HMMER Sequences in the Vpr Alignment

A_U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
B_HXB2R	HIVHXB2R	K03455	Starcich,B.	Science 227, 538 (1985)
D_ELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
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)
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)

HIV1 VPR

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" based on an HMM model (Part II). Annotation is not attempted for this alignment of paralogous sequences.

```
VPR_VPX_CONSENSUS ATGGaacaAGcc??????????????????????????????????????????????????Cca???g     16
A_U455            ------------.............................................------...-     16
B_HXB2R           ------------.............................................------...-     16
D_ELI             ------------.............................................------...-     16
O_ANT70C          -----------A.............................................--T...-     16
O_MVP5180         ------------.............................................-TG...-     16
CPZGAB            -----------A.............................................------...-     16
CPZANT            ------------.............................................--G...C     16
VER_AGM3          ----CCTC--G-AGA.........................................GAT--CAGA-     25
GRI_AGM677        ----C-TC--GAAGA.........................................GAT---AGA-     25
SAB_SAB1C         ----CCTC--GAGGGTGGCTCCCACCAGTAGGAGGGGATCCTCCCAAGGATCCCCCCAAGAAT--CAGA-     70

VPR_VPX_CONSENSUS aa???...ga?cagGGg?????????????????ccaca?AggGAgCCat?c...aatGAaTgG?ca..     53
A_U455            --.......--C------.....................-----G--------G-A-...GC-----A-G--..     57
B_HXB2R           --.......--C--A---.....................-----G---------CA-...---------A--..     57
D_ELI             C-.......--C------.....................-----G----------A-...---------G--..     57
O_ANT70C          -G......A-T-----A...................---GCT-AA-----C-T-...---------G--..     57
O_MVP5180         --......A-T-----A...................---GCT----------T-...---------A--..     57
CPZGAB            -G......--C--A--T.....................----CA--A--A--T-A-...C-G------G--..     57
CPZANT            -G......--TG-----.....................-----G--A--A---ATG...-----G---TTG..     57
VER_AGM3          --GCA...AGA-C----GAACTAGAAATATGGGATTTGAGC--------G-GG...G-C------CT-..     87
GRI_AGM677        --CCT...TTA-CA--ATGGCTGGAAATCTGGGAT-T-G-C----------GG...G-C------CT-..     87
SAB_SAB1C         --GAG...ATA-CA--ATGGCTAGAAACATGGGATTTG-CC--A-----G-T-...G-C------CTC..    132

VPR_VPX_CONSENSUS ....ctA...GAgat?tTagaaGAgcT?AA??atGAaGc?gtaagaCAtTT?...cctaggc?ttggcT?    105
A_U455            ....T--...---T-G-----------T--GC------AT--T--------T...-------AG-----C    117
B_HXB2R           ....--...---C-T-----G-----T--GA-------T--T--------T...------AT------C    117
D_ELI             ....T--...---C-T-----G-----T--GAG------T--T--------T...------ATA-----C    117
O_ANT70C          ....--...---C-C-----------A--AGCA-----A-----------C...-------C------A    117
O_MVP5180         ....--...---C-CC-G--------G--AG-A-----A-----------C...-------C----T-A    117
CPZGAB            ....T--...----CC--G-----AT-A--AA-------A-----------C...-----A-CA---T-G    117
CPZANT            ....T--...--A-CT--------AA-A--AA----G--A-----------C...--ACA--C-ACAT-A    117
VER_AGM3          ....AG-...--C--GC-------TA-C--TC-G-----CAAG-TG-----T...GGGC-CGAGCTC--G    147
GRI_AGM677        ....-A-...--C--GC-CAGG--T--A--CG-A-----CAG---G--C--T...GGA-T-AACAT---A    147
SAB_SAB1C         ....-G-...--C--G---C-G--C--C--TTCA--G--CCAGT-C--C--C...--A---AA-CTC--T    192

VPR_VPX_CONSENSUS ca?cgcttagg?????????caatatatttatgAg???????????acttatggggatac???????t    145
A_U455            --TG-A-----A.........---C----C---A-C.............--C-----------T......-    160
B_HXB2R           --TG-------G.........---C----C-----A...........--------------T......-    160
D_ELI             --TA-------A.........---C----------A...........--------------C......-    160
O_ANT70C          --CGC---G--A.........-----C-----------------------------C--T......-    160
O_MVP5180         --AGC--GT--G.........--G--C-----------------------------A--C--T......-    160
CPZGAB            --T-AA-----A.........----TC--------C......................A-----A......-    160
CPZANT            --A-A---G-AAATTGGGTATATGC-A---ACG---.............-T-C----A--GGAG......-    169
VER_AGM3          TTC-AAG--T-G.........A-C---TG-C-G---GAGGGA...GAA-GGA--C-CAC-C-CATG...C    202
GRI_AGM677        ATC--AG--T-G.........A-T--CTG-GTA---GAGGGA......-GGAGACATA----CCCATGGA    202
SAB_SAB1C         TTC--TC-TT-G.........TGGA-----GTG--AGAACCAGCCATTGA-C-----C-A--CAGG...C    250

VPR_VPX_CONSENSUS gggaaggagtt............?a?gc?at?ataAgAaT??Tg...CAAcAatt?cTgTTTat?CAT??    190
A_U455            ----G------............G-A--T--A--------TT--...--------A--------T---..    213
B_HXB2R           ---C------G............G-A--C--A--------TC--...------C-G--------C---%%    213
D_ELI             ---T-------............G-A--T--A--------AC--...--------A--------T---..    213
O_ANT70C          ---T-------............ATG--A--T--------CT-A...--------G-------CC---..    213
O_MVP5180         ----------............ATG--A--T--------CT-A...-------C-A-------CC---..    213
CPZGAB            ---T------A............G-G--C--C--------CT-A...-----C--G--------C---..    213
CPZANT            -CAGGTTT                                                                 177
VER_AGM3          TA---A-G-C-............T-TAAATATTAT-A-T-GG--...---A--GCT--C---G-G---..    255
GRI_AGM677        AT--GAT--GC............T-CAAGTACTAT-----TG-T...---A-G-CTA-----G-A---..    255
SAB_SAB1C         T------CTGG............T-TAAATATTGT-----TC--...---A-GGCTT-A---G-G---..    303
Z
VPR_VPX_CONSENSUS ?ttcagaatTgG?TGccaacatAg?????????????????????????????????????aGaatagg?    220
A_U455            .-----------G-----------C.........................................--------C    246
B_HXB2R           T-----------G--T-G---                                                    234
D_ELI             .-----------G--T--------C.........................................--------C    246
O_ANT70C          .-AT--------A-----------T.........................................--------A    246
O_MVP5180         .-AT--------A-----------T.........................................--------A    246
CPZGAB            .--T---C----G-----G-----C.........................................--------C    246
VER_AGM3          .---C-GTG---A----GCAGA--ACAACCCTTTGAACCA.........TACGAAGAAAGG--GGAT--A    315
GRI_AGM677        .------TG---T--TAG-AGG--AGGACCTTTTTCCCCT.........TACGAAGAGAGG----AT--A    315
SAB_SAB1C         .A-G-A-GG-A-A--T...TG--AGCCAAAGACTCACCCTGCGTATGGACCAGGAGCAGGAG-CCCGCCC    369
```

```
VPR_VPX_CONSENSUS att????????????a??????gAgga...aga?gA?????????ggaag?aatggatCcagtagatcc     257
A_U455            ---............-TT...CC---G...---A--........--C--G------G-----------     288
D_ELI             ---............-TT...C--CAG...---A--........-C---A------------------     288
O_ANT70C          ---............-ACCCAA-----...---G--........A----A------------------     291
O_MVP5180         ---CTCCCATCTAAC-CA...A-------...---G--........A----A------------------     300
CPZGAB            --C............CTC...CC-CA-...---A--........A--TCC----------A-------     288
VER_AGM3          CAA............GGG...G-------...C--GC-........--GC-CGTACC-C-AG-ACTTGAT     357
GRI_AGM677        CAA............GGA...G-----...GCCCC-............CCCCC-CCTC-AG-ACTT       351
SAB_SAB1C         CCA............GGTCTGG------...GCTTC-GGAGGAGCT-C-TCGGCC-CTC-AG-CCTG       420
```

HIV1 VPR CONSENSUS

```
                                                                                                <- vif cds
CONSENSUS-A   ATGGAA??A......GCCCCA...GAAGACCAGGGGCCACAGAGGGAGCCGT?C??TGAAT?G?CA?TAG     53
CONSENSUS-B   ------ca-??????------???--------A-----------------a-A-AA-----G-a--C---     61
CONSENSUS-D   ------CA-.....------...--------a--------?--------A-A-AA-----G-a--T---     60
CONSENSUS-O   ------CA-.......--?-??...--?A-T-----A---GCT-??-----?-T-AA-----G----C---     53
CONSENSUS-U   ------CA-.....------...-C------A-----------------ACA-AA-----G-A--T---     61
CONSENSUS-CPZ ------CA-.......--?--?...?-G--??-?--?----??--A--A--??-?-A?--?-G--??T---     42

CONSENSUS-A   A??TGTTAGAAGAGCTTAA?CATGAAG?TGTTAGACA?TTTCCTAGGC?GTGGCTCCATGGATTAGGACA    117
CONSENSUS-B   -GC-T-----g--------GAg-----C---------T---------ata---------a-c--------    131
CONSENSUS-D   -GC-T-----G--------GAG-----C---------T---------ATA---------A-C--------    130
CONSENSUS-O   -GC-C?-?--------?--AG?A----CA--A-----T--C-------CT---?-A--?-CC-??--?--    114
CONSENSUS-U   -AC-T-----G--------G--A----C---C-----C---------ATA---------A-T--------    131
CONSENSUS-CPZ --AC?--?-----A?-A--AA----?-CA--A-----T--C--????-C????T-?????????????--     86

CONSENSUS-A   ACATATCTAT?A?AC?TATGGGGATACTTGGGA?GG?GTT?AAGCTATAATAAGAATTTTGCAACAA?TA    180
CONSENSUS-B   ----------G-a--T----------------cA--A--gG----c------------c--------C-G    201
CONSENSUS-D   -?----t---G-A--T-----------C----?A--a---G-----?-----------?--------t--    196
CONSENSUS-O   ?T-C--T---G-G--T-----?--C-------?A--A---ATG--A--T--------C--A--------?    179
CONSENSUS-U   ----------G-A--T-----------C-----A--A---G---------------G-C--------C-G    201
CONSENSUS-CPZ -??C?-??-?-?-??----??????????????-??-???-?????????????????-?---???--?     98

                                 \/ 3' sj
CONSENSUS-A   ?TGTTT?TTCAT...TTCAGAATTGGGTG?CAACATAGCAGAATAGGCATTATT.........C?AGGGA    234
CONSENSUS-B   C-----A-----..?--------------t-g--------------------c-...........---Ca--    258
CONSENSUS-D   C-----A-----...--------------T----------------t----c-..........-G-CA--    254
CONSENSUS-O   C-----ACC---...-AT--------A--C--------T--------A---??C??????????C-A-AG    234
CONSENSUS-U   C-----A-----...--------------T----------------------C-..........-G-CA--    259
CONSENSUS-CPZ -?-???A-?---...????--???-?---C??G??---------------CC-C.........-C-CAA-    138

                              -> tat cds
CONSENSUS-A   GAAGAGGC...AGGAATGGAGC??GTAGATCCTAG                                       264
CONSENSUS-B   -g----CA???--a--------CA-----------                                        290
CONSENSUS-D   ------CA...--A------T-CA--------                                           283
CONSENSUS-O   -------AAGA--A------T-CA--------                                           266
CONSENSUS-U   ------CA...--A------T-CA--------                                           288
CONSENSUS-CPZ -----A-A...TCC------T-CAA-------                                           167
```