AT-    639
H_VI557       G-G----------------------GT-------T-CA--------------T-----A---------GT-    645
O_ANT70C      A-G--------G---TT-A---G-GC----G--AG-A--------------A-T---G--------AACT    639
CPZGAB        G-G--------G----------G-GC----G--AG-C--T-----------C--C--G------C-CTT-    669
CPZANT        G--C-------G---------GG-GC-------AG-A--------------A--T--G---------TT-    693
A_ROD         C--C-----------------G--AA-C-GG--G--T--------------A-----A------GT-CA-    651
B_EHOA        A--------G--------------TA-T-GG--G--A--T------------A-----C-----CCAACAG    636
C_2238        A--T-----G--T--T-----G--TA-T-G---A--T--------------A--C--T------GAACAC    541
D_FO784       A--------G-----------G--CA---GG--A--A--------T--G--A-----T------CA-CAG    541
SD_MM251      A--C--------G--T-----G--TA-C-G---T--T--A--------G--T-----T-----CTT-CAG    651
STM_STM       G--G---------------------TA-------G--C--A------------T--T--T-----C-T-CAG    651
VER_AGM3      A--C-------GG--GC-A-----AG-------A--C--------------A--CC-G-------TAGCT    681
GRI_AGM677    A--C-----G-GG--AT-A-----TC-T--G--AG-C--T-----A-----A-----G-----C---AC-    651
TAN_AGM17     A--C--C----G---G--------TA-T-----GG-A--T-----A--G--A--CC-G-A---CCTAACC    669
SAB_SAB1C     A-----C--G-GG--AC-A-----AG-G-----TG-T--------A-----G-----T-----CCTTAG-    684
SYK_SYK       A--T-----G-G---------GG-AA-T-----C--C-------------GG--G-CA------CTTGA-    720
```

HIV1 GAG

```
GAG_CONSENSUS CAtcca???ca?gcagg?CC?????t?ccaccaGg?cag?T?AGaGA?CCaaggGgatctGAtATaGCaG    577
A_U455        ------GTG--T-----G--T...A-T--------C---A-G-----A---------AG---C-------    700
B_HXB2R       ------GTG--T-----G--T...A-TG-------C---A-G-----A---------AG---C-------    712
C_UG268       ------GTG--T-----G--T...G-TG-------C--AA-G-----A---------AG---C-------    697
D_ELI         ------GTG--T-----G--T...A-TG-------C---A-G-----A---------AG-----------    712
F_BZ163B      ------GTG--G-----A--T...A-C--------C---A-A--G--A--T------AG---------T-    700
G_LBV217      ------CAA--G-----G--T...A-T--------C---A-A-----C---------AG-----------    706
H_VI557       ------GTG--T-----G--T...A-T--------C---A-G-----A---------AG-----------    712
O_ANT70C      --C---CCA-CG-T---G--G...T-G--------G---A-A--G--A----CA---AG---C--T--T-    706
CPZGAB        -----CACT--T--T--T--G...A-AG-C-----A---C-A--G--G--------TAG-----------    736
CPZANT        --C--CACT--T-----A--A...G-A-AGG----A--AT-A--G--A----CA---AG-----------    760
A_ROD         ------...ATAC----C--C...T-A---G-G--G---C-T-----G--------------C-------    715
B_EHOA        ------...TCGC----C--A...A-G---G----G--AC-C--G--A-----A--G--A--C-----G-    700
C_2238        --C---GTA...C----C--A...C-G---G----A---C-T--G--T--------C-----C--T----    605
D_FO784       ------...--AC----C--A...C-A---G----A---C-C-----T---C-A---------------    605
SD_MM251      --C---...--AC---CT--ACAA......-A---A---C-T--G--G--GTCA-----A-----T----    712
STM_STM       ------CAA-CAC-G--T--G...C-G---G----G--AC-T-----A-----C--G--A----------    718
VER_AGM3      --C---CCA-CA-----A--A...T-A---G----A--AC-C-----C--T--A--C-----C-------    748
GRI_AGM677    --CAG-CCA-CA--T--C--G...T-A---G----G---C-A-----C--G-CA--G--A----------    718
TAN_AGM17     --C--TCCA-CA-----A--A...T-A---G----A---T-A-----T--------T--A----------    736
SAB_SAB1C     --C--GCCG-CGCA-CAG--C...CCAG-T-AG--AGT-C-A--G--C---CAA--G--A--C-----G-    751
SYK_SYK       -----CCAA-CGCA-CAA--T...GCT----A--CAGGAT-G--G--T--CTCC-CC--A--C-------    787

GAG_CONSENSUS GaAC?ACtAGcaCa?t?ca?GAaCAaaTa?aaTGGAtgac??????????a?aatcc?cctaTcccaGT    629
A_U455        ----T--------CG-T--A---------GG---------A.........GGC-----A-----------    761
B_HXB2R       ----T-----T--CC-T--G---------GG---------A.........A-T-----A-----------    773
C_UG268       ----T-----T--CC-T--A---------A-T--------A.........AGT--C--A-----T-----    758
D_ELI         ----T-----T--CC-T--G---------GC---------A.........AGT--C--A-----------    773
F_BZ163B      ----T-----T--CC-T--G---------C----------A.........AGC--C--A---G-------    761
G_LBV217      -G--T-----T--CC-G--G---------AG---------C.........AGC-----A-----------    767
H_VI557       ----T-----T--CC-T--G---------GC---------A.........AGC-----C-----------    773
O_ANT70C      -G--A--------CCAG--A--G-----TC-C----CT--CAGG...CCCA-CC-A...-----------    770
CPZGAB        -G--C--C------C-G--G--------TGGG----CA--A.........GCA-----T--C--------    797
CPZANT        -G--A--A------G-G--G--G--G--GC--------T-AACACCTCAAC-G---GGAGGAG-------    830
A_ROD         -G--A--A------G-AG-A-----G--CC-G------TTTAGG...CCAC-A---...---G-A-----    779
B_EHOA        ----C--C------G-AG-A-----G---C-G------TACAGA...CCCC-A---...---G-------    764
C_2238        ----A--C-----GG-AG-A-----G---C-G------TATAGA...GCAC-G--C...--AG-A--G--    669
D_FO784       ----C---------G-GG-G---------C-G------TACAGG...CAGC-G---...--C--A-----    669
SD_MM251      ----A-----TT--G-AG-T--------CC-G------TACAGA...CAAC-G--C...--C--A-----    776
STM_STM       -G--T------T--CCAG-G--------TC--------TACCGG...CAGC-A---...-----A--G--    782
VER_AGM3      ----C--C------G-G--A-----GC-GG-------ATAC......ACAGCC-----CAGAG-AGAT--    812
GRI_AGM677    ----T--C---T--A-T--G---------G-G----CCTTC......AATGCC-----AAGA--AGAC--    782
TAN_AGM17     ----C------T--G-G--A--------TG-G----CCTTC......AATGCA-C---TAAGG-AGAT--    800
SAB_SAB1C     -C--T---------A-A--A--G------G------CC--CAGA...GCAC-G---...G--G--AAT--    815
SYK_SYK       ----A--A---T--A-TGCT--G------G-------T--TAGG...CAGA-C--C...--AG-A-A---    851

GAG_CONSENSUS aGGagacatcTAtagaagaTgGaT?at?cTggG?tTacA?AAa?t?GT?agaatgTAtaacCC?gt?a?c    689
A_U455        G-----------------------A--C-----A---A-T---A-A--A----------G---T--T-G-    831
B_HXB2R       ------A--T----A---------A--C-----A---A-T---A-A--A----------G---TACC-G-    843
C_UG268       G-----A-------A---------A--T-----G---A-T---A-A--A----------G---T--C-G-    828
D_ELI         ------A-------A---------A--TG----A---A-T---A-A--A----------G---T--C-G-    843
F_BZ163B      G-----A-------A---------C--C--A--A---A-T---A-A--A----------G---T--C-G-    831
G_LBV217      G-----A--T----A---------A--C-----G---A-T---A-A--A----------G---T--C-G-    837
H_VI557       G-------------A---------A--C-----A---A-T---A-A--A----------G---T--C-G-    843
O_ANT70C      ------------------A------AG-GT-A--AC--A-C---A-G--A-A------C-G---A--G-G-    840
CPZGAB        ---G--TG-T------------G-A--TT-A--G---A-C---G-G--T---------TG---A--A-GT    867
CPZANT        ---G----------AG--------C--CA----A---A-T--GG-G--C--G-N-----GT--A--C-G-    900
A_ROD         ----A-------------------CCAGA-A--A--G--G--GTGT--C--G-----C-----GACC-A-    849
B_EHOA        G---A-------------G-----TCAGT-A--GC----G---TGT--CC-G-----C--T--TACT-AT    834
C_2238        ---CA--------C-G-------CCAA--A--CC-G--A--GTGT--C-----------T--AACT-A-    739
D_FO784       T---A-T--------G--------CCAG--A--G-----G---TGT--C--------C-----AACT-A-    739
SD_MM251      ---CA----T--C--G--------CCAA-----G--G--A---TGT--C--------------AACA-A-    846
STM_STM       ----A-------C--G--------TCAA--A--GC----A---TGT--C--------C--T--A--T-A-    852
VER_AGM3      G--T-C------------G-----T--C-----G--G--A---TGT--A-A------C-----A--GTCT    882
GRI_AGM677    ---G-CACAA--C----A----G-T--TT----C-----A--GG-A--GCAG-----C--T--CCAA-AG    852
TAN_AGM17     T---CG---T--CC--G---A-G-T--C---A-GC-C--A--GTGT--A-A------------TA-CTCA    870
SAB_SAB1C     ----A-T--T----A-G-------C--C-----TC-C--G---TGT--C-A------C-----A--A-A-    885
SYK_SYK       ------A--A--------------C--T--A--GC-G--G--GTGT--GCAGG-C--C-----T--TTCT    921
```

```
GAG_CONSENSUS aTttTaGAcaTaaaaCAaGGgcCaAAaGaacc?TTcaaagacTAtGTaGAtaGgTTCTataAaac?tTaa    757
A_U455        -----G-------G------------------C----GG--T-----------A----T------TC-C-    901
B_HXB2R       ---C-G-------G------A-----G-----C--T-G------------CC-------------TC---    913
C_UG268       -----G--------------------G-----C--T-G-------------C------T------C----    898
D_ELI         -----G-------G---G--A-----G-----T--T-G------------CC-------------TC---    913
F_BZ163B      -----G-------G----------G--G--C--T-G------------C-------T------CC---    901
G_LBV217      -----G-------G------------------C--T-G---T----------------T------T--G-    907
H_VI557       -----G-------G------------------C--T-G--------------------T----G-T----    913
O_ANT70C      --C-----T--T--G--G--A-----G-----A--T-G-------------C-------C-----A----    910
CPZGAB        --CC-T-----C-G------------G--G--C--T-G---T-----G--C--A-----C-----T----    937
CPZANT        ---C----G-----------A--------G--C----G---T-----G-----A-----------AA-T-    970
A_ROD         --CC-------------G--A-----G--G--G---C--AG------------A-----C----GC--G-    919
B_EHOA        --AC-G--------G------------------C---C--AG------------A-----C----GC---C    904
C_2238        ---C-----G-G-----G--T-----------A--TC-GAGT                              781
D_FO784       ---C-G--T--------------------GA-G---C-GAG-                              781
SD_MM251      ---C----TG-------------------G--A--TC-GAG---------C--------C----GC----    916
STM_STM       ---------G----G---------------A-G---C--AGT--------C--------C--G-GT--G-    922
VER_AGM3      G-C----------G---G-----C------G-A-----------C--------------C---G-AA---    952
GRI_AGM677    G-CC-------TCG---G--A--T--------C---C-G-----------C--A---------G-CC-G-    922
TAN_AGM17     G-------T--C-G------AG-C----G---A---------------------------C--G-CC-T-    940
SAB_SAB1C     -----------T-----------T--------T-----G--T-----------A-----C--GG-TC-G-    955
SYK_SYK       --C--------------G--C---------ATT-----------------C--A-----CC-TTGC----    991

GAG_CONSENSUS Gagc?GAgCAagCt?cacagga?GTaAAaaatTgGaTgACagaaaC?tTgtTggTcCAaAATGC?AAcCC    822
A_U455        ----T---------A-------T--------C--------------C---C-------------G--T--    971
B_HXB2R       ----C---------T-------G----------------------C-----------------G-----    983
C_UG268       ----T--A------A----A--T--------------------C--C-----------------G-----    968
D_ELI         ----C---------T-------T-----------------------C-----------------A-----    983
F_BZ163B      ----T---------A-------A-----GGG------------C--C-----------------G-----    971
G_LBV217      ----T---------A-------T--------C-----------C--T-----------------G-----    977
H_VI557       ----T--A------A-------G--G---GG------------C--C-----------------G--T--    983
O_ANT70C      ----T---------A-T--A--A-----------------------TC--C-T--T--------C--T--    980
CPZGAB        -G--T---------T-------A--G--------------T--C--TC-C--A--G--G-----A-----   1007
CPZANT        ----A--A--G---T-----CCT--G---GCC--------------C-----AA----------C--T--   1040
A_ROD         -G--A--A---A-AGAT-CA-CA--G--G-----------CC----AC--C-A--A--------C-----    989
B_EHOA        -G--A------A-AGAC-CA-CA------------------C----AC--C--A-T--G-----T-----    974
SD_MM251      ----A--A---A-AGATGCA-CA-----G-----------TC----AC--C--A-T--------T-----    986
STM_STM       ----A--A-----AGAT-CATCA--G--------------TCGG--AC--C-AA-T--------T--T--    992
VER_AGM3      ----T-----G--CT--GGA--A------C-G------------T-A--AC-CA-T--G-----T--T--   1022
GRI_AGM677    ----A--A-----AC-------T--T---------------C----T---C-TA----G-----C--T--    992
TAN_AGM17     -GCGC--A---A--C-C-----C----------A------------TC--C-CA----G-----T--T--   1010
SAB_SAB1C     ----A-----GA-AGAC-CT-CT-----G--C---------C-GT-TC-----A----G-----C-----   1025
SYK_SYK       ----A-----G--GGAC-CATCA--C---GGG---C-A--TC----AC-----A-T--------C-----   1061

GAG_CONSENSUS aGAttgtaag???aT?tTaAaagcatTggga????a?gCtaC?tTaGAaGAaATG?TgacaGC?TGtCAg    880
A_U455        ---C------TCC--T----G------A--GCCAGGG-----A------------A-------A--C---   1041
B_HXB2R       ----------ACT--T---------------CCAGCG-----AC-----------A-------A------   1053
C_UG268       ----------ACC--T----G------A--GCCAGGA---T-A------------A-------A------   1038
D_ELI         ------C---ACT--C---------------CCAC-G-----AC-----------A-------A------   1053
F_BZ163B      -------C--ACC--T--------T------CCAGGG-----A                              1014
G_LBV217      ----------ACC--C-----------A---CCAGGA-----AC-----------A-------A--C---   1047
H_VI557       ----------ACT--T----G------A---CAAGGG---T-AA-----------A-------C--C---   1053
O_ANT70C      ------C--ACAG--T--G---T----A--GCCAGGA-----C--------G---A--GT---C------   1050
CPZGAB        ----------CAA--TC--------------CCTGGG--C--CC----G------A-------T-----A   1077
CPZANT        ------C--ACAC--CC-G--G--T------ACAGGA--CT-C------------T-A-----T-----A   1110
A_ROD         ---C-----ATTAG-GC------G-C-A--GATGA-CC----C--------G---C----C--C------   1059
B_EHOA        ------C--ATTAG-GC-T--G-GC------ATGA-CC-C--A------------C-A-----C--C---   1044
SD_MM251      ------C---CTAG-GC-G--G-GGC----TGTGA-TC-C--CC-----------C----G--T-----A   1056
STM_STM       ---C--C--ATTGG-A--G----GTC----CATGA-TC-C--CC-------G---T-A-----A-----A   1062
VER_AGM3      ---C-----AGTC--CC----G-GCC-----ATGC-TC-C--TC-----------T-A--T--C-----A   1092
GRI_AGM677    G--------ATTG--TC-G----G-------ATGA-TC-A--C--G--G------C-A-T---T--C---   1062
TAN_AGM17     ----CAC--ATTAG-A--G---AG-C-A--GATGC-TC-A--A------------C-A-----A-----A   1080
SAB_SAB1C     ----------ACAG-GC-C----GG--A--CATGA-TC-A--CC-------G---C-T--T--A--C--A   1095
SYK_SYK       ---A--C-GACAG--CC-C--G---A--...GTCA-AC----C--G--G------C--CA---T-----A   1128
```

HIV1 GAG

```
                                                  p24 \/
GAG_CONSENSUS GGagTaGGaGGacCagg?cAtAA?GCaagagT?aTgGc?GAagc?aTga?a?a?gc????caa?at????    933
A_U455        -----G--------C--C-----A-----G--TT----T--G--A----GTC-A-TA..----C-G....   1104
B_HXB2R       --------------C--C-----G--------TT----T-----A----GCC-A-TA...AC-A--TCA.   1119
C_UG268       -----G--------T--C--C--A--------GT----T--G--A----GCC-A--A...A-CA--....   1101
D_ELI         -----G--G--G--CA-C-----A--------TC----T--G--A----GCC-A--A...AC-A--TCAG   1120
G_LBV217      -----G--------CA-C-----A--------TT-A--T-----G----GCC-G--A...TC-GG-ACAG   1114
H_VI557       -----G-----G--T--C-----A----A---TT----T--G--A----GCC-A-TAAATAC-A--....   1119
O_ANT70C      ---------------ACT-----G--C-----GC-A--A-----A---GCTACA--C...--GC-AGATT   1117
CPZGAB        -------------TCT-----G-----G--TC-A--T-----C---TC-ATG-TT...--GA--CAAG   1144
CPZANT        -----------C----CC-----G--------GT----A-----T---GCTTCT--T...A-TA--GCAC   1177
A_ROD         --G-----T--G-----C--G--A--T---T-A-----A--G--CC---A-G-G-TCATAGG-CC-....   1125
B_EHOA        --GA-------C-----G--G--G-----GC-A-----T--G--TT-A-A-G-A--CTTGAC-CC-....   1110
SD_MM251      --------G--------A--G--G--T---T-A-----A-----CC---A-G-G--CCTCGC-CCA....   1122
STM_STM       --G--T--G--------A--G--A--T---C-G-----A-----CT---A-G-G-TCTTC---CCA....   1128
VER_AGM3      -----G---------A-TT-C--A----A---G-----A---ATG---CA-A-TATG...---AGC....   1155
GRI_AGM677    -----------G---CAA-----G--T-AGC-A----TA---ATG----GTA-T-GA...--GA--....   1125
TAN_AGM17     ---A----G--------A--C--A--C-AGC-T-----G-----C---CA-C-AATG...---GGA....   1143
SAB_SAB1C     --CA-T------G--CAA-----G---C-TC-A-----G-----T----C-GCG--CTTT--GC-ACAAA   1165
SYK_SYK       -----C--------CTTA--C--G----A-C-C-----A--G--C---GT-ATG--C...--GC-G....   1191

                                          \/ p7
GAG_CONSENSUS ??????????????c??c?at?atgaTgcag?????????????????a?aggc???ca???????a?    954
A_U455        ..............A-AAGC--A---------..................-G------...A-T...TTT-G   1136
B_HXB2R       ..............G-TA-C--A---------..................-G------...A-T...TTT-G   1151
C_UG268       ..............ATAAAC--A---------..................-G-A--...A-T...TTT-A   1133
D_ELI         TT............A-TA-AGCA---------..................-G------...A-T...TTT-A   1154
G_LBV217      CA............A-AG-C--A---------..................-A------...A-T...TTT-A   1148
H_VI557       ..............G-AGTT--G---------..................-A---T...A-C...TTC-A   1151
O_ANT70C      TGAAAGGAGGATACA-AG-AG-AT-C-----A..................-G---G...--A...AATCC   1163
CPZGAB        GG.........AGAG-AGATG-TT-CT-C---..................-AG--A...--A...GGTGC   1181
CPZANT        AG.........GGAA-CG-AG-CT-TC-----..................-G------...A-TGGAAAT-G   1217
A_ROD         ..............G-CC-T--CCCAT-CGCA.................GC--C-...--G........   1152
B_EHOA        ..............T-CA-C-ATCC-T-TGCC.................GCC-CT...--A...CCA-G   1142
SD_MM251      ..............GTGC-A--CCCTT-TGCA.................GC--C-...--G...AAG-G   1154
STM_STM       ..............GACC-AC-GCCCT-CGCA.................GC--C-...--A...CAGCA   1160
VER_AGM3      ................CAG-AC--------A.................CAG--...GGT...CAG-G   1184
GRI_AGM677    ......................---G-C--A.................GTG--ACCT--G...AAA-A   1151
TAN_AGM17     ................GTC-AC-----A---GGCCACTCAAAGGGAGGA-GG--A...AGA...GGA-G   1190
SAB_SAB1C     CA.........GTTGGGAAC--CT-TG-----CAAGGGGCA...AGACCC-GG--T...-CC...TTAGG   1217
SYK_SYK       ................T-AG-A-AT---GTA.................CA---G...-CA...TCA-A   1220

GAG_CONSENSUS ggg?c?aagaaga?????????????attaaaTGtTtcaAtTGtGG?aAagaaGG?CAcatagcaAga   1005
A_U455        ---C-CG-----...............-----G--------C-----C--------A---C----C-A-   1191
B_HXB2R       -AAC-A-----AGATT...........G----G--------------C--------G----C---C---   1209
C_UG268       ---C-CT-AG---ATT...........--------------------C--G-----G---C----C---   1191
D_ELI         ---C-C-----A-ATT...............-----G--------------C--------G----------A-   1212
G_LBV217      ---C-CG----A-AAT...............-----G------G-C-----C--G-----A--TC----C---   1206
H_VI557       ---C-A-------ATT...........--------------------C--------A---C----C---   1209
O_ANT70C      AATTAGG-A-G--ACT.............--A-----------C-----A-----G--A--T---------   1221
CPZGAB        ---C-C--A----AAA.............--A-----C--T--------G--------T---C----T---   1239
CPZANT        A--AGG--A----CCT...........C-C--------T--C--C--T-----G--C--T-CT------   1275
A_ROD         ....-AG----AGGCA...........T-------C-GG--C-----A--G-----G---TCG------   1206
B_EHOA        A-CAGGG-AG--GACA...........G-G-C---C-GG--C--C--C--G-CG--A--T-C---C--G   1200
SD_MM251      ---A-C-----AGCCA...............-----G----GG--------G--G--G--A---TCT-----G   1212
STM_STM       ---A-G---G-C-..............G-A-----C-GG--------G----C---G----CT----A-   1215
VER_AGM3      A--AAG-CC----CCCCCA.........G-A--G----A---------A---TTT--C--T--GCA----   1245
GRI_AGM677    ---C-CCC--G-GCCG...........C-A-----C--T--------C---TTT--A--T--GCA---G   1209
TAN_AGM17     A--A-C-CCTC-C.....................-----T--A--C--AC-GATT--A--TG-TCA--A-   1239
SAB_SAB1C     ---AAG-G-T---CCTCTGAACCCAAAT----------A---------A--GCCT--T---C-G--T---   1287
SYK_SYK       A--GAG-TC--TG..............------C-A---------GC--ATT--G-----GCA--A-   1275
```

```
GAG_CONSENSUS aatTGcagggcccctag?aaaaagggcTGctggaaaTGtGGaaaggaaggaCAccaaatgaaagacTGca   1074
A_U455        -----T-----------G-----------T--------C--G--A-------------------------   1261
B_HXB2R       -----------------G-----------T-----------------------------------T--T-   1279
C_UG268       -----------------G-----A-----T-------------G------------------------T-   1261
D_ELI         -----------------G-----------T----G----------------------C-A-----T----   1282
G_LBV217      -----------------G----G------T----------------G-----T----------------   1276
H_VI557       -----------------G--------A--T--------C--------------T--GC------------   1279
O_ANT70C      -----TC-A--A-----G--------T---------------C-------T--------------T----   1291
CPZGAB        -----T-AA--A--A--A-G---A-----T---CGG-----GC-A--G----------------------   1309
CPZANT        -------A------A--A-GG--A----------G-------C-------------GC-T---A----TC   1345
A_ROD         C-A---C-A--A-----A-GGC-------------G-----T---CC-------ATC----C-A-----C   1276
B_EHOA        C-G----A---------A-GGC----A---------------C--C--------ATC---TC-A-A---C   1270
SD_MM251      C-A-----A-----A--A-G-C----A-----------------AATG-AC--TGTT---GCCA-A---C   1282
STM_STM       C-G--T-AA-G---G--A-G-C-A--T--T------------C-AC-G--C--T--G---GCCA-A---C   1285
VER_AGM3      C-A---CCT-AA--A--A--G-T-A-A----T---------G--ACC---G--TTT-GCA-----T----   1315
GRI_AGM677    G-A----A---A--A--AC-G-TCAAA----TT--G--C--C--AATT--C--TATGGCA----------   1279
TAN_AGM17     G----TCC-CGAGGGG-TCC--TTAAG---CTC--------C--ACC-A-G---ATGGCA--------T-   1309
SAB_SAB1C     TT---T-A---------G-G-C-A-----------------C-GCCC--AC--T--G--------T--TC   1357
SYK_SYK       G-C----AAAAG--ACTT---GCTAAA----TT--T--------AAC--------T-GCA-GG-C---T-   1345

                                          pol orf start ->
                                                      p7 \/
GAG_CONSENSUS ca?????????????????????????????ga?agacAgGc?aaTTTTTAgGgaaaggctgg?c??cc??   1110
A_U455        -T.............................--G--------T---------------ATT---C-TT--AA   1304
B_HXB2R       -T.............................--G--------T--------------GAT----C-TT--TA   1322
C_UG268       -T.............................--G--------C--------------GATT---C-TC--AA   1304
D_ELI         -T.............................--G--------T-------------G-ATT---C-TT--CA   1325
G_LBV217      --.............................--G--------T---------------AT----C-TT--CA   1319
H_VI557       --.............................--A--------T---------------ATT---T-TC--AA   1322
O_ANT70C      G-AAT........................-GA-A------A-----------C---TA----C-TC-GGG   1337
CPZGAB        --............................-GA-------TG----------------T---C-TT--CG   1352
CPZANT        --GCAACA..................AATACAG--A-A-TA---------------CCGACCC-CA-GTG   1397
A_ROD         --.............................--T--------AGG---------ACTG---CCTTGGGGA..   1317
B_EHOA        --.............................--A--------GGG----------TTC---CCCTGGGGA..   1311
SD_MM251      --.............................--C--------GGG---------CCTT--TCCATGGGGA..   1323
STM_STM       --.............................--A-------TGGG---------CTTT---CCATGGGGA..   1326
VER_AGM3      G-GGA..............................----TA------------T-T---C--TGGATGGG   1355
GRI_AGM677    AG...........................A-TG-------A------------T-T---CATTGGGGAGG   1322
TAN_AGM17     G-...........................-GAG-------C---------A---G-ATGCCCG-AC--TA   1352
SAB_SAB1C     A-AAA..............................----TT------------TTT---CCCTGGGG-AG   1397
SYK_SYK       GGCAACCCAAGAGAAATCAAGGCCCTCCA-TTGC------T-----------------A---GGGGT-TC   1415

GAG_CONSENSUS ????aagaggaggCCagg?AAtTttct??????????????????????????????????????????   1132
A_U455        C...---G----------G-------CT.........................................   1329
B_HXB2R       C...---G-A--------G--------T.........................................   1347
C_UG268       C...---G----------G-----C--T.........................................   1329
D_ELI         C...---G-A-----G--G--C-----C.........................................   1350
G_LBV217      C...---G----------G--------T.........................................   1344
H_VI557       C....--AG----------G--------C........................................   1347
O_ANT70C      G...GGC-C---------C----A-G-G.........................................   1362
CPZGAB        C...-GCG----------C--C---G-G.........................................   1377
CPZANT        GTGGGG-T-C--A-----G--C---G-G.........................................   1425
A_ROD         ...----A-...--CC-C--C--C-CCGTGGCC....................................   1344
B_EHOA        ...----A-...--TC-C-----C-CCGTCCAGGCACCCCAGGGGATAGTGCCATCTGCGCCCCCGATG   1374
SD_MM251      ...----A-...--CC-C-----C-CCATGGCT....................................   1350
STM_STM       ...----A-...--CCAC--C--C-CCATGGCC....................................   1353
VER_AGM3      A...GC--AA...--CA-A-------CC.........................................   1377
GRI_AGM677    A...GC--AA...---A-A------G-G.........................................   1344
TAN_AGM17     A...GGTTCC-A---GC-C------T-A.........................................   1377
SAB_SAB1C     A...GG--AA...--CA-G-----C-CT.........................................   1419
SYK_SYK       C...-G----CCC----CC--C--C-CA.........................................   1440
```

HIV1 GAG

```
GAG_CONSENSUS ??????????????????????caga?cagacca??????????????????gagCCaaCaGCcCCaCCag?a?   1163
A_U455        ........................----G------................------------------C-.   1362
B_HXB2R       ........................----G------..................------------------A-.   1380
C_UG268       ........................----G------................--A----------------C-G   1363
D_ELI         ......................--A-G------..................------------------C-.   1383
G_LBV217      ........................----A------................------T----------C-C-.   1377
H_VI557       ........................----G---G--................------------------C-.   1380
O_ANT70C      ........................----GACC-G--................C-C---T----------GATG.   1395
CPZGAB        ........................----A----A--................--A-----G--------GAT-.   1410
CPZANT        ...............CAGAAGG--GAAGT-GTG...............-----------T-----CAT-.   1464
A_ROD         ...............CAAGTT-C-CAGG-G-TG...............ACA--------A--C----TG.   1383
B_EHOA        AACCCAGCATTCGGCATGACA-CTCAGG-TG-G...............ATT---T-T--A--C--T-C-.   1428
SD_MM251      ...............CAAGTG--TCAGG-G-TG...............AC------T--T--C----AG.   1389
STM_STM       ...............CAAATA-CTCAGG-G-TG...............AC------T-----T----A-.   1392
VER_AGM3      ...............GCCGCTACTCTTG-GGTG...............--------T--G--C--TCC-C   1417
GRI_AGM677    ......................--ATA----GG-GACACAGTTGGTCTG--A-----------C---ATGG   1393
TAN_AGM17     ..................GAG--AGGAG--G-C...............-TT--G--G--T--C--GATCG   1414
SAB_SAB1C     ......................-T--C-TCCATC...............AGA--------T--C---ATGG   1453
SYK_SYK       ........................GTG--GAGC...............------T-T--T--C--TCTGG   1471

GAG_CONSENSUS ??????????????????????????????????????????????????????????????????????????   1163
A_U455        ......................................................................   1362
B_HXB2R       ......................................................................   1380
C_UG268       AGCCAACAGCC...........................................................   1374
D_ELI         ......................................................................   1383
G_LBV217      ......................................................................   1377
H_VI557       ......................................................................   1380
O_ANT70C      ......................................................................   1395
CPZGAB        ......................................................................   1410
CPZANT        ......................................................................   1464
A_ROD         ......................................................................   1383
B_EHOA        ......................................................................   1428
SD_MM251      ......................................................................   1389
STM_STM       ......................................................................   1392
VER_AGM3      CGAGT.................................................................   1422
GRI_AGM677    AAACA.................................................................   1398
TAN_AGM17     CCGCGCATGGATTCCCCACAGGT...............................................   1437
SAB_SAB1C     AACGGGATTACAGCAGGCCGGAGGAAAATTGGTAT...................................   1488
SYK_SYK       AGGACATCGAAGATGGGCCATGGCTCACATGGTCAGCACAGATGAGCCAACAAGCACAAGCGAAGGCACA   1541

GAG_CONSENSUS ??????????????????????????????????????????????????????GAga??g??g?g?a???g   1172
A_U455        ......................................................--A-TCTTT-G-ATGGG-   1380
B_HXB2R       ......................................................----GCTTCAG-TCTGG-   1398
C_UG268       ...........................................CCACCGGCA----GCTTCAG-TTC...   1398
D_ELI         ......................................................----GCTTC-G-TTTGG-   1401
G_LBV217      ......................................................----GCCTC-G-TTCGG-   1395
H_VI557       ......................................................----GCTTC-G-TTTGGA   1398
O_ANT70C      ......................................................---GAG-AA-T-A-GGGA   1413
CPZGAB        ......................................................----GTTAT-G-T-CCA-   1428
CPZANT        ......................................................----TCTATCA-G-G...   1479
A_ROD         ......................................................--TCCA-CA-T-G-TCTA   1401
B_EHOA        ......................................................--TCCA-CA-A-G-GAT-   1446
SD_MM251      ......................................................--CCCA-CT-T-G-TCT-   1407
STM_STM       ......................................................--CCCA-CT-C-G-TCT-   1410
VER_AGM3      ..............................................CCATAC--CCCT-CAAA-A-GCTC   1446
GRI_AGM677    ...........................................GCTTAC--TCCA-CAAA-A-GCTC   1422
TAN_AGM17     ...............................CCGCCAGTTGCAGGAGCCTAC--TCCA-CCAA-AGACTC   1476
SAB_SAB1C     ................GCAGACAGACCTCCGACCAGAGGACCAGGGCCAGAC--TCCA-CAACAGCCCT-   1542
SYK_SYK       GAACTCTCCCTCGAAGAAACCCCCCACCAACAGGGAAGTGCTCTCGCCGAAG----GCAGT-G-A-G...   1608

GAG_CONSENSUS g?ggag?a????a?????c??gg??agaagcag???gagcaga?aga?a?g??????????????????   1202
A_U455        -AAA--......-TGACCTCCCCTGC---A---...----T--A---C-G-...................   1422
B_HXB2R       -TA---ACA...-CAACT-CCCCTC----------...----C--T---C-A-...................   1443
C_UG268       -A--G-......-CAACT-CCCCTCC--------...----C--A---C-G-...................   1440
D_ELI         -AA---......-TAACC-CCTCTC-A--A---...-------A---C-A-...................   1443
G_LBV217      -A----......-TAGCC-CCTCCCC--------...----CA......-A-...................   1431
H_VI557       -A---A......-TGACC-CCCCTCCA--A---...---AG--AG--TGCA....................   1440
O_ANT70C      CA----A-T.......................---...--A----A--GGGGCCCGAAC.............   1446
CPZGAB        -A---A...............-AGA---GC---...---A---A---GGGA...................   1461
CPZANT        -A-C-CA-GAGG-CTCAGAAG--TCTC---GG-...---G--.............................   1518
A_ROD         CT----A-ATAT-TGCAG-AA--GA-A-GA---AGA-------G---G-GACCATACAAGGAAGTGACAG   1471
B_EHOA        CTAA--A-CTAC-TGCAA-TA--GA--------AAG---A-C-G---G-GACCCTACAAG...........   1506
SD_MM251      CTAA--A-CTAC-TGCAGTTG--CA--C-----AGA--AAGC-G---G-A-CCTTACAAG...........   1467
STM_STM       CT-AGAAGTTAC-TGCAG-TG--CA--------AGA---AGC-GGA-G-CACCCTACAAG...........   1470
VER_AGM3      CT-C--C-ATATGCAGACAAG--GA--C--TT-AGG--A--A-GGA-A-AACCACCAGCAGTGAAT....   1512
GRI_AGM677    CTCC--C-GTATGCAGAGAAG--AC--CGC-T-AGA---G---G---ACA-ACAAGGAAACAG.......   1485
TAN_AGM17     CT--G-C-GTATGCAAAAAAA--AG-CC-A-TAAGGAG---A-A---G-A-....................   1527
SAB_SAB1C     TTAA--C-GTATGCTGTT-AG--GA-ACG----.......................AAACAGCAGTGGC   1588
SYK_SYK       -A---ACC.....................................-A-....................   1620
```

```
GAG_CONSENSUS ??????????????????????????????????????????????????gaa????????g?at?c???   1209
A_U455        ........................................................---.......CA-AC-C-TTT   1436
B_HXB2R       ........................................................---.......CT-T--C-TTT   1457
C_UG268       ........................................................---..................   1443
D_ELI         ........................................................---.......CT-T--C-TTT   1457
G_LBV217      ......................................................--G......CTAT--C-T..   1443
H_VI557       ........................................................---............T-T..   1446
O_ANT70C      ......................................................--G......CTCT--C-GTT   1460
CPZGAB        ......................................................--GAGCAGTCTAT--C-ACC   1481
CPZANT        ........................................................---.......CTACC-C-CTC   1532
A_ROD         AGGACTTACTGCACCTCGAGCAGGGGGAGACACCATACAGGGAGCCACCAAC-......GA-G-CTTGCT   1535
B_EHOA        ..............................................GAGGTGAC-......GA-G--TTGCT   1526
SD_MM251      ..............................................GAGGTGAC-......GA-G--TTGCT   1487
STM_STM       ..............................................GAGGTGAC-......GA-G--TTGCT   1490
VER_AGM3      ......................................CCCGATTGGACA--G......GGAT--T-T..   1536
GRI_AGM677    ......................................AAGGAGAAA---GTG...GA-G--GTT..   1509
TAN_AGM17     ........................................................---CTG...GA-G--TAT..   1542
SAB_SAB1C     AAAACCAC..................TCGCCACAACAGAGCCCCTACGAG--G......GCTT-CAGC..   1632
SYK_SYK       ....................................................AGT......CTCT--C-C..   1632

GAG_CONSENSUS ????tccctcaa?TC?cTctttgg??aa?????????gaccaa                               1235
A_U455        AGTT--------A--A--------CA-CGACCCCTTGTCA--G                                1479
B_HXB2R       AACT-------GG--A--------CA-CGACCCCTCGTCA---                                1500
C_UG268       ....C--T-A-TT--C---..................A-ATC-                                1464
D_ELI         AACT--------A--A--------CA-CGACCCCTTGTCG---                                1500
G_LBV217      .......T-A-CT--C---..................A-ATC-                                1461
H_VI557       ....C-TT-A-CT--C---..................A-ATC-                                1467
O_ANT70C      TGCC--------A--C--------GAC-........------                                1494
CPZGAB        AACC--------A--A--------CAGCGACCCCTCATC---G                                1524
CPZANT        GTAT-----G--A--C--------CA--........------                                1566
A_ROD         G...CA------T--T--------AA--........-----G                                1566
B_EHOA        G...CA------T--T--------AG--........-----G                                1557
SD_MM251      G...CA------T--T--------AGG-........-----G                                1518
STM_STM       G...CA------T--T--------AG--........-----G                                1521
VER_AGM3      .......T-G--C--C--------AG--........------                                1563
GRI_AGM677    ....---T-G-GC--C--------AGG-........------                                1539
TAN_AGM17     .............--CT-G.................AGTTCC                                1554
SAB_SAB1C     ....-------GG--T--------AG--........------                                1662
SYK_SYK       ....--T----GC--T--------AG--........------                                1662
```

HIV1 GAG CONSENSUS

```
          p17 ->           /<-       nls        ->/
              /<-    membrane binding     ->/
CONSENSUS-A   ATgGGTGCGAGAgcGTCagTatTaAgtGggggaaaATTaGATgcaTgGgAgAaaATTcgGtTAaGGCCaG     70
CONSENSUS-B   ---------------------------c------g--------ag------a-------------------     70
CONSENSUS-C   ------------------a-------A--c------------a-------a------A------------     70
CONSENSUS-D   --------------------------C----------------a------A------------C------     70
CONSENSUS-F   --------------------------C-----------------------A-----------------g-     70
CONSENSUS-G   --------------------------C---------?------?------A------?--?-G-------     66
CONSENSUS-H   ------------------------?--C--------?---------?-----------------?-------     66
CONSENSUS-O   -----------?-----T--G--G-CA--?A-?-----G-----------AC?----A------?-----     65
CONSENSUS-CPZ -----?---?-?-----?--T?-?-??--?-??--??-----???-----A-???-??--C-T?----C-     48

                                                  /<- nls ->/       p17 \/ p24
CONSENSUS-A   ggGGAAagAAAaaaTATagacTgAAaCAT?TAgTaTgGGCaaGCagggAGcTgGAaagATTcgCACTtaA    139
CONSENSUS-B   ------------------a-t-A------a----------------------A---c--------g----    140
CONSENSUS-C   -----------c-c----TGa-a-----cc------------------------------T--------    140
CONSENSUS-D   -a----------------a---a------a----------------------a---C----T--------    140
CONSENSUS-F   --------------------?----g---C----------------------A---c----T---a----    139
CONSENSUS-G   -----------?--------A-?--?-----------------------------?-?---T--------    130
CONSENSUS-H   -----------?-------?--?-----------------------?--------------T-----?--    130
CONSENSUS-O   -ATCT--A--G??------?--A-----?T----------------?-------------?----TG---    129
CONSENSUS-CPZ -?--?-??----??--?-TGA-?--------?--?--------???-------?-??-?--?--??????     93

CONSENSUS-A   cCCtaGccTTtTagAaACAgcaGaagGatGTcaacaaaTAaTggaaCAgtTacAatCAgcTcTcaagACA    209
CONSENSUS-B   t---g----g---------t-------c---aG-------c---g----c-----c--t-c--tc-----    210
CONSENSUS-C   ----g----------g---t-------C---A----------aA-----C-----C-------TC-----    210
CONSENSUS-D   T---G----------g---t-------c---A----------a-g----C-----C-----a-tc-----    210
CONSENSUS-F   t---g-----c--------t-------C----g-a-------A-G----------c--T-C--TC-----    209
CONSENSUS-G   ----G----?---------A-------?------------?--??------G---C--?-----C-----    194
CONSENSUS-H   ---?G?---------?---?-------C----T---??----A------C----?C-------T------    193
CONSENSUS-O   T---G??--A?-------T-----?--TA?-G-G?-?C-GC-AC-G------G-GC--------------    192
CONSENSUS-CPZ -?-C???--??-?----?????-----?---???A-?????-?C-?--A--????C-??-??-??-A-?-    129

CONSENSUS-A   Gga?CAgAAGAAcTtAaaTCaTTAtttAAtaCagTAGcAacCCTcTatTgcgTaCAtcaaaggATAgA?g    277
CONSENSUS-B   ---t------------g--------a------------------------T--g--------------g-    280
CONSENSUS-C   ---A----g-------g--------a---c-----------t--------T------g--?a------g-    279
CONSENSUS-D   ---t---------------------?-----------------------t------g----------g-    279
CONSENSUS-F   ---T-------G----g--------a-------------gt--------tt------------g----G-    279
CONSENSUS-G   ---A----G--G?--------?---?----?--?-----??T----??--T?----------------g-    254
CONSENSUS-H   ---A-----------C--------------CT-T-----??---------?-----------A-----T?    259
CONSENSUS-O   --GT----???C--G?-?--?C-C-GG--CG--A---?-GT?----GG--?--T--CA-C--AT?T?--A    250
CONSENSUS-CPZ ?-?T-????-?-?-????--?--G-----C--??-????GT?--G-GG---?-----A??G?????A?-?    168

              /<-     MHR      ->/
CONSENSUS-A   TAAaaGACaCCAAgGAaGCttTagAtAaAaTaGAgGaAatacAAAatAAgagcaagcaaaag????????    339
CONSENSUS-B   -------------------------?--g---------GAg-----C--a--t---a----?........    340
CONSENSUS-C   --Cg---------------c-----C--g---------GA------C--a--TC-------A------..    341
CONSENSUS-D   -------------------------a--g--g------GA------C--A--T---A-------------    341
CONSENSUS-F   ----G--------------------c--gC--------GA------C--A--Tc--------........    341
CONSENSUS-G   -------------A------c----GG--G-G--aA-??c------G--c--TC-------A........    314
CONSENSUS-H   ------------------------???--G?--?----?G?------C--?--?C--?-?-??........    308
CONSENSUS-O   -??G---T--?C--C-G--AA--C-A--GT--A-----G--....-TGG??----G-A-GTCT........    304
CONSENSUS-CPZ --?-------A?-??----??-??-A????-?A-???G?A????????-???AC-?????-AAC-G-AAG    203

                                     p24 \/        \/   'p2'   \
CONSENSUS-A   ??????????acacagcaggCaGcaGCT......g?cACA...ggaAgC???......agcagcaa?gtc    379
CONSENSUS-B   ..........g-------a---------??????-a----...----a----??????------c-g---    382
CONSENSUS-C   ..........------------aa---g...gct-A-???...----Ag..................---    374
CONSENSUS-D   -........g-------A---a----------.....-A-----...a---a-........------C-G---    383
CONSENSUS-F   ..........--------a-?g------.....-A--A-...--g.....................---    370
CONSENSUS-G   ..........-?----------??????......-A??A-...-----A-.........------C-A---    347
CONSENSUS-H   ..........--------?---A---?-......-AT-A?...-?--A?.........G?--A?--G?--    342
CONSENSUS-O   .GC-GA-GCCG-TA--G-A...............-A?---.........---........?C?--GC-G?C?    336
CONSENSUS-CPZ TA-C--A---????G???A-G??????????C???-????????????---??????????????-???    212

                                  Zn-motif                          pol cds ->
                          / p7         /<-Zn-motif ->/      /<-Zn-motif ->/ p7 \/  'p1'
CONSENSUS-A   ???????????AGcCAaAAtTACCCtATAGTgCAaAAtgcacAagGgCAAatgatAcatCAG?cctTaT    436
CONSENSUS-B   ............------------------------g--ccTc--G---------G--------g--a---    440
CONSENSUS-C   ............--T--------t-----------g---CTc------------g----c---g--a---    432
CONSENSUS-D   ............-----------t--------------CcT---g---------G----c---G--a---    441
CONSENSUS-F   ............--T----------------A--g---cTt--G--a------G-----------a---    427
CONSENSUS-G   ............--?--------t-----------G------------------?----c------a---    402
CONSENSUS-H   ............--T--------T--?-----A--?-----C-----?--G??-G----C---G--A---    395
CONSENSUS-O   ............?-T-----------------A?C------?--G--A------G-------?G--A-C-    390
CONSENSUS-CPZ ------------?-??G---------C?-C?-A????----???---?????????G--------??-??    244
```

```
                                vpr binding           (minor) \/ vpr binding     p6 terminus
          'p1'          \/ p6  /<-->/     \/ (minor)            /<- ->/     / (80%)
CONSENSUS-A   CACCtAggACttTgaAtGCATgGGTgAAaGtaaTAGAagAaAagGcTTTcAgcCcAGAaGTaATACCCAt    506
CONSENSUS-B   -------A-----A-----------A------g-------g-----------------------------    510
CONSENSUS-C   -------A-----------------a-----------g--g-----------------g-----------    502
CONSENSUS-D   -------a--------c--------a--------------?-------------------------------    510
CONSENSUS-F   -------A-----A-----------a--g--g--------g--------t--------------------    497
CONSENSUS-G   -------?-----------------A------g-------------C-----T-----------------    471
CONSENSUS-H   -----------C--A-----------A------G-------?--------T-----------------?--    463
CONSENSUS-O   -C--C--------A-----------A--G-C-G-------?--?--C--T-A---T---A-?--T--T--    457
CONSENSUS-CPZ ----?--A--C--?-----?-----A--????G----?--?--????--?-??--?--?--?--?--?--    293

CONSENSUS-A   GTTctCaGcatTaTCAGAgGGaGCcACccCaCAagAttTaAAtatgATGcTaAAcatAGTggggGGaCAc    576
CONSENSUS-B   ---T--------------A-----------------------c-Cc----------C------------t    580
CONSENSUS-C   ---ta-------------a-----------------------c-cc---t----t-c------------t    572
CONSENSUS-D   ---T--------------A-----------------------c-CC----------C------------T    580
CONSENSUS-F   ---t--------------a--g-----t--------------c-CC--------t-C------------T    567
CONSENSUS-G   ---T--------------a-----------------------cC--------?-C---------g--T    540
CONSENSUS-H   ---T--------------A-----------------?------GCC---?----T-?---------?--T    529
CONSENSUS-O   ---?ATG-----?-----?--?--??T??-CT-T---A-?----C?-----?--TGCCA-A--?-----?    514
CONSENSUS-CPZ ---T--?--T?-------?--?--????--T--?---?-?--C-CC---?-?--TGC---?----??--T    347

CONSENSUS-A   CAgGCAGCtATGCAaATGtTaAAagAtaCcATcAATGagGAaGCTGCAGaaTgGGAcaGg?tACAtCCAG    645
CONSENSUS-B   --a-----c-----------------g-----------------------------T--At-g-------    650
CONSENSUS-C   --a-----c---------------------------------g--------------t---t---------    642
CONSENSUS-D   --a-----c-----------------g-----------------------------t---c---------    650
CONSENSUS-F   --a-----c-----------------c--------------------------------aT---------    637
CONSENSUS-G   --A---------------c----G-----t--------------------G---------A--------C    610
CONSENSUS-H   --A-----C--------??----------A-----------------------T--------------    596
CONSENSUS-O   --A-G?---T-A---G--?-?--G--AGTA--------------A-?---?-----T--AACT--?---C    578
CONSENSUS-CPZ --A-G---C-----?G--C----?--AGT?--?-----------?--?--G-----T?-?T----?--CA    408

CONSENSUS-A   TaCatGCAgGGCCTattcCaCcagGcCAgaTgAGaGAACCAAggGGaAGtGAcATAGCAGGaacTACTAG    715
CONSENSUS-B   -g---------------g-------------------------------------------------    720
CONSENSUS-C   -?------------?--g----------a-----------------------------------------    710
CONSENSUS-D   -g---------------G---------------------------------T-----------------    720
CONSENSUS-F   -G--g-----------C--------------a--g-----T------------t-----t-----------    707
CONSENSUS-G   ?---g--------------------------A-----?--------------T-----------------    678
CONSENSUS-H   -G--------------------------?--------------------?--T-----------------    664
CONSENSUS-O   C-?C??T?-----GT-?--------G-----A--G-------CA-----------T--T--?--A-----    642
CONSENSUS-CPZ CT-----?--?--??-A????----A--??-A--G--?----??--?-----T--------G--?--?--    462

CONSENSUS-A   tAcCctTcaaGAAcAAaTaggaTgGATgACaagc.........AAtCCACCtatCCCaGTgGGagAcATc    776
CONSENSUS-B   ---------G----------------------aT.........-----------------a-----a---    781
CONSENSUS-C   ---------g----------c------------T..........--C--------T---------------    771
CONSENSUS-D   ---------G----------?-------------.........-----------------A-----A---    780
CONSENSUS-F   ---------G---------ca-------------..........--c------g-------------a---    768
CONSENSUS-G   ------?--G---------A----------C---.........-----------------------a--t    738
CONSENSUS-H   ---------G----------C----------?--.........-----??--------------------    722
CONSENSUS-O   C--?-AG-----G-----T???----CT--?--?...????C?--C????----------A---------    694
CONSENSUS-CPZ C--A?-G--G--?--?--????----???--?CA?????????---???????-------A--G--??-?    502

CONSENSUS-A   TATAaaAgaTggATAaTccTgGGatTAAATAAaATAGTaAGaATGTATAGcCCTgttAGCATttTgGAca    846
CONSENSUS-B   -------------------------------------------------------acc------c------    851
CONSENSUS-C   -----------------T-----G--------------------------------C-------------    841
CONSENSUS-D   -------------------------------------------------------c-------------    850
CONSENSUS-F   --------------c-----a-----------------------------------C-------------    838
CONSENSUS-G   -------------------------------------------------------c-------------    808
CONSENSUS-H   --------------------------------?-----------------?-----C-------------    790
CONSENSUS-O   ----G--A-------G-G?-A---C----C-----G----A------C--?--A--G-----C--A--T-    762
CONSENSUS-CPZ ----??------?-?--??-?--?-----?--?G-G--?--?-?----?-?--A--?--?--?C-?--?-    552

CONSENSUS-A   TAAgaCAAgggCCaAAaGAaCCcTTcAgaGAcTATGTagAtAggTTCTtTAAaactcTcAGAGCtgAgcA    916
CONSENSUS-B   ----------a-----g--------T--------------cc------A---------A-----c-----    921
CONSENSUS-C   ---a------------G--------T--------------C--------------T-A--------A--    911
CONSENSUS-D   ----------A-----g--------T--------------?C------A---------A-----c-----    919
CONSENSUS-F   -------------------------T--------------C--------------c--A-----------    908
CONSENSUS-G   ----------------??????????????????????????????????????????????????????    824
CONSENSUS-H   ---?------?--------------T--------------?------------?--T-A--------?--    855
CONSENSUS-O   -T-?G--G--A-----?-----A--?-----?---------C----?-AC-----AT-A-----------    827
CONSENSUS-CPZ -?-?------?-----?--G-----?-----T-----G--?--A----A?-----??-?--?--?--?--    609

CONSENSUS-A   aGCTaCaCAgGagGTAAAaaacTggATGACaGAaACctTgcTggTCCAAAAtGCgAAtCCAGAcTGTAAg    986
CONSENSUS-B   ----t----------------t------------------t----------------c-----t------    991
CONSENSUS-C   ---------A--T--------t-----------c------T----------------C-----T------    981
CONSENSUS-D   ----t--------t--------t--------?---------T----------------C-----T-----a    988
CONSENSUS-F   ------------A-----gggT-----------C------T----------------C-----T------    978
CONSENSUS-G   ????????????????????????????????????????????????????????????????????    824
CONSENSUS-H   ---?--------?--?--???T-----------C------T----------------------T--?---    918
CONSENSUS-O   ------T--A--A-----?--T--------------??----T--T--?---?-?--?-----T--?--A    889
CONSENSUS-CPZ ?---T-----???--G---???--------?--?--??-?T-A?-?--?-----?--?-----T--?--?    660
```

HIV1 GAG CONSENSUS

```
CONSENSUS-A   tCCATttTaAgAgcatTagGa?cagggGCtaCaTTAgAaGAaATGATGACAgCaTgcCAgggAGTggGaG   1055
CONSENSUS-B   A-t-------A------g---C---ca------c----------------------T-------------   1061
CONSENSUS-C   A--------------------C--------t------------------------T-------------   1051
CONSENSUS-D   A-t--C----A------g---C---?-----------------------------t-----------g-   1057
CONSENSUS-F   A---------A------G---C-----------c----------------------t-------------   1048
CONSENSUS-G   ??????????????????????-????????????????????????????????????????????????   824
CONSENSUS-H   A?T-------?-----------?-------T--A--------------------?--?-------------   982
CONSENSUS-O   CA?---?-G-A-?-------?C---?A-----?--------?------GT---C--T--?-----A----   951
CONSENSUS-CPZ CA?--?C-?-A?--?--G-----?--?--C?-C?----?------?-?-----T--T--A-----A----   717

CONSENSUS-A   gACCcgGcCAtAAaGCAagggTtTTgGCTGAgGCAATgAGtCAaGta.........caacAt????caAa   1112
CONSENSUS-B   --------------------A-----------a--------C------.........ac-a-----t--..   1117
CONSENSUS-C   ----Ta----C--------A--G-----------------C----C-.........A-CA--...a----   1109
CONSENSUS-D   -----a-t-----------A--------------------C----c-.........ac-a--...t--?-   1114
CONSENSUS-F   ----t--------g-----a--------------------c----c-.........ac-a--...-????   1101
CONSENSUS-G   ??????????????????????????????????????????????.........??????...-????   824
CONSENSUS-H   -?--T?----?-------?A--------------------C------???......AC-A--...-????   1031
CONSENSUS-O   -?--AACT--?--G--?-?A?-?C-A--A--A------GC-?C?-CCCAGCAAGATTT?A-AGGAGG-T-   1012
CONSENSUS-CPZ -?--A?C?-----G-----?--??-?--?--A--?---?C????-?T?A????...???--?...-???G   753

CONSENSUS-A   c??????...aTAATGATgCAgAgaGGcaAttTt...agggGccagaaaaga???aTtAAgTGtTTcAAc   1167
CONSENSUS-B   .gctacc???-------------------------...---aa---a-g--Agactg-------------t   1180
CONSENSUS-C   -.........---------------a--------...-Aa----Ct------AtTg----A---------   1167
CONSENSUS-D   T-ctgC-...-------------------------...-A-----cA-g--a-att---------------   1176
CONSENSUS-F   ?gca-cc...-------------a-a-t--c---...-A------A-G----AtTg----A-----t--T   1163
CONSENSUS-G   ?------...--?--------?-A-?--???---...-A-----C--?--?-A-T------------?--   871
CONSENSUS-H   ?GCAG--...--?----------A---?--C--?...-A---?--A-G----ATT?----A--------?   1086
CONSENSUS-O   -ACAGCA...G--T-C-----A-----GC-?AA-...CCAA?TAG?---G-?-C---A--A--?-----?   1068
CONSENSUS-CPZ A-C-G--...G-?T-??-?----??--??-?GG???????--???A---------?-?--A--?--T--?   790

CONSENSUS-A   TGTGGcAaaGAaGGaCAccTAGCcAgAAATTGcAggGCcCCTAggAAAAagGGcTGtTgGAAaTGtgggA   1237
CONSENSUS-B   --------------g---a------a-----------------------------------------A-   1250
CONSENSUS-C   --------g-----G---a-------------------------------------------------A-   1237
CONSENSUS-D   --------------g---a----a-a------------------?-----------------------A-   1245
CONSENSUS-F   ------------------a------a------------------A-----------------------a-   1233
CONSENSUS-G   --------G--------T------------------------------?------------------A-   940
CONSENSUS-H   ------------------?----?----------------------------?-----------?--A-   1152
CONSENSUS-O   -----A-----G-----TA----A-?---?--TC-A--A-----?----?---T--C-----------AC   1134
CONSENSUS-CPZ --?--?-----?--?--????--?--------?-A?--?--A--A-G?--A-----?---?G?-----?C   843

CONSENSUS-A   agGAaGGaCAcCAAaTgAAagAcTGcACt...???gAg...aGaCAGGCtAATTTTtTAggGAaAaTtTg   1298
CONSENSUS-B   ----------------------t--t---...------???-----------------------g--c--   1311
CONSENSUS-C   ?------------------------T---.......---...--g--------------------------   1297
CONSENSUS-D   ----------------------t------.......--a...-----------------------------   1306
CONSENSUS-F   ga----------------------------.......--a...-----------------------------   1294
CONSENSUS-G   ?---?-----T-----------------A.......---...----------------------?---?--   997
CONSENSUS-H   ?---------T--G?-------?-----A.......--?...---------------------------?--   1208
CONSENSUS-O   -------T--------------T----?A......A-TGGA-?------A-----------?--?TAC--   1194
CONSENSUS-CPZ -?--?--------??-?---?----??-A???---???????--?-?-T?---------------?????   884

CONSENSUS-A   GcCTTCCagCAAgGGg...AGGCCagGaAAtTTtCctCAgagCAGacca.....................   1344
CONSENSUS-B   -------ca------a...--------g-------t-------------?????????????????????   1357
CONSENSUS-C   -------cA-------...--------G-----C-T---------g---?????????????????????   1343
CONSENSUS-D   -------ca------a...--------g--c--?-t-------------.....................   1351
CONSENSUS-F   --------A-------...-----?--?-------T?------------.....................   1337
CONSENSUS-G   -------CA-------...-----?--G-------T-----?---??--.....................   1039
CONSENSUS-H   -?--?---?---A---...--------G-------T?--------G---.....................   1250
CONSENSUS-O   ----C-GG-GGGCAC-...--------C----A-GTG----?AC?-G??.....................   1236
CONSENSUS-CPZ ?--??-??-??G?---???--?-----?--C---GTG----A??????-??????...............   914

CONSENSUS-A   .....................GAgCCaaCAGCcCCACcaGCAGAg.....................a?ct   1371
CONSENSUS-B   ??????......?????????-------------------a----.....................-g--   1385
CONSENSUS-C   ?????????????????????-----------------------?????????????????????-G--   1371
CONSENSUS-D   ..........................------------------------.....................-G--   1379
CONSENSUS-F   ..........................------------------------.....................-g--   1365
CONSENSUS-G   ............................------?----------?-----?.....................?-??   1060
CONSENSUS-H   ..........................------------------------.....................-G--   1278
CONSENSUS-O   .......................??C---T----------?ATG---.........................   1257
CONSENSUS-CPZ .......................--?-----?--?-----?AT----........................?-   936

CONSENSUS-A   tt???gGgatgggggaagagataacc...tCctct............ccgaagcaGgAgCagaaagac..   1421
CONSENSUS-B   -C...a--T-t----------c---aactc-----???????????-a-----------c--t----..   1438
CONSENSUS-C   -C...a--T-C...--G----C----...C--g--............-------------c-------??   1418
CONSENSUS-D   -C...---T-t-----g---------...C-----............-a---a---------------??   1429
CONSENSUS-F   -c...---T-C?-a--G---------...C-----............----------------?---..   1413
CONSENSUS-G   ??...?????????--G------?--...?----?............-----------?-C???????..   1088
CONSENSUS-H   -?...---T-?--A--G--?--G---...C--?--............--?--?------???-?G--?..   1317
CONSENSUS-O   ..............--G-?AG-G-AG...G?ACA?............GA---T---???-----?-GG..   1289
CONSENSUS-CPZ A-...??????CA---G--??A?-AG...AG????............?A----?????????G?G???..   959
```

```
CONSENSUS-A   ....??????aaggaacag??tcc?cCtTtAgtttCCcTCAAATCACTCTTTGGCAACGACCccTtGTCA   1478
CONSENSUS-B   .......----------t-TA-??--------C--------g-----------------------c----   1495
CONSENSUS-C   ????-------g----???--???---C---Ac--------------------------------        1464
CONSENSUS-D   ????-------------t-ta-...-------C-------------------------------------   1486
CONSENSUS-F   ..........g---g--t-ta---T--c----C-------------                            1449
CONSENSUS-G   ..........-----?-TATA-??-------??-------------                            1118
CONSENSUS-H   ..........???---------?-T--?---?C-------------                            1341
CONSENSUS-O   ....G--C-G?-?--G-T?TA?...--?--T-CC-----------C--------G-CA----AA-A-      1341
CONSENSUS-CPZ ..........??????-TA--?-----???????-----?-----?----------??----??-?A--C    995

CONSENSUS-A   CAG                                                                      1481
CONSENSUS-B   --aTAA                                                                   1501
CONSENSUS-D   --A                                                                      1489
CONSENSUS-CPZ ---                                                                       998
```