HMMER Sequences in the Rev Alignment

A_U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
B_HXB2R	HIVHXB2R	K03455	Starcich,B.	Science 227, 484 (1985)
D_HIVELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
F_BZ163A	HIV1BZ163A	L22085	Louwagie,J.J.	ARHR 10, 561 (1994)
O_ANT70C	HIVANT70C	L20587	Vanden Haesevelde,M.	JVI 68, 1586 (1994)
O_MVP5180	HIVMVP5180	L20571	Gurtler,L.G.	JVI 68, 1581 (1994)
CPZGAB	SIVCPZGAB	X52154	Huet,T.	Nature 345, 356 (1990)
CPZANT	SIVCPZANT	U42720	Vanden Haesevelde,M.	Virology 221, 346 (1996)
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)
SAB_SAB1C	SIVSAB1C	U04005	Jin,M.J.	EMBO J. 13, 2935 (1994)
SYK_SYK	SIVSYK	L06042	Hirsch,V.M.	JVI 67, 1517 (1993)

HIV1 REV

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 types and subtypes are shown. An ordinary consensus sequence (lowercase signifies majority, uppercase signifies 50% or greater) was created from these sequences using MASE; this is not a "most likely sequence" generated from the HMM model (Part II). Annotation is based on HIV-1s, therefore users should be cautious about its applicability to other PIV sequences.

```
                    \/3'sj
REV_CONSENSUS ATGgcaggaagaagc?????????gaa???ga?gaaga?ct?ct?aaagca?t?ag??t?atcaag?tcc     48
A_U455        ------A--------GGA......A-CCCC--C--G--C--C--C--G---G-A--GA-C-----AC--T     64
B_HXB2R       ---------------GGA......--CAGC--C-----G--CA-C-G-A--G-C--AC-C------C-T-     64
D_ELI         ---------------GGA......--CAGC--C--G--C--C--C--G--GG-C--GC-C------T---     64
F_BZ163A      ---------------GGA......A-CAGC-GCACC--A--C--C-G----GCC--ATAC------A--T     64
O_ANT70C      ------------............AGCGAG--C--CC-G--G--GC----CA-CCAGA-C--A---A---     58
O_MVP5180     ---------------.........---GAA--CC-GC-G--G--GC----TA-CCAGA-A--A---A---     61
CPZGAB        ---------------GAACCA...C--GAC--C-C-CGG--C--GC-G---G-A-AAA-A------A-T-     67
CPZANT        ------------GAAGAGCTCGAA-G-ACA--C---C-A--G--G------G-G-AAA-A------A---     70
A_ROD         ---AAC-A---G............-C-GAC--A----GA--C-AA-G-AA-C-A--AC-C---CGTC---     58
B_EHOA        ---AAC-CTC-....................--AAG---T--T-AA----GGC-A--AC-AC-AC-TC-T-     52
SD_MM251      ---........--C--TCAC.......---AGA--A-----A--C-GA---AGGC-A--GC-A--AC-TC-T-     58
STM_STM       ---.........---..........--TCAA--A-----GT-AAGA---AGGC-A--AC-T---C-AT-T-     52
VER_AGM3      ---.........CC-.........CT-GGATCA-----A...AGA-G-TTGC-CC-CT-G---GCTT---     49
GRI_AGM677    ---.........TCT.........CT-GGAA-A-----A...AAAC-----C-C-AGA-A------AC--     49
SAB_SAB1C     ---.........TC-.........CT-GGCC-A-----G--T--A-G-AG-T-TC-CA-A------T-T-     52
SYK_SYK       ---.........CCA.........--CCAA-GAAGC--G...-AGCTC---G-ATTTC-G-GG-T-A-AG     49

                                      exon \/ exon
REV_CONSENSUS tgtAccaaagc?????????????????aAtcC?tatCCaaa?cc?g?aGGaac??g?aa?gcccGaAg     92
A_U455        -A---------...........................C--T-GC--C--T--CAG---GT-TC-GC-G---A-G-A    116
B_HXB2R       -C--T------...........................C--ACC---C--C--C-AG--G--CC-AC-G--------    116
D_ELI         -A---------...........................C--TCC---C-GC--C-AG--G--CC-AC-G--------    116
F_BZ163A      ------------..........................C--A-----C--G--C-AG-----CC-AC-G--------    116
O_ANT70C      ------------..........................------CC-A---TCA--AAG-----GCA-G--C---A-G-A    110
O_MVP5180     -------G---...........................C-G---C-CA--G-C-----GCA-A--C---A-G-A    113
CPZGAB        -A---------...........................C--T-----C-GT--A-AG-----AA-G--G--T--G--    119
CPZANT        ----T---..............................C--C-----A--A-C----T-CA-AGCA---A----    119
A_ROD         -AC----G-CA...........................------C-AGGACCG--G--AGCC-GCCAA-----    110
B_EHOA        --C------CGAGTGAGTATGGCACAACAG----C-----GC-AGGACCG--G--AGCC-GCCAA-----    122
SD_MM251      --C-T----CA...................--CT-A--C----CAGGACCC--C--TGCC--CCAGA----    110
STM_STM       --C-T----CAACA...............G----C------C-AGGGCC------AGCC--CCAAA----    107
VER_AGM3      --A--A---A-...........................------CCCGTG-A---G--AGCC-GACAA--C--    101
GRI_AGM677    -A---GGC---...........................------C-GTTCAGC-----GGCC-GACAA--G--    101
SAB_SAB1C     -C---AC--CT...........................------CCTGGACA---G--AGCC-GACAAA----    104
SYK_SYK       C-C---T-CAG..................G-A--C--C---GGG--G-A----C-CA-GC-AA----C--    101

REV_CONSENSUS ?aa?agaaGaag?agGTGG???aga?a?aggcaga?acagatcc??????tt???gca?ggat?cTt??c    140
A_U455        G--TC-------A------...---GCA-----A-G-------GATTCGC--AGT-AGC----T---AG-    183
B_HXB2R       G--T--------A------...---G-G--A----G--------ATTCGA--AGT-A-C----C---GG-    183
D_ELI         G--C--------A------...---GCG-------G--------GTGAGA--GCT-A-C----T--CGG-    183
F_BZ163A      A--TC-------G------...---GCA-------G--------GTGAGA--AGTAG-C----T---AG-    183
O_ANT70C      G--C--G--G--G------...---AGA------GCC---G-GGATACCC-CGCC---A--G-T---GC-    177
O_MVP5180     G--C--G-----A------...---AGG--A---GCC--AG-GGACAGCC--GCCA-CA----T---GCA    180
CPZGAB        A--TC--C----G-------...C--GCA--A----AG------AGTGAGA--AGT-G-A--G-G---GC-    186
CPZANT        A--T-AG-----A------...-AGA-A--A---GAT---G-GGAGGGCC--GCA---C--G-T---CG-    186
A_ROD         A--C--------AC------...-AGC-AC--TG--G------A-TG...GCCCTG--CGAT-GCA-ATAT    174
B_EHOA        A--C--------G------...-A-C-AC-AGG-CT-------TTG...GCCCTG---GATCGAA-ACAT    186
SD_MM251      GC-A-----G--AC------...---AGGC--TG-CA----C---TG...GCCTTG---GAT-GAA-ATAT    174
STM_STM       G--C--------G------...----C-G---TG--AT--AC---TG...GCCTTG---AAT-GAA-ATA-    171
VER_AGM3      A-GGGCC--G--A------...----C-AGC---AGA----C---GA...GCCTTG---GAA-GAA-CTGG    165
GRI_AGM677    G-GGGC----CAG------...----A-GCA----CAG--A--TGATAAGA--GCA-G-A-AG-C--CAA-    168
SAB_SAB1C     A-GAGC---GCAGC-----GCT-A-C-GC-A---CAGGTC--A-AC...T-GGCA-AGA-A--T---...    168
SYK_SYK       ACGA--------GCA----...---C-G--AAG--CT--ACGT-TC...TACCTTCAGCA--GAA-C...    162
```

```
REV_CONSENSUS ac?ttt???????ga????cgga?cct???cc?????t??a??t?ccacc?cttga???a?tt????t??    173
A_U455        GAT-GCCTGGGAC--CCTG----A---GTG--T...C-TC-GC-A-----GA----GAG-C--CGCC-TG    250
B_HXB2R       --T-A-CTGGGAC--TCTG----G---GTG--T...C-TC-GC-A-----G-----GAG-C--ACTC-TG    250
D_ELI         --T-A-CTGGGAC--CCTG----G---GTG--T...C-TC-GC-A-----G-----GAG-C--AATC-TA    250
F_BZ163A      T-T-G-CTGGGAC--CCTGA---A---GTG--T...C-TC-GC-A-----A-C---GAG-C--ACTA-TA    250
O_ANT70C      --TG--GTACACG--CCTCA-A-CAA-AATATT...G-GG-CT-A-----T-----GCA-C--AGCA-CA    244
O_MVP5180     --AG--GTACACG--TCTCA---CAA-AAT-TT...G-GG-CT-A-----T-----GCA-C--AATA-CA    247
CPZGAB        --T-A-CTGGGAC--CCTC--A-A---GGGGAT...T-GG-GC-A---GAG-----CAAGC--AGCC-GC    253
CPZANT        --TC--                                                                    192
A_ROD         --A---............C-T--T--G...--AGCT......GATT-G--T-----CCAGAC-...A-AC    220
B_EHOA        C-AC--............C-C--TT-G...--AACT......GAGGG---T-----TTTGGC-...A-AC    232
SD_MM251      T-A---............C-T--T--G...--AACT......GATA-G--T-----CTTGGC-...A-TC    220
STM_STM       T-A---............C-T--T--G...--AGCT......AGTA-G--T-----CTTGGC-...G-AC    217
VER_AGM3      CACAGC......A--GTGGAA--G-AA...-TG...G-GC-AGCG......A----CCA-T-GGTGC-CG    217
GRI_AGM677    --C--C...........GA---T-AACAA-TGGTGGCTC-AC-T-A-GAG--GC-GCTTGAGAATAAGG    226
SAB_SAB1C     ..................GA-ACA---GTGT-A...CAGATTGAT...-ACT-G-CTCAGGAATTTGACC    214
SYK_SYK       ..................TTT--GG-A...ATC...T-TGGAAGC......AGGACAGC-GC-...T-GG    199

REV_CONSENSUS a??a??????g??????????a??atct??ct????ag??????c?g?????????????????????cc    189
A_U455        -TTGTAGCGA-........AGCTG-GGAA--TCTGG-ACG...-A-CAGCCTCAAGGGACTGAGACTGG    308
B_HXB2R       -TTGTAACGA-........G-TTG-GGAA--TCTGG-ACG...-A-.....................GG    287
D_ELI         -TTGCAGTGA-........G-TTG-AGAA--TCTGG-ACG...-A-.....................GG    287
F_BZ163A      -TA-CAGCGA-........A-TTG-GAACAGGGGGCTGAG...GA-.....................GG    287
O_ANT70C      GGG-TCCAGAA........GGTG---AGCTATCTG--GCT...TG-.....................A-    281
O_MVP5180     GGG-TCCGGA-........GCTG---GACTACCTGG-ACT...GG-.....................A-    284
CPZGAB        -ATGTGTGGA-ACAACTCAAG-CGT-GGGA-ATCTA-TACTTCA-A-.....................--    302
A_ROD         -GC-T...CT-........C-GGGA--TA--ATCC--GAG...-TT.....................--    254
B_EHOA        -GCGG...CT-........C-GA-C--TAT-ATCA-AGAC...-TT.....................--    266
SD_MM251      -GC-A...CT-........C-GA-C--TG--ATCG--AGC...ATA.....................--    254
STM_STM       -AC-A...CTT........C-GGGCT-GT--ATCC--GAT...-TA.....................--    251
VER_AGM3      -TC-G.............C-GC----GG--ATAC-ACAG...TT-.......................--    248
GRI_AGM677    -CTTGGTGTTA........C-AC----G.......................................--    248
SAB_SAB1C     -GTTGGTTCTC........G-CA----T......C--CAA...-CT.....................--    245
SYK_SYK       -GG-CAGTCT-........C-GC-G--TCAGATCTCTGAT                                231

REV_CONSENSUS ????ga?cctcc?????????????????????c??a?t?ct??aga?????????????????????    204
A_U455        GGTG-GAGGG--T.....................-AA-TAT--GTG--A.....................    336
B_HXB2R       GGTG-GAAGC--T.....................-AA-TATTGGTG--A.....................    315
D_ELI         GGTG-GA-A---T.....................-AA-TAT--GTG--A.....................    315
F_BZ163A      GGTG-GAAGC--T.....................-AA-C-T--GGG--A.....................    315
O_ANT70C      TGTG--T---AGGGCAGAAGATAATTAATGTTTG-AG-A-TTGTGCAGC.....................    330
O_MVP5180     TGTG--T---GGG.....................ACA-AAGACAAT                         309
CPZGAB        T.........-AG.....................ACTGC-A-GGG---GACTGTGCCTGCT.........    333
A_ROD         T...--C-----A.....................ACTCA-CTACC---A.....................    279
B_EHOA        A...A-C-----A.....................ACC-G-A--CC-ACCGCTCAGGCTTCCACCTGCATA    312
SD_MM251      A...--T-----A.....................ACC-A-A--CC---GGCTCTC...............    285
STM_STM       A...--C-----A.....................-CC-ACCT-CC-A-G.....................    276
VER_AGM3      T...--C-----T.....................AGTTCAT--                             267
GRI_AGM677    T...--C-----T.....................-AT-T-CA-                             267
SAB_SAB1C     ATCGCTG-----A.....................GGACA-C-AAC---G.....................    273

REV_CONSENSUS ??????????????????????????????????????????????????????????????????????    204
A_U455        ......................................................................    336
B_HXB2R       ......................................................................    315
D_ELI         ......................................................................    315
F_BZ163A      ......................................................................    315
O_ANT70C      ......................................................................    330
CPZGAB        ......................................................................    333
A_ROD         ......................................................................    279
B_EHOA        CCTCCGATATGGGATCAGCTGGTTCCAAGAAGCAATCCAAGCAGCAGCCAGGGCTGCGGGAGAGACTCTT    382
SD_MM251      ......................................................................    285
STM_STM       ......................................................................    276
SAB_SAB1C     ......................................................................    273

REV_CONSENSUS ?????????????????????????t?tcat?c??ta?tgg?g?cagga?????????????????????    221
A_U455        ........................C-TC-G-TG--T---G-T-G---.......................    360
B_HXB2R       ........................C--C-A-AG--T---A-T-----.......................    339
D_ELI         ........................C--C-A-AG--T---A-T-----.......................    339
F_BZ163A      ........................C---CG-TG--T---G-T-----.......................    339
O_ANT70C      ........................G-A-CA-AA--C---CTA--A--.......................    354
CPZGAB        ....................GGGGGGAA-T--T-AA--T---G-AA----.....................    363
A_ROD         ........................C---G...AG-C--...............................    291
B_EHOA        GCGAGCGCGGCGAGGACCTCGTGGGGAG--C-......CA--A-AGC--CAGGAGAGATCATTGCAATAC    446
SD_MM251      .......................TGCGAC-C-A-GAAGGGTTC-AG-A-T....................    312
STM_STM       ........................G--......C-CCA-GATACT.......................    291
SAB_SAB1C     .......................AAC--GA-AGCTAACAGCT--TCT                        297
```

HIV1 REV

```
REV_CONSENSUS ????????????actaa???????ga?   228
A_U455        .............-----A......A-T   369
B_HXB2R       .............-----A......--A   348
D_ELI         .............---G-GGAACAGTGC   354
F_BZ163A      .............-----A......--A   348
O_ANT70C      .............-T-GCA......--A   363
CPZGAB        .............G----A......A-T   372
A_ROD         .............G--G-G......ACT   300
B_EHOA        CCAGGAGGATCAG--A-GG......-GC   468
SD_MM251      .............C--C-G......--C   321
STM_STM       .............G-AG-G......A-T   300
```

```
       tat ->       \/ 3′ sj
CONSENSUS-A   ATGGCAgGAAGAAGCGGAgaCagCGACGAG...GAaCTCCTC...AAaGCAaTAAGGATCATCAAAaTCc     64
CONSENSUS-B   ----------------------------a....-g------...--ga--G-c--ac-------gt-t-     64
CONSENSUS-D   ----------------------------a...-C------...--G--gg-C---C-------G----     64
CONSENSUS-F   ------------------A------?-ACC...----------...-G----G?C--ATA------G---T     62
CONSENSUS-O   ----------------A?--?...---......C-G-?G--G???C----?--CCA---?--A--G----     53
CONSENSUS-U   ---------------------GC-------...----------...-G--TCG-C----------------     64
CONSENSUS-CPZ ------------???-A?C???AA-?????G?????-??--G????-?---G-?-AA--A-----G--?-     44

                    exon \/ exon
CONSENSUS-A   TATACCAAAGCAAcCCaTaCCCCAAaCCCAagGGg...TCtCGaCAGGCCcGaAAGAATCGaAGAAGAAG    131
CONSENSUS-B   -c--t-----------?cct----gc---g-----...A-C-------------G---------------    130
CONSENSUS-D   ----------------TCCT----gc---G-----???A-C-------------G----A----------    131
CONSENSUS-F   -G-----------------T-----G---G----A...A-C-------------G?-----------G--    128
CONSENSUS-O   -G-----?-----T--C???--??C---???A--A...AGCA-?A-C---A-G-----CA-G--?--?--    108
CONSENSUS-U   -G--------------T-----ACC---AG-----...A-C------A-------------------G--    131
CONSENSUS-CPZ -?--?---???--?--?--?--?-??--AG??--A...?-?A-????--??-?-GA---????----?--     84

CONSENSUS-A   GTGGAGAGCAAGgCAgAGACAGATCGATTCGcTTAGTgAGCGGATTCTTAGCacTTGCCTGGGACGACCT    201
CONSENSUS-B   --------ag--a------------Cgg---a-------at---------------aT---------t--    200
CONSENSUS-D   ---------g---------------c-----A--G----A--------c-------aT------------    201
CONSENSUS-F   ----------------??--?-----CG-G?-?-----??A------------T----T------------    191
CONSENSUS-O   -------AG?--?---GCC--?G-G--?A?C--?GCC?C?A--?-----GC?--?GTTG-ACACG--?--    166
CONSENSUS-U   ----C-------A------------AGGG--A-----------T-----------------------T--    201
CONSENSUS-CPZ ----?????---A---?A?---?-????G???--???-?A?--G-?---??----??T?-----------    128

CONSENSUS-A   GCGGAgCCTGTGCCTCTTCAGCTACCACCGcTTGAGAGACTTCatCTTGATTGtaGCGAGgaCtGTGgAA    271
CONSENSUS-B   ------------------------------------------Ac-----------a------t-------    270
CONSENSUS-D   -a---a-----------------------------------A-----A----C--------T-------    271
CONSENSUS-F   -A---A--------------------?--A-?----------A?-A--A--A?C------?-T----A-.    255
CONSENSUS-O   CA-?-CAA-AAT?T-G-GG-CT-------T------CA----A??A-CAGGGA-CC?--?-?TGA-C??C    227
CONSENSUS-U   ------------------------------A-----------TG---------C--T----G--------    271
CONSENSUS-CPZ C--A-A----G-GA-T-GG--------GA------C-AG---AGC--GC-A---GTG---ACAAC-CA-G    198

CONSENSUS-A   CTTCTGggACgCAGcag?ctca?ggg?ctgagactGGGGTGGGAaGGccTCAAgTAtCTGTGGAAtCTtC    338
CONSENSUS-B   --------------.................???-----------c------a---tg---------C-    319
CONSENSUS-D   --------------....................-----------c------A--------------C-    320
CONSENSUS-F   ..CAG---G-T?--G--.................-----------C------ACT----G-------CA    304
CONSENSUS-O   TA?---?-?-T?G-ACTGTGG-TCCTAGG-CAGAA-ATAATTA-T-TTTG--GAAT-TGTGCAGC-G-AA    293
CONSENSUS-U   ----------A------C----A---A-------------------C----C-A------G---------    341
CONSENSUS-CPZ ACGT------ATCTA-TA--TCACA-C--C------CTAC----GA-A--GTGCCTG---G--GGAA--A    268

CONSENSUS-A   TGcTgTaTTGGGGTCGGGAACTaAAAAt?G                                            367
CONSENSUS-B   ---a-------a---a---------G-a-AG                                           349
CONSENSUS-D   -a-A-------a---a------G-gG-A-AGTGC                                        353
CONSENSUS-F   C--------------A---------G-ATAG                                           335
CONSENSUS-O   CA-??-????-?T?-??-?-????-G-?-ATTGA                                        310
CONSENSUS-U   -A--A--C-------A---------G-ATAG                                           372
CONSENSUS-CPZ -T-AA--------AAA---G--------                                              296
```