HMMER Sequences in the Env Alignment

A_U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
B_HXB2R	HIVHXB2R	K03455	Starcich,B.	Science 227, 538 (1985)
C_UG268A2	HIVUG268A2	L22948	Louwagie,J.J.	JVI 69, 263 (1995)
D_ELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
E_TN2432	HIVTN2432	L03703	McCutchan,F.E.	ARHR 8, 1887 (1992)
F_BZ163A	HIV1BZ163A	L22085	Louwagie,J.J.	ARHR 10, 561 (1994)
G_LBV217	HIVLBV217	L11778	Louwagie,J.J.	AIDS 7, 769 (1993)
O_ANT70C	HIVANT70C	L20587	Vanden Haesevelde,M.	JVI 68, 1586 (1994)
O_MVP5180	HIVMVP5180	L20571	Gurtler,L.G.	JVI 68, 1581 (1994)
O_VAU	HIV1VAU	X80020	Charneau,P.	Virology 205, 247 (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)
SYK_SYK	SIVSYK	L06042	Hirsch,V.M.	JVI 67, 1517 (1993)

The following alignment was generated using the HMMER program as described in the Introduction to this Part and in Part III. For simplicity, only representative PIV types and subtypes are shown. An ordinary consensus sequence (majority for lowercase, 50 percent or higher for uppercase) was deduced using MASE; this consensus is not a "most likely sequence" derived from an HMM model(see Part II). Annotation is based upon HIV1s, therefore the user should be cautious about its applicability to other PIV sequences.

```
ENV_CONSENSUS atGa?a???gtg???????????a??g???????g???aa????c??????t?????????????????     15
A_U455        ----G-.....-.............-TG-GGATACA-AGG--TTAT-CTTGCT-GTGGAGATGG.......     48
B_HXB2R       ----G-.....-.............-AG-AG..........--ATAT-AGCACT-GTGGAGATGG.......     39
C_UG268A2     ----G-.....-.............-TG-GGATACA-AGG--TTGT-AACAATGGTGGATATGG.......     48
D_ELI         ----G-....C-.............-GG-GGATAGA-AGA--TTGT-AAAACTGGTGGAAATGG.......     48
E_TN2432      ----G-.....-.............-AG-AGACACA-ATG--TTGG-CAAACT-GTGGAAATGG.......     48
F_BZ163A      ----G-.....-.............-GG-GGATGCA-AGG--CTGG-AGCACT-GGGGAAATGG.......     48
G_LBV217      ----G-....C-.............-AG-GGATACT-AGG--TTGG-AACACT-GTGGAAATGG.......     48
O_ANT70C      ----T-.....-ACTATGAAAGCA-TG-AG...AA-AGG--C...AAGAAGT-ATGG.............     48
O_MVP5180     ----C-.....-ACT......ATG-AA-TAATGAA-AAG--CAACAGGAAGTCATGG.............     48
O_VAU         ----C-....C-ATTATGAAAGCA-TG-GG...AA-AGG--C...AGGAAGT-AGGGATCTGG.......     54
CPZGAB        ----A-.....-A............-TG-AGAAGAA-AAG-GA......................GACTGGA     34
CPZANT        ----GG...AA-.................................-CGATACATATTATTTGG.......     27
A_ROD         ----TG...AAT.................................-AGCTGC-TATT.............     21
B_EHOA        ---GC-CAT--T.............................--TAATTACCTAC-TGTT.............     30
SD_MM251      ---GG-...TGT......................CTTGGG--T...-AGCTGC-TATC.............     30
STM_STM       ---GCC...TGC......................CCTGGA--T...-AACTGC-TATC.............     30
VER_AGM3      ----AG...C--.................................ACATTAC-GATA.............     21
GRI_AGM677    ---GGG...AGA.....................................TTGC-TATA.............     18
TAN_AGMB14    %%-CA-...-CA............TTTTGTATGAC-CCC--T............................     25
SAB_SAB1C     ----AG...C-T.................................................................      9
SYK_SYK       ---GC-...-CT.................................................TTTAGAA     16

ENV_CONSENSUS ??????????????????????????????????????gg?at??????????????t??t??t??g     23
A_U455        ..........................................--A-CT............A-GA-CT-GG-     65
B_HXB2R       ..........................................--GTGGAGATGGGGCACCA-GC-CC-TG-     68
C_UG268A2     ..........................................--A--C............T-AGGCT-TT-     65
D_ELI         ..........................................--C--C............A-GC-CC-TG-     65
E_TN2432      ..........................................--G-CT............T-GA-CC-TG-     65
F_BZ163A      ..........................................--CC-T............T-AT-CC-GG-     65
G_LBV217      ..........................................--A--T............T-GA-CC-TG-     65
O_ANT70C      ......................................ACCT-G............TACT-AGCCAT     65
O_MVP5180     ......................................A-CT-A............TACA-AGCCAT     65
O_VAU         ......................................T-CT-G............A-TT-GGCTTT     71
CPZGAB        ACAGC.........................TTATCCATAATT-CA............A-CA-AACAAT     65
CPZANT        ..........................................--TC-G............GCTT-GC-AAT     44
A_ROD         ......................................CC--T............T-AT-AGCTA-     38
B_EHOA        ...............................................................ACAC-CCT     38
SD_MM251      ......................................-CC--C............T-GC-TT-AA-     47
STM_STM       ......................................-CT--C............T-GC-AT-AA-     47
VER_AGM3      ..........................................--G--A............C-AT-AA-AG-     38
GRI_AGM677    ...........................................-AA............A-AC-AA-AAT     32
TAN_AGMB14    ................................................................................     25
SAB_SAB1C     ......................................CTT-CA............G-AC-GC-TT-     26
SYK_SYK       CTTACATAGTTTGTCTTTTTAGCTTAATAAGTTTAGGATTT...........................     57

                   signal peptide/gp120
                       <- vpU end
ENV_CONSENSUS g?t?tt??taat????tgta??????????????a??tgt?tgtcacagt?tattatggggtacctgt?     65
A_U455        -T-GA-AA----T...----AT...GCA...CAAC-AT---GG-----G--C--C--------------G    126
B_HXB2R       -A-G--GA-G--C...----GT...GCTACAGAAA-AT---GG--------C----------------G    132
C_UG268A2     -A-A--AA-G--T...----AT...GTGATGGGGA-CT---GG--------C----------------G    129
D_ELI         -A-A--GA-G-CC...----GT...GCTGCAGACA-TC---GG--------T-----------G-----A    129
E_TN2432      -T-GG-GA----T...----GT...GCCTCAGACA-CT---GG--T-----T-----------T-----G    129
F_BZ163A      -A-A--GT----C...----AT...GCT...GAAA-CT---GG--------C----------------G    126
G_LBV217      -T-GG-AA----T...----AT...GCCTCAGGTA-CT---GG--------C----------------CG    129
O_ANT70C      -GCT--GA---CCCCA---TTG...AGCCTTAGAC-GC-A-A--CA-----C---GC------G-----A    132
O_MVP5180     -GCT--GC----CCCA---TTGAGTTATAGTAAAC-AT---A--C------T----C------------A    135
O_VAU         -A-AA-CCC-TGT...-TG-GC...TGT...AACC-AC-A-A--C------C----C------------A    132
CPZGAB        CA-T--GC---CCCCA---TTG...ACCTCTGAG...T-A-GG--A-----A--------A--------T    129
CPZANT        CCAG--TA--GAG...AAGGGG...ACGAATGAA...GAC-A---A-----A-TC-----A--C-----C    105
A_ROD         TGCT-GCT--G-A...-A-......TGC...ACC...CAA-A---A--T--T-TC-----C-----CACG     93
B_EHOA        -C-TA-AAGT--C...-A-GGG...TATATGGGCA-GAAC-T------T--C-TC-----TA----C-CA    102
SD_MM251      TG-C-ATGGG--C...-A-......TGT...ACT...CAA-A---------C-T------T-----A-CT    102
STM_STM       TGCT-GCT-G-CC...-A-......TGC...ACT...CA--A------G--A-T------T-----A-CA    102
VER_AGM3      -A-AGGAG--G-G...CT--AT...ACA...AGGC-ACAA-GG--------A-T------A-----A--A     99
GRI_AGM677    AGCAA-AGGG--A...A---TA...GGAATAGGTA-CC---A---G-----G-T------AA-C--A--A     96
TAN_AGMB14    ...............................................................          25
SAB_SAB1C     -C-AAGTGGGTGT...--G-GCTTAGTGTGGTTAGTCCA--A---T-----G-T------TA-C--A--A     93
SYK_SYK       .......................ATGGAAAAAC-ACAA-A---A-----A-T-------A----ACAT    102
```

```
ENV_CONSENSUS tggaaagatgcaac?ac??c?ctattttgtgc?tcagatgctaa???a??????ac?ga????cataat?    115
A_U455        ------------GTT--CA-C-----------A-----------AGC-TATGATG-A--AGTG------G    196
B_HXB2R       -----G--A-----C--CA-T-----------A-----------AGC-TATGAT--A--GGTA------G    202
C_UG268A2     --------A-----A--TA-T-----------A-----------AGC-TATGAG--A--AGTG------G    199
D_ELI         -----G--A-----C--CA-T-----------A-----------ATC-TATGAA--A--GGCA------A    199
E_TN2432      ----G-------GAT--CA-C-----------A--------C--AGC-CATGAG--G--AGTG--C---G    199
F_BZ163A      --------------C--TA-T-----C-----A-----------ATC-TATGAG-GA--GGTA------G    196
G_LBV217      ---G----C---GAT--CA-G-----------A--T--------AGC-TATAGT--T--AAGC------G    199
O_ANT70C      ---G----------AC-AGTA-----C-----T-----------CCT-ACAAGC--T--AAAG------A    202
O_MVP5180     ---G----G---G-AC-AGTA-----C-----T-----------CCT-ACAAGC--T--ACAG------A    205
O_VAU         ---G---------AAC-AA-AT-G--C-----T-----------CTTGACAAGC--T--ACAG------A    202
CPZGAB        ---C-T-----TGACC-GGTA--C--------C-----C-----GGC-CATAGT--A--GGCT------A    199
CPZANT        ----G-A----G--AC-TA-T-----------CA--A----CTCCATGACAAGT--A--GGTG--C---G    175
A_ROD         ------A-------C-TTC-C--C--------AA-CAGA...--TAGG...............G--...A    142
B_EHOA        ------A-----T-A-TTC-C--C--------TA-CAGA...--CAG-...............G--...A    151
SD_MM251      ----GGA----G--A-TTC-C--C--C-----AA-CA-G...--TAGG...............G--...A    151
STM_STM       ----GGA----G--G-TTC-C--C--C-----AA-CA-G...--TAGG...............G--...A    151
VER_AGM3      ------A-CAGCT-AGTACAGGCT--C--CATGA--CCCA-C-CCAG-......................C    148
GRI_AGM677    ------A--T----AGTTCAGGC------CATGA-GCCC...--TACC....................--A    145
TAN_AGMB14    ................................................................ACC---T     32
SAB_SAB1C     ------A--AGTT-AGTCCAAGCC------AAGA-CCCC...--TACG...................--CC    142
SYK_SYK       ---G--------TATG-AC-T--C------A-CA-CTCCCA---AGG-......................G    151

ENV_CONSENSUS t?TGGGc?aCacatgccTgtgt?cc???????acaga?ccc??????aatcca?aa???Gaa?ta???cT    160
A_U455        -C-----T--------------A--C.....------C--------..--C---C--..---A--GATT-    251
B_HXB2R       -T-----C--------------A--C.....------C--------..--C---C--..---G--GTAT-    257
C_UG268A2     -C-----T--------------C--C.....------C--------..--C---C--..---A--GTT--    254
D_ELI         -C-----C--------------A--C.......--G--C--------..--C---C--..---A--GCA--    254
E_TN2432      -C-----C--------------A--C.....------C--------..--C---C--..---A--TAC--    254
F_BZ163A      -C-----T--------------A--C.....-----T--------------C--..---G--GTT--    251
G_LBV217      -C-----T--------------T--C.....------C--------------C--..---A--TCT--    254
O_ANT70C      -T-----AT----A--------T--T.....-----C--------.-C----T-T..---TATCCAT-    257
O_MVP5180     -T-----AT----A-----C--T--T.....-----T--------------C-T..---T-TCCA--    260
O_VAU         -T-----A-----A-------T--C.....-----C--------.-G----A-T..---TATGAG--    257
CPZGAB        -T-----C-----G--A-----A--T.....------T--------.-G---TC-G...---G--TTT--    254
CPZANT        -A-----A--TACCAGT-----G--A......-T---T--A......G----TATT...-T-G-TAGG--    230
A_ROD         CT----GA--CATACAG--C...TTG......C-T--CAAT......G--GATT-TCAG---A--ACTT-    197
B_EHOA        CC----GA--TGTACAA--C...-TC......C----TAAT......G--GACT-TACC---A-CCAAT-    206
SD_MM251      CT----GA---AC-CAG--C...-TA......C----TAAT......GG-GATT-TTCA---T-GGCC--    206
STM_STM       CT----GA---AC-CAA---...-TA......C----CAAT......GG-GATT-CTCA---T-GGCAA-    206
VER_AGM3      -A-----A--TAC-AA--CG...ATA......C----TGAT......C--GACT-CACA--GG--CCAT-    203
GRI_AGM677    -G-----A--CACCAA---CA-A--A......GAT--T-AT......G--AATAC-...--GG-GCCT--    200
TAN_AGMB14    -A-----C---ACAAAT---A-A--A......GAT--T-AT.......---TACAC-...---G-GCAAT-     87
SAB_SAB1C     -A-----CT-TAC-AAT--CA-C--C......GAT--TGAACCAGAAGGAA--AT-GCA---G--CCCA-    206
SYK_SYK       GA-----T--TA-GAAT-----G--GTCAGCA..................GATC--ATA---G--AGGG-    203

ENV_CONSENSUS ????AAtgt?aca???Gaaaa?TTt?a?at?Tgg??????aaa??????Aat?a????atg?????????    192
A_U455        GGTC-----G---...---G-A---A-C--G---.........---.........---A-C...---.....    294
B_HXB2R       GGTA-----G---...-----T---A-C--G---...................---G-C...---.....    300
C_UG268A2     GGAA-----A---...----GT---A-C--G---.........---.........---G-C...---.....    297
D_ELI         GGAA-----G---...-----C---A-C--G---.........---.........---A-C...---.....    297
E_TN2432      GGAA-----A---...-----T---A-C--G---........--C........---A-C...---.....    297
F_BZ163A      GGAA-----A---...-----T---G-T--G---.........---.........---AGC...---.....    294
G_LBV217      GAAC-----A---...-----T---A-C--G---.........---.........---A-C...---.....    297
O_ANT70C      GCAC-----G---...--TG-C---A-T--A---.........---.........---T-C...---.....    300
O_MVP5180     AGGC-----G---...--T--C---G-T--A---.........---.........---T-C...---.....    303
O_VAU         AAAA-----G---...-GT--A--CA-T--A---.........---.........---T-T....--A.....    300
CPZGAB        TCCA-----A-T-...---TCA---A-C--G---.........---.........---A-T...---.....    297
CPZANT        C...---ACCT--...-TCTGG---A-TGCT-AT.........---.........---T-T...---.....    270
A_ROD         G...-----A---...--GGCT---G-TGCA---........--T..........---.....-CA.....    234
B_EHOA        A...---A-A---...--GGCT---G-TGCA---.......G-T..........---.....-CA.....    243
SD_MM251      T...-----T---...----GC---G-TGCT---.......G-G..........---.....-CA.....    243
STM_STM       A...-----C---...--GGCT---G-TGCC---.......G-C..........---.....-CA.....    243
VER_AGM3      A...--CA-C--T...---CCA---G-AGCA---GCTGAC-G-........--CCCC...T-A.....    249
GRI_AGM677    A...--CA-T---...---GCT--CG-GGCT---.......G-T.........---CCG...C-.....    240
TAN_AGMB14    A...-----CT--...--G--A---G-AGCG---........---GACAGG--CCCT...T-.....    133
SAB_SAB1C     ACCT---A-C---...-----A---G-TGCT---.........---.........CGGAACCCT    252
SYK_SYK       C...--CA-T---GGA---T-C---CCAGCA---........--T.......-GCAGTCAC---.....    249
```

```
ENV_CONSENSUS ???gTa???gaacaaatgca?ga?gAtaTaat?agttTaTggga?caaag?ctaAagCCaTGTGtaaAat    251
A_U455        ...---...--T--G-----T--A--------C---C-------C-----C-----------------G-    358
B_HXB2R       ...---...-----G-----T--G--------C-----------T-----C-------------------    364
C_UG268A2     ...--G...--T--G-----TC-G---G----C---C-------T-----C-----------------G-    361
D_ELI         ...--G...-----G-----T--G--------C-----------T-----C-----A-------------    361
E_TN2432      ...---...--G--G-----G--G---G----C-----------T-----T-----------------G-    361
F_BZ163A      ...---...-----------TACA--------C-----------T-----C--------------G--G-    358
G_LBV217      ...---...-----G-----T--G--------C--C--------TG----C-----------------G-    361
O_ANT70C      ...---...-----------G--A--C--T--T-----------C--G--T--T--A--T-----TC--A    364
O_MVP5180     ...--G...--C--------T--A--C--C--T-----G-----A--G--TT-------T----AG---A    367
O_VAU         ...---...--C--------C--A--C--T--AGA---G-----C--G--TT----A--T-----TC--A    364
CPZGAB        ...--G...--C--------T--A--C--T--C---C-C-----T--G--T--C-----T--------G-    361
CPZANT        ...---...---AGT---ACA--A-----GN-ACAA----TCC-A-----C-AT---------------C    334
A_ROD         ...---ACA------GCAATA--A---G-CTGGCA-C---TC--GAC-TCAA----A--------C---C    301
B_EHOA        ...--GACA--T--GGCAACAA-G---G-GTGG---C-C-TT--GACCTCAA----A--------T----    310
SD_MM251      ...--CACA-----GGCAATA--G--CG--TGGCAAC-C-TT--GACCTCAA-------T----------    310
STM_STM       ...--CACA------GCAATA--G---G-CTGG-ACC-T-TT--AAC-TCCA-T--A--T----------    310
VER_AGM3      ...---...-C----GCAGGAAGTA-----CACCTGC---TT--G--G-CT--G----------------    313
GRI_AGM677    ...---...A-----GCAG-GAGTA-----CATCTAC-C-TT--A----CGA-G-G---T-----T--GC    304
TAN_AGMB14    ...--G...-C----GCAG-AAGCA-----CATCTGC-T-TT--AAGT-CCT-G-----T-----C--G-    197
SAB_SAB1C     CTA---...-GC---GCAG-GAGTA-C---CACCT---G-TT--AAGT-CTT-------------T--GC    319
SYK_SYK       ...A--AGGC-G-----ATTA--G--C--GTCAGC-----TCTTG---GCCAAT-GA--C-----G--GC    316

ENV_CONSENSUS TaaC?cc??T?TGtgTtacaaTgaa?TGta?taa?????a?????????????????????????????    280
A_U455        ----C--TC-C--C-----GT-AG-T--CCA---CATCACC.............................    399
B_HXB2R       ----C--AC-C-------GTT-A--G--C-C-G-TTTGA-G.............................    405
C_UG268A2     -G--C--AC-C--------TT-A--T----CA--TGTTA-T.............................    402
D_ELI         ----C--AC-C-----C--TT-A--C----G-G-TGAATTGAGG..........................    405
E_TN2432      ----T--TC-C--C-----TT-A-TT----CC--TGCTA-GTTGACC.......................    408
F_BZ163A      ----C--AC-C--------TT-A--T----C---TGCCATT.............................    399
G_LBV217      ----C--TC-C--------CT-A--C----C---TGCAATC.............................    402
O_ANT70C      -G--TTTCC-G-----ACA----G-G----CA--CATAGCT.............................    405
O_MVP5180     -G--TTTCT-A-----ACA------C---GTAG-TCTGC-A.............................    408
O_VAU         -G--TTTCT-G-----ACA------T----CAG-TATCA-A.............................    405
CPZGAB        -G--C--TT-A-----A--CT--C-G--C-G---GGCTA-CTTTAGC......................CA    410
CPZANT        ----A--TA-G---A-A-A------T----CAGGATAC................................    372
A_ROD         ----A--TT-A-----AG-------A--C-GC-GCACAG-GAGCAGCACAGGG.........AACAACAC    362
B_EHOA        ----C--AC-G-----G--------G----A---GACGTGGAGCTCA.......................    357
SD_MM251      --T-C--AT-A--CA----T----GA--C-A---AAGTG-GACAGATAGATGGGGA...TTGACAAAATC    377
STM_STM       -G--T--AC-A---A--G------GG----A---GAATG-AACAGACAAATGGGGATTGACAGGGAAAAC    380
VER_AGM3      --T-A--TT-G--CA---A----TCC---GTAG-ATTGA-CTCCTCTGAGCCTACC......ACCACTCC    377
GRI_AGM677    -CT-C--CA-A---A---A----TCC---GTAG-GCTG................................    342
TAN_AGMB14    ----A--AA-G--CA-A-G------T--C-C--GGCTTCCCTCCCCTACTCCC.........TCCTCCTC    258
SAB_SAB1C     -CT-A--AA-G--CA-A-A------C--CTA-CGCTTAG-AGGAGGA.......................    366
SYK_SYK       --G-A--CA-G---A---GG---TTG----CATTGGAC................................    354

ENV_CONSENSUS ???????Aa????ac????a?????ac?ac?a???c?ac????a?????????????????????????    294
A_U455        ....ATT--CAAC--C...-ATAAC-AC--C-ATATC--AGATGGAGTG.....................    441
B_HXB2R       ........--TGAT--T...-AT...--C-AT-GTAGT-GCGGG-GAATG...ATA...............    444
C_UG268A2     ....GTT--TATT--C...-AT...-ATG-T-ATG-T--CAAT-GCCCCTACGAA...............    447
D_ELI         ....AAC--TGGC--T...-TG...GGG-AC-ATGTC--TACA...........................    438
E_TN2432      .AATGCT--TTTG--C...-AT...GTC-AT-ACATA--CAATGTCTCTAACATA...............    456
F_BZ163A      ........--TGCC--T...GCC................................................    411
G_LBV217      ....GCT--TGTA--C...-AC...-GC-AT-AGA-T-GTAAC-ACAGCAGCCTT...............    447
O_ANT70C      ....GGA-CA...--A...-AT.................................................    417
O_MVP5180     ....ACA--TAAA--A...GGCCTATTA-ATGAGA-A-TAAAT...........................    444
O_VAU         ........--TAGT-TT...-AT...--C--A-ACAGTC-CTTA-ACTCA....................    441
CPZGAB        GGCAAAA--CCTA--A...-ACCAG--AT-TTCTC-GC-TCTC...........................    450
CPZANT        ........--TGGA--A...CCT...--A--ACCAAGT--AACA-CAAGTACAGTAACACCAAAG......    423
A_ROD         AACCTCA--GAGC--A...-GC...--A--C-CAA-C--ACCC-CAGACCAGGAGCAAGAGATA......    420
B_EHOA        .GCTAGC--AGAG--C...-CT...--GT-CTCCG-GT-CCTC-GATCTTCTACTCAGACCCTACTC...    417
SD_MM251      ATCAACA-CAATA--AACAGCAGCAC-A--ATCAG-AC-AGTATCAGAAAAAATAGACATGGTC......    441
STM_STM       AGTAACA-CAGTG--A...CCA...--AG-AGCAG-AG-AGCA-CAAAGCCAGAG...TTAGTA......    435
VER_AGM3      TAAAAGT-CC...--G...GCCTCA--A--C-ATATC--AGCCTCAACAACCACTTTGCCGTGTGTCCAG    441
GRI_AGM677    ........--TGGT--A...GCC...--G--A-AGG-C--CACT-CTGCAACTACAACAATGACT......    393
TAN_AGMB14    TAGTACA-CAAGG--A...-CA...-GA-ATCCCTGCC-AGGC-CAACA.....................    301
SAB_SAB1C     .GCAGCA-CAACA--A...TCA...C-AT-A-CAT-A--AGCACGCCCCGAGGTAGTGAGTGTG......    423
SYK_SYK       ........--CAGCC-T...GCT...--A-GT-CCC-T--TACA-GTCCCCCTACTACACCTCCA......    405
```

```
ENV_CONSENSUS ??????????????????????????????????????????a??g?a??a???????????gaagaaa    305
A_U455        ..................................................AGG............-------    451
B_HXB2R       ..............................................-TG-AGAA-.............-G---G-    460
C_UG268A2     ..............................AATGGGAAACTC-TG-A-CAG............-G-----    475
D_ELI         ...............................................-AGGAG............A---G--    451
E_TN2432      ....................................ATAGGA-ATAT-AC-.............--T---G    478
F_BZ163A      ..................ATT...CAAGACGGGGACCTGAAGGAA-G-CC-.............-GG-C--    448
G_LBV217      ........................GGGAATAGCACTGTG-GTAGCAT-............-------    478
O_ANT70C      ...................................................GA-............A-CCTT-    427
O_MVP5180     ......................................................................--G-    448
O_VAU         ..........................................-ACAATAC-............A----GG    457
CPZGAB        ......................................................................----    454
CPZANT        ACAACAACACCA..................................AT-GT-.............--T-GC-    448
A_ROD         AGTGAGGATACTCCATGCGCACGCGCAGACAACTGCTCAGGATTG-G-GAG............-------    478
B_EHOA        AACGAAGATAGCAAATGTATTCAAAATGACAGCTGTGCAGGG-TA-G-CT-............-------    475
SD_MM251      AATGAGACTAGTTCTTGTATAGCTCAGAATAATTGCACAGGCTTG-A-CA-.............--GC---    499
STM_STM       AATGAAACAAGCTCCTGTGTAAGTAACAACAACTGCACAGGCTTG-A-GA-...............---TC-C    493
VER_AGM3      AACAAGACAAGTACTGTGTTAGAATCATGTAATGAAACAATC-TA-A-AAGGAATTAAATGAA--GCCTG    511
GRI_AGM677    ACCCCCTGTCAGAATTGCAGTACAGAGCAGATAGAAGGAGAA-TG-C-GAG...............---CC-G    451
TAN_AGMB14    ......AACGATGAGTGCAATAGCCACTTGGTGACAAATAGC-TG-A-TTT...............---A-C-    353
SAB_SAB1C     GGCTTTAATGACTCAGTGATAGAACAAGAG............-TG-A-AA-...............---C-GG    469
SYK_SYK       AATGAAACGTGG.....................TGGGGAGAC-ACAGTAC-.............--GCC--    442

ENV_CONSENSUS tga?a??????aa?TGt?c?TTtaatatgac?acag?a?taagaGataagaa??agaaa?a??a?gca?t    355
A_U455        ---A-........--T--CT-T--C--------C----A-C-------------AC-----GTAT-TT--C-    515
B_HXB2R       -A-A-........--C--CT-T--C-----C-GC---AGCA------G----GTGC-----G-AT-T---T-    524
C_UG268A2     -A-A-........--T--CT-T--C---G-A--C----A-A-------------AC---C-GCGC-T---C-    539
D_ELI         ---A-........--C--CT-T--C---G-A--C----T-C---A---------GC--C--GTAT-T---C-    515
E_TN2432      -A-G-........--C---T-T-----------C----A-C-------------GC----GGTCC-T---C-    542
F_BZ163A      -ACA-........--C---T-T--C---G----C-TG-A-G--%--------C-GC-----GTAC-T---C-    511
G_LBV217      ---A-........--C--CT-T--C-----A--C----A-G------------GTA-----G-AT-C--TC-    542
O_ANT70C      ---AG.......--G---GAG------G-A--C--T-TTA-C-A---C--A--GG-----A-AC-G--TC-    491
O_MVP5180     ---G-........--T---AGT------G-A--T----TCC-C-C---C--A--GG--C--A-AC-G--TC-    512
O_VAU         ---A-......C-G---GAC------G-A--T----TGC-C-A---C--AC-GG-----A-AC-G--TC-    521
CPZGAB        ---A-........--C---T-A------G----A----A-T-------------GA-AC-GGTCT-TT--C-    518
CPZANT        ---AGCTACAAG-A---AAC------CA--GC----G-T-T-A---------AC-A---ATGA-A--CA-    518
A_ROD         C--TC........--T--CCAG--C--------AGG-TT-GA------------AA-AC-GT-TA-T-A-AC    542
B_EHOA        ---T-......G-C---CAA--C--A-----AGG-TT-AA-------G--TCAA-AC--T-TA-G-ACAC    539
SD_MM251      ---T-......-GC---AAA--C-CC-----AGGGTT-AA------C----CAA--G-GT-CA-T-A-AC    563
STM_STM       -AGT-......GGT---AAA--C--------AGG-CT-AA---------A--GAGAG-GT-CA-T-A-AC    557
VER_AGM3      CTTCT........--T---A-A---GCA---G-AGGGTATG--------C----AA----GT-TTCA-TGG-    575
GRI_AGM677    CATCC........--C--CA-T---GCA--TG-AGG-TATCA-------GTA--AA----TT-TAGCATGAC    515
TAN_AGMB14    GTTCT......-TG--CA-C---GCA---G-AGGGTACA-G--G---C----AA---C-T-TA-CAGCAC    417
SAB_SAB1C     CC-TG........--C---T-C---GC----G-AGG-TATAG---G---GTA--GA-A--TT-CTCAA--G-    533
SYK_SYK       GATTC........--T--CT-T-----CT----AGGG-GGT-C-AG-----A--AC--C--T-TAGA--TT-    506

ENV_CONSENSUS ?Tt?tAt??a???GAt?T?gt?????t????ga????aa?aa?????????????aata??Ag???????    380
A_U455        T--T---AG-CTT---A-A--ACAAA-C...A-TAAA-CTG-T.............--C-AT--T......    564
B_HXB2R       T--T---AA-CTT---A-AA-ACCAA-A....-T.....-TG-T............-C--CC--C......    570
C_UG268A2     T--T---AA-CTT---G-A--ATCAC-T....-A...GGG--T............-G--AC-CT......    585
D_ELI         T--T---AG-CTT---A-A--ACCAA-A....-C...--TG-TAGTAGT...ACC----GT-CCAAT...    573
E_TN2432      T--T---AAGCTT---A-A--ACAAA-T...-GA...G-T--G............----GT--TGAG...    591
F_BZ163A      G--T%--AG-CTT---A-A--ACCAA-T...AGC...--TGGAAGA............---GGC--TGGAGAC    565
G_LBV217      T--C---AG-CTT---G-G--ACCAA-T...A-TGGT--TG-A.............----GT--TACAACT    597
O_ANT70C      A--C---GT-TCA---T-GA-GGAAC-G...A-TGAG-CA-GCAGCACAAATAAG-CA-AC--CAAA...    555
O_MVP5180     A--C---GT-TCA---C-GAGTAAGG-TAAT--C...TCA--TGCA......GTA---GGA-CAACA...    570
O_VAU         A--C---GTGACA---T-G--TAAGA-T...A-CGCC-CATCA...............---GAA-CAATG...    573
CPZGAB        A--T---GTGGAG---G-G--AAACC-A...-GG...--TG-G.............--C-AC-CA......    564
CPZANT        A--T---AA-GGA---C-TA-GAAGTGTCAG--C...--C--T............G-G-CT-ACTGC...    570
A_ROD         A-GG--CTC-AAA---G-G--TTGTGAGACAA-T...--T-GCACA...........---CAG-CCCAG...    597
B_EHOA        C-GG---AA-CAA---T-A--GTGTGAAAAG-GGACA-GG-GT...............---GAA--CAAA...    594
SD_MM251      T-GG--CTCTACA---T-G--TTGTGAACAA-GG...--T-GCACT......GAT---GAA--CAGA...    621
STM_STM       G-GG--CTCTAGC---C-AA-CTGTGAACAAA-TGTC-CTGGC............G-AGAG--CAGA...    612
VER_AGM3      G-GGA--GATGCA--AA-CA-GTGTAAGAAG-GT...--C--T............TC--AC--AGAA...    627
GRI_AGM677    C-GG---GATCAG--GT-A--CTGCAAT...A-T...--A-CAGGAAGT...GAA--GGGA--TAAG...    573
TAN_AGMB14    A-GG---GATGCA--AC-CA-GTGTGAACCG--GAGC--A--G..............--CTCA--GGGG...    472
SAB_SAB1C     A-GGG--GATCAA--GG-G--TTGTGAAGAG-GAAGGG-A--ATCA..........--CGCA-CACATACT    594
SYK_SYK       C--C---AAGGAT--CT-GA-GAAAGAGGAA-GC.......................----GT--CTAT...    552
```

```
ENV_CONSENSUS ??????Tataa??TaataaatTGtAAtacctcagtcAT?a?acaagC?TGtccaAAg???tctTttga?c    435
A_U455        .....----GAC--------------------AC---T-C---G--T-----C---GTG--C-----G-    628
B_HXB2R       .....-----GT-G-C--G------C-----------T-C---G--C---------GTA--C-----G-    634
C_UG268A2     .......--C-GAT---------------------C---A-CG-----C---------GTCA----G--C-    649
D_ELI         .....----GGT---------------------C---T-C---G--T---------GTA--C-----G-    637
E_TN2432      .....----GGT----------------T--------T-AG--G--T---------ATA--C-----T-    655
F_BZ163A      ...TAC--C-GGC----------C---------AC---T-C---G--T--------AGTA----GG-GT-    632
G_LBV217      ...GAT----GGC--------------G-----AC---T-A------T---------GTGA-------C-    664
O_ANT70C      ...ATG----CAT---CT--------CT--A--AC---C-CG-----C---------GTA--------A-    622
O_MVP5180     .....----TGT---CT--------CT--A--A-T--C-AG--G--C-----G---GTAAG------G-    634
O_VAU         .....----GAT----T--------CT--A--AC---C-GG--G--C---------GTA--------G-    637
CPZGAB        .....----GGA----T-----C-----TA---C---A-C------A--------AACG--C-----G-    628
CPZANT        .....---T-CT--TGGC-C--C--C---A--ACT--C-C----T-C---GA----TCTA-------A-    634
A_ROD         .......-G-T-CA-G-ACC----C--C--A--------C-C-G--T-A---GAC---CAC-A--GG--TG    661
B_EHOA        .......-G-T-TA---A--CC--------A---A-T--CCA-G--T-A---GAC--ACAT-A--GG--TA    658
SD_MM251      .......-GCT-CA-G-ATC-C-----C--T--T--T--CCA-G--T-T---GAC--ACAT-A--GG--TA    685
STM_STM       .......-GCT-CA-G-G-C----C------AGT--A--TCA-G--T-C---GAC--ACAT-A--GG--TG    676
VER_AGM3      .......-G-T-TA-G--TC--------GAT-----T--A-A-G----C---GAT--AACA-A--GG--TG    691
GRI_AGM677    ...GAT-G-T-CA-G---C--------GAT-----G--A-A-G----T---GAT--AACA-A--GG--TA    640
TAN_AGMB14    .......-G-T-CA-G--CC-------CGAT--------A-AGG----T--CGA---AACT-AC-GG--TA    536
SAB_SAB1C     GTGGGA-G-T-CA-G--CC----------------A--C-A-G-G--T---GAT--AACT-A--GG--TA    664
SYK_SYK       .....---T-CC-GT--C----------AAGT--T--ATCCGC---A---GAG--ACAGA-C---C-A-    616

ENV_CONSENSUS c?aTtcc?aTaca?TAtTGtgCtCCagC?GgttatgC?aT?cTaAaaTGtaatgatac?aa?TttaaaGG    497
A_U455        -A-----C-----T--------C-----T----T---G--T-----G-----G---C-TG-G--C--T--    698
B_HXB2R       -A-----C-----T--------C--G--T----T---G--T------------A---AG-CG--C--T--    704
C_UG268A2     -A-----T-----T--------------T--------G--T-----G------A---AA-CA-----T--    719
D_ELI         -A-----C-----T--------C-----T----T---G--T-----G----GA----AG--G--C--T--    707
E_TN2432      -A-----T-----T------A-------T--------G--TT-T--G----------AG--T--C--T--    725
F_BZ163A      -A-----C-----T--------------T--------G--TT----G----------AA--G--C--T--    702
G_LBV217      -A-----C-----T-----C--------T--G-T---G--T----C-----GG----AG--G--C--T--    734
O_ANT70C      -A-----A-----C--------------A--A-----T--CT-T--G-----CAGC--AG-A-----T--    692
O_MVP5180     -C-----C-----C------------A-A--A-----C--CT-T--G--------C--AG-C-----T--    704
O_VAU         -C-----C-----C---------------G--A-G---C--CT-T--G--------A--AGGA-----T--    707
CPZGAB        -T-----A-----C--------A-----A--G-T---A--T--T-----C-------AAG-C---TC---    698
CPZANT        -A-----T-----T--------------A--A-----T--AT-G-G----G-A---GAGG-T----C---    704
A_ROD         -T--AAGGT-TAGA--C-----A---C-G--------CC-AT---G------------C--T-A-TC---    731
B_EHOA        GTT-AAGAT-TAGG-----------CC-A--A-T---TT-G----G---C--------T--A-A-TC---    728
SD_MM251      -T---AGAT-TAGG--------A--TC-A--------TT-G--T-G---------C--A--T-A-TC---    755
STM_STM       -CT-AAGGT-TAGG--------A---C-T--------TT-G--C-G---C--------A--T-A--C---    746
VER_AGM3      AGT-AAGAT--AGG--C-----C--G--A--G-T---TT-AT-------C--C---TATG-T-A-GC---    761
GRI_AGM677    -TT-AAGAG--AGA--C-----A-----A--G-----TT-G----------------AGG-T-A--G---    710
TAN_AGMB14    -CC--AGGT--AGA--------A-----T-C-GT---T--CT-G--------A-----T--T-AC-C---    606
SAB_SAB1C     -CT--AGGC-TAGG--------------A--------TT-G----G----GCA-----AG-T-A-TC---    734
SYK_SYK       -CT----G--T--A--C---------C-G-----CT-AT-GT-G-----C--------A--T---G----    686

ENV_CONSENSUS aa?a???gg???aTGta??AAtgT?a??gtagt?ca?tgtacacatggaat?aagccaac?gtaac?act    550
A_U455        --A-...--GCC---C-GG-----C-GCAC---A--A--------------C------GTG---T-A---    765
B_HXB2R       --C-...--ACC-----CA-----C-GCAC---A--A--------------T-G----GTA---T-A---    771
C_UG268A2     --C-...--ACC-----AT-----C-GCAC---A--A--------------T--A---GTGA--T-A---    786
D_ELI         --C-...--CCC---C-CA-----C-GCAC---A--A--------------T-G----GTG--GT-A---    774
E_TN2432      G-C-...--GCC-----AA-----C-GCTC---A--A--C------------T------GTG---T-A---    792
F_BZ163A      G-C-...--GCC---C-AG-----C%GCAC---A--G--------------T--A---GTG---T-A---    768
G_LBV217      --C-...--ACCG----AA-----C-GTAC---G--A--------------T------GTG---T-A---    801
O_ANT70C      --C-...--CAC---C-GA--CA-A-CG-----TACT-----------C--T-G------A----GT---    759
O_MVP5180     --C-...--CCT---CCAC---A-TTCA--G--TACT-----------C--C--------A----GT---    771
O_VAU         --C-...--TCTC----AA--C--T-CA-----TACT-----------C--C--A-----A----GT--C    774
CPZGAB        --A-...--TAA-----CT-----A-GTAC---G--T--------------T--A---GTG--T--C---    765
CPZANT        -GT-...--GATG----AA-----CTCA-----A--T--C--T--------A-GC----TG--GG-A--A    771
A_ROD         CTTTGCACCCAAC---TCT--A--AGTA-CTTCTACA--C--CAGG...--G-T-GA---GCA---TT-C    798
B_EHOA        CTTCATGCCCAAC----GT--G--AGTA---TCT...CTGTACAGA...--G-T-GA---ACAG--CT--    792
SD_MM251      CTTTATGCCTAA----TCT--G--GGTG--CTCTTCA--C---AGG...--G-T-GAG--ACAG--TT--    822
STM_STM       CTTTGCACCTAAT--CTCT--G--AGTA--CTCCTCG-----TAGA...--G-T-GAG--ACA---TT-A    813
VER_AGM3      GTTTAAGACAAAC---TCT-----TTCA--G--G--T-----TA-CTTG--A--TA----A--G--T---    831
GRI_AGM677    CTTTGCTCCAAAG--C-AG-----TTCA-----G--T-----TAGATT---C--TA-T--TA----T--A    780
TAN_AGMB14    CTTT...--CAC-----GA-----GTCA-----AAGT--C---GGCTT---G--TA-T--A----GCT-A    673
SAB_SAB1C     -CAT...AAGGC-----GG-----A-CA---TCAGCT--C--TAGGTT---C--TA-T--T----GC--A    801
SYK_SYK       -GAC...-ATGTC----CA-----A-CT-C---ATCA--C-----G......................    729
```

```
ENV_CONSENSUS caa?t??t??t????aatgg?agt?tagcagaagaaaaaataa?aat?ag?t???aaaatat??caga?a    600
A_U455        ---C-GC-GT-A...-----C---C--------AG-G------GG--T--A-CTG-----T-CA--A-C-    832
B_HXB2R       ---C-GC-GT-A...-----C---C-----------G-GG--GT---T--A-CTGTC---T-CA-G--C-    838
C_UG268A2     ---C-AT-GT-A...-----T---A-----------G------T---T--A-CTG-----C-GA--A-T-    853
D_ELI         ---C-GC-GT-G...-----C---C-----------G-GG-C-T---T--A-CCG-----C-CA--A-C-    841
E_TN2432      ---T-GC-GT-A...-----C---C-----------G-G----T---C--A-CTG-----C-CA----C-    859
F_BZ163A      ---T-GT-GT-A...-----C--CC-----------G-T----T---C--A-CTC-------CT----T-    835
G_LBV217      ---T-AC-GC-G...-----C---T-----------G------TG--T--A-CTG----CT-CA--A-C-    868
H_VI557                                        A-CC-TG---T---T--A-CTA-------CA----C-     37
O_ANT70C      --GC-AA-AT-A...-----G-CAC-CT-TA---G--------G---G-TGGCAA--G----TTTG--AG    826
O_MVP5180     ---C-AA-AC-G...-----G-CAC-CT-TAG------G----G---T-TGGGAA-------TA----AT    838
O_VAU         ---C-AA-AC-A...-----G-CAC-CT-TA---G---T----C---C-TGGGAA-G-----TT----C-    841
CPZGAB        --GT-AC-TA-T...-----A---T----T----G---C----CTG-A--AGTGG-G----AGT--A-A-    832
CPZANT        TGGT-AC-AT-A...-----A-C-.................................................    792
A_ROD         AC-TGGT-TGGCTTT-----C-C-AG------GA-T-G--C-TAT--CTAT-GGC-TGGC-GA...--T-    865
B_EHOA        AC-TGGT-TGGCTTC-----T-CAAGG-----A-C-GG-C-TAT--CTAT-GGC-TGG--AA...--C-    859
SD_MM251      ACTTGGT-TGGCTTT-----A-C-AG------A-T-G--CTTAT--TTAC-GGC-TGG--GG...--T-    889
STM_STM       ACTTGGT-TGGCTTT-----A-C-AG------A-T-G--C-TAT--CTAT-GGC-TGGC-GG...--T-    880
VER_AGM3      GG-C-GT-GT-G...-----G--CTACT----GA-TCG--CCCAG--AT-GCAGA--C---GAGT-AGC-    898
GRI_AGM677    GGGA-AGGAT-A...-----T---AG-T-----A-T-G--C-GAG--AT-GCAGA--GGAGGAAAT--T-    847
TAN_AGMB14    GCTT-TGGTC-A...-----T---CA------GA-T-G--C-GA---AT-GCAAA--C--GGAGT-AGT-    740
SAB_SAB1C     GG-A-AGGAA-T...-----T---TAT-T--C-A-T-G--C-GA---CT-GCAGA-----GGCAATTCC-    868
SYK_SYK       ...................................---TTC-AT-C-T-GGCTAGCAC-TGGT-C.......    759

ENV_CONSENSUS atg??aa?a??at?atagt?ca?ct?aataaatct?????????????????????????????????????    626
A_U455        ---CA--A-CC--A-----A--G--TGTC--TC--......................................    867
B_HXB2R       ---CT--A-CC--A-----A--G--G--C-C-----.....................................    873
C_UG268A2     ---CC--A-TA--A-----A--G--T--C-------.....................................    888
D_ELI         ---CT--A-AC--A----CA--T--T---G-----......................................    876
E_TN2432      ---CC--A-CC--A-----G--C--T---------......................................    894
F_BZ163A      ---CA--A-CC--A-----A--CT-T---G-----......................................    870
G_LBV217      ---CC--A-AC--A-----G--AT-T---------......................................    903
H_VI557       --ACC--A-AC--A-----A--G--T--A-GTC--......................................     72
O_ANT70C      G--GA--A-AT--C-----GACC--A--CTCTA-C......................................    861
O_MVP5180     CA-CA--G-AT--C-----AACC--A--C-CTC--......................................    873
O_VAU         G--GGG-G-AC--CC--A-AACC--A----CTAA-......................................    876
CPZGAB        --ACTG-TGTCTGG-----A--G--AGTAG--G-A......................................    867
CPZANT        .................TAC--AACA--C-CT--AGTAGTAATGAATGGTCGCAAAAATGAATCTGTG..    843
A_ROD         --...-GA-CT--C--CAGC...T-A--C----A-TAT...................................    897
B_EHOA        --...-GG-CT--C---AGC...T-A---TC--A-TAT...................................    891
SD_MM251      --...-GG-CT--A--TAGT...T-A-----G-A-TAT...................................    921
STM_STM       --...-GG-CT--T--TAGC...T-G--C--G-A-TAC...................................    912
VER_AGM3      --...G-CTCAG-GT----GTTAT-T------CA-TAC...................................    933
GRI_AGM677    --...G-T-CAG-T---A-AA-GT-G-----G-T-TAC...................................    882
TAN_AGMB14    ---ACTCTGTC--T---AAG...--T-----GCA-TAC...................................    775
SAB_SAB1C     ---ATTCTGTT--T--TAGG....-C----G--A-TTT...................................    903
SYK_SYK       ....................--A--G---GGGA-CTACAAAGCAAAAGATAAAGTTAGATTTATCAAACA    809

ENV_CONSENSUS ?????????????????????????????????????????????????????Taa?aaT?a??TGta?aAGa    641
A_U455        ...................................................G---A---C-AT---TCC---    888
B_HXB2R       ...................................................G--GA---T-AT----C----    894
C_UG268A2     ...................................................G--GA---T-AT---GCG---    909
D_ELI         ...................................................G---A---T-CC---GC---G    897
E_TN2432      ...................................................G--GG---C-AT----CC---    915
F_BZ163A      ...................................................G--CAG--T-AT----C----    891
G_LBV217      ...................................................A--GAT--TGTG----CC---    924
H_VI557       ...................................................G--CC---T-AC----CC---     93
O_ANT70C      ...................................................C---AC--G-CC---GA----    882
O_MVP5180     ...................................................A---AC--G-CC--C-T----    894
O_VAU         ...................................................A---C---AGCA---GAG---    897
CPZGAB        ...................................................G-T-GTC-A-AT---CAT---    888
CPZANT        ...................CTTGTAAGATTTGGAAAAGAATTCGAAAACT---C---T-CA----T----    894
A_ROD         ...............................................AATC-C-GTT-GCAT----AG--G    921
B_EHOA        ...............................................AATC-G-C---GCAC----A---G    915
SD_MM251      ...............................................AATC---C---G-AA----G----    945
STM_STM       ...............................................AATT---C---G-GC----G----    936
VER_AGM3      ...............................................AATC---C-G-T-CT--C-A----    957
GRI_AGM677    ...............................................AACT-G-C-G-G-GA--CCG----    906
TAN_AGMB14    ...............................................AATC-G-C---ATTG----G----    799
SAB_SAB1C     ...............................................AATT---CT--C-GA----G----    927
SYK_SYK       AAAGGATAAAAATGAGTCAGTAATTATTTTAGTTCCAGAAGCACTTAGAT-GCAG--T-TA---GA----    879
```

```
ENV_CONSENSUS ccaggaaat?????????aa?aca?ta??a???ataa??ata??????gg??cagga?tagc?t??tat?    680
A_U455        --TTAC---ACAAGAAAA--T-T-CGT...AGGTAT-GT---.......--AT-----CA---A-TC---G    949
B_HXB2R       --CAAC--C...........T---AG-AA-AGA--CCGT--CCAGAGA--AC----GAG---A...-T-G    952
C_UG268A2     --CAAC---...........T---AG-GA-AGT----GG---.......--AC-----CA-A-A-TC---G    964
D_ELI         --CTATC-A...........T---AG-CA-AGA-C-CCT---.......--ACT---GCA-T-ACTC---A    952
E_TN2432      --CTCC--C...........T---AGRCC-AGT----CTG-R......--AC-----CA--TA-TC---A    970
F_BZ163A      --CAAC--C...........T---AG-AA-GGT---CAT---.......--AC-----CG---A-TT---G    946
G_LBV217      --CAAC--C...........T---AG-AA-AGT---CAC---.......--AC-T---CA---A-TA---G    979
H_VI557       --CAAC---...........T--GAG-AA-AGT----GT---.......--GC-----CA---A-TC---G    148
O_ANT70C      ---CA--TA.........G-C-T-CA-...GAG--G-GA---.......--T...CC-A-G--C-GG--CA    931
O_MVP5180     GA-----T.........GCAGAGG--CA-GAT---TAT-C-......--TC--ATGAG-TGGCGCAG-A    949
O_VAU         ---------.........C-G---A--CA-AAG----TGGC-......--T...CC-A-G--T-GG--CA    949
CPZGAB        --------C...........T---AGGGG-GAGG-GCAG---.......--GC-----A-GA-C-TT---A    943
CPZANT        ---------.........-GG---G--AG-AATC--CAA---.......--AC-----A-GA-T-TC---.    948
A_ROD         -----G---...........-G-T-G-GAA-CAA----TGC-T......ATGT-----CAT-TG-TTC-CT    976
B_EHOA        --------C...........-A-TGG-TGTGCCA----GA-CC......-TGT----TA-TCTC...-TC.    966
SD_MM251      ---------.........--G---G-TTT-CCAG-C-CC--T......ATGT-T---T-G-TT...-TC.    996
STM_STM       -----G--C...........--A---G-CTT-CCAG-C-CC--T......ATGT----CT-G-TC...-TC.    987
VER_AGM3      --------C...........-A---G-CTT-CCAG---CA--C......ATGG----GC---TG...-T-.   1008
GRI_AGM677    --T--T---...........-A---G-GTTGCCAG---CA--C......ATGG----GT---TA...-T-.    957
TAN_AGMB14    --------C...........-A---G-CTT-CCAG---CC--C......ATGG-----A-G-TG...-TC.    850
SAB_SAB1C     ---------.........--G---G--TTGCCTG---CC--C......ATGG-T--CT-G-TC...-TC.    978
SYK_SYK       ---------.........G-GT-TA-CAA-AAT--TCAAT-GGCTGCA--C.................--T    925

ENV_CONSENSUS ??a???????????????a????a?a?????????????????atagg????aa??taagacaaGC?ta    703
A_U455        TA-CA.........GGTA-A...-T-.....................-----G...G-TA----------AC-    986
B_HXB2R       TT-CA.........ATAGGA...-A-.....................-----A...-TA-G--------AC-    989
C_UG268A2     CA-CA.........GGAG-C...-T-.....................-----A...G-TA----------A--   1001
D_ELI         CT-CAAGA......TCAAGATCA-T-.....................-----A.................-AC-    986
E_TN2432      GA-CA.........GGAG-C...-T-.....................-----A...G-TA----RAG---A--   1007
F_BZ163A      CA-CA.........GGAG-C...-T-.....................-----A...G-CA-C---A-G--CC-    983
G_LBV217      CA-CA.........GGCGCC...-T-.....................-----C...G-TA----------AC-   1016
H_VI557       CA-CA.........GGTG-C...-TC.....................-----A...G-TA-----A----A--    185
O_ANT70C      GC-TGGGA...ATAGGGGGA...-C-..................GC---AAAC-GCTC---GGC---T--    977
O_MVP5180     TG-CACTTAAAAGAAGTA-C...-AT..................-C-TCACCA-GATC---GGT---T--    998
O_VAU         GC-TGGCC...CTTAGTA-T...-C-..................-AG--G...G-TAC---GGC---T--    992
CPZGAB        AT-TA.........GAAA-T...GT-..................G----A...G-TACC---TCT--C--    980
CPZANT        ..-AC.........GTAG-A...-T-GCA...............-C---A...G-CACT--GA----G-T    986
A_ROD         CCCAC.........TACC-G...CCG..................--CAATAAA-GACCC--------A-G   1016
B_EHOA        ..CAT.........TCAC-G...CCT..................--CAATAAA-GACCT-A------T-G   1004
SD_MM251      ..CAC.........TCAC-A...CC-..................--CAATGAT-GGCC--AG--G--A-G   1034
STM_STM       ..CAT.........TCGC-A...CCC..................--TAATGAG-GACC--A------T-G   1025
VER_AGM3      ..CAT.........TCTC-G...-GG..................TACAATACA-GGC-G--------T-G   1046
GRI_AGM677    ..CAC.........TCTC-G...-A-..................TATAATACC-GGT---A------G-G    995
TAN_AGMB14    ..CAC.........TCAC-A...-A-..................TACAATACA--AT-G-----G--A-G    888
SAB_SAB1C     ..CAT.........AGCC-G...-A-..................TATAACACC-GGT-G-A------T-G   1016
SYK_SYK       TC%%C.........TCCC-G...TA-TACAAG%GGAAATTGAAA-CC--GAGGG-CGC--A--GT--C-T    980

ENV_CONSENSUS ?TGtaa?ttta?tgga????aaTGGaaa?aagc?ttaaAa?aggtaa??????????g?aaaa??????    745
A_U455        T-----TG-C-G-A--AGGG-C-----TAG-A-TA--C--C-----GCT.........-A-C-.......   1041
B_HXB2R       T-----CA---G-A--GCAA-------TA-CA-T------C--A--GCT.........AGC---.......   1044
C_UG268A2     C-----CA---G-A--AATG-------TAT-A-T---C--TG-----GA.........-A----------   1056
D_ELI         T-----TA---G-A--GCAC------GTA--A-T---C--C-A---GCT.........AG----------   1041
E_TN2432      T---G-GA---A----ACAA-------TAG--TT------C------CT.........-A----------   1062
F_BZ163A      T-----TG---G----ACAC-------TA--A-A---G--C-A----GG.........-C---G.......   1038
G_LBV217      T-----TG---G--A-ACAG-C----G-G-CATG---C-GA------AA.........-C-C-.......   1071
H_VI557       T-----TA---C-A-GGAAG-T------AGGA-T---C-TG-----GTT.........CAGC-.......    240
O_ANT70C      T--C--G-A--A--CCACTG-T---GG-A--ATA------C-AAC-GCTGAAAGGTATTT-G-.......   1041
O_MVP5180     T----CA-A--A-AAGACTGT----G--A-T--CC--C--C-AAC-GCTATAAGGTATTT---T......   1062
O_VAU         T-----T-A--G--CCACTG-C-----CA----C------A-CA---CTGAAAGATATTT-G-.......   1056
CPZGAB        C-----GA-A-A---GACCAC------TCG-A-GG-GG--G-A----AA.........AAGGCT......   1035
CPZANT        C----CAG-C-A-AAGACGCT----G--C----ACGT--CA-AAC-GAG.........CACGTT......   1041
A_ROD         G--CTGG--C-AA--C...A--------G-C--CA-GC-GG----G-AGGAAACCCTT-C----------   1077
B_EHOA        G--CTGG----AA------...A-C----C-G----CA--C-GG----G-AAGAGACCATTAA---T.....   1065
SD_MM251      G---TGG---GGA------...A-------GG-T--AA-----G----G-AACAGACCATT-TC-------   1095
STM_STM       G--CTGG---GGA------...G------G-GG---AA-----G----T-AGGAAACCTTG-TC-------   1086
VER_AGM3      G---C-C--CCAG--C...A-C----G-GG---C-GG---G-A----AAAATGAAATA-T----------   1107
GRI_AGM677    G--CC-C--CCAA------...G-T------GGG--A-GG---G-A--C-GAGAAGAAGTAAAG---GTGAAA   1062
TAN_AGMB14    G--CC-T--CCAA--G...G-C------GG---T-GG--GG----G-GAGAAACGATA-TT-------   949
SAB_SAB1C     G--CTGG---GGA--T...A-T----G-GG---C-GG--GG-A--C-AAGAAACAATA-TG-G-.....   1077
SYK_SYK       C---CGAG-A-CA------...A-T----C-G--TTT--T--GC-----CAT.........-A-C-G.....   1032
```

```
ENV_CONSENSUS ??????????????????????????????????ta?????????a?a?a??a?t??a??????????a    754
A_U455        ................................T--..........-AGA-AA-G-TT-AT.........-   1060
B_HXB2R       ................................T--..........-G-G-AC-A-TTGGAAAT.......-   1066
C_UG268A2     ................................T--..........-A-AGGC-C-TCCCT.........-   1075
D_ELI         ................................T--.........GG-ACCCTTCTT............-   1057
E_TN2432      ................................T--..........-A-G-GC-C-TT-AT.........-   1081
F_BZ163A      ................................T--..........-AGTCTC-T-TTCCT.........-   1057
G_LBV217      ................................C--..........CA-GGAATC-AC............-   1087
H_VI557       ................................T--..........-G-G-AC-C-TT-AT.........-    259
O_ANT70C      ............................................................CTAGTA.........-   1048
O_MVP5180     ................................C-T.........GT-A-CC-AACAGAG..........-   1081
O_VAU         ................................C-T.........GT-G-AT-TAATCAAACT......G   1078
CPZGAB        ................................T--..........GC-ACCTCA-CA............-   1051
CPZANT        ................................C-T.........GCGG-GC-T-GG-AAAAAGTAGAC-   1069
A_ROD         ..............................CATCCC.........-GGT-TAGAGGA-CC.........-   1099
B_EHOA        ..............................CATCCC.........-G-T-TTCAGGA-CAACA......-   1090
SD_MM251      ..............................CATCCC.........-GGT-TACTGGA-CT.........-   1117
STM_STM       ..............................CATCCC.........-GGT-TACAGGA-CC.........-   1108
VER_AGM3      ................................T--CCAAAAGAT-G-T-CC-AGGA-CC.........-   1135
GRI_AGM677    AATCTTACAGAAGTAAGCATAGAAAATATACATC-G..........-G-AGGATA-GG.............   1110
TAN_AGMB14    ................................T-GCCAAAACAT-A-T-TC-AGGG-CT.........-    977
SAB_SAB1C     ................................C--CCACCAAAA-AGT-TAGTGGC-CA.........-   1105
SYK_SYK       ................................GCC.........-CGA-AACC-GG-AA.........-   1051

ENV_CONSENSUS at?a?aca???????????ataa??tt??????a?cc??a???ggaggaGAtccaGAagt?acaa??c?    792
A_U455        -CA-A---.............----TC--TGCTAGCT--TC-...-----G---AT----A-T----CA-A   1115
B_HXB2R       --A-A---.............----TC--TAAGCA-T--TC-...-----G--C------A-TGT--CG-A   1121
C_UG268A2     --A-A---.............----AC--TACTCA-C--TC-...-----G--CTT----A-T----CA-A   1130
D_ELI         -CA-A---ATA.........----AG--TAAACC-T--TC-...-----G--C------A-T----CA-A   1115
E_TN2432      --A-G---.............----TC--TCAACC-C--TC-...----------TG---A-T----TG-A   1136
F_BZ163A      --GCA---.............----AA--TAATTC-T--TC-...-----G--C-T----A-T----TG-A   1112
G_LBV217      -CA-A-GC............----CC--TAACTC-T--TC-...-----G--C-T----A-T---GCA-A   1142
H_VI557       --C-A---.............----TT--CGAACC-T--TC-...-----G--CATG---A-T----TG-A    314
O_ANT70C      -CA-T---GGTAGTATTAAC--G-CA--CAATCACAG-AGC...--T-------T---G--A--CCATTT   1115
O_MVP5180     --GTT--C.............----TA--CAGCAG-A-TAGT...--T------G-------A-GCCATTT   1136
O_VAU         --GTT--C............--G-AA--CGGTAATCA-AGT...--T-A----G-------A----ATTT   1133
CPZGAB        --CGA---GCAGCC...AAC----CAC-CAACAG-G--TCT..............G--C---CAT-A   1115
CPZANT        -CA-A--CAATGCGAAAACA---TGGACATTCCA-GAT.........--------T-----A-A-GTG-A   1130
A_ROD         --G-C---AGG......AAT--T-GC--TGCAGCGC-AGG-AAA--CTC---C--------AG--TACAT   1163
B_EHOA        --ATCT--CAG.........----GG--AGCAGAGCA-GCGAGAA-CTC------------A-G-TATAT   1151
SD_MM251      -CA-T--TGAT......AAA--C-AT--AACGGCTC-TGG-...------------G-----T--CTTCAT   1178
STM_STM       --G-C--GGCA......AAA----GGA-AGTGGCTC-TGGG..............G--C--CTTCAT   1169
VER_AGM3      --G-T--TGAA........GAG-TT-ATCTGCAGAGACT-...TTT-------------CAG---ATTT   1193
GRI_AGM677    ..........................................------------TCAG-G-ATTT   1133
TAN_AGMB14    -CG-T--TAAG......AAA---TGGC-AAGGAG-CAA......T-G-----C-----G-CAG-G-GCAT   1035
SAB_SAB1C     --G-C---AAC.........-AG-TT--TCTACA-AGACA-...T-G------------TCTGA-TTCTT   1163
SYK_SYK       --GTA--C............-AC-CCACATGGAGGT--CAGCCA--G-------T---G--ACG--CA-A   1109

ENV_CONSENSUS ?tg?tttaAtTGt?gaGGaGAaTTttTcTAtTGtaataca?ct?gg?tgtTtaat?????ta????????    843
A_U455        TA-C---------G-------------------C--C---T-AG-CC--------...AGC-TTTGG...   1179
B_HXB2R       CA-T---------G----G-----------C------T--A-ACAAC--------...AG--CTTGGTTT   1188
C_UG268A2     TA-C---------A-------------------C------T-AA-CC--------...AG--GTGATAAT   1197
D_ELI         CA-T---------G----G-----C-----C---------T-AG-AC--------...AG--CATGG...   1179
E_TN2432      TCAT---------A----G--------------C------A-AC-AC--------...AA--CTTGCATA   1203
F_BZ163A      TA-T---------A----------------C--C------T-AG-AC-------CGAC............   1170
G_LBV217      TA-T---------A--------------------------T-AG-AC--------...AG--GTATATTA   1209
H_VI557       TACT---------A-------------------C------T-AAAAT-------C...AG--CTTGGGTT    381
O_ANT70C      ACAC-----C---CAT--------C--T--------C---G--AA-A--------...TA--CCTTTTCA   1182
O_MVP5180     ACAT-----C---CAT--------C--T--------C---T--G--A-------C...TA--CTTTTATC   1203
O_VAU         C-TT-----C---CAT--------C--T--------C---AA-C--C--------...CA--CCTTTTCC   1200
CPZGAB        TATG---------G-------------------------TAGCCAAA-A----C-GAC............   1173
CPZANT        T--G--------CCA---------C--T------G---T-A-ACCTTG---C---...GCC-CATACACG   1197
A_ROD         G--GAC---C--CA-------G---C----C--C--C-TGA--T--T-CC-C---...TGG-TAGAG...   1227
B_EHOA        G--GAC---C---A-G-----G---C----C-------TGA--TTTT-C--A--C...TGGGTAGAA...   1215
SD_MM251      G--GACA-----CA-------G--CC----C-----A-TGAA-T--T-TC-A---...TGGGTAGAGGAT   1245
STM_STM       G--GACA-----CC-T---------C-T-----C--A-TGAA-T--T-C--A---...TGG-TAGAAAAC   1236
VER_AGM3      A--G---------CAG--G-----C-----------A-TGGA-T--T-TC-A---...TACCTGAAT...   1257
GRI_AGM677    T--G--C------CA---T-----------------G-TGGACT--T-TA-C---...TA-CTAAAC...   1197
TAN_AGMB14    A--G--------CCTG-------------------CAC-TGA-T--T-TC-A---...TA-TTGAAC...   1099
SAB_SAB1C     C-TT---------CA---T-----C-----------A-TGGA-T--T-CC-C--C...TA-TTGAAT...   1227
SYK_SYK       T--G--CC-A---G-G-----G--C-----C------GT-T--AA-C-C---GC-...AAC-TTACG...   1173
```

```
ENV_CONSENSUS ?????????aatag?ac??????????????????????????????????????????aat???????    853
A_U455        ............---G-C-GCATGTCA.....................................---GAC....   1200
B_HXB2R       ..............-----T--TTGGAGTACTGAA.......................GGGTCA---AAC....   1221
C_UG268A2     ......AAT-----T--C..................................................   1209
D_ELI         ..............----TT-GTGCATGG.....................................---AATATT.   1203
E_TN2432      ......GGA---GAA--CATG.............................................---.......   1221
F_BZ163A      ................--A...............................................   1173
G_LBV217      ......TCT--C-AT-ATGCACCA.....................................TC-.......   1230
H_VI557       ...........--C--T--ATCA.....................................---GATATC.    402
O_ANT70C      ......TGT--CG-A--CACCTGTAGT..............................GTT-G-AAT....   1212
O_MVP5180     ...........--CT-T--AAAGTCCGGATGCCAGGAG............ATCAAAGGGAGC---.......   1245
O_VAU         TGCAAGAAG----TG--CAATAACAAGATC......................AATTGTACT---ATTAGC.   1248
CPZGAB        ...........--C-TT--A...............................................   1182
CPZANT        ......GGA--CCTC-TC.................................................   1209
A_ROD         ..............----AG--A...........................................   1236
B_EHOA        ..........-----A--T................................................   1224
SD_MM251      ......AGGG--GTA--TACCCAGAGG..............................CCA--G.......   1272
STM_STM       ..........-GA--T--ATCAGAAATGAGAGAT..................TGGAACAAA--CAAA....   1275
VER_AGM3      ...........---C-T--AGTAGATCCGGACCAT..................AATCCGTGT---GGT....   1296
GRI_AGM677    ...........---C-A--AGAAGAT................................GCAG-AGGT....   1221
TAN_AGMB14    ...........---GAGT-CAGTGAAGGAAGCTTCACTGACGTAGAAGGCAATAGATGTTCA--CATAACAA   1160
SAB_SAB1C     ..............----AGT-AGTAGACCCAGACCATAAC............AACTGTGCAAAA---AAT....   1272
SYK_SYK       ...........---G-A-ACGCAAGT................................AAG---AAT....   1197

ENV_CONSENSUS ??????????????ga??????????????????????????????????????a?t?a?a??at?a?    862
A_U455        ...........ATG-GGCCA....................................-A-GGC-CT--A-C   1223
B_HXB2R       ...........ACT--AGGA....................................-G-G-C-CA--C-C   1244
C_UG268A2     ..........................................................-TC--C-C   1217
D_ELI         ...........ACA--GTCAAATAAT........................AGCACA-ACACA-AC--C-C   1238
E_TN2432      ...............-GGTGT...................................-A-GGC-CT--C-C   1241
F_BZ163A      ...............-GATCC...................................-A-GGC-CT--C-C   1193
G_LBV217      ...............A-TAGC...................................-G-A-G-AT--C-C   1250
H_VI557       ...........TCAA-TGGC....................................-CAACA-TA--TGA    425
O_ANT70C      ...........GTTAGTCAA....................................GG-A-C-ATGGC-C   1235
O_MVP5180     ...............--GACC...................................-A-A-A-ATGGT-C   1265
O_VAU         ...........AATA-TAGC....................................-A-GGC-CTCAGGC   1271
CPZGAB        ..........................................................-A-GGC-TT--A-T   1196
CPZANT        .................ACA....................................-ACGGAGCCC-C-T   1226
A_ROD         ..........................................................C-CCGC-ATTA   1247
B_EHOA        ...............-GG.......................................CTCA-G-GA-ATTA   1241
SD_MM251      ...............--ACGG...................................CA-AGA-GG-ATTA   1292
STM_STM       ...........AAG--ACAA....................................CAAA-G-GG-ATTA   1298
VER_AGM3      ...........ACGA-GGGAAAAGGTAAGGCA............CCAGGACCCTGTGCAC-A-GA-CATA   1343
GRI_AGM677    ...........ACTA-TAGGACCTGTGACAAAGGGAAGCCAGGACCAGGACCATGTGT-C-G-GA-CTTA   1280
TAN_AGMB14    GTGGAGGACTGACA-GGTCCACCAGA.......................AAGTGTCT-A-A-GG-CCTA   1206
SAB_SAB1C     ...........ACAA-GCCATGT................................TGGC-G-GA-CCTA   1298
SYK_SYK       ...........TAT-CC......................................AGT-ACC-TCG   1214

ENV_CONSENSUS tct?ccaTgcaa?aTaagacaaaTa?TaAa?a??tgGcaga??gtaggaaaag?aat?TAtgc?cCtCCc    921
A_U455        ---C-A---T-GA----AG-----TA----T-TG-------GA------C---C---G-----C------   1293
B_HXB2R       C--C-------GA----A------TA----C-TG-------AA----------C---G-----C------   1314
C_UG268A2     A--C-------GA----A------TA----C-TG------GGG------CG--C---G-----C------   1287
D_ELI         A--C-A-----GA----A------TA----G-TG%%-TG-CAG-C----%---C---A-----C-----T   1305
E_TN2432      A--T--------G----AG-----TA----C-TG------GGA-C----C---C---G-----T------   1311
F_BZ163A      ---C-T---TCGA----A------GG----C-TG------GAA--G---CG--C---G-----CG-----   1263
G_LBV217      A--C-----T-GA----A------TG---GA-TG-------GG--G---C---C---G-----C------   1320
H_VI557       A--G-------GA----A------TA----C-TG-------GA------C---C---G-----C------    495
O_ANT70C      ---A--T-----AC-G-----GG-GG---GGTCA---ATA-GG-G-CAGTCG-G-C-C-----A------   1305
O_MVP5180     -A-A--T-----GT-------GC--G---GATCA---AT--AG-G--AGTCGAG---C-----A------   1335
O_VAU         AA-A--T----GGT-G-----G--G---GGGAC---AT--GG-G----TCG-G-C-T-----A------   1341
CPZGAB        A--G-------GA--T-----G---G-T-GTTCA---AT--GG-------G--G---A-----A--C--A   1266
CPZANT        AGCA-AT----GA--T-AG--G---G-T--TCAT---GGC-TA--TTC-----GC--T--CTTAG-C--T   1296
A_ROD         -GCA--G---C-T----AG------A-T--C-CA-----T-AG-----G-G-AATG-A---TTG------   1317
B_EHOA        -GCGT-----C-C--C-----G---G-C--C-CG-----C-AGA-T----G-AATG-G---TTG-----A   1311
SD_MM251      CG-G--G--TC-T--T---------A-C--C-CT-----T-AA-----C---AATG-T---TTG-----A   1362
STM_STM       -G-A------C-C--C------G--A-T--T-CA-----T-AA-----G---AATG-T---TTG-----A   1368
VER_AGM3      -G-TG-T---C-T---C--TCTG-CA-T--TGAT---T-C-CAC--TC--GGAA--CC-----A--G--A   1413
GRI_AGM677    -G-TG-C---C-T---C-----G--G----TGAT---T-C-CT--CTCT---AAGG-A-----T--A--A   1350
TAN_AGMB14    CG-TGGC-TGC-T---C-GTC-G-TG-C--TGAC---T-T-CAC--TTGC--AA--GA-----G--G--A   1276
SAB_SAB1C     -G-T--C---C-C-----G---G--G----TGAT---T-C-CCT-GTCT---AAG-CA--C--C------   1368
SYK_SYK       ---GT----TGCC--TC-C------A-T--TGAT---AGATAT---A-----TTG--A--CCTG--C--T   1284
```

```
ENV_CONSENSUS At??aaggaaac?T?ac?TGtaa?TCaaa?at?AC?ggacTa?ta?t?a?aatagat?g???t???aat?    970
A_U455        --CC-----GTAA-A-GG---G-A-----C--T--A------C--T-A-C--G----G-GGG-...-C-.   1359
B_HXB2R       --CAGT---C-AA-T-GA---TCA-----T--T--A--G--GC--T-A-C--G----G-TGG-.....-A   1381
C_UG268A2     --TA-------AA-A--A----GA-----T--C--A------C--C-G-C-CGT---G-AGGGGAG-CA.   1356
D_ELI         --CG--A-----A-TCTA---TCA-----T--T--A--G---C--T-G-C--G----G-TGG-ATA---.   1374
E_TN2432      --CAGT-----AA-T-AT---GTA-----T--T--A---A--C--T-G-C--G----G-TGG-GCT---.   1380
F_BZ163A      --CTC--A----A-T--C-----C--G--T--T--A--T---C--T-G-C--G----G-TGG-CAG---A   1333
G_LBV217      --TGC-------A-T--A-----A-----C--T--A------A--T-A-C--G----G-TGGG...---A   1387
H_VI557       --CA-------AA-T--A---G-A-----T--T--A--G---A--T-G-C--GT---CATGT-...----T    562
O_ANT70C      --CA----T--TC-A--A----TG-----C--A--T---A-GA-CC-ACA---G---AACACA...TGG.   1371
O_MVP5180     --CCCC--C---T-A--A---C-T--C--C--A--T---A-GA-TC-ACAGT-----CAACCA...TGG.   1401
O_VAU         --CCC-------C-AGTA--C-GG-----C--A--T---A-GA-TC-ACA-T-G--CACGCCATGG---A   1411
CPZGAB        --CAG-------A-C--T--C--T--T--C--T--T-----TC-CT-A-CTTC---CACTCC-GTG-C-.   1335
CPZANT        -GGAG---G--TG-TT-C----CT--C-GC--A--T---A-TA-GT-GGA-.........GG-CAA-TAT   1357
A_ROD         -GGG----GG-GC-GT-C--C--C----CAG-A--CA-CA--A-TGCT-AC--T--CT-GCAAAAC---.   1386
B_EHOA        -GAG-G--TG-AC-CT-C-----T--C-CTG-T--CA-C--CA--GCC-AC--T--CT-GATAGAT--G.   1380
SD_MM251      -GAG-G---G--C-C--G-----C--C-CAG-G--CA-T--CA--GCA-AC------T-GAC-GATGGA.   1431
STM_STM       -GAC-----G--C-T--T-----T--C-CTG-G--CA-TA--A--GCA-AC--T--CT-GAC-AAC---.   1437
VER_AGM3      -GAG----GC--T-GCAA--C-CA--C-CGG-A--G--TA-GTC-G-GGAGC--A--TACAA-AGT--G.   1482
GRI_AGM677    -GGG----TC-TT-GGAG-----C---TCAG-C--G-C----TACG-GGC-------TATAAC...--C.   1416
TAN_AGMB14    -GAG-----C--T-GGAG--C-GA--C-CGG-G--ATC-A-GC-GG-GTCGC-G---TATAACGAC--G.   1345
SAB_SAB1C     -GAG----CC--T-GGAG-----C--G-CAGCC--A-CG--GTATG-AGA-T-GA-CTATAACAGT--A.   1437
SYK_SYK       -CGGC---GC-TA-T-AA----CT--C--TG-C--T-C-G--T-GACGGAC-----ATACTACCCAGGA.   1353

ENV_CONSENSUS ??????????aacaa?ac?aa???????gt?acctt?agacCt??agg?ggaga?atga??gacaattg   1012
A_U455        ...........-----T--A--A...AAT-AG-----C------GG---A-----T----GG--------   1415
B_HXB2R       GC..........-----TGAGTCC.......-AG-T---C------GG---A-----T----GG--------   1436
C_UG268A2     ...........-GTG-G--A--TAGCACA-AG--A--C------GG---A-----T----GG--------   1415
D_ELI         ............--T-GT--T--C.......-AG-----T------GG---A-----T----GG--------   1427
E_TN2432      ...........-CT-CG--T--C.......-AG-----C------GG---A---A-T--A-AG--------   1433
F_BZ163A      AT..........--TC-G--T.........-AG-----C------GGG--T--T--T----AA---T----   1385
G_LBV217      GT.........-C---TGGA--T.......-AG-----C------GC---A-----C----GG--------   1442
H_VI557       CT.........GCAG-A-GT--T.......--T-----C------GG---A-----T----GG--------    617
O_ANT70C      ...........----GC-GC--CAACAAT--A--A--T-----AAT---G-----C----AA--T-TA--   1430
O_MVP5180     ...........--TTCC--AGGT...GAAAAT--AC-T-----AGT---G-----T----AA--T-TA--   1457
O_VAU         AAACACATCCT----GC--C............---C-T-----AGG---G-----T----AA--T-TA--   1469
CPZGAB        ...........-----C-GTGGT...AACC-G--A--T------AC---A---A-C----AA--T-T---   1391
CPZANT        AT..........--TG-A--T..........--T-AAG-GTC-----...-CT-C-AGAG-AGCA---C-A--   1406
A_ROD         ............--TC-G--A--C......A-T-----T--TG-A...-AG-TG-CAGAA...CTAT-C..   1431
B_EHOA        ............---CTT--T--T......A-T--TG-G--TG-A...-AA-TGTCAGAA...CTGT--..   1425
SD_MM251      ............---C-A--T-GT......A-C---A-G--TG-A...-AG-TG-CAGAA...CTGT--..   1476
STM_STM       ............--TG-G--T--C......A-C--TGCA--TG-A...-AG-TG-CAGAA...CTGT--..   1482
VER_AGM3      ............----GG--T--T.......--A--A--A--T--C.........C-G--AGAAAC--TC--   1526
GRI_AGM677    ............--GTCTGGCCCAATAAAT--G---C-A--T-----------C-GG-ACGCAG--TA--   1466
TAN_AGMB14    ............----TG--A--T........--C--A--A-C-G-A.........A-TT--GAAA---TA--   1389
SAB_SAB1C     ............----GG--C--T........--G--G--ATC------------C-G--TGAAAG--T---   1481
SYK_SYK       ...........TCT-CGTTG--C............--T-CT--AAC--CAAAT-TGGA-...--TGT---   1397

ENV_CONSENSUS gagaa?tgaatTgt??aa?TAtaAagTagTagaaaT?aaaCCa?TtggagTaGCaCCcACaaa?g?aaaa   1074
A_U455        --A--G------A-AT--G------------A----TG-----C-A--------------C-GG-C---G   1485
B_HXB2R       -----G------A-AT--A------------A----TG-----T-A--------------C--G-C---G   1506
C_UG268A2     -----G------A-AT--A--------G------G-T--G---C-A--------------T--G-C----   1485
D_ELI         -----G------A-AT--A-----G------C----TG-----C-A--------------C-GG-C---G   1497
E_TN2432      -----G------A-AT--A------------C----TG-----C-A---A----------C-GG-C---G   1503
F_BZ163A      -----G--T---A-AT--G-----------------TG-----C-A--------------C-GG-C----   1455
G_LBV217      -----G------A-AT--G---------A--A----T------C-A--------------C--G-C--GG   1512
H_VI557       -----G------A-AT--A------------AG-G-CG-----C-A---T----------C-GG-C--GG    687
O_ANT70C      -----C--------TC--C--C---------AGGG-A-----TT--A-T--G---------CGTATTGC-   1500
O_MVP5180     -----C-A------AC--C--C---------C-G--A-----TT--A-T--------T-----AATGTC-   1527
O_VAU         -----C-C------TC--A------------AG-G-A-----TT--A-T--------A-----AATTGC-   1539
CPZGAB        -----G---GC-C-AT--G-----------TCG---AG-G--TT-ATC---G-----A-----A-C--G-   1461
CPZANT        ----GCG--G----CC-GG--CC-G--G-----G--TGRT--CT-GTC------C--A----CA-GN---   1476
A_ROD         .---TTG--G---GGAG-T------T-G--------A-C----A----CT-C-----T-----A-A----   1500
B_EHOA        .-A-TTG---C--GGGG-C--C---T-------G--A-C----A----CT-T-----T----GTAT----   1494
SD_MM251      .C--TTG--G---GGAG-T------T-------G--C-CT--GA----CT-G--C------G-T-TG--G   1545
STM_STM       .C--TTG------GGAG-T--C---T----------A-CT---A----CT-G--C--A--G--T-TG--G   1551
VER_AGM3      -GC-GCA------GGC-GG--C---T----------T-C----A----CT-C---------G-A-T--G-   1596
GRI_AGM677    -GCGTAC---C--GGAG-C------T-------G--A-C----A----CT-T--T--T---G-T-T--G-   1536
TAN_AGMB14    -GCTTA-------GGC-GG-----GC-CA-T--G--C------A-A--GT-T--T--A---G-G-TG-GG   1459
SAB_SAB1C     -GCT-A-------GGAG-T--C---T----G-----C--G---A-----T-T--T--A--C--A-T--G-   1551
SYK_SYK       ----GCA--TC---TC--T------T-GA-TC----T------A----CT-T--G-----TG-TCA-CGT   1467
```

HIV1 ENV

```
                                                   gp120 / gp41 in HIV1s
ENV_CONSENSUS aGa??ag??gt?g?????????????????AGagaaAaaaGaGcagt???????ggactaGga????t?t   1110
A_U455        ---AG--TG--G-AG.............-----------------T.....---------GCTA-C-   1534
B_HXB2R       ---AG--TG--GCAG.............-----------------G.........---A-----GCTT-G-   1555
C_UG268A2     --GAG--TG--G-AG.............-----------------G.........---A-----GCTG-G-   1534
D_ELI         ---AG--TG--G-AA.............--------------A-A.........---T-----GCTA-G-   1546
E_TN2432      ---AG--TG--G-AG.............-----------------G.........---A-----GCTA-GA   1552
F_BZ163A      ---CA--TG--GAGG.............-----------------G.........---A-G---GCTT-G-   1504
G_LBV217      ---AG--TG--G-GG.............--------------A-T.........ACT---GCTG-C-   1558
H_VI557       ---AG--TG--G-AA.............--------------T--G.....-----G-A-GCTG-T-    736
O_ANT70C      --GCC--TCA-AAGCACTAGA...ACTCAT-----------------A.........---T-G---ATGC-A-   1561
O_MVP5180     ---CC-ATAA-AAACATTCACACCCCTCAC--G--------------A.........---T-G---ATGC-A-   1591
O_VAU         --GCC-ACTA-A-GAACTAGA...TCTCAT-----G----------CA.......--TT-G-C-ATGC-A-   1600
CPZGAB        --GCATACA--G-CAAGACAG...AAAGAC---C-G----------CCTTC...--T--G---GCAC-G-   1525
CPZANT        --GCC--AAA-AAAACAACAC......TCC---C---------GCA-T.........---A----GCTGT-C-   1534
A_ROD         ---TACTCCTCT-CTCAC.........GGG---C-T-C-----GT--GTTC...-TG-----G......-   1552
B_EHOA        ---TATTCCTCA-TGACA.........CCG--GA-T-------GT--ACTT...-TG-----G......-   1546
SD_MM251      --GTACACTACT-GTGGCACC......TCA---A-T-------GG--CTTT...-TG-----G......-   1600
STM_STM       ---TACACAACTAGTACC.........TCA---ACT--G----GG--CTTC...-TG-----G......-   1603
VER_AGM3      --GTATACG-GA-GTCAT.........GAC---AC---GC---TCCCGTTC...-TG-----G......-   1648
GRI_AGM677    ---TATACT-GCCCCACA..........------------G-TGCCATTC...-TG-----G......-   1585
SAB_SAB1C     ---TACACT-GCCCCGAA..............---C-------G-TGCCATTC...-TG-----G......-   1600
SYK_SYK       C-GTAT-AGT-GCCTAAT.........ACC--G--------G----CACCATTG-C---T--G......-   1522

            <- RRE                                / RRE primary binding site
ENV_CONSENSUS TccT?GGgtTctT?ggagCaGCaGG?agcaC?ATGGGcGcagcgtCaataaCgcTgaCggtaCaG?cccg   1175
A_U455        ----T--------A--------T--A-----A--------G------------------------G--A-   1604
B_HXB2R       ----T--------G-----------A-----T-----------------G---------------G--A-   1625
C_UG268A2     ----T--------G-----C-----A-----T-----------A---------------------G--A-   1604
D_ELI         ----T--------G-----------A-----G---------CG----G-G---------------G--A-   1616
E_TN2432      --T-T--------A-----------A-----T--------G------------------------G--A-   1622
F_BZ163A      ----T--------GA-C--G-----A-----T--------G------------------------G--A-   1574
G_LBV217      ----T--------A--------T--G-----T--------G------C--G--------------GT-A-   1628
H_VI557       ----T--------G-----------G-----T--------G------G-G---------------G--A-    806
O_ANT70C      --T-G---G-TC-AA-T--------T-----T------------G---C-------G--------A---A   1631
O_MVP5180     --T-G---G-GC-AA-T--------T-----T------------G---C-G------------G-A---A   1661
O_VAU         --T-G---A-TC-AA-T--------A-----T------------G---C-G------------G-A---A   1670
CPZGAB        ----T------C-G--T--------A-----T------------G--G----------------G----   1595
CPZANT        --T-G--TC-TC-CA-T-----T--C--T--A------------------G---------C----A--A-   1604
A_ROD         --T-G--T--TC-C-C-A-------TTCTG-A--------G-----CC-G--CG--T---CT---T----   1622
B_EHOA        -TT-G--A---C-T-CGA-G-----TTCTG-A--------G-----CT-G---T--T---CT---T----   1616
SD_MM251      --T-G--T--TC-C-C-A-G-----TTCTG-A--------G-----GT-G--------C-CT---T----   1670
STM_STM       --T-G--T--TC-C-CGA-------TTCTG-A--------------GC-G----------CT---T----   1673
VER_AGM3      ----A--C-----A-----T--T--G-CTG-A-----A------G-G-C-G-C--------C---T-T-A   1718
GRI_AGM677    -T--A--C-----G-----T--T--A-CTG-A------------G---C---------A--C---T-T--   1655
SAB_SAB1C     ----A--T-----G--------T--AGCTG-A-----T------G-G-C-G-C--------C---T-T-A   1670
SYK_SYK       ----G--A--AC-CTCG--T--G--C-CTG-A-----G-GC--TG-G-C-G----A--AC-C---T-T-A   1592

RRE primary binding /
ENV_CONSENSUS ?cattTa?Tg?ctGGtATagTgCAaCAgCAgaa?aAtcTgcTg???Gc?aT?gaggc?CAaCA?cA?cTg   1233
A_U455        A--A---T--T---------------------GC---------AGG--T--A-----T-----G--T---   1674
B_HXB2R       A--A---T--T-------------G--------C---T-----AGG--T--T-----G-----G--T---   1695
C_UG268A2     A--A---T--T-------------------A--C---T-----AGG--T--A-----G-----G--TA--   1674
D_ELI         A--A---A--T-------------------A--C---T-----AGG--T--A-----G-----G--T---   1686
E_TN2432      A--A---T--T-------------------R-GC---T-----AGG--T--W-----G-----G--T---   1692
F_BZ163A      A--A---T--T----A----------------GT---------AGG--T--T--A--G-----G--T---   1644
G_LBV217      A--A---T--T----C--------------A--C---T-----AGG--T--A-----G-----A--T---   1698
H_VI557       A--A---T--T-------------------A-GC---T-----CAA--T--AC----A-----G--TA--    876
O_ANT70C      CAC---GC--AAG------------------G-C--C-----AAGA--A--AC----C--G--G--AT--   1701
O_MVP5180     CAG-G--C--AAG------------------G-C--C------AGA--G--AC----C--G--A--CT--   1731
O_VAU         G---C-GA-AAAG------------------G-T--C-----AAGA--A--AC----C--G--A--CT--   1740
CPZGAB        A--G--GC--T-G-----------G--A-----T------T--AAA--A--A-----T-----G--C--A   1665
CPZANT        GA----GY-CCA------T--A---------GCC---------CAA--C--A---A-A--G--A--T---   1674
A_ROD         GAC----C--G-C--G--------G--A---C-AC-G---T--GAC-TGG-CA--AGA-----AG-A---   1692
B_EHOA        GAC----C--G----G--------G------C-AC-G--CG--GAC-TGG-CA-AAGA-----AG-A---   1686
C_2238                                                   GAC-TGG-CA-AAGA-----GG-AA--     27
D_F0784                                                  GAC-TGG-CA--AGA-----AG-AT--     27
SD_MM251      GAC----T--G----G--------G--A---C-AC-G---T--GAC-TGG-CA--AGA-----AG-AT--   1740
STM_STM       GAC----T--A----G--------G--A---C-AC-G---T--GAC-TGG-CA--AGA-----AG-AT--   1743
U_SMCI2                                                  GAC-TGG-CAC-AGA-----AG-A---     27
VER_AGM3      G------C-TG----G---C----G--------G---------GCG--TG-G-----T-----G--GA--   1788
GRI_AGM677    G-----GC-TG----G---T----G--------G--CT-----GCG--TG-G--ACAG-----A--GT--   1725
SAB_SAB1C     G--G---C-CG----G---C----G--------A--------TGCG--TG-A---CAG--G--G--AA--   1740
SYK_SYK       GAC---GT--G----G--------G------C-A--GT-----GAG--AG-T-----A--G--A--C--C   1662
```

```
                                                                  RRE ->
ENV_CONSENSUS tTgcaact?aC?gTcTGGGG?at?AaacA?CTccaggC?aGagTc???GCt?T?GAgA??taccTa?agg   1289
A_U455        ---A----C--T--------C--T-----G--------A------CTG---G-G--A-GA------C---   1744
B_HXB2R       --------C--A--------C--C--G--G--------A---A--CTG---G-G--A-GA------A---   1765
C_UG268A2     --------C--G--------C--T--G--G------A-A------CTG---A-A--A-GA------C---   1744
D_ELI         --------C--G--------C--T-----G--------A---A--CTG---G-G--A-GA------A---   1756
E_TN2432      --------C--A--------C--Y--R--G--------A------YTR---G-G--A-GA------A---   1762
F_BZ163A      -----G--C--A--------C--T-----G--------A------CTG--CG-G--A-GA------C---   1714
G_LBV217      --------C--A--------C--T-----G--------A------CTG---G-A--A-GA------C---   1768
H_VI557       ---A-G--C--G                                                             888
O_ANT70C      C--AGG--AT-TR-A-----T--C-G---A----GA--TC-CC-GCTA--CT-A--A-CC-TA---C--A   1771
O_MVP5180     C--AGGT-AT-T--A-----T--T-G---A----GA--TC-CC-GCAA--CT-A--A-CCCTTA--C--A   1801
O_VAU         C--AGG-CAT-T--A-----T--T-G---A----GA--TC-CC-GCTA--CT-A--A-CC-TTA--C--A   1810
CPZGAB        --------AT-AA-------AG-A-----A--A--A--C---T-GCTT---G-A--A-GG-----GC---   1735
CPZANT        C-A--G--CT-G--------AG-A-----A--------A---A-GCTT--AG-C----AG------AGA-   1744
A_ROD         ----G---G--C--------A-CG---A-C--------A------ACT---A-A----AG------C---   1762
B_EHOA        ----GG--G--C--------G-CG---A-C--------A------ACT--CA-C----AA--T--CA-A-   1756
C_2238        ----GG--G--C--------A-CT---A-C------A-T------ACT--CA-C----AG------A---     97
D_F0784       ----G---G--C--------G-CT---A-C------A-T------ACT---A-C----AG------A---     97
SD_MM251      ----G---G--C--------A-CA--GA-C------A-T--G---ACT--CA-C----AG---T--A---   1810
STM_STM       ----G---G--C--------A-CC--GA-C------A-T------ACT---A-C----AA------A---   1813
U_SMCI2       ----G---G--C--------A-CT---A-C------A-T------ACT--CA-C----AA------A---     97
VER_AGM3      ---A-G--G--CA-T-----TG-G---A-C---A-T--CC-C---ACA---C-T----AG------G---   1858
GRI_AGM677    ---A-G--G--CA-T-----TG-G---A-C---A-T--CC-C---ACA---C-C----AG------G---   1795
SAB_SAB1C     C-AA-G--C--TA-T-----TG-A---A-C---A-T--CC-C---ACT---C-T----AA------G---   1810
SYK_SYK       --AGGCT-A--T--------TG-C---A-C--TA-C--CC-GC--ACA---T-A----CG--T--CAGA-   1732

ENV_CONSENSUS AtCAg?agct?cTaaa??ttTGGGG?TGcgc??g?Aaaca?gtcTGccacACtactGTgcc?TGG?????   1344
A_U455        ----AC----C---GGAA-------C---T-TG-A--G-TCA-----AC------------C---.....   1809
B_HXB2R       ----AC----C--GGGGA-------T---T-TG-A----TCA-T---AC----G-------T---.....   1830
C_UG268A2     ----AC----C---GGGA-------C---T-TG-A----TCA-----AC----G-------T---.....   1809
D_ELI         ----AC----C---GGAA-------T---T-TG-A-----CA-T---AC-----A------C---.....   1821
E_TN2432      ----AA----C--RGGAC-------Y---T-TG-A---ATCA-----AC----G-------C---.....   1827
F_BZ163A      ----AC----C---GGAC-------C---T-TG-A----TCA-----AC-----A------C---.....   1779
G_LBV217      ----AC----C---GGGA-------C---T-TG-A----TCA-----AC-----A------T---.....   1833
O_ANT70C      -----C-A--C----GCC-A-----C--TAAAG-A--G-TA------T----AT-A--AAAA---.....   1836
O_MVP5180     -----C-A-GC-----CC-A-----C--TAAAG-A----TAA----TT----AT-A--AAAA---.....   1866
O_VAU         -----C-A--C--T--CC-G-----C---AAGAAT-G--TAA-----T----AT-A--AAAG---.....   1875
CPZGAB        -----C--A-T--GGGCC-G-----C---T-TG-A--GGCT-----TT-T--C--G-----T---.....   1800
CPZANT        ----AC-A--AT-G-GCC-C-----T--T--TGAC--GGTGAC---T--------G-----T---.....   1809
A_ROD         -C---GC--GG-----TTCA-----A--T--GTTT-G---A-----------------A--A---GTT..   1830
B_EHOA        -C---GCA-AA-----TTCA-----A--T--TTTC-G---G-----T-----G-----A--A---GTA..   1824
C_2238        -C--AGCAAAA-----TTCA-----A--T--CTTT-G---G--T-----------------A---GTG..    165
D_F0784       -C---GC--AG-----TTCA-----A-----GTTC-G---G-----------C-----AGAA---CCA..    165
SD_MM251      -C---GC--AG-----TGC------A--T--GTTT-G---A-----------------A--A---CCA..   1878
STM_STM       -----GCA-AG-----TTC------A--T--GTTT-G---G-----------C-----A--A---CCA..   1881
U_SMCI2       -----GCA-GG-----TGC------A--T--ATTT-G---A--------------G--A--T---GAA..    165
VER_AGM3      -C---GC--GGT-G--TGC------G-----AT-G--G--A-----T--T--A--C--A--G---CAGTG   1928
GRI_AGM677    -----GCA-GG-----TTCA-----A--T--GT-G-----A--A--T-----C--A-----A---AAGTA   1865
SAB_SAB1C     -----GCA-GGT-G--CA-------A--T--TTTC-GG--G--G--T-----G--G--C-TT---AAGTA   1880
SYK_SYK       -C---GCCA-CT-GTCAAA------A-----ATTC--G--GA-------T--AG-A---A-G---GAG..   1800

ENV_CONSENSUS ?AAt???acttgg?????????aa????????t??a??????????????????????????????c??att   1361
A_U455        .--CTCT-G---......AGT--TAAATCTCAGG-G...........................GAC--A   1845
B_HXB2R       .---GCT-G---......AGT--TAAATCTC-GG-A...........................-AG---   1866
C_UG268A2     .--CTCT-G---......AGT--TAAATCTC-AGGG...........................GAT---   1845
D_ELI         .--CTCT-G---......AGT--TAGATCTC-AA-T...........................GAG---   1857
E_TN2432      .--CTCC-----......AGT--TAGATCTT-TG-A...........................GAG---   1863
F_BZ163A      .--CTCT-G---......AGT--TAAATCTCAGG-G...........................GAG---   1815
G_LBV217      .--CACT-G---......AGC--TAAATCTT-TA-T...........................GAG--A   1869
O_ANT70C      .---AGA--A---ATA...GGA--CGAAAGC.....................................---   1866
O_MVP5180     .--CACAT-A---......TCAGGAAGATATAATG-TGAC........................AGT---   1905
O_VAU         .---AAA--A---GGAGGAGAT--TGAATCA.....................................---   1908
CPZGAB        .--CAACT-----CCTGGGAGC--TTCCACA...G-T...........................GAC---   1839
CPZANT        .---AATT-C---......GTA--CTTCACGCAAACATGTGCAAAGAACAGCAGTGATATACAATGT---   1872
A_ROD         .---GATT-C.....................T-AGCA...........................-CTGAC   1851
B_EHOA        .---GAAT-C.....................C-TA-G...........................-CAGAC   1845
C_2238        .--CGAC-G-.....................A-GC-G...........................-CAGAG    186
D_F0784       .---AGT---.....................C-CACA...........................-CTGAC    186
SD_MM251      .---GCA-G-.....................C-AACA...........................-CAGAC   1899
STM_STM       .---GATT--.....................T-GGTA...........................-CGGA-   1902
U_SMCI2       .---AAC---.....................T-GC-A...........................-CTGA-    186
VER_AGM3      G---AAT-GG.....................ACC...........................-CTGA-   1947
GRI_AGM677    T---AAC---.....................................................-CT-AG   1881
SAB_SAB1C     T---AAC---.....................................................-CCGAC   1896
SYK_SYK       .--AGCATG-G-AAACAACTCC--CTTCTGTCCCA-A..........................-CACAG   1842
```

```
ENV_CONSENSUS TGG?a?aacatgAC?TGGcaggaaTGGgA?aga?aaaTt?acaattt?a?ag??aa?aTatat?a??ta?   1416
A_U455        ---A-T--------C----T-C-------G-A-G-----AGT-G--AC-C--GC-TA------C-AC--A   1915
B_HXB2R       ---A-TC---C---C---AT---G-----C---G-----A------AC-C-AGCTTA---C-CTCCT--A   1936
C_UG268A2     ---G-T--------C---AT-C-G-----T---G-----AGT----AT-C--GA-CA-----CAGGT-GC   1915
D_ELI         ---C-G--------C---AT---G-----A---G-----G------AC-C--GCTTA------AGCT--A   1927
E_TN2432      ---A-C--------M---ATA--------R---G-----AG-----AC-C-AACC-A------G-GA--C   1933
F_BZ163A      ---G-G--------C---AT-------C-A-A-G-G---AG-----ACTC-AATG-AG----CAGGT--A   1885
G_LBV217      ---G-C--------C---ATA--------A--GG-----A----C-AC-C-CACC-A-----CTCCC-GC   1939
O_ANT70C      ---G-C-C-T-A--A--------------TC-GC-G--AAG---CA-A-GCTCC-CC------G-GGA-A   1936
O_MVP5180     ---G-C---C-T--A------C-------CCA-C-C--AA-----G-A-GCTCC-TT------G-TGA-A   1975
O_VAU         ---G-TG-GT-A--A------C-G-----TCA-C-G--AA----CG-A-GCTCCTTC------G-AAA-A   1978
CPZGAB        ---GGG--TC-A--A------C-------T-A-TT-G-GTCT--C-AC-C--GG--A--T-T-GGTC-GT   1909
CPZANT        ---G-A--T-----A-----A--------C---TT-G-AC-G----CA-C--GAC-G------A-TA-CT   1942
A_ROD         ---G-C--T-----G-------------A-A-C--G-CCG-T-CC-GGAG-CA--T--CAG-A-AAGTT   1921
B_EHOA        ---A-T--------A-----AC-------G--GC--G-CCG-TTC--GGAT-CA--T---ACAA-AT--C   1915
C_2238        ---C-A--------T-----AC-------A-A-C-G---GC-TTC--GGAG-AT--T--CACAG-GC-GT    256
D_F0784       ---A-C--T-----T--------G-----A---C-GG--G-TTTCC-AGAG-CA--T---ACAC-AT--T    256
SD_MM251      ---A-C--TGAT--T-----A--G-----GC--A-GG--G--TTC--GGAG-AA--T---ACAGCCC-CC   1969
STM_STM       ---A-C--T-----A-------------G---A-GG--G--TT-C-TGAG-CA--T---ACAC-AC-GT   1972
U_SMCI2       ---A-T--T-----T-----A--------A-T-A-G--CAGAG--C-GGAG-CA--T--C-CAG-ATC-T    256
VER_AGM3      ---A-T--T-----T----T---------A---C-G--ATCGT----GGA--GT--C---ACAACACA-T   2017
GRI_AGM677    ---G-C--T-----T---TT---G-----G---C-----A-TGCC--GGA--GC--C---AC-C-AC--T   1951
SAB_SAB1C     ---G-A--------G-----A--------G--GC----AG-A--G-ATGA--CT--C--CAG-AGGA-TT   1966
SYK_SYK       ---A-A--T-----A-----CAG------ACA-G-GG-AG-------G-CT-ACC-T--TG--GGAT-GC   1912

ENV_CONSENSUS TagaagaagCacaaa??CAgcAggaaaagAAtgtg?a?ga?ttac?ggaatTagataaaTGGgc???ttT   1477
A_U455        -T--G---T-G--G-AC----------------AACTA--C---TT--C---G--C--G-----AAA-C-   1985
B_HXB2R       -T------T-G----AC-----A----------AAC-A--A---TT------------------AAG---   2006
C_UG268A2     -T-----CT------AC--------G--A----AAA-A--T----TA-C---G--C------CAAAA-C-   1985
D_ELI         -T--G---T-G--G-CC-----A----------AAA-A--A--GTT------G--C--G-----AAG---   1997
E_TN2432      -TAC----T-G--G-AC--------C-G-----AAA-G--T--GTTA-----G-----------AAGCC-   2003
F_BZ163A      -T---A--T----G-AC----------------AAC-A-GA---TT--C---G--C--------AAG-C-   1955
G_LBV217      -C------T-G--G-TC----------------AAC-A--C---TT--C---G--C-------AAAGC--   2009
O_ANT70C      --C--A-G------GTA---------C-A----A-A-AA-G--G-T---G------G-------CTC-A-   2006
O_MVP5180     --C---C-------GAC--A--------------AA-A-CA--GTT---GC-----G-------CTC-C-   2045
O_VAU         --C----G------GAA--A-----G--A----A-A-A--A--G-T---G------G-------CTC-A-   2048
CPZGAB        ----G---------TCA--A--A--G--A----AAAGG--CC--TT------G---C-------CAGC--   1979
CPZANT        --C--AT------TGAG--A--A--G-GA---AAAA-G--A---TAT---C----C------AGCTCA--   2012
A_ROD         -----C-G-------TT-----A--G--A---A--T-T--AC---AAA-----A---GC----ATAT---   1991
B_EHOA        -------G-----G-TA--A--A--------CA--T-T--G----A-A-----A--C------ATAT---   1985
C_2238        ----GC----T----TA--------G--A---A--T-T--G--G-A-A-----A---GC----ATGT---    326
D_F0784       ----G----------TT--------G--A--CA--T-T--G----AAA--C--A----C----ATATA--    326
SD_MM251      -------G-------TT--A--A--G-----CA--T-T--A----AAA-G--GA---GC----ATGTG--   2039
STM_STM       ----------T---GTT--A------------A--T-T--G--G-AAA--C--A---GC----ATGTA--   2042
U_SMCI2       -G--G--G-------TA--A--A---------A--T-T--G--G-AAA-----A---GC----ATGTG--    326
VER_AGM3      ----G-----CAG-GCA---G----G------T--G-T-CA-AC-AAA-----AG--GT---T-AGA---   2087
GRI_AGM677    -G-------------AT---G-ATC-------C--G-TCTG-AC-A-A----G---G-T---T-AGGG--   2021
SAB_SAB1C     ----GC-------TGAA---G-AC--------T--G-TAGT-AT-AAA----G-TGTCT---T-AGAC--   2036
SYK_SYK       -CAG----------GAA--A-----G-GA------C-T--TC--ACAA-G---C-GG------ACTCG--   1982

ENV_CONSENSUS ?tGgaattGGtTtgacaTaaCaaa?TGG?TgtggtataTaaaaatag??aTaat?atagTaggagg??Ta   1539
A_U455        G--A---------A----CT----T---C----------T-G-C--TTTG----A-----------CT--   2055
B_HXB2R       G------------A----------T---C--------------T--TTC-----G-----------CT-G   2076
C_UG268A2     G----G------------------A---C-----------------TTC-----G-----------CT--   2055
D_ELI         G------------AG-------C-A---C-----------------TTC-----G---A-------CT-G   2067
E_TN2432      G----G------------------T---C-----------------TTT-----G-----------TT-G   2073
F_BZ163A      G------------AG----T----C---C-----------------TTC-----G-----------CT-G   2025
G_LBV217      G----G-------AG-------C-A---C-A---------------TTT-----G-----------TT--   2079
O_ANT70C      T---------C----------T--A---T------------------CA-----C----------CAC--   2076
O_MVP5180     T--------------------T--A---T------------------CT-----C-----G----CAC--   2115
O_VAU         T---------C----------T--A---T------------------CT-----C----------CAC--   2118
CPZGAB        G-----------------------A---C----------------TTTCC-T--GGC---------AA-C   2049
CPZANT        A---------------------C-A---C-A---------------TTT--T--G----------CTA-T   2082
A_ROD         TG-C------------T----CTCC---G-CAA------TC--TAT-GAG-GC-T---A---T--CAG--   2061
B_EHOA        CA-T------------T-C--CTCC---A--GCA--C--C-GGT---GAT--TAT------AT---AA--   2055
C_2238        TG-C------C----TC-C--CTCT---G-AAAA-----CT-TT---GAT-TTAT-------C---AG--    396
D_F0784       TG-C--C----------C-T--TTCT---A-TAA-------C-GTAT-GAG-CT-C-------T---TA--    396
SD_MM251      TG-C------------C-TG-TTCT---A-AAA-------C--TAT-GA--TTATG----T-T---AG--   2109
STM_STM       TG-C------------C-T--CTCT---G-AA-A--C---C--TAT-GAG--TATT----TAT---AT--   2112
U_SMCI2       TG-C--CC-------TT-T--CTCT---G-CAAA-----CT-T----GAT-TTATG------T---AA-T    396
VER_AGM3      C---TC------C--TT-CT----G---C--AACAT-C---------GAT-TT-GGAT---CT---TA-T   2157
GRI_AGM677    C---TCA-----CTCAC-GT---CT---T-AG-C---G-T-------GAT-TT-AG-GA---TGATTA-T   2091
SAB_SAB1C     T---TCG---------C-C-----A---T-TG-A-GG--G--G----CC--C--GG-TA---C---CA-C   2106
SYK_SYK       G----GC---------T-GT-G--A---T-C-TT---C----G----GAT-TTATG--A------CTT--   2052
```

```
                                                        3′ sj \/  \/ 3′ sj
ENV_CONSENSUS aTagg?ttaAgaaTagtt?t???tgta?t???tata?Taa?tAg??TtaggcagGgaTAt?caCCt???c   1592
A_U455        -----A------------T-TAC---GC-TTC----A-C-A---AG-------------CT-----...-   2122
B_HXB2R       G----T------------T-TGC----C-TTC----G-G-A---AG--------------T----A...T   2143
C_UG268A2     -----T---------A--T-TGC---GC-CTC----G-G-A---AG-----A-------CT----A...T   2122
D_ELI         -----T------------T-TGC---GC-TTC-T--G---A---AG-------------CT-----...-   2134
E_TN2432      -----T---------A--T-TGC---KC-TTC----G-G-A---AG-------------CT-----...T   2140
F_BZ163A      -----C------------T-TGC---GC-TTC----G---A---AG-----A---C---CT-----...T   2092
G_LBV217      -----T----A-------T-TGC---GC-TTC----G-G-A---AG-------------CT-----...-   2146
O_ANT70C      G----GG-G---G-TA-CA-GATA---C-TAA----G-G-AA-ACA--------------CA---C...-   2143
O_MVP5180     -----TA-----G-TA--A-GATAA--C-TAA-C--G-G-AG-ACA--------------CA---C...-   2182
O_VAU         -----TG-----G-----A-GATA---C-TAA-C--G---AG-ACA--------------CA---C...-   2185
CPZGAB        -----G--------TA--A-GACA--TT-CTCAG--G-C-GG--AG----------C--CT-----...-   2116
CPZANT        G----AC-------TT-GC-TGTAT--G-TAG-TGCT---GA-AGG--------------CAT---...-   2149
A_ROD         ----CT-----------GA-ATA----G-ACAA--GT---G---GC----AA----C---AGG---GTTT   2131
B_EHOA        G---TA----------CAA-ATACA-TA-ACAG--GC--GCA--GC-----A----C---AGG--AGTAT   2125
C_2238        ----TG----------C-A-CTA----G-ACAA--GT---TG--AC-----A----C                 453
D_F0784       ---CTG---------A-AA-CTA----G-ACAA--GT--GC--AGT-A--TA--                    450
SD_MM251      ---CTG----------GA-CTA-A--G-ACAA--GC--GC--AGT-A--------G---AGG--AGTGT   2179
STM_STM       G--ATG------G---CAA-CTA-A--A-GCAAT-GT--GC---GC-----A-A--T---CGG---GTGT   2182
U_SMCI2       -----G---------A-AG-CTACT--A-ACAAT--T--GG--AGC-----A--                    450
VER_AGM3      -----A------T-GC--TATACA---TATTC-TGCA--GC---GG----------T--CT-T---...-   2224
GRI_AGM677    C----A------T-T-CATGGGTAT--TGGGGATGTA-C-GA-ATA--------------AAT---...-   2158
SAB_SAB1C     ----TAGCT---G-TC-GC-AGTAA--A-AGGA--TC-T-GG-AGT----AA-A--G---G-----...-   2173
SYK_SYK       G-GCTT-----GC-----AGTTT-AGTG-AGGA--CA---AG-ATT-GTT-GGA--C---GTT---ATTT   2122

ENV_CONSENSUS Tctc?tt??????????????????cagAccC?tatCCaaaa?cag??gGaa???ccagaca?gccaGa   1634
A_U455        -G--G--T..................-------T-GC--C--TC-CAGA--GT...-TC-G--G-----G   2171
B_HXB2R       -A--G--T..................-------ACC---C--CC-C-AG--G-...--C----G---C--   2192
C_UG268A2     -A--G--T..................-------T--C--C--CC-C-AG--G-...--C----G---C-G   2171
D_ELI         -G--G--T..................-------TCC---C-GCC-C-AG--G-...--C----G---C--   2183
E_TN2432      -G--TC-C..................-------C--C---TC-T---AG----...-TC----GA--C--   2189
F_BZ163A      -G--A--T..................-------A-----C--GC-C-AG----...--C----G---C--   2141
G_LBV217      -A--A--C..................-------T-GC---CC-T---AG----...--C----G---T--   2195
O_ANT70C      ----G--A..................----T--CC-A---TC-C--AGA----...G---GA-C-----G   2192
O_MVP5180     ----G--G..................----T--C-G---C-C-C-G-CA----...G----A-C-----G   2231
O_VAU         ----G--A..................----T--CC------C-A--AGC----...GT--GA-C-----G   2234
CPZGAB        ----G--G..................-------T-----C-GTC---AG----...-A--GA-G--TC-G   2165
CPZANT        -G--A--T..................----T--C--C------C---CA---T...-----GCA------   2198
A_ROD         ----T-CCCCCCCCGGTTATATCCAA----T--A------C--GG-CCG--G-CAG----C--ACGA---   2201
B_EHOA        ----C-CCCCTCCCTCTTATACTCAA----T--C-----GC--GG-CCG--G-CAG----C--ACGA---   2195
SD_MM251      ----T-CCCCACCCTCTTATTTCCAGT----T-A--C----C-GG-CCC--C-CTG---AC--GAGA--G   2249
STM_STM       ----T-CCCCTCCCTCTTGTCGTCAG----T--C------C--GGGCCA----CAG---AC--AAGA--G   2252
VER_AGM3      -T--TCCA..................----T--A------CCCGTG-AA--G-CAG-------ACG----   2276
GRI_AGM677    --C-C.....................----T--A------C-GTTCAGC----CGG-------ACGG--G   2207
SAB_SAB1C     ----C-CCCTCCCCTCTTCTCATTAA                                               2199
SYK_SYK       -GCAGAAC..................-CT----AGGG--GG--GG-CCCA-GCAAA--C-CAGACGA---   2174

ENV_CONSENSUS a??aa?agaagaag?aggtggaga??aaggcagagacAgatcc?????????????tgc??ttg????tg   1679
A_U455        -AG--TC-------A---------GC------A---------G............A-T-GC--A...G--   2226
B_HXB2R       -GG--T--------A---------GAG--A-------------............A-T-GA--A...G--   2247
C_UG268A2     -GA--TC-------A---------GC---A-------------............G--AGA--A...A-C   2226
D_ELI         -GG--C--------A---------GCG-----------------............G--AGA------C--   2238
E_TN2432      -AG--TC-------G------C--AC----------A------............G---GC--A...G--   2244
F_BZ163A      -GA--TC-------G---------GC------------------............G--AGA--A...G-A   2196
G_LBV217      -AG--TC-------G------C--GC---A-------------............G---GC--A...G--   2250
O_ANT70C      -AG--C--G--G--G---------AG-------GCC---G-GG............A-A-CC-C-...CC-   2247
O_MVP5180     -AG--C--G-----A---------AGG--A---GCC--AG-GG............ACAGCC---...CCA   2286
O_VAU         -GG--C--G-----G---------CG---A---GCG---G-GG............ACT-CA------...CC-   2289
CPZGAB        -GA--TC--C----G------C--GC---A----AG------A............G--AGA--A...G--   2220
CPZANT        -GA--T-AG-----A-------AGAA---A---GAT---G-GG............AG-GCC------...CA-   2253
A_ROD         -...C--------AC------AGCA-C--TG---------A-............TG--CC-G-CCGA-A   2256
B_EHOA        -...C--------G-------A-CA-C-AGG-CT-------T............TG--CC-G-CAGA-C   2250
SD_MM251      C...A-----G--AC--------AGGC--TG-CA----C--------------TG--CT-G-CAGA-A   2304
STM_STM       -...C--------G---------CAG---TG--AT--AC--------------TG--CT-G-CAAA-A   2307
VER_AGM3      -...GGGCC--G--A---------CA-GC---AGA----C--------------GA--CT-G-CAGAAA   2331
GRI_AGM677    -................--GCA---CAG---TG----A--CAG-AGCAGCAAATTGA-AAGA------...CA-   2259
SYK_SYK       -......................--AGGCA-TG--------GAA............GGA-TCAAC...G-C   2208
```

```
ENV_CONSENSUS ?aa???gg??????????????????????????????????????????????????????ttctt??cac   1691
A_U455        AGC...-A.....................................................-----AG-GA   2242
B_HXB2R       A-C....-A....................................................-C---GG---   2263
C_UG268A2     AGC...-A.....................................................-----AG---   2242
D_ELI         A-C....-A....................................................----CGG---   2254
E_TN2432      AGC...-A.....................................................-----AG---   2260
F_BZ163A      G-C....-A....................................................-----AG-T-   2212
G_LBV217      AGC...-A.....................................................-----GT-GA   2266
O_ANT70C      C--....-G....................................................-----GC---   2263
O_MVP5180     CC-....-A....................................................-----GCA--   2302
O_VAU         C--....-G....................................................-----GCAT-   2305
CPZGAB        G--...-G.....................................................-G---GC---   2236
CPZANT        C-C....-G....................................................-----CG---   2269
A_ROD         GC-TATATA....................................................CAT--CCTGA   2275
B_EHOA        G--TACATC....................................................CA---CC-GA   2269
SD_MM251      G--TATATT....................................................CAT--CCTGA   2323
STM_STM       G--TATACT....................................................CAT--CCTGA   2326
VER_AGM3      G--TCT--CACAGCAGAGTGGAAGAGCAACTGGTGCAAGCGATTGACCAATTGGTGCTCGA--AGCAGCA   2401
GRI_AGM677    G--...-AG....................................................-C--CAA---   2275
SYK_SYK       TCT...ACC....................................................---AGCAG-G   2224

ENV_CONSENSUS tt???tg??acga?ct?cgga?cctg??c?t?tg?a??taccaccgcttgagagactta?ta????????   1737
A_U455        --GCC--GG----C--G----A----TG-C-C-TC-GC--------A-----------CGCC........   2304
B_HXB2R       --ATC--GG----T--G----G----TG-C-C-TC-GC---------------------C-C........   2325
C_UG268A2     --GCC--GG----T--G----G----TG-C-C-TC-GC----G----------------C---------.   2304
D_ELI         --ATC--GG----C--G----G----TG-C-C-TC-GC---------------------A-C........   2316
E_TN2432      --GCC--GG----T--A----G----TG-C-T-TC-GC--------------------CA-C........   2322
F_BZ163A      --GTC--GG----C--GA---A----TG-C-C-TC-GC-------A-C-----------C---------.   2274
G_LBV217      --GCC--GG----CT-GA---G----TG-C-C-TC-GC-------AA-----------CA-C........   2328
O_ANT70C      -GTTG-ACACG--C--CA-A-CAA-AATAT-G--G-CT-------T------CA-----GC--------.   2325
O_MVP5180     AGTTG-ACACG--T--CA---CAA-AAT-T-G--G-CT-------T------CA-----A---------.   2364
O_VAU         -GTTG-ACACG--C--CA---CAA-AAT-T-G--G-TT-------T------CA-----GCC........   2367
CPZGAB        --ATC--GG----C--C--A-A----GGGA-T--G-GC-----GA------C-AG----GCC........   2298
CPZANT        -CTTG                                                                   2274
A_ROD         -C...C-CC-GCTGA-T--CCT-T--AC-...A-ACTA---AG-AT--GC--G------C---------.   2331
B_EHOA        --......CG-C-A--GA--GA---CTTGA-T--GCTA---AG-G---GC---AC----T---------.   2325
SD_MM251      -C...C-CC-ACTGA-A--CCT-T--ACT...--GCTA-T-AG-AA--GC---AC---GC---------.   2379
STM_STM       -C...C-CC-GCTAG-A--CCT-T--ACT...--GCTG---A--AA---C--G-CT-GTC---------.   2382
VER_AGM3      -C.............................--GCTA---A--A-T-GCCTGAC-C-CC--GTTCATCT   2441
GRI_AGM677    C-.......................................TCGA-GA-C-ACA---GGTGG........   2298
SYK_SYK       AA...........................................................           2226

ENV_CONSENSUS ????t?????at?????????????????gag??t?????aa?t?ct?ggac?????a??????????   1754
A_U455        ....-TG...-TGTA............GC---AGCTGTGG--C-T--G----GCAGC-GCCTCAAGGGA   2355
B_HXB2R       ....-TG....-TGTA...........AC---GA-TGTGG--C-T--G----GC...............   2361
C_UG268A2     ....-TG....-TGCA...........GC---AGCAGTGG--C-T--G----GCAGC-GTCTCAGGGGA   2355
D_ELI         ....-TA....-TGCA...........GT---GA-TGTAG--C-T--G----GC...............   2352
E_TN2432      ....-CG....-TGCA...........GC---GGCTGTGG--C-T--G----ACAGC-GTCTCAAGGGA   2373
F_BZ163A      ....-TA....-AACA...........GC---AA-TGTGA-CAGG......................GGG   2304
G_LBV217      ....-TG...-TGTA............AC---GACAGTGG--C-T--G----GCAAC-GCCTCAAGGGA   2379
O_ANT70C      ....-CAGGG--CCAGAAG........GT--TCAGCTATCTGAGG--T----TGTGG-TC......CTA   2376
O_MVP5180     ....-CAGGG--CCGGAGG........CT--TCGACTACCTGGGA--G----TGTGG-TC......CTG   2415
O_VAU         ....-CAGAG--CCAGAAG........TT--TCAGACACCTGGGA--T----TATGG-TC......ATA   2418
CPZGAB        ....-GC...-ATGTG...........TG---ACAACTCA-GACGT-G----ATCTA-TACTTCACAGC   2349
A_ROD         ....-CCAGG-GCTTC...........CT--CCC-CCAACTCA-CTACCAGAAT...............   2370
B_EHOA        ....-CAAAG-CCTTC...........CAA-CCC-CCAACC-G-A--CCA--CG...............   2364
SD_MM251      ....-CGAGAGCATAC...........CA--TCC-CCAACC-A-A--CCAGAGGCTCTCTGCG...ACC   2430
STM_STM       ....-CCAGG--CTAC...........CA--CCC-CCACCC-ACCT-CCA-AGGATCTCCAGG...ATA   2433
VER_AGM3      TAGGAGC...GCTTTC...........CA-TACA-ACAAT-TGGG--T--GGAACTC-AAGCC...GCA   2493
GRI_AGM677    ....CTC...-ACTTCAAGAGCTGCAGCTT---AA-AAGGACTTGGTGTTACAACATCTGCCTGACCCTC   2361
SYK_SYK       ............................................................           2226
```

```
ENV_CONSENSUS ct????aggg???gggaag?cctcaga??t?tg?g???t?t??tgcaatat??????????????????   1787
A_U455        --GAGACT--GGT----G-G-----A-TA-C--T-GAA-C-TC---TG---....................   2406
B_HXB2R       ..........-GGT------C-----A-TA-TG-T-GAA-C-CC-A--G---....................   2406
C_UG268A2     --ACAG----GGT------C---T-AGTA-C--G-AAGCC-TG----G---....................   2406
D_ELI         ..........-GGT----CAT-----A-TA-C--T-GAA-C-CC-A--G---....................   2397
E_TN2432      --AAGAC---GGT------G-----A-TA-C--G-GAA-C-TC--TT----....................   2424
F_BZ163A      --GAGG----GGT------C-----A-CT-C--G-GAA-C-CAC--TG---....................   2355
G_LBV217      --GAGACT--GGT----G-G-----A-TACT--T-GAA-C-CC---TG---....................   2430
O_ANT70C      GGGCAG-A-ATAATTA-T-TTTG----AT-TGTGCAGC-G-AACA-----C....................   2427
O_MVP5180     GGACAA-A-ACAATT----CTTGT---CT-TGTG-AGC-G-AA--------....................   2466
O_VAU         GGGCAG---ACAATT----CTTG----CTCT-TAAAGC-A-AA-A-----C....................   2469
CPZGAB        --CAGACT-CTAC----GAGA--GT-CCTGC--G-GGGAA-TA-T------....................   2400
A_ROD         ...CTC--A-ACT--CTGAGA--T---ACA......GCCT-CT--------GGGTGC..............   2418
B_EHOA        ......................-----GCT-CCACCTGCATACC-C-G----GGGATC.............   2400
SD_MM251      --ACGA----TTC-A----T------GAC-GAACTGACCTACC-A------GGG..................   2484
STM_STM       --GCAG--AATTA-A----T-G-----CTAGGAGCAGCCTATT-A--G---GGGTGC..............   2490
VER_AGM3      GCACAA.........----CAG-TGTCGC-T--GCACGCC-TGCA---A-C....................   2535
GRI_AGM677    --CATATTCATCA--AC--CAG-GG--TAC............C----G---....................   2400
SYK_SYK       ....................................TC-T-GAG-----C-TTGGAAGCAGGACAGCAGC   2260

ENV_CONSENSUS ??????????????????????????????????????????????????????????tgG?g????????   1791
A_U455        ..............................................................---G-T........   2412
B_HXB2R       ..............................................................---A-T........   2412
C_UG268A2     ..............................................................---G-T........   2412
D_ELI         ..............................................................---A-T........   2403
E_TN2432      ..............................................................---G-C........   2430
F_BZ163A      ..............................................................---G-T........   2361
G_LBV217      ..............................................................---G-T........   2436
O_ANT70C      ..............................................................---CTA........   2433
O_MVP5180     ..............................................................---CTA........   2472
O_VAU         ..............................................................---CTA........   2475
CPZGAB        ..............................................................---G-A........   2406
A_ROD         ..........................................................GAG---ATC........   2427
B_EHOA        ..........................................................AGC---TTC........   2409
SD_MM251      ..............................................................---A-CTAT.....   2493
STM_STM       ..........................................................ATC---ATC........   2499
VER_AGM3      ............................................................GC-G-CTAC.....   2544
GRI_AGM677    ............................................................G--CTC........   2406
SYK_SYK       TTTGGAGGACAGTCTGCAGCAGCTTCAGATCTCTGATTAGACAGCTTACTATCACC---G-ATTCATCAG   2330

                                          <- rev cds end
ENV_CONSENSUS ?????????caggaact?aaaa?tagtgcta??ag??tgcttga?gc??tagca?t?gcagtagc?ga?   1838
A_U455        ..........-G------A----T--------TT-CTT-------C--TG-----G-A--------T-GG   2472
B_HXB2R       ..........--------A--G-A-------GTT--CT----CA-T--CAC---CA-A--------T--G   2472
C_UG268A2     ............--A-----A----AG-------TT--TT-------TA-CA-----A-A------T-T--A   2472
D_ELI         ..........---------G-GG-AC-------GT--CT--T----T--CA-----A-A--------T--G   2463
E_TN2432      ..........--------A----T--------TTTCTT-----A-T--TAC----A-A--------G-GG   2490
F_BZ163A      ..........--------A--G-A--------TTG-CT--T--A-CA-CAC----A-A-T------T--G   2421
G_LBV217      ..........--------A--G-A--------TT-ATT-------TA-AG-----A-A---AC---TA-C   2496
O_ANT70C      ...........--A---T-GC-G-A--------CA--CT----A--CA-AC-T---G-G--------CA-T   2493
O_MVP5180     ...........--A---T-G----A--------CA-ACC-------TA-TA-T---G-GT----T--CA-T   2532
O_VAU         ...........--A---T-GC---C--------CA-ATC-A--A--TA-TG-T---G-G-----T--TA-T   2535
CPZGAB        ..........A----G--A----T--------TA--CC-------T--AAC---TA-T--------A--A   2466
A_ROD         ...........--A---...GC-TTCCAG--CGCCGCGAG-GC-ACAAGAGAGA-TC-T--G-GC--GTGC   2484
B_EHOA        ...........--A---...GC--TCCAA--AGCAGCCAG-GC--CG-GAGAGA-TC-T--GAGC--G-CG   2466
SD_MM251      .......TTC--T--G...GCGGTCCAA--CGGCT-GAGATC--CGA-AGA-A-TC-T--G-GC--GTGG   2553
STM_STM       ...........--A---...GC-GCGCAA--AGCAT-GAGAGC--CG-GAGAGA-TC-T--GAGC--G-GG   2556
VER_AGM3      ...........---AT-..........................................................   2550
GRI_AGM677    ...........---C----CC--GAGGCA--A-CAG-GC-TGC-C-A--TC-G--GAGG--T-CGAGG--A   2466
SYK_SYK       CTATGGGTTCA-C--G--C----TCGCA--AGCTTCGC-CGGAAGG-AAA--CGCGATTGG--G--G-CT   2400

ENV_CONSENSUS ?ggac?ga?ag??t?at?gaagt??t?caaagagttgg?agagc?aTtctcaacat?CC?agaaGaaT?a   1893
A_U455        T---TA--T--GG-T--A---A-AGGA----C-A----T-----T-----------A--T--------C-   2542
B_HXB2R       G----A--T--GG-T--A-----AG-A---G---C-T-T-----T----G-C----A--T--------A-   2542
C_UG268A2     G-A--A--T--AA-T--A-----GGGA---G--A----G-----T--C---C----A--T--------A-   2542
D_ELI         G----A--T--AG-T--A---A-AA-A-------C-T-C-----TG----T-----A--C--------A-   2533
E_TN2432      T----A--T-AGG-T--A-----AGCA---G---C-T-G-----C------                     2541
F_BZ163A      G-A--A--T--AA-T--A-C--CTT-G------A----T-----T---T-G-----A--T-C------A-   2491
G_LBV217      T----A--T--GG-C--A-----AGCG-------C-T-C-----T--------T--A--T-----G--A-   2566
O_ANT70C      T----T--CG-CA-A--C-C--GGA-A------A-A--A-C--GA----GT-----C--A--G-----T-   2563
O_MVP5180     T----T--CG-CA-C--CTT--GTC-A------A-A--ACA--GAT-C--TC----C--A--------T-   2602
O_VAU         T----T--C--CACA--CTT--GCA-A-----CA-A--G----GG-----T-----A--A-----G--T-   2605
CPZGAB        G-T--A--T--AA-A--A----CTT-T---GTTACACTT---ATT---AGA-----C--T--GC-C--A-   2536
A_ROD         A--...-GCTTG...TGGAGG--AT-GG--C--A-C--G--G-GA--A---GCGG-T--A-----G--C-   2548
B_EHOA        A--...ACCTCG...TGG-G---CC-CAGG----CG-CAG---AG--CA-TGCA--A--C--G--G--C-   2530
SD_MM251      A-A...--CTTA...TGG--GACTC-TAGG----G---A---TGG--C---GCA--C--T--G--G--T-   2617
STM_STM       A-A...--CTTA...TGG---ACTC-TGG---G-----G---AGG--CGGGGC---C--A--GC-C--C-   2620
VER_AGM3      ...............TGGCTT-CTTGCAG-TCC-C-TAT--G--A--CA-----TCT--A------G-GC   2605
GRI_AGM677    GCCTGG-GC--AC-GGGT-CTA-TG-C-G-TCC-C-TATC-G--AG-CA------GT--A------G-GC   2536
SYK_SYK       ATTTGGC-AGCAA-ATAC-CC-CTACCAG-C------TCGAG--CG--GCAGC-T-G--GC-G--GC-TC   2470
```

HIV1 ENV

```
ENV_CONSENSUS GaCAaGGc?t?GAaag??c?tT??t?                                  1911
A_U455        --------T-A-----GG-T--GC-A                                  2568
B_HXB2R       ----G---T-G-----GATT--GC-A                                  2568
C_UG268A2     --------T-T---GCAG-T--GCAA                                  2568
D_ELI         ----G---T-A-----GT-T--AC-T                                  2559
F_BZ163A      ----G---T-C-----GG-T--GC-A                                  2517
G_LBV217      --------T-G-----AG-T--GC-A                                  2592
O_ANT70C      ----G---T-A-----AAGT--AT-G                                  2589
O_MVP5180     -------TGCA-----AATC--AG-G                                  2628
O_VAU         ----G---C-T---C-ACTCC-GT-A                                  2631
CPZGAB        -------GC-G-----AG-T--AC-T                                  2562
A_ROD         ----G--AGCA----TCG-CC-CC-G                                  2574
B_EHOA        -------GGCT---CTCG-CC-CT-G                                  2556
SD_MM251      -G-----GC-T--GCTCA-GC-CT-G                                  2643
STM_STM       -------GC-T---CTCA-AC-CT-G                                  2646
VER_AGM3      --------C-T---G-AATCC-TAAT                                  2631
GRI_AGM677    -G------C-T----AAGTCC-GGGG                                  2562
SYK_SYK       --------T-G--G-TCTACC-TAAC                                  2496
```

```
CONSENSUS-A         ATGAgAGtGAtGggGatacagaggAatTatcaacacTtg...TgGAga??????.........     45
CONSENSUS-B   ATGGCA----------a------cag--a---------g------???-----------?????????     51
CONSENSUS-C         ----------?--------?--------g------a-g-...----T-------?????????.     43
CONSENSUS-D         ----------g--?----g-------------g------...------------?????????     44
CONSENSUS-E         ----------A--a--c-----t-----gg-c-a-----...----a-------.........     45
CONSENSUS-F         ----------G------G----------GG--g------...G---A-................     45
CONSENSUS-G         ----------a-----------------GG---------...----a-................     45
CONSENSUS-O         ----?-----cTAT--a-G?a-T-g-GA-GA?gA--A?-...AA-Tt-................     41
CONSENSUS-U         -------------a--------------g--c-----?-...----?-................     43
CONSENSUS-A         ----------------------------------------...------------.........     45
CONSENSUS-CPZ       ----?????-?-??--???-??????????????????...?T????------???.......     10

                                                      signal peptide \/ gp120
                                                                   <- vpU
CONSENSUS-A   ..??????TGG...gG?actatgaTctTtgGgatgataataATttgT...a?Tgct????cagaa...aA     92
CONSENSUS-B   ??---------...--c--c---c--c--------t-g--g--c---...-g-------g-----...--    101
CONSENSUS-C   .........---...--c-TCT-AGG----t-----t----g------...-a--tg----tg-gg...--     92
CONSENSUS-D   ??---------...--C--c---C--C--------T-g--G--a---...-g--t-...g-----???--     94
CONSENSUS-E   ..---------...--g---t-----C-----t--g-G---------..?-g---c...T----c...--     95
CONSENSUS-F   .........---...--CCT-T-AT--C-G-----AT-g-----C---...-A----...-??---...--     92
CONSENSUS-G   .........---...--a---T-----C-----T--G-?---------...-g---C...t--a-t...--     94
CONSENSUS-O   ........---AtCTtgT?C?-????a-G-CTT-------CCCA---...TTGAGc...t?TA?c???C-     86
CONSENSUS-U   .........---...----T?t-a------------?----g------...-A---?...at---????--     88
CONSENSUS-A   ..---------...--------------------------------...---------------...--     92
CONSENSUS-CPZ ........????????-????-???T--G......C---?CC??-?-????-GA??...-????G?????     26

CONSENSUS-A   ?tTGTGGGTcaC?GTcTacTATGGGGTACCTGTgTGGaaaga?gcagag......ACcACccTaTTttGt    153
CONSENSUS-B   a-----------a-----t-----------------------a---acc.....-----t---------    165
CONSENSUS-C   C-----------a-----t-----------------------a---A-a??????--t--t---------    156
CONSENSUS-D   t-----------A--t--T----------------------A---acc.......--t--t---------    158
CONSENSUS-E   C--------T--A--t--T--------t----------g---t-----t.....----------------    159
CONSENSUS-F   c-----------A-----T-----------------------a---ACc......--T--T-----C---    156
CONSENSUS-G   C-----------A-----T------------------G----t-----t.....-----t---------    158
CONSENSUS-O   ac-?-AT-C?--A-----T?C------?-----A---G----t---a?a......C-AgtA?----C---    144
CONSENSUS-U   c-c---------A-----t-----------------------t-----?.....-----t---------    151
CONSENSUS-A   ------------------------------------------------.....----------------    153
CONSENSUS-CPZ -???-??--A--A--A-??-----A--?-----?---????-T--????......C-????--?------     67

CONSENSUS-A   GCATCagATGCtAaAGCatAtgatacaGAaat?CAtAATGTcTGG..?GctACaCAtgCCTGTGTACCcA    219
CONSENSUS-B   ---------------------------gg-a--------t---...--c--------------------    232
CONSENSUS-C   -----------------------g------G-g------------...---------------------    223
CONSENSUS-D   ---------------t-----a-a-----ggca------a-----...--------------------a-    225
CONSENSUS-E   -----------C------C----G------g-G--C---------...--c-------------------    226
CONSENSUS-F   ---------------T-------a-a---gG-A------------...----------------------    223
CONSENSUS-G   -----T---------------ag---T----gc------------...----------------------    225
CONSENSUS-O   --T-----------CcT-ACAAGC--T---cAG------A-T---...--At----A--------T--t-    211
CONSENSUS-U   ---------------t-----a-?------?ca------------...----------------------    216
CONSENSUS-A   ----------------------------------------------..----------------------    219
CONSENSUS-CPZ --C?--?-?--??????????AG------GG?---?---?-?---...--?--???????-----?--?-    106

CONSENSUS-A   CAGACCCcaacCCAcaAGAaaTacattTggaa...AATGTgACAGAAaagTTtaAcATgtGGAaAAATAa    286
CONSENSUS-B   --------------------g--gta------???--------------t--------------------    299
CONSENSUS-C   ----------------------ggt-------...-----a--------t------------------G-    290
CONSENSUS-D   ------------------------g-ac-----...-----c--------c--------------------    292
CONSENSUS-E   -------------------------cc-----...-----a--------t--------------------    293
CONSENSUS-F   ----T-----T---------G--gt-C-----...-----A--------t---g-T--------------    290
CONSENSUS-G   -------------------g---?c-C--?--...-----A--------t-----?--------------    289
CONSENSUS-O   ----?----cT-----T---TaT-ca?-aa-?...-----------T--c--?--T--A----------T-    274
CONSENSUS-U   ---------g-------------?T-------...-----A--------?--------------------    281
CONSENSUS-A   ------------------------------...----------------------------------    286
CONSENSUS-CPZ ?---T--???T--T???-?-G-????C-????...---?????--??T??-----????-??------?-    141

CONSENSUS-A   cATGGTaGAgCAgATgcAtgaaGATaTaAtcAGtcTATGGGAc...cAaAGcCTAaAgCCATGTgTaaAg    353
CONSENSUS-B   ---------a-----------g------------t-------t...------------------------a    366
CONSENSUS-C   ------g--t-----------g------------t-------T...------------------------    357
CONSENSUS-D   ------G--------------g------------t-------t...-----------a-----------a    359
CONSENSUS-E   ------------------g--g---g--------T-------T..?-----t------------------    360
CONSENSUS-F   ---------A--A------AC-------------T-------T...---------------------G---    357
CONSENSUS-G   ---------A-----------G-----------?t-------T...G-----------------------    355
CONSENSUS-H                                                                      ------      6
CONSENSUS-O   ---------?--A-----g-----C--t--t---T-?-----?...--G--Tt----a--T-----tc-A    338
CONSENSUS-U   ?--------?-----------G------------T-------T...-----------a-----------?    345
CONSENSUS-A   ----------------------------------------------..----------------------    353
CONSENSUS-CPZ T-----?--????---???-----?--??-?????-?-???-?...--?--?-??-----?--------?    181
```

```
                                <- tev or tnv in LAI/IIIb
                               \/ 3'sj in LAI/IIIb
CONSENSUS-A   tTAaCCCCtCTCTGcGTtACttTaaatTGt...a?t?????????????????????????????aa??    387
CONSENSUS-B   --------a-----T--------------c...-c-gat--g------------------------??--    403
CONSENSUS-C   --g-----A-----T--c--------?---...-ga-------------------..............--tg    393
CONSENSUS-D   --------a-----T--C--------c--c...-c-ga------g---------------------??--    395
CONSENSUS-E   -----T-------------------------...-ccaatgct--------------------------tt    403
CONSENSUS-F   --------A-----T---------?-----???-c-aatgccac----...................--tg    401
CONSENSUS-G   ?-------------T-----------C---...-c-------------...................--tG    391
CONSENSUS-H   C-------------T-----------A---................................AAA--TT     43
CONSENSUS-O   A-G--TTTC?-g--T--ACAAA-G------...-ca-------------------.................A    373
CONSENSUS-U   -----------?--T--?--------C---...-c-gatg-aaag---------.........-----Ta    388
CONSENSUS-A   ------------------------------...-------------------------------------    387
CONSENSUS-CPZ ?-?--?---?-?--T?-A-???-G?-?--?...--?-----------------------..........    200

CONSENSUS-A   tcac?????????????????????????????aa?a?cac?a?????????????????????aa?a?    401
CONSENSUS-B   ????---------------------------------tg-t--t?-t-------------------t--    416
CONSENSUS-C   ?t-----t--------------------...-----tg-t--c?-ta-------------------t?-    410
CONSENSUS-D   ??---------------------.............--tg----t?----------------------t?-    408
CONSENSUS-E   -g--c---------------------aatg-ca-t--c-ta--c-atgtc-ctaaca-aatagga--t-t    448
CONSENSUS-F   ?----------------------a--g-cact---??-?-???-?--caa-ac-g-.........??-?-    416
CONSENSUS-G   -a--ctatact------..................??-?-???-.......................-A    404
CONSENSUS-H   AT--TAGACCT.......................--TTGT--TGGAACT................--T-G     74
CONSENSUS-O   -a--a------------------.............--tga?--a--a---...............??-?-    388
CONSENSUS-U   ga--tAac---------------.............--T?-?--tgaAg-a...............??-?-    407
CONSENSUS-A   ------------------------------------------------------------------------    401
CONSENSUS-CPZ ....................................-A???A-CA---------------........    206

CONSENSUS-A   ?ac??????????.........????????????????????ga?.......................    405
CONSENSUS-B   --?t---------?????????-----------a--------a?-??????.........?????????    420
CONSENSUS-C   t--c---------???.........------------------??-........................    414
CONSENSUS-D   -------------???.............----------------??-?????????????????????????    410
CONSENSUS-E   a--a------.................................................?????????    452
CONSENSUS-F   -??----------.............................................???accctg    422
CONSENSUS-G   TtgtACT------???g?gaac------aacAAtagcactgTg??-.....................g??    432
CONSENSUS-H   T--TGTAAGTAATAGTACTGAAACTAATTGTACTGGAAAGAACA-TAGCTTTGTGGATGATCGTATC...    141
CONSENSUS-O   -??----------?????????---------------------??-??????..................    388
CONSENSUS-U   -??-------................------------------.........................    407
CONSENSUS-A   -------------.........------------------------.........................    405
CONSENSUS-CPZ ........................---A-----A---C----?C-CC?????????????.........    212

CONSENSUS-A   ...atgaaa???...g?agaa...............ataAaaAAcTGctCTTtCAATAtgACcaCAgaa?    449
CONSENSUS-B   ???---g-gaaa???-g----???????????????-----------------------c------agca    469
CONSENSUS-C   ?????????---???------......?????????--------t------------gca----------    452
CONSENSUS-D   ??????g-?---...a---g-................--G--------------------a---------g    451
CONSENSUS-E   ?????????......-at---................G---g------T-----T---------------c    492
CONSENSUS-F   aaggaag-?cca...-gg-c-................---C-------T---------------------g    473
CONSENSUS-G   ??c-aa???......-a----................--g--------------------A---------a    476
CONSENSUS-H   ................CA----................T----------------------A---------A    181
CONSENSUS-O   ................aat--g??????.........?-G---?-g--Tgag--T---g-A--?----Ttc    425
CONSENSUS-U   ......???---???-----?................--g-----?------------?-A----------    442
CONSENSUS-A   ...---------...------................------------------------------------    449
CONSENSUS-CPZ ?????????---????T??-?...............?-??--?-?--T???--T---??--??----?-T    235

                -> tev or tnv in LAI/IIIb
              5'sj \/
CONSENSUS-A   TAAgagATAagaaacagaaagtatattCaCTtTTtTAtAgacttGATgTaGTaCaAAttaat????????    511
CONSENSUS-B   -----------gtg-------a----g------------a--------------c---ag--aat-at--    536
CONSENSUS-C   ----------------------g---g--------------------a------c-c-------------    514
CONSENSUS-D   -------------ga--c-----c--g------------a---------?--------ag-?--------    511
CONSENSUS-E   -------------g-----G--cc--G------------ag------a-----------g-agataa---    559
CONSENSUS-F   -----------C-G-t-------C--g--------------------A------c---c-gcaat---a-    539
CONSENSUS-G   --------------a---?--A---CG-g-----C----a--------------C---------------    537
CONSENSUS-H   -------------GA---C--A---CG-T-----C-------A-----------C-------GGAAAG..    249
CONSENSUS-O   -C-a---c--A--GG-----AA-C-GG-T--A--C---GT-tCA---t-g?-ta-?------gacacaa-    492
CONSENSUS-U   -------------Ga---?--a-?--g-?-----?----a---a----------c-------gataat--    506
CONSENSUS-A   ------------------------------------------------------------------------    511
CONSENSUS-CPZ -?-?---------??-A?-??-??-??-??-A------???G??---?-??-?A-??????G---.....    269
```

```
CONSENSUS-A   ?????????????aata?ta?t????????????.............aata?ta?t???????????????    525
CONSENSUS-B   -------------????-??-?------------???????????--------?---------------    542
CONSENSUS-C   ---------------?-a----a-----------??????......--??---g-------.....---    528
CONSENSUS-D   ---------------------?-----------............-------c-------...------    525
CONSENSUS-E   -------...--------g--g-------------???.........????-??-?................    568
CONSENSUS-F   -----------------ac-g----agagaa---???...............?-?---............    554
CONSENSUS-G   -...................g--g----.....................---?g----------------a-t    555
CONSENSUS-H   ..................................................--G-A--G-GAT............    261
CONSENSUS-O   -------.....---?-a-CA---------.......................................    499
CONSENSUS-U   -............----a--a----------...............????-??-?---............    515
CONSENSUS-A   ----------------------------------.............------------------------    525
CONSENSUS-CPZ .............--??-?.........................--??A?-C?...............    276

CONSENSUS-A   ????a?TATagatTAATaAatTGTAataccTCAgccatTAcAcAgGCtTGtCCaAAggTatCCTTTGAgC    590
CONSENSUS-B   ----?------g--g----g------c-------t------------c----------------------    606
CONSENSUS-C   ----?-----------------------------------A-----a--C-----------c--t-----c-    592
CONSENSUS-D   ---a?-----------------------------------------g------------a---------    590
CONSENSUS-E   ----?------g-----------------T----T-----aG---------------a----------T-    632
CONSENSUS-F   .......--C--gC--------------------A----------------------------T-GG--T-    618
CONSENSUS-G   agtg-t-----gc--------------gt----A------A------------------ga?T-----c-    624
CONSENSUS-H   ......----T-C--------------GT------------A------------------GAGT-----T-    325
CONSENSUS-O   ...atg----c-----?T--------CT--A--A----c-aG--?--C-----?--------t-------    563
CONSENSUS-U   ...-?------g?---------------------------?-----------------------------    577
CONSENSUS-A   ----------------------------------------------------------------------    590
CONSENSUS-CPZ ----?----????--????-?--C--?--?A--?-?--?-----A?-?---??---??C??-?-----?-    316

CONSENSUS-A   CaATTCCCATAcaTTATTGtgCcCCaGCTGGtTttGCgATtCTAAagTGtAa?gataaggagTTcaatGG    659
CONSENSUS-B   -------------------------g---------------------------t------a----------    676
CONSENSUS-C   -------T--------------t----------a------------------ta-----aca--------    662
CONSENSUS-D   ------------------------------a-----a--------a------------A----------    659
CONSENSUS-E   -------t------------a-t----------a-------t----------T------a-t--------    702
CONSENSUS-F   ----------------------T----------A------------------T-----aA----------    688
CONSENSUS-G   ----------------------T------------------t---------gg------a-?--------    693
CONSENSUS-H   ----------------------T--------G---------A--------GG------A----------    395
CONSENSUS-O   -?-----?-----c--------T-----a--A-a---c--CT-T---------?-c-CA--a--T-----    629
CONSENSUS-U   ----------------------------------------------?-----t------A----------    646
CONSENSUS-A   ----------------------------------------------------------------------    659
CONSENSUS-CPZ -?-----?-----?--------?-----A--?-?---?--??-?-?A--??-----?-?--?--T?CA--    368

CONSENSUS-A   AacAGGgccatGcaagAATGTcAGcaCaGTaCAATGcACacATGGa...ATcAagCCAGtagTatCAACt    726
CONSENSUS-B   ------a-----t-ca--------------------t---------...--t-g----------------    743
CONSENSUS-C   g-----a------c-t--------------------t---------...--t--------g---------    729
CONSENSUS-D   g-----?--------a--------------------t--------g...--t-g---------g------    725
CONSENSUS-E   g-----------t--A---------T--------------------...--T--------G---------    769
CONSENSUS-F   G--g------------------------------T---------...--T--A-----g---------    755
CONSENSUS-G   ------A-----T--a--------T-----------T---------...--T--------g---------    760
CONSENSUS-H   G--G--AA----T--A--------T-----------T---------...--T--------G---------    462
CONSENSUS-O   ------c??------a--ca-t-CaGT---TACT--T--------C...----------AC----AGt---    694
CONSENSUS-U   ?-----a--------A--------?-----------?---------...--t--------?--?------    708
CONSENSUS-A   ----------------------------------------------...----------------------    726
CONSENSUS-CPZ -??---?A??--T-??-----??????---?--T--?--?------...--?-??---?-G--??-?--?    411

CONSENSUS-A   CAaCTgcTGtTaAATGGcAGtcTAGCAgaAgaa???gaggTAatgaTtagaTCTgAaaataTcacAaAcA    793
CONSENSUS-B   ------------------------------...------g-a---------------t----gg---    810
CONSENSUS-C   -----a-----------t--c------------...---a----a---------------c-g-------    796
CONSENSUS-D   ------?----g---------------------...---a----A---------------c---------    791
CONSENSUS-E   ---t-----------------------------...---A----A--c------------C---------    836
CONSENSUS-F   ---T--T-------------C------------...--ta----A--c------c--------t--g-t-    822
CONSENSUS-G   ---t-a---c-g---------t-----------...--aA----a--------------c?-----g---    826
CONSENSUS-H   ---T-A---C-GTCAAATG-CAGTTT--C---------?a----a--------------c------g---    528
CONSENSUS-O   --?--AA-A?-------G-CA--CT-TA---g-...A-aA---gA--?-TGGgAA-------Tt-gG--?    757
CONSENSUS-U   -----------?----------------------...---?----A---------------------g---    773
CONSENSUS-A   ----------------------------------------------------------------------    793
CONSENSUS-CPZ ??GT-A--??-?-----A-?-T???????-???---????--????-???????-?????-A?????-A?    436

CONSENSUS-A   ATgccaaaAccaTAaTaGTacAgcTtg??aagcctGTAa?aATtaatTGTaccAGACct...aacaacaa    857
CONSENSUS-B   ----t--------------------gaa-g-at-----ga------------a-----c???--------    876
CONSENSUS-C   ---t------a-----------t---aAtg-at-----ga----gtg-----a-----c...-----t--    863
CONSENSUS-D   ----t----?----------------AAtG--t-----?c------------a--g--c???t----?--    855
CONSENSUS-E   -------------------G--C---AAT--At-----Ga---C--------------c...Tc------    903
CONSENSUS-F   ----A-----------------?---AATg-At-----ca------------A-----c...--------    888
CONSENSUS-G   --------gt---------g------AAt--a?--a--ga-----??------------c...--t-----    890
CONSENSUS-H   --a-----gt----------------AAt--a?-----g-------?-----------c...--------    592
CONSENSUS-O   g--gg----At--C------ACC--AAAttcTa--a---a?--g-Cc---gag-----A...gga--tc-    823
CONSENSUS-U   ---------?----------------aatg--?------a------------------c????-------    837
CONSENSUS-A   ----------------------------------------------------------...---------    857
CONSENSUS-CPZ ?-?????TGT??G??-?-?-?-???A?T-G-A????-?--??-?-??---???-----A...GGA--?-?    469
```

HIV1 ENV CONSENSUS

```
CONSENSUS-A   tacaAga...Aaaagt?????????gtacgtata?????????GgacCAGGaCaA??????gcatTctaT    900
CONSENSUS-B   -------...-------------...a---a--------------------gag-......-----t---    919
CONSENSUS-C   -------...-----------------a--a-g---.......................a--------    906
CONSENSUS-D   -------...c----?.........---ac--c----.............----t---g---.......---c-----    897
CONSENSUS-E   -------...-C----.........a--ac----.......................-------t-------    946
CONSENSUS-F   -------...-----------------A---a----.......................g-......-----t---    931
CONSENSUS-G   -------...-----------------A--a??t-c.......................--------g------    931
CONSENSUS-H   -------...-------------...A--a-g--c.......................--------g------    635
CONSENSUS-O   ?---gT-...C--gAg.........A--a??---...........--T---...aTG.......--?-GG--c    859
CONSENSUS-U   -------???-----------------a--------.......------------------.---------    880
CONSENSUS-A   -------...------------------------------------------------------------    900
CONSENSUS-CPZ ?---???...?G-?A?.........?-?-A?---...........--?------ATG......A-?--?---    499

CONSENSUS-A   gcA...acaggtgacata...............atAGGg...gAtAtAAgacAaGCacAtTGtaatgTcA    949
CONSENSUS-B   a--...-----a?-a---?????????.....-----a...------------------------ca-t-    967
CONSENSUS-C   ---...-----a-----.................-----a...--c--------------------cA-T-    955
CONSENSUS-D   A--???---???aga---?????????.....-----a???------------------------a-T-    943
CONSENSUS-E   ag-...-----A-----.................-----a...---------a-----T-----G-gA-T-    995
CONSENSUS-F   ---...-----a--a---.................-----a...--c--c---a-G-----------c--t-    980
CONSENSUS-G   ---...------------.................-----a...---------------------------t-    980
CONSENSUS-H   ---...--------?---.................-----A...---------------------------T-    683
CONSENSUS-O   AGc...-TG--c...c--gag?ga???aa?aca-a??????agctC---Gg?---TT-------?TAT-    910
CONSENSUS-U   ---...--------a---.................-----a...------------------------?a-t-    928
CONSENSUS-A   ---...------------.................------...---------------------------    949
CONSENSUS-CPZ AA?...?T--AAA??G?-................??---A...--?-C?--????--?T?C----???-?-    529

CONSENSUS-A   gtaga?cagaaTGGaAtaaaacttTacaa?aggta............gcta?acAatTAagaaaa.....    999
CONSENSUS-B   -----g--a----------c------a--c--a--............-t----A--------g--?????   1019
CONSENSUS-C   ---a-ga?a--------?-----------a-----............ag--a-a-----gc-g--.....   1006
CONSENSUS-D   --g?-g----------------------c-----............----a-A-----g--g-c?....    995
CONSENSUS-E   A-g--A--a--------g--g-----a--c-----............a--ga-a------a-g-g.....   1048
CONSENSUS-F   --g--a--C----------?--?---g--a-----............a?ggc-a-g----agtct.....   1030
CONSENSUS-G   -----a?-a-t-----?g-g-tG-----ga-t--?............a??gc-?--C---a-g--......   1027
CONSENSUS-H   -----g?-a-?------g-g-tg---?-?a-----............-?-?c-------?a----.....    729
CONSENSUS-O   a-gccA-T--t---g-a---g-c---a--c-aAc-............---ga-AGg-ATTT-g--.....    963
CONSENSUS-U   --g--a-------------?---------c-----............------?-----------.....    978
CONSENSUS-A   ------------------------------------.............------------------.....    999
CONSENSUS-CPZ A-??GA-???----?-?C?-?-?????-?--A??-??????????????AG??T-????--??.....    552

CONSENSUS-A   .tacTtt?????????aacaaaaca............??????ataatcTTtgctaac...?cctcaGGa   1037
CONSENSUS-B   ?c-a---...g-g-----t------.......................g-----aa-c-a???t--------   1060
CONSENSUS-C   .c----ccct--------T------.....................----aa-----acca...t--------   1048
CONSENSUS-D   .cTtc--......------------..............aca-----t---aaacCa...t--------   1037
CONSENSUS-E   .C---------...a-t--T--G---.....................---------caaCCA???c--------   1089
CONSENSUS-F   .c-t--c----...-----tgc----.......................aa---aactcA...t--------   1069
CONSENSUS-G   .at--------------------?c.........???----------c----aaCtCA...t-tg-----   1065
CONSENSUS-H   .?---------...a----t----??.......................c----aaacca???t--------    766
CONSENSUS-O   .CTtg-A.........---?-t---?????aata?tgTtAcc--g-?A--Caa-c-?agcAgt???--t   1010
CONSENSUS-U   .c-----......aat--t-??---.......................---------?---aa??cA???t--------   1015
CONSENSUS-A   .------------------------...........---------------------...---------   1037
CONSENSUS-CPZ .??A???..........--???---?.........???GC-AA--?--?A???A?????...-?-???-??    569

CONSENSUS-A   GGGGAt?TaGAAaTtacAAcacAtAgttTTAaTTGTggAgGagaattt..?TTcTAtTGcaataCAtCag   1103
CONSENSUS-B   -----cCc-------gt--tg--c----------------g------...-----c--t------a--c   1127
CONSENSUS-C   -----cc-------------------c---------a-----------...------------------a   1115
CONSENSUS-D   -----ccc---------------c----------------g------...-----C------------a   1104
CONSENSUS-E   --a---C------------tg---ca----------A----g------...---------------a--c   1156
CONSENSUS-F   -----CC------------tg---------------a-----------...-----C-------------   1136
CONSENSUS-G   -----cC-----------------------------a-----------...--------t----------   1132
CONSENSUS-H   -----Cc------------------?----------a-----------...--------t------a--a    832
CONSENSUS-O   --A---gc---gg-A--ccATtT?ca------C---CAt--------C...--T-----t--C-----T?   1075
CONSENSUS-U   -----------------------c----------------------...------------------?   1080
CONSENSUS-A   ------------------------------------------------..-------------------   1103
CONSENSUS-CPZ --A---CC?--?G-?-?-???---??G--------???---------?...--?-----T?---??A??C    617

CONSENSUS-A   g?cTgTTt...aatagcacttgg?a?????????????????.............aatagcact?????   1132
CONSENSUS-B   aa------...-----t-------?-------------------???????????---g?t--------   1155
CONSENSUS-C   ac------...-----t--a-ac-?-------------------...?????????---g-t--a-a---   1145
CONSENSUS-D   -a------...-----t--a----?-------------------????????????----?t--?-----   1131
CONSENSUS-E   aa------...----at-----catagga------...................??---gaa--catg--   1194
CONSENSUS-F   aa------...---ga---a???-?-----------------------------?a---------   1161
CONSENSUS-G   -a------???----at-g-?t?-?-------.........................----at---ac---   1161
CONSENSUS-H   aa------...-----T-------t---t------.......................--?--?----a---    862
CONSENSUS-O   -ga-----...---tAT--c-TTtcatgt---------------???????????---?-t--ca--t-   1111
CONSENSUS-U   aat-----...-----t-?-----?----...........................--????--?-ggga   1107
CONSENSUS-A   --------...-----------------------------------.............--------------   1132
CONSENSUS-CPZ ?-???--?...-?-G?-??????-?-------.........................??C-T?--A.....    629
```

```
CONSENSUS-A   ??????????????ca?????????aat?acac?????????????.....???????a?tcaaatgac   1151
CONSENSUS-B   --------------??------------a-?--t------------??????-------?-????-?--?   1166
CONSENSUS-C   --------------??-----------?-??---------------??????-------?-?????????   1149
CONSENSUS-D   -------...----??---.....---a?t--a------------??????-------?-????-????   1140
CONSENSUS-E   ----------gaggggtgt....-----gg---t..........................????????   1212
CONSENSUS-F   -----.....................???-????----------.........-------?-?????????   1161
CONSENSUS-G   --------------??---...-------??-----------.....................?-?????????   1166
CONSENSUS-H   ----...aat----??---.....---g??---.......................-?-??????a??   872
CONSENSUS-O   --------------??-aa---a-c---gt?--taAt--t....................a??----g-   1134
CONSENSUS-U   ------.......-??---...-------?--?-------...---??????-------?-???-?--?t   1116
CONSENSUS-A   ------------------------------------------.....------------------   1151
CONSENSUS-CPZ ..............................................................--?-G?   633

CONSENSUS-A   actATaact......ctccaa...TGcAgaATA...AagcaaaTTgTaaatATg...TGgCAgagagtaG   1206
CONSENSUS-B   -----c--a??????----c-...---------...--a------a----c---...------ga-----   1221
CONSENSUS-C   -?c--c--a......?---c-...---------...--a------a----c---...------gag----   1202
CONSENSUS-D   --a--c--a.....----c-...---------...--A------a----c----..?------g------   1195
CONSENSUS-E   ???--C--A.......--T-C-...----ag---..?---------a----c---...------g---c--   1264
CONSENSUS-F   -----c---.....----c-..?--T------...--a-----------c---...------ga---g-   1216
CONSENSUS-G   ???--c--a.....----c-...--T--g---...--A----------ga---...----------g-   1218
CONSENSUS-H   -??--t--a.......--g-c-...---------...--A--------g-ga---...----------g-   925
CONSENSUS-O   ---?a????......a-A-CT...-----gt--...-gA---g-g----Gatca...---AT---G-G--   1184
CONSENSUS-U   -?---?--a.....----C-...---------...--a----------g?---...-------------   1168
CONSENSUS-A   ---------.....------...---------...------------------...-------------   1206
CONSENSUS-CPZ ????-?-TA......???-??...---------T...-??--G--A--T-?-???...---???-??--??   665

CONSENSUS-A   gacaA...gCaatgTAtgCccCtCCcATccaa...gGagtaaTaaggTgTgaatCAAacATtACaGGacT   1270
CONSENSUS-B   --a--...---------------------ag-...---ca---t--a---tc------t--------g--   1285
CONSENSUS-C   ---g-???--------------------tg--...---aac----ca---a-------t--c--------   1266
CONSENSUS-D   --a--...---------------------g--...---c----c-aa---tc------t--------g--   1259
CONSENSUS-E   -----...-----------T---------agt...---ag---T-At----T------T---------A-   1328
CONSENSUS-F   ---g-...------------g-------tgc-...---aac--T-cc---A?c-----t--------t--   1279
CONSENSUS-G   -----...---------------------GC-...---AaC--t-ca---a-------------------   1282
CONSENSUS-H   -----...---------------------gc-???---AAc--T-cA---a-------------------   989
CONSENSUS-O   -gTCG...-G-c-c-----a----------?-...--tAAcc---cA---a??--------A--T---A-   1245
CONSENSUS-U   -----...--------------------tgc-...---a?---t-??--c?c------t-----------   1228
CONSENSUS-A   -----...------------------------...-----------------------------------   1270
CONSENSUS-CPZ ?-A?-...-G?--?--???A?-C--?-??AG-...--?AA??-??C?--?A?T--?-?---?--T---?-   706

CONSENSUS-A   acTatTaACaagaGAtggtGgg?at???............aataata?ta?a???.........??????   1304
CONSENSUS-B   g--------------------ta-----???????????--?--?-cc?-?---???...???------   1317
CONSENSUS-C   ------g---c-t-----a--aa?a---???????????-----?-ca?-?---?????????------   1298
CONSENSUS-D   -t----g--------------ta?a---???.............----?----?-?---.......???------   1290
CONSENSUS-E   ------g--------------tgc----???.....------------cg-c?---..............   1364
CONSENSUS-F   ------g--------------tc-gaat??????.........---?-?-c-?-?---???...........   1314
CONSENSUS-G   ----------------------a-----...................-------ca?-?.................   1316
CONSENSUS-H   -a---------------------??-a-???...............--??c-?c?---................   1020
CONSENSUS-O   GA-tc--CA-?Tg---a??...cCATGG.............--c-?c-gc?-?---................   1275
CONSENSUS-U   -?----g--------------ta-----?.............----?--g-?-?---................   1260
CONSENSUS-A   ---------------------------.............---------------.........------   1304
CONSENSUS-CPZ T?-?--???????--???-C??-T?--T..............--??-?---.....................   721

CONSENSUS-A   ?????????aatgagAccTTcAGaCCtggaGGagGAgatAtgAgggAcAAtTGGAgAAgTGAatTATAtA   1365
CONSENSUS-B   ----------c?----t-------------------------------------------------------   1377
CONSENSUS-C   ---------???-----a-------------------------------------------------------   1356
CONSENSUS-D   ----------------------------------------------------------------------   1351
CONSENSUS-E   .....------c------------------------A----a-A--------------------------   1425
CONSENSUS-F   ...------???----------------g------a------aa--------------------------   1372
CONSENSUS-G   ...a-------------------------------------------------------------------   1378
CONSENSUS-H   .....---???------------------------------------------------------------   1078
CONSENSUS-O   -------a---c???--ac-T-----a-t---G----------aA--T-TA-------C------g-tC-   1334
CONSENSUS-U   ...----a--?-----------------------------------a-------------------------   1321
CONSENSUS-A   ----------------------------------------------------------------------   1365
CONSENSUS-CPZ .....---????T?-?A?-???----???-??-?-A???-???A--????------???--G?-?-??-   749

CONSENSUS-A   agTAtAAaGtaGTaaaaaTtgaaCCacTaGGagTAgcACCcacCcgg...GCAAaaaGaAGAGtGGTG..   1430
CONSENSUS-B   -a-------------------------t-----------------aa-...-----g------------..   1442
CONSENSUS-C   -a--------g---g-----a-g---t-g--------------ta?-...--------g---------??   1420
CONSENSUS-D   -A-------------g---------------------------a--...-----g------------..   1416
CONSENSUS-E   -A------------C-----------------a-----------a--...-----G------------..   1490
CONSENSUS-F   -a------------g----------------------------A--...----------ca-------..   1437
CONSENSUS-G   -------------------A---------------------a--...----GG------------..   1443
CONSENSUS-H   -?-----------------c?--?--------------------a--...----gg------------..   1140
CONSENSUS-O   -a--c----------ggg-AA----TT-TA-T--g-----t--AaaaATt----GgCC--tcA-A-gcAc   1404
CONSENSUS-U   -a-------------------------?-----------------a--...-----g------------..   1385
CONSENSUS-A   ----------------------------------------------...------------------..   1430
CONSENSUS-CPZ ?---??-?--?--????--?-??--?T-?TC---?--?--A--AA?AG??A?--GG????C--????AA?   790
```

HIV1 ENV CONSENSUS

```
                                gp120 \/ gp41                          <- RRE or CAR
CONSENSUS-A   ..........gagAGAGAaAAAAGAGCa??????gTtGgA...cTgGGAGctgTctTCcTTgGGTtcTTa   1481
CONSENSUS-B   ..........c------------------.......--g---???a-a------a-g--------------g   1493
CONSENSUS-C   ????......------------------...-----g---???a-a--------g--------------g   1471
CONSENSUS-D   ...........--A---------------......A-A---...--a------a-G--------------G   1467
CONSENSUS-E   ..........------------------------.....--G---...A-A------A-Ga--t-----------   1541
CONSENSUS-F   ..........a-------------------------g---...a--------t-g--------------g   1488
CONSENSUS-G   ..........-G----------------......-------...?--------------------------   1493
CONSENSUS-H   ..........------------------.....-------...---------------------------   1191
CONSENSUS-O   T?g?aCt???C-t---------------.......--A---...T-----ATGC-A---T-G---g-tC--   1460
CONSENSUS-U   ..........------------------...-----g---...a--------?-G--------------G   1435
CONSENSUS-A   ..........-----------------------------...---------------------------   1481
CONSENSUS-CPZ ACA??A?...??C---C-?-------??...---?-?--?...?-?--?????-?---?-?--??-?C-?    823

                                  /<-      RRE primary binding site      ->/
CONSENSUS-A   GGAgCaGCaGGAagcactatgggcgcggcgtcaataacgctgacggtacaggccagacaattattgtctg   1551
CONSENSUS-B   --------------------------a--------g----------------------------------   1563
CONSENSUS-C   --------------------------?----------------------------------?-------   1539
CONSENSUS-D   -----------------g--------A------?-g-----------------------g----------   1536
CONSENSUS-E   ------------------------------------------------------------------------   1611
CONSENSUS-F   ------------------------------------------------------------------------   1558
CONSENSUS-G   --------------------------a-------------------------T-----------------   1563
CONSENSUS-H   -------------------------------------------------------?----------------   1260
CONSENSUS-O   A-T--------t--------------A---G---C-g--------------a--CAcacT--ga--AAG-   1530
CONSENSUS-U   --------------------------?---------------------------------------------   1504
CONSENSUS-A   ------------------------------------------------------------------------   1551
CONSENSUS-CPZ ?-T-----?--?--?--?--------A---?--?--?---------?----?--?-??-?--G?-????-    874

CONSENSUS-A   gtatagtgcaacagcagagcaatttgctgagggctatagaggctcaacaacatctgttgaaactcacggt   1621
CONSENSUS-B   -------------------a------------------t-----g-----g---------c-------a--   1633
CONSENSUS-C   ----------------a--------------?-----------g-----g---a-----c----------   1608
CONSENSUS-D   ----------------A-A------------------------G-----g---------C-------?--   1605
CONSENSUS-E   ----------------a--------------------------g--g--g---------c-------a--   1681
CONSENSUS-F   -?----------------a------------------t--a--G-----G---------C-g-----a--   1627
CONSENSUS-G   -C--------------A--------------a-----------G--g--g---------C-------A--   1633
CONSENSUS-H   -c--------------a--------------------------?--?--g---------?-------a?.   1325
CONSENSUS-O   -----------------GA---CC----a--A--a---C----C--G-----aT--C---GG--AT-T--   1600
CONSENSUS-U   ----------?-----A?A?----------------------G-----G---?-----C-------A--   1570
CONSENSUS-A   ------------------------------------------------------------------------   1621
CONSENSUS-CPZ ----?--?--?--?---???---C--?--?AA--?------?-?--?--?--?--??-?C-?--?T-??-    922

                                                     RRE or CAR ->/
CONSENSUS-A   ctggggcattaaacagctccaggcaagagtcctggct?tggaaagatacctaa?ggat...caacagctc   1686
CONSENSUS-B   ---------c--g------------------------g---------------a----..?---------   1700
CONSENSUS-C   ------------g---------a--------------a-a-------------a----...---------   1675
CONSENSUS-D   ----------------------------?--------g---------------a----...---------   1671
CONSENSUS-E   -------------------------------------g---------------a----...---a--t--   1748
CONSENSUS-F   -------------------------------------g---------------a----...---------   1694
CONSENSUS-G   -------------------------------------g-A-------------a----...---------   1700
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   A-----T--?-G---A----GA--TC-CC-G--A--CT-A----CC-T?a--CA-A--...--G--A---   1665
CONSENSUS-U   ----------------------------?--------G--------------?A-??-...---------   1633
CONSENSUS-A   ------------------------------------------------------------...---------   1686
CONSENSUS-CPZ ------AG-A-----A--?--?--?---?-G--T--?G-?--?-?G-----??-?---...--?--??-?    973

CONSENSUS-A   ctaggaatttggggctgctctggaaaactcatctgcaccacta?tgt?ccctggaactctagttgg....   1750
CONSENSUS-B   --?--g--------t-----------------t---------gc---g--t-----tg--------....   1765
CONSENSUS-C   -----g-----------------------------------gc---g--t--------?------....   1740
CONSENSUS-D   --------------t----------?--a---t--------------g------------------....   1735
CONSENSUS-E   ------c--------------------a--------------gc---g-----------c-c----....   1814
CONSENSUS-F   -----gc-------------------------------------a---g------------------....   1760
CONSENSUS-G   -----g-----------------------------------A---G---------a--------....   1766
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   ---AaCC-?--------tAA?----??--A------TA---aTCA--AaAa-----?AaA-CA---acAg   1730
CONSENSUS-U   -----??-------?---?---------?---?----?---------G--------------C---....   1691
CONSENSUS-A   ------------------------------------------------------------------....   1750
CONSENSUS-CPZ ?-G?-CC-?-----?--??---??--GG????---T?A?--?-CG--G--T-----?AA?TC?---????   1021
```

```
CONSENSUS-A   .......................agt...aataaa?ct?a?a?tga?atatggga?aacatgacctggct   1788
CONSENSUS-B   .......................---???------t--ctgga-?-g--t---?-t------------a-   1807
CONSENSUS-C   .......................---???----?-t--c-a?-?--t--t---?-t------------a-   1779
CONSENSUS-D   .......................---...----g-t--cta?a---g--t---c--------------a-   1777
CONSENSUS-E   .......................---.......-g-t--tttgaa--g--t---a-c--------a---a-   1858
CONSENSUS-F   .......................---.......------t--c-ggag--g--t----gg------------a-   1804
CONSENSUS-G   .......................---.......-------t--TtT-A---g--t-----t------------A-   1810
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   ga....................????????????ga-a-tGAaagt--T-----c---?-a--A----A   1766
CONSENSUS-U   .......................---...------T-?-T-G-???-?-T---?----??--------A-   1722
CONSENSUS-A   .......................---...--------------------------------------   1788
CONSENSUS-CPZ ??????????????????????-?C...-?????-??A?A?A???---T----?---T?-?--A----A   1046

CONSENSUS-A   gcaatgggataaagaaattagcaattacaca?a?ataatatat??t...ctaattgaagaatcgcagaac   1851
CONSENSUS-B   -g-g-----a-g-------ga----------agct-------cacc...t--------------------   1874
CONSENSUS-C   ---g-------g---------t-----?-------c------cagg..?t-gc-------c-----a---   1843
CONSENSUS-D   -g-g-----a-g-------ga----------gg-t-------?agc...t----------------a-?-   1841
CONSENSUS-E   ag-------g-g-------------------a-cca-------gag...a--c--ac-------------   1925
CONSENSUS-F   -g-g-----a-----g------------t--a-cga------cagg...t----------?---------   1870
CONSENSUS-G   Ag-------?-Gg-------A----------c-cCa------CaaC...--Gc---------------?-   1875
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   -?--------C-?C-g--A-a---c?TA-GctcC-?c------GA-...gA---AC--?-gG-A--AG?a   1827
CONSENSUS-U   -G-G-----?-G-------?----?-------?-G-?-----?--?...T-??-------?--?--A-??   1772
CONSENSUS-A   ------------------------------------------...--------------------   1851
CONSENSUS-CPZ ??-------?-?-TT-G-????--?-??---GG-?A?--?-?----...?-?T-A?-???-G-A--????   1085

CONSENSUS-A   cagcaggaaaagaatgaacaagacttattggcattggacaagtgggcaa?tctgtggaattggtttgaca   1920
CONSENSUS-B   --a--a-----------------a-------a---?--t--?-------g-t--------------a---   1942
CONSENSUS-C   --------?--a------a----t---?-a----------??---?a--a--------?-------??--   1905
CONSENSUS-D   -----a-----------------a?-----?a---------a-------g-t----------------a?--   1908
CONSENSUS-E   --------c-g-------a-g--t--g--a-a------t--a-------g--------------------   1995
CONSENSUS-F   -----------------------a-----------------a-------g--------------------   1940
CONSENSUS-G   --------------------------------------------------G--------g-------Ag--   1945
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   --?------?-a-----ga--?-?--Gc---AG--A--TG-A-----CTC-a-T---------c------   1893
CONSENSUS-U   -----?-------?----?----?---?-??A------?--?-------G-?-------?-------?--   1830
CONSENSUS-A   ------------------------------------------------------------------   1920
CONSENSUS-CPZ --A--A--G-?A---?--A?G--??---??-A-?-?--??-A---??C?-?T-?----------------   1138

CONSENSUS-A   ta?caaa?tggct?tggtatataa?aatatttataatgatagtaggagg?ttaatagg?ttaagaatagt   1984
CONSENSUS-B   --a----------g----------a------c-----------------c--gg----t-----------   2011
CONSENSUS-C   --a----------g----------a------c-----------------c--g-----t---------a-   1974
CONSENSUS-D   --a----a-----g----------a------c-----------------c--g-----t-----------   1978
CONSENSUS-E   --a----t-----g----------a------------------------t--------t---------a-   2065
CONSENSUS-F   --t----------g----------a------c-----------------c--g-----c-----------   2009
CONSENSUS-G   --t----c-----A----------g------------------------t--------T----a------   2015
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   --A-T--A---T-G----------A----GC?-----C----------CAC--?----t?-----G-t?-   1959
CONSENSUS-U   ------?----?-G----------A------C-----------------C--G-----------------   1895
CONSENSUS-A   ------------------------------------------------------------------   1984
CONSENSUS-CPZ --A--?-A----------------A---?--??-T---??--------?-A-??-----?-------T?-   1194

                                 3'sj (minor) \/  \/ 3'sj (minor)            \/ 3'sj (major)
CONSENSUS-A   ttttgctgtgctt?ct?taataaatagagttaggcagggatactcacct?tgtc?tt?cagaccc?tacc   2048
CONSENSUS-B   ---------a---t--a--g-g--------------------------at-a--g--t--------cct-   2080
CONSENSUS-C   -------------?t--a--g-g--------------------------?t-?--g--t-------t--?-   2040
CONSENSUS-D   -------------t--t--g----------------------t------c----a--t-------tcct-   2048
CONSENSUS-E   -------------t--a--g-g---------------------------t----t--c-------c-?--   2134
CONSENSUS-F   -------------t--a--g--------------a--------------t----a--a-------a--t-   2079
CONSENSUS-G   -------------T--A--G-----------------------------t-a--a--C-------T-?--   2084
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   ?A-GATA?-A---AA-c--G-g--?-ACA-------------TCA---CC-C--G-------t--Cc-t-   2025
CONSENSUS-U   ----?-----?--T-----?-?----?---------------------?T---------------T???-   1953
CONSENSUS-A   ------------------------------------------------------------------   2048
CONSENSUS-CPZ ??-???A?-??-?-??-???-?-G?-??-----------?--????---C-?----------?-----?-   1233

CONSENSUS-A   c???acc??aggg?????c?cgacaggcccg?aagaatcgaagaag?aggtggagagcaagacagagaca   2105
CONSENSUS-B   -cagc--cg----ga...-c-----------a-g------------a----------ag-----------   2147
CONSENSUS-C   -caa---cg----ga----c--------t--g-?------------a-----------------?-----   2105
CONSENSUS-D   -cagc--cg----ga...-c-----------a-g----a-------a-----------?--g-----?--   2113
CONSENSUS-E   -atc-t-ag----aa...-c------a----a--------------g------c-------g--------   2201
CONSENSUS-F   -caag--cg----aa...-c-----------a-g------------g--------------g--a-----   2146
CONSENSUS-G   -ACC---AG----aA...-C---------?-A-g------------G-----------------------   2150
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   -aaC---a-g?--AA...GcA-?A-C---A-G-----CA-G-----g---------aG-------GCc--   2089
CONSENSUS-U   -CAA?--CG----GA...-C?--------?-G-?---?--------A---------??------???---   2010
CONSENSUS-A   ------------------------------------------------------------------   2105
CONSENSUS-CPZ --A-?--AG??--A-...--A-????--??---GA---????-----------?????-------?A?--   1274
```

HIV1 ENV CONSENSUS

```
CONSENSUS-A   gatc?at?cg?ttagtgagcggattcttagcacttgcctgggacga?ctgcggagcctgtgcctcttcag   2171
CONSENSUS-B   ----cggt--a-------at---------------at---------t----------------------   2217
CONSENSUS-C   ----c--t--a-------?-------------------------------------------?-----   2172
CONSENSUS-D   ----c--t--a--g----a--------c-------at-------------a---a---------------   2182
CONSENSUS-E   ----cg-g--a-----------------------------------t--a--------------------   2271
CONSENSUS-F   ----cg-ga-a-------a------------t----t---------c---a---a---------------   2216
CONSENSUS-G   ----C--t--a--------------------?--------------c----------------------   2219
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   -g-gG-?a-Cc-?GCC-CaA--?-----GC?---gtTG-ACACG--c--CA---CAA-AAT-T-G-GG-?   2154
CONSENSUS-U   ----C??T--A----?--?-----------------?------?--C---------------?-------   2073
CONSENSUS-A   ----------------------------------------------------------------------   2171
CONSENSUS-CPZ -?-?-??G??---???-?A?--G-?---??----??T?--------C--C--A-A----G-GA-T-GG--   1325

CONSENSUS-A   ctaccaccg?ttgagagactt??tcttgatt...g?a...gcgagg...?ctgtggaacttctgggac?c   2226
CONSENSUS-B   ---------c-----------ac--------?..-t-?..------?..at-----------------g-   2278
CONSENSUS-C   ---------a--------------a------...-c-...-----a...g-?-----------------?   2228
CONSENSUS-D   ---------c-----------aa----a---...-c-...-------...at-----------------g-   2243
CONSENSUS-E   ---------c-----------ca--------...-c-...-------...a------------------a-   2332
CONSENSUS-F   -----g--ac-----------ca-a--a---...-c-...-------...at----???............   2262
CONSENSUS-G   ---------------------CA--------...-c-...-------...A-a----------------G-   2279
CONSENSUS-H   ......................................................................   1325
CONSENSUS-O   T-------TC-----CA----Agca-CAGgG...ATCCaGAa-gt-ATCagc?AcCTGgGa--t----Tg   2220
CONSENSUS-U   -----?---------------A--?------...---...??????...-??-???-????---?--?-?   2109
CONSENSUS-A   ------------------------------...---...------...---------------------   2226
CONSENSUS-CPZ ------GA-C----C-AG---AGC--GC-A-...-TG...TG---A...CAAC-CA-GACGT------AT   1386

CONSENSUS-A   a?????????????????t??g???ggggtggga?g??ctcaa?tatctg...tGgAAtCTtcTg?taTatT   2268
CONSENSUS-B   -gg...----------...-?---?--------a-cc-----a---tg-...--------c---cag----   2324
CONSENSUS-C   -gcag-ctcagggga--acagag---------a?-c--t--g------...G-A-g----G--Cag----   2291
CONSENSUS-D   ......................ag---------a-cc-----a------...--------C---cAg----   2289
CONSENSUS-E   -gcagtctcaagggac-ga-acg---------a-gc-----a------...G-----------t------   2399
CONSENSUS-F   .......-acagggggc-ga-gag---------a-cc-----act----...G-------CaC-C-G----   2322
CONSENSUS-G   -aCAttCTCaAGGGAC-GA-ACT---------G-GC-----A------...------------t-G----   2346
CONSENSUS-H   ......................................................                 1325
CONSENSUS-O   TGGATCcTaGGg.........CA-Aa-A?AATT?a-GcTTGcAGA?-tTgTgcAGC-g-AA?aCA---c-   2276
CONSENSUS-U   ?---------------?--?--G---------A?C---?--A--?---...G-?-?----G--C?G----   2147
CONSENSUS-A   ------------------------------------------------...-------------------   2268
CONSENSUS-CPZ CTAATACTTCACAGCC-CA-ACT-CTAC----GAGA--GTGCCTG---...G--GGAA--A-TCA-----   1453

                                <- rev
CONSENSUS-A   GGGgtCgGGAaCTaAAaA?TAGTGCTattAaTTTGcTTgAtaCcatAGCAaTAGcAgTAGCTggctGGAc   2337
CONSENSUS-B   --a---A---------g-a-------g---gc-----ca--g---c---t-------------Agg----   2394
CONSENSUS-C   ------t---g-------aG----------g-c------------------------------AaG-A--   2361
CONSENSUS-D   --a?--a------g--g-a-----------Gc--------------------------------Agg----   2358
CONSENSUS-E   ----c-A-----------T----------TC-------a--g-t-C---------------G--G-----   2469
CONSENSUS-F   ------A---------G-A-----------GC------A------C--------t--------AGg----   2392
CONSENSUS-G   ------a---g-----G-A----------------?-------A------------------AA------   2415
CONSENSUS-O   --cTA-AA---T-Gc---a--------CA-??c-?--a-----tc-T---G-G-----t--cAAT-----   2343
CONSENSUS-U   ------?---------?-A?----------G-?--?--?--G---C--------?--------A?G-A--   2209
CONSENSUS-A   ----------------------------------------------------------------------   2337
CONSENSUS-CPZ ----AAA---G-------T---------A-GCC--------G-A-C---T--T--------A-AAG-T--   1523

CONSENSUS-A   AGAtAGggtTATAGAAaTAGgACAaagAattggTAGaGCtaTTctcaACATACCtagAAGAATCAGaCAg   2407
CONSENSUS-B   ----------------g---t-------gc-t?-------------c---------?------a------   2462
CONSENSUS-C   -------a--------t--at----------t----------ctg------------------a------   2431
CONSENSUS-D   ----------------?---t-------gc-t?-------g----Tc-------C-c-c----A------   2426
CONSENSUS-E   ----------------G---C----g--gc-T-G-----c------C-----------------------   2539
CONSENSUS-F   ------A---------GCTTTG------gc---------------t-----------------A------   2462
CONSENSUS-G   ----------------g---t-------gC-T----------------------------g--A-----A   2485
CONSENSUS-O   t--Cg-CA-a--Ctt-GGgaT---------A--a----Ga----?T-----c--A--------T------   2412
CONSENSUS-U   ----?-???-------?-???----------?----?---?-???---?--------------?------   2263
CONSENSUS-A   ----------------------------------------------------------------------   2407
CONSENSUS-CPZ ------AA-A------GCTTTT---GTT-CACT----AT----AGA-----C-----GC-C--A-----A   1593

CONSENSUS-A   GGctt?GAAAggGctTTgctATAA                                                2430
CONSENSUS-B   -----g------------------                                                2486
CONSENSUS-C   -----t---gCA-------a----                                                2455
CONSENSUS-D   -----g-----------a------                                                2450
CONSENSUS-E   -----A------------------                                                2563
CONSENSUS-F   -----g------------------                                                2486
CONSENSUS-G   -----a-----A------------                                                2509
CONSENSUS-O   -----a---?-Aa??--at-g---                                                2433
CONSENSUS-U   --?------???-----?-??---                                                2279
CONSENSUS-A   ------------------------                                                2430
CONSENSUS-CPZ --GC-G-----A-----A--T                                                   1614
```