

```

B_FR.HXB2R      |-> 5' LTR U3 start
.TGGAAGGGCTAATTCACCTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTG      129

A_UG.U455      .....0
C_ET.ETH2220   .....0
D_ZR.84ZR085   .....0
FL_BE.VI850    .....0
G_SE.SE6165    .....0
H_CF.90CF056   .....0
J_SE.SE92809   .....0
CRF01_AE_TH.CM240 .....0
CRF02_AG_NG.IBNG .....0
CRF03_AB_RU.KAL153 .....0
CRF04_cpx_CY.94CY03 .....0
N_CM.YBF30     .....0
O_MVP5180     C---T---T---T---TAAG---GC---A---G---C---A---T---T---G---A---T---G---TGT---G---A---ACCT---TC---A---A---130
O_CM.ANT70     .....129
HIV2_A_ST      .....0
HIV2_A_BEN     ..-C---A-GT--T--AGTAGGA-G---TAGA---A--CA-A-ACC-AG-AA-AGAGG---GATAA-A--A---G---T---T---T-AT-----AG-A-G-C--A--TA---C---129
HIV2_A_ALI     ..-----A-GT--T--AGTG-GA-----TAGA---T-A--CA-A-ACT-AG-AA-GGA-G---GATAA-T--A---G---T---T---AT-----A--A-G-C-GA--TT---129
HIV2_B_D205    ..-A-GTA-T-TAGTG-GA-----CAGA--AT-A--CACA-ATT-TG-GA-TGA-G---ATTG-GT--GC-G--A---T---T-AT-----A-GC---CAAATA---129
HIV2_B_EHO     ..-----A-TT-CT-TAGTG-GA---G--TA-A--A--A--CACA-ACT-AG-AA-TGA-G---ATTG-GT--GA-G-----AT-----AG-A-----CAA-TT---129
STMM83293     .....0
SMM251        ..-----A-TTA-T--AGTGC-A-----TAGA---T-A--CA--ACT-AG-AA-GGA-G---AT-A-A--A---G---G-T---CT---A---A-T---C---AA---A---129
SMM9          .....0
SIVCPZANT     .....0
SIVCPZGAB     ..-T--G--T--AGGA-----G-----C--C--G---T-----T-----C-G-----A---A---A-CA---TC---CTGT---129
SIVCPZUS      ..-T--G--T-T--AG-A-----G---A---C--G-T---T-----AAT-----C-G-----T---C-----T--AG-A---C--A--A-A---129
SIVSTAN1      ..-G--CA--TA-TGG--A-C-AAG--G-GC-A---GA---C-ATGC-CTAA-TGA-TGG--AATA-AGA---G--AGCA--TT-C--T-----C-C--GA-G--C-GATG---C---129
SIVGRI677     ..-T--A--TA-T---TG--A---G--A-G---GA--T---ATGC--TGA--GAGTGG--AATAA-AGA---G--AGCT--T-----C-G-----A-G-C-GAGAGT---129
SIVVER9063    ..-T--A--TTA-T---G--A-G--GA-TAGA---CA-C---ATGCTCTTA-TGA-TGG--AAT-A-TGA---G-A-CGC--GGT--A---A---CG-A-----C-CCTG---129
SIVVER155     ..-T--A--TTA-T---G--TA-G--A-TA-G---A-----ATGC-CT-A-TGA-TGG--AAT-A-TGA---G-A-CGCA-GGT--AA--A--T---A---C-GAG-TG---129
SIVVERTYO     .....0
SIVVERAGM3    ..-----TA-TGG---GA-----G--C---T-AAT--C--ACTGTG-GA-TGAGTGG--TCT-A-TGGA--C-TCAT-----GAT--A---A-CA---C-T---A---129
SIVLHOEST     ..-----CTGGT---AG-ACC--G---A---T-G---T---ATGCAG-AA-TGAGTGG---T-A-AA--GC-G---AG-T---CAA-----T--AG-G-----AA-G---129
SIVSUN        .....0
SIVMNDGB1     ..-----A-TTA-T--AGTG--A-G--A-GA-G---T-G---C-ATGCTCTTA-TGAGTGG--ATAG-AGA--GA-G---T--T--GAT-----T-C--G--C--AA-TG---129
SIVSAB1C      .....0
SIVSYK173     .....0

```

B_FR.HXB2R	GATGGTGCTACAAGCTAGTACCAGTTGAGCCAGAGAAGTTAGAAGAAGCCAACAAAGGAGAGAACCACCAGCTTGTACACCCTGTGAGCCTGCATGGAATGGATGACCCGGAG.....AGAGAAGTGT	252
A_UG.U455	0
C_ET.ETH2220	0
D_ZR.84ZR085	0
FL_BE.VI850	0
G_SE.SE6165	0
H_CF.90CF056	0
J_SE.SE92809	0
CRF01_AE_TH.CM240	0
CRF02_AG_NG.IBNG	0
CRF03_AB_RU.KAL153	-G-----C-C-----CTG-AG-----G-----CAG-G-----A-----C-A-----A-AT--AA-C-----TGA--A.....-A-----	237
CRF04_cpx_CY.94CY03	0
N_CM.YBF30	0
O_MVP5180	----TG-TT--A--G-----GTCAG--AG--GC--GAG--TGGGT--T-C-A-TG-AG-TGC-AG-CTT-TACA-CCAG--TGTAATC----GCTGAG-AT-CACAC..G-G--GA-AC	256
O_CM.ANT70	----TG-TT--A-----GTCAGA--AG--GC--AG--T-GG-G..T-C-TGTG-GAG-GC-AA-CTC-TGCA-CCAG-ATGTGCC-----TTGAA-ATACACAT..-A---A-AC	255
HIV2_A_ST	0
HIV2_A_BEN	-G---CTG-GG-----T---A--A-TCTCAC-AGAG-C--G-AAG-TG.....-GC--A-T---AG-----A-CACAAACAAGCA--CAT----TGA-C-T.....G-G--GACA-	246
HIV2_A_ALI	-G---CTG-GG-----A--TGTCCAC-AGA--GG--G-A-CTG.....-CTCA-T-C-C-----A-ACAAACAAGCA-GCAT----A--C-T.....G-G--GACA-	246
HIV2_B_D205	-T---CTG-GG-----G-----A--GTGCCAGCAGCAGCCCG--AGG-GG-G-A...A-C-CATT--C-AA-G---G-CACAGA-CTCCTC-TG-----ATCC-T.....G-G--GACTC	249
HIV2_B_EHO	-C---CTA-GG-----G---A-AA-TATGATAGCAGA-CC--G-A-G-GG.....A-C-CATT-TC--G-G--T--A-CACAGACCTCC-C-TG-----CC-C.....GA---GACCC	246
STMM83293	0
SMM251	-C---CTA-GG--AT---C--T--AA-TGT-TCAG-TGAG-C-C-G-AGG-TG.....-GGCATTAT--AA-G--G--A-CTCAAACCTCCAAGTG-----TTG.....G---G--TC	246
SMM9	0
SIVCPZANT	0
SIVCPZGAB	-----T--A---G--CC-GACAGAG--C--G---C-----TG-----T---A-T-C---G--T--CA-AT-T-A-----G---AA-AT-AGGAC..-A---G--C-	252
SIVCPZUS	-GC-----TT--T---C--TC-CACAGA---G-AG---GCGG---T-----C-A--TAC-C-G---CA--T---A-----AA-AT-AACAT..G-C-----C-	252
SIVSTAN1	-C-TT---TTG--T---G--G--A--TGTCTCCC--GA--CCC-G-ATG.....-GACA-T-T--C-C-----CCCAGA-TG-AT-GGAAAGC---TTG.....A---GACA-	243
SIVGRI677	-C-TC---TT-----C---G--C-TGC-TG--GAG-C-CGCAA-TGTG.....-GACA-T-TC--A-G--T--A-CACAGA--GGG-A-GA.TCCTGATG-A-TAGATCATG-----C-	252
SIVVER9063	-C-TT---T-----G--A---ATGC-TG-AGAG-C--GA--TGTG.....-GACATTA-----GCG-T---CACAAA-AA-A-AGGA.CCCCGATG-C-TCAGTCATG---GACC-	252
SIVVER155	-C-TC---T-----G---CC-TGC-TG--GA--C---A-ATGTG.....A-GGCATT---G-----A-CACAA-----A-GA.CCCTGATG-T-TAAATCATG---A-A-	252
SIVVERTY0	0
SIVVERAGM3	0
SIVLHOEST	----TG-GGC--T--A-----G-CTTGT--TG--ATA---TC--TCAG-T-A...-CACAAAT--C-T--G--TT-AAGCCAG--GGG-TCC---A---CTG.....G---GAG-C	249
SIVSUN	-C---CTA-GG-----C-----CACCATT--TG-AGACAGG...ATCCT--TCATCCTTGTCAAGC-C-AC-G--TT-CAGCCAG-AAGGG-TT-AT--A-----TTG.....G---GAGAC	249
SIVMNDGB1	0
SIVSAB1C	-----TT--A--G---T--G--C-TGTCAG--GAG-CTA-GAAT-GTG.....-CA-T-C-C-G--T---CTCAAG--GCCTATGA-----G-CTG.....A---GACCC	246
SIVSYK173	37

	TCF-1 alpha	
B_FR.HXB2R	TAGAGTGGAGGTTTGACAGCCGCTAGCATTTCATCACATGGCCCAGAGCTGCATCCGGAGTACTTCAAGAAGCTGCTGACA.....	334
A_UG.U455	0
C_ET.ETH2220	0
D_ZR.84ZR085	0
FI_BE.VI850	0
G_SE.SE6165	0
H_CF.90CF056	0
J_SE.SE92809	0
CRF01_AE_TH.CM240	0
CRF02_AG_NG.IBNG	0
CRF03_AB_RU.KAL153	--AT----A-----T-----G--TC-AACA---GA-----T-A---AG-----CAGACT.....	325
CRF04_cpx_CY.94CY03	0
N_CM.YBF30	0
O_MVP5180	--A-A---CA-----T--ATCAT--GC--AACA--T-A---TGC-AAA---C--A---CT---CCC-G-AACTGACACTGCGG.....	345
O_CM.ANT70	-GAT----A-----T--ATCT---GCAACACC--TG-T--TATGATAACT--C--A---CT---C---GGA--A-A-ACTGCTG.....	344
HIV2_A_ST	0
HIV2_A_BEN	--T---CA-----TC-ATG--G--C-A-A-CT---A---TTCACT---T-C--A--AG-G--TGG-C--AAG-C-GGATTGCCAGAGAAAGAATGGAAGGCAAAACCTGAAAGCAAGAGGGGATACC.....	376
HIV2_A_ALI	--TT---A-----CCTAAG--G--TCA-G--T--AA---TTTATT--A--C--A--G-A--TGG-T--AAG-C-GGCCTGCCAGAAAGATGAGTGAAGGCAGAGACTGAAAGCAAGAGGGGATACC.....	376
HIV2_B_D205	-TATC--CA-----TTC--T--G---A-G--T-TG---TTTCA-TAG-TT--A--AG-G--GG-T--CAG-C-GGGCTACCAGAGGAGGAGTGAAGGCTAGACTAAAAGCAAGAGGGGATACC.....	379
HIV2_B_EHO	-T-TC--CA-----TC--T-----A-G-CT-TG---TTCAGCAG-TTC--A--G-G--TGG-T-TCAG-C-GGAATGCCAGAGAAAGAGTGAAGGCTAAACTGAGAGCAAGAGGAATACC.....	376
STMM83293TGGAAGGGATTTATTACAGTGAGAGAAGGCTAACCGC.....	37
SMM251	---C---A-----TCCAACT---C-ACAC-T-TGA--ATAT-CTAGAT-C--A--AG-G--GGA-G-AAG-C-GGCCTGCAGAGGAAGAGGTTAGAAGAAGGCTAACCGC.....	363
SMM9	0
SIVCPZANT	0
SIVCPZGAB	-G-TC--C-C-----A-G-G--C--AAGA--T--T--A---A-AA-----A--AG---ACTTT-GACTGGCGCATGCGC.....	349
SIVCPZUS	-ATC--CA-----CTGAA--G--TCGGAGA--GA-TAA-----A-----CG---A--G-TGTAACCGCGCAGGCG.....	349
SIVSTAN1	-G-T---AA---A-TCCA-T-----G---GT--ACC-GGACAGCT-TA-GGACAT-C-TGGTCTTGTTAAGA-GA-GTAGCCTAACCGCAGGCTTGCGGT.....	349
SIVGRI677	-G-TC--A-----CCGAAGT-G--GG-GG-GT--CGCC-GGACAT-T-TA-GGACAT-C--GAAC-TGCAAAAGC-CT-GTTTCGCTGGTTAGCCTAACCCGGTTG.....	361
SIVVER9063	-G-TC--A-----CCTATGT---G-G--GT-TGACC-AAACAGA-A-T--TTT--AG--A-GC-TGCAATAGTGA-GAGGAAGTAGCTAACCCGAGGCTTGTG.....	361
SIVVER155	-G-CA---A-----TCCAATGT-G--TG---GT--GACC--TC-AG-GA-T-CTTTACAG---AT-TTCAACAGTTGGTACAGGAAACTAGCCGACCACAGGCTTG.....	362
SIVVERTYO	0
SIVVERAGM3	0
SIVLHOEST	-TATC--CAC--CA--TCCAATGT-G---G-GG-CTTT--A---TCA-AAA---G--T-C-A-GA-AC-A--TACTGCTTTCGCTTTTGACTGCAAGAGGAAGTAGCCAAACC.....GC	365
SIVSUN	-TATC--CA-----TCC-ACT-G--C-A-G--TTT-G---AT-C--AA---C--T--AG-G-----C-TGTTACTAGTCTGCAGTGGGAGGTGTAGCTTAACCCGAAACCGC.....	366
SIVMNDGB1	0
SIVSAB1C	-G-T---A-----CCA-TA-----G-GG-CT-TG-T--TT-GAG-----C--A---C-GG-GCC-TCAGCACA-GGCTAACCCACAAAACCATCCTATG.....	352
SIVSYK173	G-ACTA-TTCCAATGAGATA-AGGGATGC--GGAGTATGA-GAGA-T-CACT-C---TCCA-C-AG-CTCTGGGA-A-GGGAGTTCGAACATGGGGAGAACCTCATGTGGAACCTCAACCCCACTTG.....	167

see Pohlmann, J Virol 72(7):5589-5598 (1998) and similar publications for information on this enhancer region
NF-kappa-B-II NF-kappa-B-I Sp1-III

B_FR.HXB2RTCGAGCTTGCTACAAGGGACTTTCCG.....CTGGGGACTTTCCAGGG.AGGCG	382
A_UG.U455	0
C_ET.ETH2220	0
D_ZR.84ZR085	0
FI_BE.VI850	0
G_SE.SE6165	0
H_CF.90CF056	0
J_SE.SE92809	0
CRF01_AE_TH.CM240	0
CRF02_AG_NG.IBNG	0
CRF03_AB_RU.KAL153GCTGACAT--AG----GAC-----GC-----G--T-	379
CRF04_cpx_CY.94CY03	0
N_CM.YBF30	0
O_MVP5180A.....GACTGCTGACACTGC-----C-TG--A-	390
O_CM.ANT70ACCTGAAGATT--GA-CTGT--A-----AGCAAAGACTGCTGACACTGC-----T-GGA-G-	418
HIV2_A_ST	0
HIV2_A_BEN	ATATAGTGA...ATAACAGGAACAACCATACTTGGTCAAGGCAGGAAGTAGCTA.CTAAGAAAC.AGCTGA-GCTGC-----A.....GAA--G--G-AACCAAGG-AG-	483
HIV2_A_ALI	ATTTAGT.AAGAACAGGAACAGCT.GATTTGGTCAGGGCAGGAAGTAACACTG...AAAAC.AGCTGA-CTGC-----A.....GAA--G--G-AACCA-GG-AA-	480
HIV2_B_D205	TACAGATTAGGCAAGAGACAGCAGCATAAACAGGAACTAGCTGACACTGCACAA.GAAGGAAACTAGCAGACTGC-----A.....AAA-A-G--G-AACA.TGG-AG-	489
HIV2_B_EHO	TACAGAGTAGACAGGAAATAACAGCACAAAGAGGAACTAGCTAACACTGCATAGAGAAGGAAACTAGCTGATACTGC---A.....GAA--G--G-AACA-TGG-AG-	487
STMM83293AAGAGGCCTTCTAAAGATGGCTGACAAGAAGGAAACAAGCTGA-AC-GC-----A.....-AA--G--G-AA.CA-GG-AG-	120
SMM251AAGAGGCCTTCTAACATGGCTGACAAGAGGAAACTCGCTGA-A-GC-----A.....-AA--GA-G-TAT--G--A-	447
SMM9	0
SIVCPZANT	0
SIVCPZGABACAAGAACT--GA--CTGC-----A.....A-----G--AGA-GT	401
SIVCPZUSCAA-AA-A--GCTG--TGA-----TA.....A-----A--G-C-TT	399
SIVSTAN1TAAGCTAACTAGGCCGTTGCCCTAGGAGATG.ACATAAAGAA-T--GATC-GC-----A.....--TG-TGC	422
SIVGRI677	364
SIVVER9063GTTAAGCCGTTGCCAGGGAGATGACATT-GA-A--GCTG--G-----A.....A-----CG--C	436
SIVVER155CGGTTTCTGTTGCTTAGGAGATGACATTAAGAACT--GA-.....A.....G-AC-----A--CG--AC	440
SIVVERTYO	0
SIVVERAGM3	0
SIVLHOEST	AAACCGCAGCACATCTCTTGCAGCCCGGTTGCTAAGGCAACCGGGCTAGCGCATGCGCAATGG-TT--GA-G-GC-----A.....GGACGGGCGG--AGG--	473
SIVSUNATCCTCTTGCATCGCCTAGGCAACGGGGCTAGC.GCAT.....-C--GCTAG-GT-----	427
SIVMNDGB1GGAGTCTCTACTACAGAGGCTAAGGGTTGTA--TCTGAGCAG-TCCCCTTAGAG-AAGGACCAGAGTCTGAGTGACT---T--GAG--CCTC-CT--	98
SIVSAB1CGAGTTGTCATGGTGTGACATTAAGAACT--GA--GAG-----A.....-GGT-GA	410
SIVSYK173	CCTATACACCAGGCTGGGAGATGCCCGCTTGC...AGTTAGAGAGACAACAGGAAAACCA-A--AGCTGC-GTCA-C---AGTAAGAACATCTCCTAGGAGA--CCATGG-GA-A--CTCG-C-C	292

	Sp1-II	Sp1-I	TATA Box	5' LTR U3 end	-> +1 mRNA start site	
B_FR.HXB2R	TGG.CCTGGGCGGACTGGGGAGTGGCGA.....GCCCTCAGATCCTGCATATAAGCAGCTGCT.TTTTGCCTGTACTGGG.....TCTC					460
	HIV-2 and SIVs have extra Tar stem-loop(s)					
	see berkhou, Nucleic Acids Res 20(1):27-31 (1992)					
A_UG.U455					0
C_ET.ETH2220					0
D_ZR.84ZR085					0
FL_BE.VI850					0
G_SE.SE6165					0
H_CF.90CF056					0
J_SE.SE92809					0
CRF01_AE_TH.CM240					7
CRF02_AG_NG.IBNG					0
CRF03_AB_RU.KAL153	..TT-----AGT-----T.....A-----G-----C-----					457
CRF04_cpx_CY.94CY03					0
N_CM.YBF30TTCTCGTTGTACTGGG----					21
O_MVP5180	G-ATAAG----TT-G---AGTG-C.....TAA-----G-----CC--T---C-----T					468
O_CM.ANT70	AC..AGG----TT-G---AGTG-C.....TAA-----AG-----C---C---T---C-----					494
HIV2_A_STA-TCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGT----					46
HIV2_A_BEN	GAC.ATG--AG-A-CTG-T--G-AAC.....GCC-T-AT-CT-A---T---A-T-T-C-C---TC--GCAT---T-CA-TCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGT----					601
HIV2_A_ALI	GAC.ATG--AG-A-CTG-T--G-AAC.....GCC-T-AT-CTC---T---A-T-T-C-C---GC--GCAT---T-CA-TCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGT----					597
HIV2_B_D205	A-C.AAG--AG---TTG-TT-G-AAC.....GCC-AT-A-ATCT---T---A-T-T-C-C---TC-CGCAT---T-CA-TCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGT----					607
HIV2_B_EHO	GAC.ATG--AG-A-CAG--A-G.AAC.....GCC-ATCA--CT--T---A-T-T-C-C---T-G-GCAT---T-CA-TCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGT----					604
STMM83293	-AC.TGG-A-GA-CTGGT--G.AAC.....GCC-TT-T-CT-T---T---A-TA-A-C---TA-AGCT---T-CA-TCACTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGT----					237
SMM251	CC-.GTC--AACA.....CC-A--TTCT-GA--T--A-TATCA---AT--CGCT---T-CA-TCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGT----					552
SMM9CA-TCACTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGT----					47
SIVCPZANT					0
SIVCPZGAB	G-T..GG-A-GC--GG-T--G.....TGTGGCTTT-----G.....C-----					480
SIVCPZUS	CCA.AGG--T--T-A--CG-AACA-GGCGTGGTTTTA---A-GA-----A-----C---C-C--T-----					487
SIVSTAN1	GA-.GGC--TACTTGGGA-.T-GCTT.....CTCAGAG-----A-----A---CGC-G--T---ACTCAG.....					497
SIVGRI677	ACC.GGG--T-AA-GACTCCTT-GCTT.....CATATAG--CA---A-----					405
SIVVER9063	AT-.GGC--T.AC-GGGA-.T-GTTT.....TA-CTCAGAG-----A-----A---CGC-G--T---ACTCAG.....					511
SIVVER155	AT-.GGC--T.AC-GGGA-.T-GCTT.....TA-CTCAGAG-----A-----A---CGC-G--T---ACTCAG.....					515
SIVVERTYOCAG.....					7
SIVVERAGM3CAG.....					7
SIVLHOESTG-C--CT-G-----C-CC--T---T-G--AG-----CT					522
SIVSUNGACA-G-A-GGG-A.....GCTTA-C-CT-A-----C--CGC-TGT-.G--AG.....					495
SIVMNDGB1	G-CTGATCA-CTCGAG-TA-T-GAA-T.CCTTGCTTGCTTG--ATTG--T-CA--A-GTA-CT-AGAA--AGAG-AAGTGA-TAAGTGTATCCATTGTGCGCC.....					207
SIVSAB1C	GACTGGGC--TACTGGGA-T-GC-T.....---CGT-C-G--CA-----T-GCGC-----AAACAG.....					486
SIVSYK173	A--GGAG--A-CT-GGC---GAA.....G-TCACCC-G-----A---C-A---C-G-A--TATCCGC--AGTTGCAACTAGCTCTCTCCAGCGGCCCTGGGAGCCCTGAGAGACT					407

B_FR.HXB2R	TCTGGTTAGA...CCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTA.....AGCCT.....	525
	TAR element stem bulge loop stem TAR element stem	
A_UG.U455	0
C_ET.ETH2220	0
D_ZR.84ZR085-CGG-----	39
FI_BE.VI850	0
G_SE.SE6165	0
H_CF.90CF056	0
J_SE.SE92809	0
CRF01_AE_TH.CM240	--T-----G-----C-----G-A-----A-----	72
CRF02_AG_NG.IBNG-AGATCT.....G-GGA-----	54
CRF03_AB_RU.KAL153-G-----G-A-----	521
CRF04_cpx_CY.94CY03	0
N_CM.YBF30	--T-C-G-----TA-----G-AG--A-C--ACTG-----	85
O_MVP5180	AG-TAGAG--...-G-----C-----C-----CTCTAGCT-----G-----A-----	529
O_CM.ANT70	GG-TAGAG--...-G-----C-----C-----CTCTAGCT-----G-----CG--	559
HIV2_A_ST	--CA-CACT...G--G-A-----T-T-C--CTAG--CTCACCAGT--TGGCCGGCACTGGGCAGA-.GGCTCCACGCTTG.	132
HIV2_A_BEN	--CA-CACT...G--G-A-----T-T-C--CTGG--CTCACCAGT--TGGCCGGTACTGGGCAGA-.GGCTCCACGCTTG.	687
HIV2_A_ALI	--CA-CACT...G--G-A-----T-T-C--CTAG--CTCACCAGT--TGGCCGGCACTGGGCAGA-.GGCTCCACGCTTG.	683
HIV2_B_D205	--CA-CACT...G--G-A-----T-T-C--CTAG--CTCACCAG--AGGCCAGTGTGGGCAGA-.GGCTCCACGCTTG.	693
HIV2_B_EHO	--CA-CACT...G--G-A-----T-T-C--CTAG--CTCACCAG--TGGCCGGTGTGGGCAGA-.GGCTCCACGCTTG.	690
STMM83293	--CA-CACT...G--G-A-----A-T-T-C--CTAG--CTCACCAG--TGGCCAGTGTGGGTAGAGTGGCTCCACGCTTG.	324
SMM251	--CA-CACT...G--G-A-----T-T-C--CTAG--CTCACCAG--TGGCCAGTGTGGGCAGAGTGGCTCCACGCTTG.	639
SMM9	--CA-CACT...G--G-A-----T-T-C--CTGG--CTCACCAG--TGGCCGGTGTGGGCAGAGTRGCTCCACGCTTG.	134
SIVCPZANT	0
SIVCPZGAB	-T-CAC-G--...-T-----AC.C-GTG-A-----A-----	540
SIVCPZUS	-TC-C-AG-.--...-AA--C-----CT-G-GTA-----	550
SIVSTAN1	-TACTAGGAGA.----C-TGAG.-----T-G---TG-G-CTAACCTGGTCTG--CATCCAGGGTAAG----C.....	574
SIVGRI677	405
SIVVER9063	-TACTAGGAGA.----C-T.-----T-T-G---T-G-CTAACCTGGT.TG--CA.CCAGGGTAAGGA-TC.....	585
SIVVER155	-TACTAGGAGA.----C-T.-----T-T-G---T-G-CTAACCTGGT.TG--CA.CCAGGGTAAGGA-TC.....	589
SIVVERTYO	-TACTAGGAGA.----C-T.-----T-T-G---T-G-CTAACCTGGT.TG--CA.CCAGGGTAAGGA-TC.....	81
SIVVERAGM3	-TACTAGGAGA.----C-T.-----T-T-G---T-G-CTAACCTGGT.TG--CA.CCAGGGTAAGGA-TC.....	81
SIVLHOEST	-GCTC-GG-GA.C-AGC-GTC-CT----T.GAGCACACTGGT-AGCTA-G-AGT-A--CCTGGGAGGCTACACTTAGCACCCACCTGGGCCTGATCA.....	619
SIVSUN	-GCTCAG--GC..TAG-CGAC--C---A---G---ACCTAG-GCT--CTTGT-G--C--G-GTGAGCACTTGTCTGAGCACTCACCCGAGCCTGATC.....	592
SIVMNDGB1	-TCTAAACCTG.....	218
SIVSAB1C	-GCTTGG--GC.TG-CAGAT-----T-T-----TA-GTCT-AACCAG-TTGAGCCTGGGTGTCTGCTG-GGAAGGCAAGCCTAAGAGCACTCTGGTGAGGCTTGGTAGACCCTGCCTGGCG	615
SIVSYK173	GG-A-AGC-GCCCGGAGC-CTT---A-TA-AGG--AGT-AGT-CG--TTTGG-CGCC-GGTCTTCGGCAGCTCT--GTTGCTTGTCTTG.....	498

Poly-A signal secondary structure in this region, see Rizvi, J Virol 67:2681-2688 (1993)
5' LTR R repeat end \ / 5' LTR U5 start

B_FR.HXB2R A_UG.U455 C_ET.ETH2220 D_ZR.84ZR085 FI_BE.VI850 G_SE.SE6165 H_CF.90CF056 J_SE.SE92809 CRF01_AE_TH.CM240 CRF02_AG_NG.IBNG CRF03_AB_RU.KAL153 CRF04_cpx_CY.94CY03 N_CM.YBF30 O_MVP5180 O_CM.ANT70 HIV2_A_ST HIV2_A_BEN HIV2_A_ALI HIV2_B_D205 HIV2_B_EHO STMM83293 SMM251 SMM9 SIVCPZANT SIVCPZGAB SIVCPZUS SIVSTAN1 SIVGRI677 SIVVER9063 SIVVER155 SIVVERTYO SIVVERAGM3 SIVLHOEST SIVSUN SIVMNDGB1 SIVSAB1C SIVSYK173C.AATAAA.GCTTGCCCTGAGT....GCTTCAAGTAGTGTGTGCCCGCTCTGTT.GTGTGACTCTGGTAACAGAGATCCCTCA -A-----G-----AT------T-----TT-----G.T-AG------A-T-----NNAA-.G-CA-C.-A-----C-----CGCTC-----GAGAAGCA...GTGTGC-CAT-TGT-CAACC.C-GTGT.....GAGAAGCA...GTGTGC-CAT-TGT-CAACC.C-GTGT.....CTTG.CTT.AAAAGACC.TCTT.....AGTTAGAAGCAAG-T-.GT-----CT--A--C-CCTA-CG-CGCCTGGT-AT.TCGGTGT--CTTG.CTT.AAAGACC.TCTT.....A.-TTAGAAGCAAGT-AGT-----T--A--C-CCTA-CG-CGCCTGGT-AT.TCGGTGT--CTTG.CTT.AAAAGACC.TCTT.....AAT.TAGAAGCAAGT-AAG-----T--A--C-CCTA-CG-CGCCTGGT-AT.TCGGTGT--CTTG.CTT.AAAAGACC.TCTT.....A.ATTAGAAGCAAGT-.AGT-----T--A--C-CCTA-CG-.GCCTGGT-A.TCGGGT--CCTTG.CTT.AAAAGACC.TCTT.....A.ATTAGAAGCAAGC-AGT-----T--A--C-CCTA-CG-CGCCTGGT-AT.TCGGTGT--CCTTG.CTT.AAAGACC.TCTT.....AATTTAGAAGTAAAG.C-AGT-----T--A--C-CCTA-CG-CGCCTGGT-A-CTCGGTA---CTTG.CTT.AAAGACC.TCTT.....A-TTTAGAAGTAAAG-C-.GT-----T--A--C-CCTA-CG-CGCCTGGT-A-CTCGGTA---GCTTG.CTT.AAAAGACC.TCTT.....AGTT.AGAAGTAAAG-C-.GT-----T--A--C-CCTA-MCG-CGCCTGGT-ATCTCGGTA--GGCCTC-----TA-TTGAGCAGTGTGTG-TATCTA-TCA-AC-----CTTGCTTA...GAAGGCTATTG-T-----GCTTAAGAGT-ATCT-AG.CAA----TCA-TGC....CGC-CTTCTCT-AG-GA-C---GC-GCT-GCTTAGTACGCTA-ATTG-AGTCAA---TCA-TGCT-C-CCGAG-CTC--GAGGTGA-C-T---ACT-----GC-T-AGAGCT-ATCT-AGTCAA----TCA-TGAC-CC-C----T-G-ACG-GAGAAGT-C-T-----GC-T-AGAGCT-ATCT-AGTCAA----TCA-TAAT-CC-C-----T-G-ACG-GAGAAGT-C-T-----GC-T-AGAGCT-ATCT-AGTCAA----TCA-TGAC-CC-C----CT-G-ACG-GA---T-CT-----GC-T-AGAGCT-CTCT-AGTCAA----TCA-TGAC-CC-C----T-A-CGGGT-A-A-.T-CAAG--TCATT--TAGAAAAG-AC-AGCTA----CT-ATAG-AG-CTCT-TCTC--TCGCC--G-TTCAGTGTACAGCTTGGGGAGCTAGAGGCTCCTTGCCTTTC...AGTACAAGCCTTC.....-C-----AGT-TTACTGCAAGCAA...T-----G-TC-CTCT-TCTC--TC-----GT-----AGT-TACAAG-AA-GCA-----A-T-A--CC-CA-GACAACCCT.....A-----G-GTGG-AA-CGCAAGTAAAGTCTGCTCTCA-CTGTTCTGC-GG--CACCTC-C-AAGGTGCTCAGGGAGCGC	601 71 0 115 0 0 0 0 148 131 598 0 160 604 630 228 783 778 785 785 420 736 231 0 621 625 666 476 680 682 175 176 744 697 218 730 595
---	--	--

	Gag and Gag-Pol start (p17) ->	
B_FR.HXB2R	GGAGG.....CTAGAAGGAGAGAG.....	790
GAG		GAG
A_UG.U455	-----A.....	257
C_ET.ETH2220	-----	170
D_ZR.84ZR085	-----	307
FI_BE.VI850	-----	145
G_SE.SE6165	-----G-----	193
H_CF.90CF056	-----	137
J_SE.SE92809	-----	113
CRF01_AE_TH.CM240	-----	355
CRF02_AG_NG.IBNG	-----	330
CRF03_AB_RU.KAL153	NNNNN.....NNNNNNNNNNNNNNNNNGCGGAGGCTAGAAGGAGAGAG.....	841
CRF04_cpx_CY.94CY03	-----	156
N_CM.YBF30	CC--A.....AAGT-G-AGAGAG-----	348
O_MVP5180	--T--C.....C--CCT--G-GAAGGGCGAAGTCCCTAGGGGAGGAAG.....	817
O_CM.ANT70	--T--C.....C--CCT--G-GAAGGGCGAAGTCTCTAGGGGAGGAAG.....	841
HIV2_A_ST	.A---.....TA--TGCTACACCAAATACAGTAGCCAGAA.GGGCTTGTATCCTACCTTTAGACGGGTAGAAGATTGTGGGAG.....	548
HIV2_A_BEN	.A---.....TA--T-CCTACACCAAACCTGTAGCCAGAAAAGGCTTGTATCCTACCTTTAGACAGGTAGAAGATTGTGGGAG.....	1103
HIV2_A_ALI	.A---.....TA--T-CCTACACCAAAT.TGTAGCC.GAAAGGGCTTGTATCCTACCTTTAGACAGGTAGAAGATTGTGGGAG.....	1096
HIV2_B_D205	A.....TA--T-CTCC-A.....TTTTGCATTGTCTACTCTAAAGAGGGAGTAGGGCATAAAGTGGGAG.....	1079
HIV2_B_EHO	AA---.....TA--T-CTCC-A-TTT.....GCAATTCGTCTACTCTAGGACAGGGTACGGAAATAAGTGGGAG.....	1080
STMM83293	-C-A-.....TG--TGCA-C-AGTAGCCAGA.....GGGCTGATTGTCTACTCTTGGGGAAGAGGTAGAAGGTGGGAG.....	709
SMM251	-C-.....TA--TGCA-C-C-AAAAAGAAA.....TAGCTGTCTTGTATCCAGGAAGGGATAATAAGATAGAGTGGGAG.....	1041
SMM9	-C-.....TA--TGCA-C-A-A.....GAGTCATAGGACTGAGTTCCTACTTTTGGAGAAAAGTAGGAGAGTGGGAG.....	535
SIVCPZANT	TTGCTAGTGAGC-ATCCT--G-GAAGGTCGAAGTCTCTAGGAACAGGAAG.....	154
SIVCPZGAB	--T-.....--CCT--G-GAAGGGCGAAGTCTCTAGGGGAGGAAG.....	827
SIVCPZUS	A-T-.....-G--CCT--G-GAAGGGCGAAGTCTCTAGGGGAGGAGG.....	832
SIVSTAN1	.TGCA.....GGTAG-TCTAG-T-GGCGACAGATCAGCACCTGGGTGAGCA.....	902
SIVGRI677	AACCA.....ACCC-GTTG--CGAAGGGTTGGTAGGGGACGGTCCGGAGCA.....	727
SIVVER9063	TAGCC.....A--ACTACTCT-G-CAAGTAGGGCAGGCCGAC.GGTACGCA.....	942
SIVVER155	TAGCC.....G--ACTACTCT-G-CAAGTAGGGCAGGCCGACGGGTACGTA.....	931
SIVVERTYO	TAGCC.....G--ACTACTCT-G-CAAGTAGGGCAGGCCGTT.GGGTACGCA.....	432
SIVVERAGM3	TAGCC.....A--ACTACTCT-G-CAAGTAGGGCAGGCCGACGGGTACGCA.....	431
SIVLHOEST	-C--A.....AGGT-----TAAGACCTCACCGGGAGAGTGGCAGCAGGCCGGGGCAGCCTTGTGAGAAGTGTGTAATTTCTGATCTAGTGTAACTGGGACACTAGTGGTAGCTTCACT.....	1076
SIVSHUN	CAGAA.....AGTCCTA-GCT-G-GGGAGGTCGCCGTAGGTTAGTACTGTCAAGTAGGCAGACTAGTAGTACGTGGAATAAAAATAAGAAGTGTGTATAACTCTGTCTCGGAAGAGTGTCTCAAAC.....	1063
SIVMNDGB1	AAGAATTAGTTGTACCCT--T-TAAGGGGCAGCATAGTCAGAGCA.....	450
SIVSAB1C	A----AAAAAGAGA-GTC--C--GGCTGACAAAGGCTGGCTTAGCAGTGGCAGAAGGCTAGGTCGAGCGGAAAAGGAAA.....	1017
SIVSYK173	--GACC.....	795
GAG		GAG

B_FR.HXB2R	TGGGTGCGAGAGCGTCA...GTATTAAGCGGGGAGAATTAGATCGATGGGAAAAAATTCGGTTAAGGCCAGGGGAAAGAAAAATATAAAATTAACAATATAGTATGGCAAGCAGGGAGCTAGAACG	917
GAG	M_G_A_R_A_S...V_L_S_G_G_E_L_D_R_W_E_K_I_R_L_R_P_G_G_K_K_K_Y_K_L_K_H_I_V_W_A_S_R_E_L_E_R	GAG
A_UG.U455	-----AA-A-----TC-----G-----C-----G-C-G-----T-----G--AA	384
C_ET.ETH2220	...A-----A-C-A-A-----GCC-----AA-----C-C-TGC-G-----CC-----C-----A-----G--AA	297
D_ZR.84ZR085	-----G-----GCT-----C-----G-----GC-----G-----A-----G-----	434
FL_BE.VI850	...A-----A-----GA-----A-----G-----G-----A-G-----C-A-----	272
G_SE.SE6165	-----C-----A-----GCT-----G-----G-----G-----TC-----A-----C-----G--GA	320
H_CF.90CF056	-----C-----A-----GCT-----G-----C-----G-----G-----G-----G-----C-----G--A	264
J_SE.SE92809	...A-----T-----A-----GAT-----G-----G-----G-----G-----G-----C-----G--CA	240
CRF01_AE_TH.CM240	-----T-----A-----GC-----GC-----GA-----GGC-G-----T-----A-----T-----A	482
CRF02_AG_NG.IBNG	-----T-----A-----GC-----G-----GC-----G-----G-----C-----C-----G-----A	457
CRF03_AB_RU.KAL153	-----A-----A-----GC-----G-----G-----G-----G-----CC-----G-----A	968
CRF04_cpx_CY.94CY03	-----T-----A-----GC-----G-GG-----G-----G-----C-----C-----A-----T-----G--A	283
N_CM.YBF30	-----G-----CA-----A-----A-----TC-----TAT-----G-----A-----C-----G-----G-----T-----G--A	475
O_MVP5180	-----T-----T-----G-----G-----CA-----A-----TA-----G-----GC-----CG-----A-----ATCT-----A-----GGC-----G-----G-----T-----G--A	944
O_CM.ANT70	-----T-----T-----G-----G-----CA-----AA-----CA-----G-----GC-----C-----A-----A-----ATCT-----A-----G-----G-----C-----CT-----A-----G--A	968
HIV2_A_ST	-----C-----AAC-----C-----C-----G-----A-----AA-----A-----GC-----CGA-----TA-----G-----A-----C-----C-----C-----GGC-----T-----G-----GCG-----AT-----AT-----G--CA	675
HIV2_A_BEN	-----C-----AAC-----C-----C-----G-----A-----AA-----A-----GC-----CGA-----TA-----G-----A-----C-----C-----C-----GGC-----T-----G-----GCG-----AT-----AT-----G--TAA	1230
HIV2_A_ALI	-----C-----AAC-----C-----C-----G-----A-----A-----A-----GC-----CGA-----TA-----G-----A-----C-----C-----C-----GGC-----T-----G-----GCG-----AT-----AT-----G--CA	1223
HIV2_B_D205	-----C-----GC-----C-----CC-----GTCA-----AAGA-----AC-----GA-----TA-----G-----A-----C-----C-----C-----TGC-----G-----GTG-----AC-----AT-----CA	1206
HIV2_B_EHO	-----C-----GC-----C-----C-----GTCA-----AAGA-----GAC-----GA-----TA-----GG-----A-----C-----C-----C-----G-----C-----TGC-----G-----GTG-----AT-----AT-----GA	1207
STMM83293	-----C-----AGC-----C-----C-----GTCA-----AA-----A-----GC-----GA-----TA-----G-----A-----C-----C-----C-----G-----G-----G-----G-----GCA-----AC-----AT-----G--CA	836
SMM251	-----C-----AAC-----C-----C-----GTCA-----AAGA-----GC-----GA-----TA-----A-----C-----C-----A-----C-----C-----G-----C-----TG-----G-----G-----G-----GCA-----AT-----AT-----TA	1168
SMM9	-----T-----AAC-----C-----C-----GTCA-----AAGA-----GC-----GA-----TA-----A-----C-----C-----C-----G-----C-----TGC-----G-----G-----GCA-----AT-----AT-----G--CA	662
SIVCPZANT	-----A-----G-----G-----T-----T-----G-----G-----A-----AGA-----GC-----AC-----GT-----CA-----C-----TC-----C-----T-----C-----G-----G-----C-----TGA-----C-----G-----T-----ATC-----GC-----G--A	281
SIVCPZGAB	-----TC-----CA-----A-----A-----C-----C-----T-----A-----C-----GA-----G-----TGA-----G-----T-----A-----G-----A-----G-----A	954
SIVCPZUS	-----TC-----G-----CA-----CG-----GCT-----G-----G-----C-----T-----A-----T-----G-----G-----TGA-----G-----T-----T-----G-----A	959
SIVSTAN1	-----A-----G-----CAC-----G-----C-----C-----GTCA-----A-----GA-----T-----G-----CAC-----TT-----G-----A-----C-----C-----C-----GAAC-----G-----G-----C-----CT-----A-----G-----AAA-----AA-----G-----G--	1029
SIVGRI677	-----C-----G-----G-----TCAC-----C-----C-----GTCA-----A-----AGCC-----C-----CAGC-----TC-----G-----G-----A-----C-----C-----T-----GAAC-----G-----A-----G-----G-----CC-----A-----T-----T-----A-----G-----A-----AA-----AA-----G-----G--	854
SIVVER9063	-----G-----GCTA-----C-----C-----C-----AA-----A-----C-----C-----A-----TC-----G-----A-----A-----C-----C-----GAAC-----C-----A-----T-----G-----T-----A-----G-----T-----AA-----A-----G-----G--	1069
SIVVER155	-----G-----GCTA-----C-----C-----C-----G-----ATA-----A-----C-----GA-----TT-----GC-----T-----A-----AC-----TC-----C-----GAAC-----G-----C-----A-----T-----T-----A-----G-----A-----A-----A-----G-----C--	1058
SIVVERTYO	-----G-----GCTA-----C-----C-----C-----ATA-----A-----C-----CAA-----TT-----GC-----T-----A-----AC-----TC-----C-----GACC-----G-----CC-----A-----T-----T-----A-----G-----A-----A-----G-----G--	559
SIVVERAGM3	-----G-----GCTA-----C-----C-----C-----ATA-----A-----C-----CAA-----TT-----GC-----T-----A-----AC-----TC-----C-----GACC-----G-----CC-----A-----T-----T-----A-----G-----A-----AA-----G-----G--	558
SIVLHOEST	-----AT-----G-----TAAC-----T-----CC-----C-----TA-----CA-----ATTGA-----A-----AGAT-----TTTGT-----GTG-----CA-----C-----A-----T-----A-----A-----G-----CT-----C-----AA-----G-----G-----A-----CG-----AA-----AT-----CA	1203
SIVSUN	-----A-----G-----AAC-----C-----AC-----G-----GGATA-----A-----TTG-----GAGGA-----C-----TT-----G-----G-----AGCTC-----G-----AA-----T-----A-----G-----A-----G-----CC-----CC-----GGC-----G-----G-----T-----AA-----AT-----TA	1190
SIVMNDGB1	-----AATG-----GAAC-----T-----CC-----GTTA-----ACT-----T-----G-----AA-----TT-----G-----G-----AA-----A-----A-----AG-----T-----T-----A-----TGT-----G-----G-----CC-----CTGT-----TGT-----AAG-----T-----AT-----TA	577
SIVSAB1C	-----TAAC-----C-----T-----TA-----GA-----C-----GC-----TC-----TC-----G-----CA-----C-----CAAC-----G-----G-----G-----G-----C-----A-----A-----T-----	1144
SIVSYK173	-----A-----AGCG-----G-----GGCAA-----CA-----TA-----GC-----CA-----AT-----G-----C-----CC-----A-----GAAAA-----G-----GCGC-----CCT-----G-----TCG-----T-----T-----AG-----A-----T-----G-----CA	925
GAG	M_G_A_A_G_S_A_I_L_T_G_R_E_L_D_R_Y_E_K_I_R_L_R_P_K_G_K_K_R_Y_L_V_R_H_L_V_W_A_K_K_E_L_D_R	GAG

B_FR.HXB2R	ATTTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAAGCTTAGATCATTATATAATACAGTAGCAACCCTCTAT	1047
GAG	F A V N P G L L E T S E G C R Q I L G Q L Q P S L Q T G S E E L R S L Y N T V A T L Y	GAG
A_UG.U455	---A-C---C---T---G---A---CAG---AT---G-T-C---A---GT---	514
C_ET.ETH2220	---T-C---C---A---T---T---C---A---A-TAA---G-T---A---G---A---T---G---T---	427
D_ZR.84ZR085	---T-C---T---T---A---A---A---G-T---A---	564
FL_BE.VI850	---T-C-G---TC---CA-A---AA---G---AG---T---GT---	402
G_SE.SE6165	---T-C---C---A---T---G---T---CA---A---A---T-G---T-C---A---G---A---A---G---G---G---	450
H_CF.90CF056	---T-C---C---C---T---C---CT---G---A-A-A---A---G---G-TA---A---A---A---T---CT---GT---	394
J_SE.SE92809	---T-C---C---TC---GT---G---A---CA---A---T---C---G-T-C---A---C---A---A---G---	370
CRF01_AE_TH.CM240	---C---C---A---T---T---G---A---CA---A---A---T---GT---A-T---CA---T---A---T---GG	612
CRF02_AG_NG.IBNG	---C---C---T---G---G---A---CA---C---A---A---T---T---G-T---CAG---T---A---T---A---T-GG	587
CRF03_AB_RU.KAL153	---GC---C---A---T---A---CA---A---A---T---A-T---CA---	1098
CRF04_cpx_CY.94CY03	---C---C---T---G---A---CA---T---A---A---T---T---A-T---CA---A---TA---A---GG	413
N_CM.YBF30	---TTG---C---A---T---CA-G---C---G-G---C---GCCA-GT---AAAT---AT---G---G-T---CA---G---GC-C---T---CG-TC---GTT---T---	605
O_MVP5180	---A---TG---T---A---TG---TAC-GAG---C-G---ACAG---T---G-G---G-T---CA---G---G---C---G-A---TC-C-GG---CG---A---GTA---GG	1074
O_CM.ANT70	---TG---AG---AC---TG---G---TAA-GAGA-GC-G---ACAG---T---G-G---G-T---CA---G---CAGC---GCAG---C-C-GG---CG---A---T-GTG---GG	1098
HIV2_A_ST	---G-T-GGCAGAGA---G-GT-AA---G-T-CCA-A---T-TAC-GTTT---G-T---TAG-A-C---G---A-TT-A-A-AGCC-T-T---T---CTGCCT-A-T-GG	805
HIV2_A_BEN	---G-T-GGCAGAGA---G-GT-AA---T-CCA-A-G---T-CA-GTTT---G-T---TAG-A-CA---G---A-TT-A-A-AGCC-T-T---C---CTGCCT-A-T-GG	1360
HIV2_A_ALI	---G-T-GGCAGAAA---G-GT-AA---T-CCA-AG---T-TAA-GTTT---G---TAG-G-CA---G---A-TT-A-A-AGCC-T-T---T---CTGCCTAG-T-GG	1353
HIV2_B_D205	---T-G-T-AGCAGAGA---G---T-AA---A---CA---A---T-TAA-GTCT---GCT---TAG-A-CA---G---A-TT-A-A-AGCC-T-T---TC---CTGCCT-A-T-T-	1336
HIV2_B_EHO	---T-G-T-GGCAGAGA---G---G-T-AA---A-CC-GA---T-G-AA-GTCT---GG---TAG-A-CA---T---A-TT-A-A-AGTC-T---C---CTGCCT-A---T-	1337
STMM83293	---T-G-T-GGCAGAAA---G-GT-AA---G---CA---A---TA-AACGGTCT---G---TAG-G-CA---T---A-TT-A-AGAGTC-T-T---T---CTGCCT-A-T-GG	966
SMM251	---T-G-T-AGCAGAAA---G-G-ACAA---A---CA---A---TTCGGTCT---GCT---TAG-G-CA---C---A-TT-A-A-AGCC-T---T---CTGCCT-A---GG	1298
SMM9	---T-G-T-AGCAGAAA---G-G-ACAA---T---CA---A---T-TTC-GTTT---GCT---TAG---C---T---A-TT-A-AGAGCC-T---C---T---CCGCGT---T-GG	792
SIVCPZANT	T-T---GC-C-GCT-CTC---TC---T---GA-A-GGCTA-CCAT---AT-GAGC---T---A-AG-A-T-A---CCCT---A-A-T---T-G-T---C---CA-TTGTGTT---G-GG	411
SIVCPZGAB	---TG-G-C---C-G-AA-G---GTAAG---A---CTA-T-GT-ACA---AT-G-G---G-T---CA---A---C---G---GC-G---C-G-T---C---TC-G---GTA---G-GG	1084
SIVCPZUS	---T---TTG---C---A---T-AA-G---G---C---CTC---T---T-AAA---T---G---G-T---AA-A---G---GC---GC---GC-C-T---C---CT---GGT---T-GG	1089
SIVSTAN1	C-T-GGC-CC-CGAGAAA---C---AA---G---CA-A-G---CA-A-A-GTAT-GAGC---GCTAGAA-CA---G---A-A-AGTC-G-T---TTGTGCTGCGT-A-T-GG	1159
SIVGRI677	---T-GGT-AC---GAGAAA---T---AA---CA---A---CA-A-A-GTTT---ACC-G-TGGAA-C---C-C-G-GG---A-A-G-TC-G-T---TTGTGCTGCGT-A-T-GG	984
SIVVER9063	C---GCC-CC---GA-A-GT---G---GA---T---A---A-A-G---CA-A-A-GTC---TTTT---T-TAGAA-CA---C-G-G-GCT-A-A-AGTC-G-TC---CTT---GTGCGT---A---	1199
SIVVER155	---GCC-CC---GAGAAGT-A---G-G---GAG---T---A-A-G---CA-A-A-GTT---CTCT---CCTAGAA-CA---G-G---G-A-G-A-AGTC-G---CTG---GTGCGTAT-GCT-	1188
SIVVERTYO	C---GCC-CC---GAGA-GT---G-G---GAG---G-G---A-AG---CA-A-A-GTC---CT-C---CCTAGAA-CA---G-G-G-GCT-A-A-AGTC-G-TC---CTT---GTGCGTAT-A---	689
SIVVERAGM3	C---GCC-CC---GAGA-AT-AC---T---GA---A---AGA-G---CA-A-A-GTA---CT-C---GCTAGAA-CA---G-G-G-GCT-A-A-AGTC-G-T---CCTT---GTGCGTAT-G-T-	688
SIVLHOEST	---T-G-T-GGGAT-ACAA---C---G---G---G---AGA---C---ATCTGTCTGCTGG---T-TATA-GCA-T-T-TA-GA-C-A-AGG-C---GTAGGG---TGTGT-A---GC	1333
SIVSUN	G-T-G-C---CAAG-AA---AG---GTG---CAGA-G---T---TC-GTCTGTGG---T-TGTA-GCAT-T-T-TA-GA-C-G-AGG-C---GTAGG---TGTGTG---TGCC	1320
SIVMNDGB1	---T-GCT-ATCGGA-AAA---CC-T---CAGC---A---GA---A---C---CTC-GTATGTTGG---TATA-G-CCA---TA-T-A-A-G-T---GGTAGGG---CTGTGTTG-AGCC	707
SIVSAB1C	---TAGCC---TCAG-AAAT---C---G---AAG---GGTCGTCA---T---TA-TGTA---C-TT---CCTAG---CT---G---GA-TT-A-TTG-TC-C-T---TT-TGCTGCGTGT-AGC-	1274
SIVSYK173	---GCC---TCGGACCAA---CA-G---GT-TAA---GA---A---T---CAC-GTC---TT---TAGAAGCT-AT---G---A-TT-A-A-AGCC-T-T-GGC-T-A-TT-GGT-G-A-GG	1055
GAG	F G L S D Q L M E S K E G C E K I L T V L L P L E A N G S E N L K S L F G I I S V V W W	GAG

B_FR.HXB2R	TGTGTGCATCAAAGGATAGAGATAAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAA.....AACAAAAGTAAAGAAAAAAGCACAGCAAGCAGCAGCTGACACAGGA.....	1158
GAG	<u>C V H Q R I E I K D T K E A L D K I E E E Q N K S K K K A Q Q A A A D T G</u>	GAG
A_UG.U455	----A-----TG-----A-T-A-----AT-----T-G-AC--C--GGA--G-----A-----	625
C_ET.ETH2220	----A-----A-----C-----A-----G-----C-C--A-----G--G--A-CTGACA-----	538
D_ZR.84ZR085	----A-A-----A-----A-----C-----GG-----G-----	675
FL_BE.VI850	-A--A-----GC--G-G--C-----C-----A-----C-C--GA-----G-----A--G-----	513
G_SE.SE6165	-A--A-----G-----A-----C-----AG-AG-G--AA-GATA-----G-----C-G--AT-----G-----ATG--T-A-----	561
H_CF.90CF056	--C-A--G-AA--TG-G-----G-----T-----ATA-----C-C--A-----T-AG-A-----	505
J_SE.SE92809	--C-A--G--A-----A-----T-----G-----ATT-----G-AC--AC-GC-G--A-----A-A-----A-AA-----	481
CRF01_AE_TH.CM240	--C-A-C-----G-----T-A-----TA-----T-G--CC--CG--GA-----G-----G-----	723
CRF02_AG_NG.IBNG	--C-----C-----C-----T-A-G-----ATA-----T-G--C--C--G-T-----GA-----C-----	698
CRF03_AB_RU.KAL153	----A-----C-----T-A-----ATA-----T-G--C--C-----GA-C--A--G-----A--G-----	1209
CRF04_cpx_CY.94CY03	--C-A--A-----TG-C-----T-A-----ATA-----GT--G--C--C-----GA-----G-----C-G-----	524
N_CM.YBF30	----C--AGT--C-----C-CA--AC-----G-----A-A--GA-A--GGAACAGCACAAGCC-G-GCCA--A--CCC--A-GCAGGG--G--A-CA--T-AT-----	728
O_MVP5180	--C-T-CA-C--AT-T--C--CCG--T--AC--C-G--AA--C-A--T--A-----TAATG.....GCA-GC--G--TCTGC--AGGCCGCTAAG-A--AAACA-GCCCT.....	1185
O_CM.ANT70	--T-CA-C--ATATA-A--TGG--T--GC--C-G--AA--C-A--T--A-----TAATG.....GGG-GC--G--TCTGCG-ACGCCGCTAAG-A--ACACA-GC-C-----	1209
HIV2_A_ST	--A-A--CGC-GAAGAGA-AGCG--T--TG-A--AAA-C-A--G--C--AG-----CATCTAGTGGCAG-A-C--AA-CT-C-G--AA-AT--C-AGTA--AG-AGAC--AC-GCA....	925
HIV2_A_BEN	--CT--CGC-GAAGAGA-AG-G--T--TG-----AAAGA-ACTAGC-C-AG-----CATCTAGTGGCAG-A-CTG-A-CTGC-G-GAA-AT--C-AATA--AG-AGAC--AC-GCA....	1480
HIV2_A_ALI	--C--CGC-GAAGAGA-AG-G--T--TG-A-G--AAA-C-A-TA--C--AG-----CATCTAGCGGCAG-A-T-GAA-CAGC-G-GAA-AT--C-AGCA--AG-AGAC--AC-GCA....	1473
HIV2_B_D205	--CC--CGC-GAAGAGA-AG-G--T--AG-----AAA-A-G-TAGC-C--AG-----CATCTAGCGGCGG--C...G-A--TGC--GCTAC-AATAA-C-AACAG-----	1447
HIV2_B_EHO	--CC--CGC-GAAGAGA-AG-G--T--TG-----AAA-A-G-TAGC-C--CG-----CATCTAGCAGCGG--C...G-A--TGC--GCTATGAGTAA-C-AAGT-A-----	1448
STMM83293	--A-C--CGC-GAAGAGA-AG-G--C-T--TG-----AAAGC-AGT-G--A--AG-----CATCTTGTGGTAG-G-CTG-A-CTGC--CAA-AT--CT--CA--AG-AGAC--AC-GCA....	1086
SMM251	--CA-T--CGC-GAAGAGA-AG-G--C--TG-----AAA-C-G-TAG-GC--AG-----CACCTAGTGGTGG-A-C-G-A-CAGC-G--A-TAT--C-AA-A--AG-AGAC--AC-GCA....	1418
SMM9	--CA-T--CGC-GAAGAGA-AG-G--C-T--TG-A--AAA-C-A-TAG-GC--AG-----CATCTAGTGGTGG-A-C-G-A-CAGC-G-CAA-AT--C--A--AG-AGAC--AC-GCA....	912
SIVCPZANT	--C-A--A--G--GA-A-----AG-AC--CG-TA-A-CAG-GA-AATGA-AGT-----TGC-G-CAC-AGC-G--A--GGAAGTAGCCA-A-C-CA-GCA-----	522
SIVCPZGAB	--CA-A--AGTGAC--CACTG-G-----AC-A--C--AC--C--A--CGCG-T--T-----GGAG--CAAC--GC--A-TG-AAGTAAGT--GAAG-CGT-A-----	1195
SIVCPZUS	----C-CAGT--G-GACAG-GG-G--T--GC--C-G--C--TG--AC-GA-A--G-TAGTG.....C-GTTGCAG--AC--C-G-A-G-AA--A-CA-CAGC-ACAG-AGGCAAGT.	1206
SIVSTAN1	--A-T--CGC-GAACAGA-AG-G--T--AG-----AG-G-TATTAG--A-AC--CGTGCCATCTAGTGGAA--AG--AA-CTGC-GC--G-CATCT-GT-GCCAGC-GCA-AATTACAACA	1289
SIVGRI677	--CA-T--CGC-GAACAGA-AG-G-----AG-----G--TA-CAG-TA--C--C-CT-CCATCTAGTGGAC--A--TGAG--AGC-GCTAA-A--A--AAT-AGA-AACAG-G.....	1104
SIVVER9063	--CA-C--C--CGAACAGA-AG-G-----AG-----GG--CA-CAG--AGAC--TGCTGCCATCTAGTGGAA--GG--AA-GAGC--C--AG-CATCTAGT-G--AAA-G-A-A-TAACAGGG	1329
SIVVER155	--C--C--C--GAA-AGA-AG-G-----AG-----CA-TAG--AGAC--TGCTGCCACCTAGTGGAC--AG--AA-CTGC-GTTA-G-CA-CT-GT-G-CAGC-G-A-AATAACACAG	1318
SIVVERTYO	--CT--CA--GGAACAGA-AG-G-----AG-----AG--CA--TAG--AGAC--C-CTGCCATCTAGTGGAA--AG--AA-GTGC--C--AGACATCTAGT-G-CAAA-G-A-AATGACAAGG	819
SIVVERAGM3	--C--A--CA--GAT-AG--AG-G-----AG-A--AG--CA--TAG--AGAC--TGCTGCCATCTAGTGGAG--AG--A--TGC-G--AG-A-TAC-A--AGA-ATCT-GT--CAAAAAGA	818
SIVLHOEST	--CTGT--C-T-G-AG--AGA--TC-----TC-----CA-CAGT--AG--A-AAT--A-C-GCACCAGAAGCAGCAGGG.....	1417
SIVSUN	----TGT-----G-C--TCCAG-----AC--G--C--A-A--G-CAGACT--C-GCAAAAAATGAAG-AGC-CAAGCTTCTC--AATTTT.....	1422
SIVMNDGB1	--CA-A--CGC-G-T-----A--T--GAG--AC-A--T-----A-A--AT--A-A-TCATAAC-AGAAAAGGAAGAA--GC-GGAGG-TG-----	800
SIVSAB1C	--A-T--CGC-GAA--A-AG-G--T--AG-----AAA--CG--AG--A-A--G--AGTGCCAGCAGAAAATG-CAG-----GCC-C--GCGA--TCTAGT-GCCA-A--A--GGA-CTGCAGGCAA	1404
SIVSYK173	GCC--T--GCC-AA-AG--AG--G-----TG--C-----AAA-C-G--AG--A-A--CATGTAATTGGAAGATG--CC-CCAGC--C-TCT-GTGA--AGT-A-AA-AG--GTCA-AACATGG	1185
GAG	<u>A V H - K K E V E D T E Q A K Q K V K E - C N W K D D P P - T S G G Q S E N S S Q N M</u>	GAG

	Gag p17 \ / Gag p24	
B_FR.HXB2RCACAGCAATCAGGTCAGCCAAAATTACCCTATAGTGAGAACATCCAGGGGCAATGGTACATCAGGCCATATCACC	1235
GAGH_S_N_Q_V_S_Q_N_Y_P_I_V_Q_N_I_Q_G_Q_M_V_H_Q_A_I_S_P	GAG
A_UG.U455AG--T-----C-----A-TGCA--A-----CCA----C-----T-----	690
C_ET.ETH2220GGA-AGG-C.....-T-----T-----T-G-----	609
D_ZR.84ZR085A----GC-----T-----A----C-A-A-----C-----C-----	752
FL_BE.VI850A----C-A-----T-----A----TC-T--A-----C--T-TC-----	578
G_SE.SE6165A----C-A-----T-----A----TGCA--A-----C-----A-----	638
H_CF.90CF056A-AGA--CA-----T-----T-----A----TGCT--A-----G-----C-----	582
J_SE.SE92809G--A--G-----T-----T-----TC-G--A-----CC-----C-----C-----	558
CRF01_AE_TH.CM240AG--GCA-A-----T-----A----TGCA--A-----C-----C-TT-----	800
CRF02_AG_NG.IBNGAGCAG-----T-----A----TGC-A-A-----AC-----T--G-----	766
CRF03_AB_RU.KAL153AG--GCA-----T-----C-----A----TGCA--A-----AC--C--T--G-----	1286
CRF04_cpx_CY.94CY03GGT--GCA-T-----T-----A----TGCA--A-----AG--T-----	601
N_CM.YBF30-ATC...-TAGG---T--C--C--CTGCT--A--A-----C-GC-GA-----	799
O_MVP5180-GG--ACA--T-----AACA--TGCA--A-----A-----C--C-----	1256
O_CM.ANT70-GG--CGG-T-----ATCA--TGCG--A-----C--C-----	1280
HIV2_A_ST-CACCT-GCGG-AA-G-AGG--C-T--CG--CAA--AGTGGC--...-CA-CTATACC--GT-C-AC--AGT--	999
HIV2_A_BEN-CACCT-G-GG-AAA--AGG--C--CG-GCAA--AGCGGT...-CA-CTAT--C--GT-C-AC--GAGC--	1554
HIV2_A_ALI-CACCT-G-G-ACAGG-GGG--C-T--CG--CAA--AGTAGC--...-CA-CTACACC--GT-C-GC--GAGC--	1547
HIV2_B_D205CCA-CTAG-G--GG-----T--AG-GCA--ACTGGCT...-CA-CTAC--C--C-T-C-GC--AGC--	1515
HIV2_B_EHOCCA-CTAG-G--GG-----T--AG-GCA--ACTGGCT...-CA-CTATCC--C-T-C-GC--AGC--	1516
SMM83293-CACCT-G-GGCAGAG-AGG-----AG--CA--AGTAGG...-CA-TTAT--C--TAC-AT--AGC--	1160
SMM251-CATCT-GCGCAGAG-AGG-----AG--CAA--A-TAGGT...-TA-CTAT--C--C-T-C-AT--AGC--	1492
SMM9-CACCT-G-GGCAGAG-AGG-----AG-GCA--AGTAGGT...-CA-TTATACC--C-TAC-TT--AGT--	986
SIVCPZANTGGCATGCTTCTGCGGCTGCTCCTGTTAAACAAACAGTGGTGTGTCAGCGA--TCTT-G-GGC-AAG-GAG--C--CA-AGT-G-TGCAGGA--AAT-GCAAGG--C-AC-GA-----	647
SIVCPZGABGGGGAGCCAGTCAAGGCGCTAGTGC--TCTGC-GGCA-T--TGG-----CC-C-A-A--TGCT--A--G-----G-----C-----	1296
SIVCPZUSGG--T--CATT-GT--AGC-----AG--A-A--GCT--A--G-----A--G--T-----	1283
SIVSTAN1CACCT-G-GGCAGACATGGT-----G--A--AC-G-AT...AAT--TG-----ACTC-TT-G-----	1371
SIVGRI677-CACCTGG-GGC-AATCAAG-----AG--AA-TC-G-AT...AATGCCTG-----C--C-TT-G--T-----	1178
SIVVER9063CCACC-GGT-G--T-----T--AGCGCAA--AC-AGGA...AATGCCTG-----GTTT-AC--C-----	1411
SIVVER155GAGGAACAGCG.....-CACC-GGT-G--T-----TT--CGC-CAA--C-AGGG...AATGC-TG--G--GTAC-AC-T-----	1400
SIVVERTYOGAATAGCAGCG.....CCACC-GG-TG--T--G--TT--AGCGCAA--AC-AGGA...AATGCCTG-----GTAC--T-G-----	901
SIVVERAGM3AAAATGACAAGGGAGTAACAGTG.....CCACC-GGT-G--T-----T--AGC-CAA--AC-GGGA...AATGC-TG-A--GT-C--T-G-----	912
SIVLHOESTA-G-AAC-A--ACAG-GGG-----T--CC-GA-AAGAG-A-AT...AG-TG-----ACAC-TT--G-----	1491
SIVSHNCAGTGC-GAGA-AGG-A-----...TCAT--CCTC--TC-CCTAG--CA--G--A.....	1482
SIVMNDGB1-TA-G--T-----G--CAAAG-G-TGCAGCA--A--GTATCAGT--ACTC-A--AGT--	862
SIVSAB1CAAAAGAAAATGAGCCAACAGTG.....ACACCT-G-GGT-G-TCG-G-----C--T--CAGTGT-AT...AAT--TG--C--C--C-AT-G--C--	1501
SIVSYK173CTAGTGAGACATCTAGTGGCCAAAAGGTAGTACAGCAGGAAAAACAAAAGGCAGCA--CACCG-CACCT--AGG--C--T--AC-GC-CAGA--TCCT--AAAT--GTG-A-T--ACA-GAG-CC--GT	1315
GAGA_S_E_T_S_S_G_Q_K_V_V_Q_Q_E_K_Q_K_A_A_T_P_P_P_R_G_N_Y_P_L_L_R_N_P_Q_N_Q_W_I_H_T_G_V_P_V	GAG

B_FR.HXB2R	TAGAACTTTAAATGCATGGGTA AAAAGTAGTAGAAGAGAAGGCTTTTCAGCCCAGAAGTGATACCCATGTTTTTCAGCATTATCAGAAGGAGCCACCCACAAAGATTTAAACACCATGCTAAACACAGTGGGG	1365
GAG	R T L N A W V K V V E E K A F S P E V I P M F S A L S E G A T P Q D L N T M L N T V G	GAG
A_UG.U455	---G-C-G-----G-----C-----A-----G-----T-TG----G-TGT-----	820
C_ET.ETH2220	-----G-----G-A-----G-A-----A-T-----T-----	739
D_ZR.84ZR085	-----G-C-----A-----A-----A-----A-----T-----	882
FL_BE.VI850	-----G-GA-----A-----T-----A-----G-----T-CAC-----T-----	708
G_SE.SE6165	-----G-----A-----C-----T-----A-----TCT-----T-----	768
H_CF.90CF056	---G-C-----A-----T-----A-----C-----TG-T-----T-----	712
J_SE.SE92809	-----G-----GA-----A-A-----G-----G-----T-----A-A-----	688
CRF01_AE_TH.CM240	-----G-----G-----A-----G-----T-A-----A-----C-----G-----T-TG-----T-T-----	930
CRF02_AG_NG.IBNG	-----G-----G-----A-----A-----G-----A-----G-----G-----T-TG-----T-----	896
CRF03_AB_RU.KAL153	---G---G-----G-----A-----A-----A-----A-----G-----G-----TG-----G-T-----	1416
CRF04_cpx_CY.94CY03	---G---G-----G-----A-----A-----A-----A-----C-----G-----G-----TG-----T-T-----	731
N_CM.YBF30	C---C---T---G---GA---G---C---T---T---A---A---ATG---C---G---G---A---G---CTC---C---T---T---T---A---	929
O_MVP5180	C---G---G---C---C---T---A---T---A---T---T---ATG---C---G---G---A---G---CTC---C---T---T---T---A---	1386
O_CM.ANT70	C---G---G---C---A---A---C---T---A---T---A---C---T---CATG---G---G---A---T---T---T---T---TG-CA-A---	1410
HIV2_A_ST	CC---CC---T---C---G---A---AAG---G---GG---AG---G---AGGA---CAG---C---C---CTG---G---CT---A---T---TCAA---T---TTGT---C	1129
HIV2_A_BEN	CC---C---G---T---T---G---G---A---AAG---G---GG---AG---G---AGGA---CAG---C---C---CTG---G---CT---A---T---TCAA---T---TTGT---C	1684
HIV2_A_ALI	CC---C---G---T---T---A---AAG---G---GG---AG---G---AGGA---CAG---C---C---CTG---G---CT---A---T---TCAA---T---TTGT---C	1677
HIV2_B_D205	CC---C---G---T---GT---A---AAG---G---GG---AG---AGGA---CAG---C---TG---TT---T---A---TCAG---T---TTGT---A---A	1645
HIV2_B_EHO	CC---C---C---C---T---A---AAG---G---GG---AG---AGGA---CAG---C---GTG---TT---T---A---T---TCAG---T---TTGT---A---A	1646
STMM83293	A---C---G---GT---G---G---A---AAA---TG---TG---G---G---AGGA---CAG---C---CTGT---CT---T---CA---C---TCAA---T---TTGT---A	1290
SMM251	G---A---C---C---T---GA---G---A---AAA---TG---AG---AG---G---AGGA---CAG---C---G---CTG---CT---T---CA---T---TCAG---T---TTGT---A	1622
SMM9	A---A---T---T---G---AAA---TG---GG---G---AG---G---AGGA---CA---GC---CTG---T---CT---Y---A---C---TCAA---T---TTGT---A---A	1116
SIVCPZANT	A---C---C---C---GTGT---AAA---AT---C---C---T---T---T---G---A---T---T---T---G---A---T---T---T---T---TG---T---A	777
SIVCPZGAB	C---C---G---G---G---A---A---T---T---G---A---A---A---TC---G---TTA---T---G---G---T---T---TG---A---A	1426
SIVCPZUS	C---C---C---C---CG---A---C---T---A---G---C---C---G---CATG---C---G---G---T---C---G---C---T---C---TG---TA---A	1413
SIVSTAN1	AC-T---CC---T---AC---T---G---A---AAGG---G---GG---A---AG---CCA---CC---C---GTG---TTAAGTT---T---CA---T---TCAG---T---TGTC---A---A	1501
SIVGRI677	GC-C---G---C---TGC---G---G---A---AAGG---GGG---AG---AG---C---CCA---C---C---G---TGCT---T---CT---T---G---TCAG---C---TGT---A---A---A	1308
SIVVER9063	-C-C---A---G---C---G---A---AAA---TG---AG---GA---AG---G---A---CCA---C---C---GTGT---CT---T---CA---C---TCAA---T---TGTGT---A---A	1541
SIVVER155	-C-C---CC---C---C---A---AAA---TG---GG---A---AG---C---CCA---CC---C---CTG---T---T---CA---C---TCAG---T---TGTCT---A---A	1530
SIVVERTYO	GC-C---C---G---C---G---A---AAA---TG---AG---A---AG---C---CA---CC---C---CTG---A---CT---T---CA---T---TCAG---T---TGTGC---A---A	1031
SIVVERAGM3	AC-C---C---G---C---G---A---AAA---G---AG---A---AG---G---CCAG---T---G---G---A---CT---T---CA---C---TCAA---T---TGTCC---A	1042
SIVLHOEST	C---AA---C---AA---G---A---G---C---GA---G---GG---AA---ACAG---G---A---C---A---AA---TTG---TG---T---TGTT---T---TG---A---T---A	1621
SIVSUNA---G---GA---G---GA---GG---AA---ACAG---GG---A---G---T---ATGAC---A---TT---G---CA---C---TGTG---G---TG---CA---A---	1599
SIVMNDGB1	-G-T---A---C---GA---AC---G---A---AAG---GG---AA---G---G---C---C---TC---A---C---GA---A---T---AGT---T---G---T---T---G---TG---A---A---A	992
SIVSAB1C	A---G---AC---T---T---GA---T---G---A---AAG---TG---G---AG---AGC---CC---G---T---T---TT---T---A---C---TCAA---C---TG---A---A---A	1631
SIVSYK173	---G---C---AA---GG---G---CT---GA---CTCC---AAAA---TGATG---TCCA---AG---G---AT---A---CAG---TC---A---TTT---TT---TT---T---C---C---TGG---T---A---A	1445
GAG	R T L K T W V E A V N S K K F D A S I V P L F Q A L T E G F I P Y D L N G M L N A V G	GAG

HIV-1/HIV-2/SIV
Nucleotides

B_FR.HXB2R	GGACATCAAGCAGCCATGCAAAATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAAATGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACA	1495
GAG	<u>G H Q A A M Q M L K E T I N E E A A E W D R V H P V H A G P I A P G Q M R E P R G S D</u>	GAG
A_UG.U455	----C-G---T-----T-----G---C-GT-A-----C-----	950
C_ET.ETH2220	-----C-----G-----C-GT-A-----G-----A-----C-----	869
D_ZR.84ZR085	-----C-----G-----C-GC-A-----G-----G-----T-----	1012
FL_BE.VI850	-----C-----C-----C-T-A-----GCC-----G-----T-----G-----T-----	838
G_SE.SE6165	--G-----T-----C---G---T---T-----G---C-GA---CAA---G---T---C-----A---A-----T-----T-----	898
H_CF.90CF056	-----G-----T---A-----C-G-A-----C-----A-----C---T-----	842
J_SE.SE92809	---C---T-----T---T-----C-G-A---A-----G-----G-----T-----	818
CRF01_AE_TH.CM240	---C-G---A-----A-----C-----G-A-C---A-----C-----G-----G-----	1060
CRF02_AG_NG.IBNG	---C-G---A-----G-----T-----G-A---A-----C-----G-----G-----	1026
CRF03_AB_RU.KAL153	---C-G---T-----G---T-----T-----GT-A---CA-G-----T---C-----G-----	1546
CRF04_cpx_CY.94CY03	---C-G---A-----T-----C-----C-GACA---A-----C-----T-----	861
N_CM.YBF30	---G---A---G---C-G---G---AGT-----A---C-----GACA---C-C---TG---A---AC-AC-C---G---AC---C---T---A-----T-----	1059
O_MVP5180	---C---GG---TT-A---G---G---G---AGTA-----A-----ACT---CCAGCAATG---GT-AC---G---A---G---CA-----	1516
O_CM.ANT70	---G---TT-A---G---C---G---AGTA-----A---T---G---ACT---C---CCA-CG-T---GT-GC---G---A---G---CA-----	1540
HIV2_A_ST	-AC-----T-----AA-C-GG---A---TT---T---A---A---T---CGC-CAA---C---A---CA...-C---CT-AC-G-G-G---C---C---G---TC-----	1256
HIV2_A_BEN	-AT---C-----T-----AA-C-G---TT---T---A---A---C---TCGCA---C---A---CA...-C---CT-AC-G---A---C---C---G---A---GTC-----	1811
HIV2_A_ALI	-AC-----G-----AA-C-GG---TT---A---A---C---GTT-CA---CA---A---CA...-C---CT-AC-G---G---C---T---A---GTC-----	1804
HIV2_B_D205	-A---G-----TA-T-G---A---TA-----A---C---CCAGCAA---C---GTCA-CA...-C---A---GC-GG---A---AC-T---G---C---A---GTCA---T-----	1772
HIV2_B_EHO	-A---G-----TA-T-GG---TA---T-----A---C---CCA-CA---TC---CA...-C---A---GC-G---G---AC-C---G---A---GTCA---	1773
STMM83293	-AG-----G-----TA---T---A---T---A-----G---T---C-TGCA---CAA-CAC-G---T---GC-GC-G---G---AC-T---C---GTCA---T-----	1420
SMM251	-AC-----G---T---G---TA-C-G---T---TT---A---G-----G---T---CTTGCA---C---CAA-CA---T...-A...CA---A---C---T---G---G---GTCA---TCA---T-----	1746
SMM9	-A---G-----TA-T-G---A---TT---A---A---G---T---T---TT-CAA---C---RCA---CA...-T---A---AC-G---A---AC-T---G---R---TCC-----	1243
SIVCPZANT	-AC-----G-----GG---C---AGTA-----A---T---G---GT---A---C---CACT---A---AG-ACAGG---A---AT-A---G---CA-----T-----	907
SIVCPZGAB	-----G-----G---C---G---AGT---T-----C---C---G---C-CT-A---CACT---T---T---G---A---C---A---G---G---T-----T-----	1556
SIVCPZUS	-----G-G---A---G-T---G---AG-----A---T---A---ACA---TC-----C---A---G-----C-----	1543
SIVSTAN1	-AC-----GG---A---G---TA-C---G---AGTA-----T---A---CC-G---T-ACA---C---CCCA-CA---A---GC-AC---G---G---C---A---G---C---A---TTCA-----	1631
SIVGRI677	-AC-----G-GG---AT-A---TC-T---G---AGT---T---A---A---G---C---GACA---CAG-CCA-CA---T---C---GT-AC---G---G---C---A---C---G---CA---GTCA---T-----	1438
SIVVER9063	-AC-----GG---AC-A---AG---A---T-----A---CC-G---GT-ACA---C---CCA-CA---A---T-AC---A---C---T---C---T---TTCA-----	1671
SIVVER155	-AT---G-GG---T---AG-G---A---TA---T-----A---CC-G---GT-ACC---C---CC---CG---C---CT-GC---G-G---A---C---C---G---T---GG---TCA---T-----	1660
SIVVERTYO	-AT---GG---AT-A---AG-G---T---T---A---A---CC-G---GT-ACA---C---C---A---CC---A---CC-AC---G---A---C---C---G---C---TC---C---CTCA---T-----	1161
SIVVERAGM3	-AC-----GG---GC-A---AG---A---T-----A---CC-G---T---CT---C---CCA-CA---A---AT-AC---G---A---AC-C---C---T---A---CTC-----	1172
SIVLHOEST	-AT---G-GC---AA---TCA---AGTG-----G---G---CA---C---GCAA---ACAGCCAG---CAAC---GG---GGGT---A---ACT---T---AT---C---CA---TG-----	1751
SIVSUN	-AG-----G---T---G---A---TCAT---AGTG---AG-----C---A---C---TCA---C---GCAGCCTG-ACAGC---GG---AGGCC---A---GACT---T---AT---T---CA-----	1729
SIVMNDGB1	-AT---G-G---A---G---C---TGTA---T-----C---A-----CT-ACA---TCAA---ACA---CCAG-ACAAC---GG---AGGAT---A---GACC---TTCA---CTC-----T-----	1122
SIVSAB1C	-A---C---G-GG---AC-A---AG-G---TGTT-----A---G---G---T---CCTTAGA---C---GCC---CGCA-CA---CCCA---T---AG---AGT---C---A---G---C---CAA---GTCA-----	1761
SIVSYK173	-AT---G-G---GG-AA-T---C-T-----GG-G-C---CTT-AA---CCAA-CGCA-CAA---GC-C---A---CAGGAT---G---T---CTCC-CCTCA-----	1575
GAG	<u>D H Q G A M Q V I K D I I N E E G A A W D L E H P Q P Q Q P A P Q A G L R D P S A S D</u>	GAG

B_FR.HXB2R	TAGCAGGAACACTACTAGTACCCCTTCAGGAACAAATAGGATGGATGACAAAT.....AATCCACCTATCCCAGTAGGAGAAATTTATAAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAAT	1616
GAG	I_A_G_T_T_S_T_L_Q_E_Q_I_G_W_M_T_N.....N_P_P_I_P_V_G_E_I_Y_K_R_W_I_I_L_G_L_N_K_I_V_R_M	GAG
A_UG.U455	-----C--G--A-----GGC.....-G---C-C---G-----	1071
C_ET.ETH2220	-----A-----C-----GGG.....-C---G-T---G---C-C-----GC-----	990
D_ZR.84ZR085	-T-----GC-----C-----	1133
FL_BE.VI850	---T-----CA-----GGGC.....-C---G---G---C-C---C---A-----	959
G_SE.SE6165	-----T--G---ACT---CGGC.....-C---G---G---G---G---G-----	1019
H_CF.90CF056	-----G---C-----GGC.....-G---G---C-C-----G-----	963
J_SE.SE92809	-----C-----GGC.....-G---G---T---C-----	939
CRF01_AE_TH.CM240	-----A-----C-----C-----G---C---G-----G-----	1181
CRF02_AG_NG.IBNG	-----A-----GC-----G---C---G-----G-----G--	1147
CRF03_AB_RU.KAL153	-----G-G-----A-----GC.....-C---A---G---C-C-----G-----	1667
CRF04_cpx_CY.94CY03	-----A-----A-----GC.....-C---G---G---C-----T---G-----C	982
N_CM.YBF30	-----A---C---GGCA---GG-G-CT---TGC-----T--G-T---G---T---G-----G---G---C-G---T-G---	1180
O_MVP5180	-T--T---A---C---A-AG-A-G---TAT---CT-T-GAGGG.....GC-AACT---C-C---G-A---G-G-A---C---C---G---A---	1640
O_CM.ANT70	-T--T--G--A---C---AG-A-G---TCAC---CT-C-GGCC.....-C-A---C-C---G-A---G-GT-A---C---C---G---A---	1664
HIV2_A_ST	-----G-A-A-C---AG-AG-A-G---CCAG---TTT-GGCCA.....C-AAAT---G-A---AGC-C---G---CCAGA-A---GC-C-G---GTGC-C-G---	1380
HIV2_A_BEN	-----A-A-C---AG-AG-T---G---CCAG---TAT-GGCCA.....C-AAAT---CG-A-G---GA-C-C-C-G---CCAAA-A---GC-GC-A---GTGT-C---A	1935
HIV2_A_ALI	-----A-A-C---AG-AG-A---G---CCAG---TTC-GGCCA.....CGGAAT---G-G---GA-C-C-G---CCAGA-A---GC-C-G---GTGT-C-G---	1928
HIV2_B_D205	-----C-C-C---AG-AG---G---CAG---TAC-GGGCC.....C-AAAT---G---G---A-C---G---TCAAT-A---GC-G---TGT---CC---	1896
HIV2_B_EHO	---G---C-C-C---AG-AG-A---G---CAG---TAC-GACCC.....C-AAAT---G---G---A-C-C---G---G---TCAGT-A---GC-C-G---TGT---CC-G---	1897
STMM83293	---G---G---CT-A-CAG---TCA---TACCGGCAG.....C-AAAT---A-G---A-C-C-C-GG---TCAA-A---GC-C-A---TGT---C---	1544
SMM251	-T-----A---T-AG-AG-T---CCAG---TAC-GACAA.....C-GAAC-C-A---CA-C---C-GG---CCAA---G-GC-A---TGT---C---	1870
SMM9	-T-----C---AS-AG-C---TCA---TAY-GGCAG.....C-AAAC---A---CA-C-C-C-GR-G---CCAAT-A---GC-GC-G---TGC---	1367
SIVCPZANT	-----G-A-A-C---AG-G---G-G-GCA---T---CACCTCAACAG---GG-GGAG---G-C-C---G---C---A---GG-G---C---G-N	1037
SIVCPZGAB	-----G-C-C-C---A-G---T-G-C---CA---GCA.....T-C---G---G---G---G---TT-A-G---C---G-G---T---	1677
SIVCPZUS	-T-----C-C-C---T-G-A---G-G-C---GCA.....-C---C---T---C-A---G---G-GG---A-GC-T---G---G---A---	1664
SIVSTAN1	-----G---A-C---CT-TG---GCT-G---AG---CTTTT---GCT.....AGAG-AGAT---CGT-C-C-G-G---G-T---C-CC-A---TGT---A---	1755
SIVGRI677	-----C---CT-AA---AG---CCTTC---GCC.....AGA-AGAC---G-C-CAA-C-G-A---G-T---TT---C-C-A---GG---GCAG---	1562
SIVVER9063	-----C---AG-G-A---T---A---TTAC-CAGCC.....-C---CAGAG-AGAT---T-G-CT-C---G---T---A-GC-GC-G---GTGT---A---	1795
SIVVER155	-----G-C---AG-G-A---G-GC---AG---CTAT-C-GCT.....-C---AGGG-AGAT---G-CC-C---CG---C---A-G---C-A---TGT---A---	1784
SIVVERTYO	-----G-C-C---CT-AG-A---A---GT---A---CTAT-C-GCT.....-C---GGG-AGAT---T-CC-C---CCGG---T-T-A---C-TC-A---GTGT---C-A---	1285
SIVVERAGM3	-----C-C-C---AG-G-A---GC-G-A---ATAC-CAGCC.....CAGAG-AGAT---G-T-CC-C---G---G---T---G-GC-A---TGT---A---	1296
SIVLHOEST	-G---GGTC---CTCA---AG-AG-A---C---CC---CT---GCAGAC.....A-T---G-AGAT---A---G---G-C-T-A---CTA-GG-A---G---C---	1872
SIVSUN	-----GGTA---ATCA---AG-AG-A---G-G---CAG---CTATGGCCCTTAATGGAGCAA---AGAG-TG---CA---C---G-C---G-C---AGA---CA-GG-A-G-G---T---	1859
SIVMNDGB1	-----TC---AG-GG-A---T-G-C---AT-TGCCA.....C-AAATG-A---AAT---AC-C---G-T---T-A---CA-G---G-T-G---A-G---	1246
SIVSAB1C	-----G-C---C---AA-A---A---G---A---CC-C-GAGCA.....C-GAATG-G---AAT---A-T---G---C---TC-CC-G---TGT---C-A---	1885
SIVSYK173	-----A-A---CT-AA---GCT-G---A---T-T-GGCAG.....-CAAC-AG-A-A---A---G---C-T-A---GC-GC-G---GTGT---GCAGG---	1699
GAG	I_A_G_T_T_S_S_I_A_E_Q_I_E_W_I_T_R_Q.....N_N_P_V_Q_V_G_E_I_Y_R_R_W_I_I_L_G_L_Q_K_C_V_Q_V	GAG

B_FR.HXB2R	GTATAGCCCTACCAGCATTCTGGACATAAGACAAGGACCAAGGACCCTTTAGAGACTATGTAGACCGGTTCTATAAAAATCTAAGAGCCGAGCAAGCTTCACAGGAGGTAAAAAATGGATGACAGAA	1746
GAG	Y S P T S I L D I R Q G P K E P F R D Y V D R F Y K T L R A E Q A S O E V K N W M T E	GAG
A_UG.U455	-----GTT-----T-----G-----A-----C-G-T-----TA-A---T-----C---T-----A-----T-----C-----	1201
C_ET.ETH2220	-----GT-----T-----A-----T-----T-----CT-----T-A-----A-A-T-----C-----	1120
D_ZR.84ZR085	-----GT-----T-----C-----C-----A-----A-----A-----A-----A-----A-----A-----A-----	1263
FL_BE.VI850	-----T---GT-----T-----A-----G-----A-----A-----A-----T-----GTC-----T-----C-----GGG-----C-----	1089
G_SE.SE6165	-----GT-----T-A-----G-----A-----TA-----T-----TG-T-G-----T-----T-----GGC-----C-----	1149
H_CF.90CF056	-----T---GT-----T-----A-----G-----A-----A-----T-T-----T-----T-----CA-----T-G-G-----C-----	1093
J_SE.SE92809	-----GT-----T-----T-----T-----A-----T-----G-----T-----A-----A-----T-----T-----	1069
CRF01_AE_TH.CM240	-----GTT-----T-----G-----A-----C-----TA-----T-----C-----G-A-----A-----C-----	1311
CRF02_AG_NG.IBNG	-----GT-----T-----G-----A-----T-----TA-----T-----T-G-----T-A-----A-G-----C-----	1277
CRF03_AB_RU.KAL153	-----GTT-----T-A-----G-----A-----C-----TA-----T-----T-----T-----A-----T-----G-C-----	1797
CRF04_cpx_CY.94CY03	-----C-TT-----CT-----A-----C-----T-----TA-----T-----TG-----C-----A-A-----A-C-----G-----	1112
N_CM.YBF30	-----T---GT-----A-G-C-A-----A-----C-----A-----C-----A-----G-AA-----A-----G-----	1310
O_MVP5180	---C---T---AGTG-----CT-A-T-T-G-G-----A-----A-----C-----T-----T-C-----AT-----T-----A-T-A-A-----G-----	1770
O_CM.ANT70	---C---AGTG-----CT-A-T-T-AG-G-----A-----A-----T-----C-----AT-----T-----A-T-A-A-----	1794
HIV2_A_ST	---C-A---A---A---C-A---A---G-----G-A---CCAGAGT-----TA-A---C---G-GCT-G-G-A-A---A-AGAT-CA-CA-----CC-----	1510
HIV2_A_BEN	---C-A---A---T-A---CT-A---A---G-----A---G---CCA-AG-----A---C---GCT-G-G-A-A---A-AGAC-CA-CA-----CC-----	2065
HIV2_A_ALI	---C-A---A---A---C-A---A---G-----G-A---CCA-AG-----TA-A---C---GCT---G-G-A-A---A-AGAT-CA-CA-----G-----TC-----	2058
HIV2_B_D205	---C-AT-----A---AT-A---AG-G-----G---CCA-AG-----TA-A---C---GCT---C-G-A-A---A-AGAC-CA-CA-----G-----C-----	2026
HIV2_B_EHO	---C-AT-----T-AT-A---AG-----G---A---CCA-AG-----TA-A---C---GCT---C-G-A---A-AGAC-CA-CA-----C-----	2027
STMM83293	---C-AT---AGTT-A---T-A---G---AG-----G---A---A-G---CCA-AGT-----A---C---G-G-T-G-----A-A---AGAT-CATCA---G-----TCGG-----	1674
SMM251	---A---A---A---A---A---TG-A---G-----A---G---A---CAGAG-----A---C---GCT-----A-A---A-AGATGCA-CA---G-----TC-----	2000
SMM9	---AT---A---A---A---T---TG-G-A---A---GT-A---CA-AG-----TA-A---C---G-G-----A-A---A-AGAC-C-CAS-G-----TC-----	1497
SIVCPZANT	---T---AGT-----A-G-A---A---G-----A---G---C-----T---G-TA-A---AA-T---A-A-G-----CCT-G---GCC-----	1167
SIVCPZGAB	---T---AGTA---T-C-T---C-----G---G-----T---G---A-A---C-----T---G-T---A-G-----T-C-----	1807
SIVCPZUS	---CT-T---CGTTG-T---CT-A---T-C-AG-G---T-A-G-----A---C---AT-----A-A---T-CA---A---CC-----G-----	1794
SIVSTAN1	---C-AT-----T-TCAG-C-----C-----GG-C---G-T---AG-----C---A-A-T---C---G-C-T---A-A---A-AC-C-A---T-----	1885
SIVGRI677	---C-AT---CCAA-AGG-C-A---TC---G---T-A---CCAG-----A-A---G-C-G---A-A---AC-----T-T-----C-----	1692
SIVVER9063	---C-A---AGT-TCAG-AT-A---C---G---T-A---T-C-AG-T---TA-----C---G-AA-C---A-A-G---C---GGA---A---C-A-----	1925
SIVVER155	---C-AT---AGTGTCTG---T-A-T-C---G---C-A---A---C-A---T-----G-AA-----A-A---GGA---A---C---C-A-----	1914
SIVVERTYO	---C-A---AGTATCAG-C-A---T-G-G---T-A-G---C-AG---T---G---A-A---T-C---GG-AA-T---A-A---C---GG---A---G---C-A-----	1415
SIVVERAGM3	---C-A---AGTGTCTG-CT-A---G-G-C-A---G-A---C-A---C---TA-----C---G-AA-----T---G---C---GGA---A---C-G-----	1426
SIVLHOEST	CC---CAA---AGTA---TG-CA-----C-A---C-A---G---T---A---A---C---TA-----T---G-A-G---A-AGG---GC---T---T---A---GG-G-----AG-G-----	2002
SIVSUN	CC---CAGA-----TCAG-AG-A-A-T---G-----A---G---T---C---AC---TA-----TC---GG---T---A-ATC---GC---G---A---G-G-----A-G-----	1989
SIVMNDGB1	TC---T-T---A-TA---T---AACA---TG-----G---G---T---A---C---TGTAAG---A-A---GGA---A---G-TG-----CAGC-G-----	1376
SIVSAB1C	---C-A---AGTA-A---T-A---T-A---G---T-A---T---C-AG---T---TA-A---C---GG---G---A---GA-AGAC-CT-CT---G---C-----C-G-----	2015
SIVSYK173	C---C-A---GTTTCT---CT-A---A---G---C---A---ATT---C-A-----A-A---CC-TTGCT-----A---G---GGAC-CATCA---C---GGG---C-A---TC-----	1829
GAG	Y N P V S I L D I K Q G P K E I F K D Y V D R F Y H C L R A E Q A D P S V K G W L T Q	GAG

B_FR.HXB2R	ACCTTGTGGTCCAAAATGCGAACCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGCGGCTACACTAGAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGACCCGGCCATAAGGCAAGAGTTT	1876
GAG	T L L V Q N A N P D C K T I L K A L G P A A T L E E M M T A C Q G V G G P G H K A R V	GAG
A_UG.U455	-----C-----T-----C-----T-C-----G-----A-G---G-----T-----C-----G-----A---G---	1331
C_ET.ETH2220	-----C-----A---C-----G-----A-G---G-----T-T-----G-----T-C---C-A-----G---	1250
D_ZR.84ZR085	-----A-----C-----C-----G-----A-G---G-----T-----G-G-G---A---A-----A---	1393
FL_BE.VI850	--A-----C-----A---G-----G-----G-----TA-----	1219
G_SE.SE6165	-----C-----C---C---G-----A---A-GA---T-----G-----A---A-----A---	1279
H_CF.90CF056	-----A-T-----C-----A---G-----A---A-G---T-A-----G-----TA-T---A-----	1223
J_SE.SE92809	-----C-----A---T-----C-----C-----A---T---GA-----G-----T---T-----G---	1199
CRF01_AE_TH.CM240	-----C-A-----C-----T-----C-----T-C-----G-----A---A-GA---T-----C-----G-----TA---A---G---	1441
CRF02_AG_NG.IBNG	-----C-----C-----T-----C-----G-----A---A-G---T-----G-----T-----A---G---	1407
CRF03_AB_RU.KAL153	---C---C-----A---T-----A---C-----G-----A---T---G-----T-----T---C-----G-----A---G---	1927
CRF04_cpx_CY.94CY03	---C---C-----T-----C-----T-C---C-----A---A---G---T-----G-----A---A-----A---	1242
N_CM.YBF30	--AC-C-A-A-----A-----ACAGC-CC-----A-G---GA---CT-----G-----G-C---C-----G-G---A-CA-----GC	1440
O_MVP5180	---C-T-T-G---T-A-----ACAA---C-G-----A---AA---TT-----GT---C---A-----G-AACT---C---A-A-AC	1900
O_CM.ANT70	--TC-C-T-T-----C-T-----C---ACAG---G---T---A-G---GA---CT-----G-----GT---C-----AACT---C---GC	1924
HIV2_A_ST	--AC-C-A-G-G---C-----C-----TTAG-AC---G-C-A---GAT-AATC---CT-----C-A-C---C-----G---T---A---G-A-C---T-AA	1640
HIV2_A_BEN	--GC-C-AA-A-G---C-----C---C---TTAG-AC---G-C---GATGAATC-C-C-----G-C-A-C---C---G---C---A---G-A-C---GC-AA	2195
HIV2_A_ALI	--AC-C---A-G---C-----C---C---ATTAG-GC-G---G---A---GATGAATC---CT-----G-C-A-C---C---G---A---G-A-C---T-AA	2188
HIV2_B_D205	--AC-C-A-T-G---T-----C---TTAG-GC-T---G-GC---ATGAATC-C-CT-----G-C-A-G-C---C-A-GA---C-A-G-G---GC-AA	2156
HIV2_B_EHO	--AC-C-A-T-G---T-----C---ATTAG-GC-T---G-GC---ATGAACC-C-T-----C-A---C---G---A---C-A-G-G---GC-AA	2157
STMM83293	--AC-C-AA-T-----T-T-----C---ATTGG-A---G---GTC---CATGAATC-C-C-----G-T-A---A-G-T-G---A-A-G-A-T---C-GA	1804
SMM251	--AC-C-A-T-----T-----C---CTAG-GC-G---G-GGC---TGTGAATC-C-C-----C-G---T---A---G---A-A-G---T---T-AA	2130
SMM9	--AC-C-A-T-----C---G---C---ATTGG-GC-C---G-GTC---YAT-AATC-C-TT-----T---C-----R-G-A-A-A-AS-T---C-AA	1627
SIVCPZANT	-----AA-----C-T-----C---ACAC---CC-G-G-T---A---GA---CT-CT-----T-A---T---A-----C-A-C-----G-	1297
SIVCPZGAB	--TC-C-A-G-G---A-----CAA---C-----T-G---C-C---G---T-A---T---A-----ATCT-----G-C-	1937
SIVCPZUS	---C-T---A---C-T-T-----A-----G---C-----T-GA-C-C---G-----G-----A---G-----CT-----G-C-	1924
SIVSTAN1	--AC-C-CA---G---C-----C---TTGG---G---G-C-T---AT-CATC-A-T-----G-C---T-----A-A---A---AGC-CA	2015
SIVGRI677	--T---C-TA---G---C-T-G---ATTG---C-G---G---ATGAATC-A-CT-G-G---C-A-T---T-C-----G-ACAA---T-AGC-AA	1822
SIVVER9063	T-A-AC-CA-A-----T-----C---AGTC---CC-G---G---A---ATGCACC---CT-----C---T---C---A---G-C-AA-T-C-A---A	2055
SIVVER155	T-T---C-CA-T---G---C-----C---AGTA---G---G-GCC-A---GATGCACC-C-T-T-----C---C---A-G-G---C-AA-TT-C-A---CA	2044
SIVVERTYO	T-A-AC-CA-T-----T-T-----GTC---CC-G---G-GCC-A---ATGCACC-C-C-T-----T-A-G-T---G---C-AA-TT-C-A---AA	1545
SIVVERAGM3	T-A-AC-CA-T---G---T-T-----C---AGTC---CC---G-GCC---ATGCATC-C-T-----T-A-T---C---A---G---AA-TT-C-A---GA	1556
SIVLHOEST	-AAA-C-T-G-G---A---T---C---G-TTAG-A-----C-A---GA---G---AT-T-----G-----A---T-----T-CT---A-G-A-A-A-	2132
SIVSUN	-AAA-C-A-G---A-----ATTGG-AA---G---C---GA-A-TC-AT-T-----G-G---G---C---A-G---T---G-G---GA---A-G-A-GA-CC	2119
SIVMNDGB1	CATC-C-TA-AG---A-----A---C---CAG---G-G-AGC-A---GAA---GA---A-TT---G---T---GA-----T-G-ACAA---A-C---T-AA	1506
SIVSAB1C	T-TC---A---G---C-----AG-GC-C---GG-A---CATGAATC-A-C-----G-C-T-T---C-A---CA-T---G-ACAA---C-TC-AA	2145
SIVSYK173	--AC---A-T-----C-----A---C-GACAG---CC-C---G---A---TCAA...C---CT-G-G---C-CA---T---A---C-----TTA-C---A-C-CA	1956
GAG	T L L I Q N A N P E C R Q I L K A M V K . P T L E E M L Q A C Q G V G G P L H K A K L	GAG

see Kaye, J Virol 72(7):5877-5885 (1998) Gag p2 role in RNA encapsidation

	Gag p24 \ / Gag p2	Gag p2 \ / Pol signal start peptide	Gag p2 \ / Gag p7 NC	
B_FR.HXB2R	TGGCTGAAGCAATGAGCCAAGTAACAAATTCAGCTACC.....	ATAATGATGCAGAGAGGCAATTTTGGAAACCAA.....		1947
GAG	L_A_E_A_M_S_Q_V_T_N_S_A_T.....	I_M_M_Q_R_G_N_F_R_N_Q.....		GAG
A_UG.U455	-----G-----T-----CA-C-GA--AGC.....	-----GG--CG.....		1399
C_ET.ETH2220	-----G-----AC--A--A-C.....	-----A-A-----A-GG--CT.....		1318
D_ZR.84ZR085	-----G-----C-----G-G-CT--G-TGCC.....	-----A-A-A-----A-GG--AC.....		1467
FL_BE.VI850	-----C-G-----T-----C--ATTGAG-C.....	-----A-A-T-----A-GG.....		1284
G_SE.SE6165	-A-----G-----G-CCT--GG-G--AG-AGCC.....	-----A-----A-GG--C.....		1353
H_CF.90CF056	-----G-----A--AA--AGCC.....	-----A-----C-----A-GG.....		1297
J_SE.SE92809	-----G-----C-----A-CAAC.....	-----A-----T--C-----G--T.....		1267
CRF01_AE_TH.CM240	-----C-C-CA-C--G--A--.....	-----C-A-GG--G.....		1509
CRF02_AG_NG.IBNG	-----G-----T-----CA-C-GG-CAAC.....	-----G-----GG--G.....		1475
CRF03_AB_RU.KAL153	-----C-G-----T-----CA--G--AAC.....	-----A-A-T-----GG--C.....		1995
CRF04_cpx_CY.94CY03	-A-----G-----G-C-T--G--AG-AGCC.....	-----A-A-----A-GG.....		1316
N_CM.YBF30	-A--A--G--T--TCA--G--GCAGC-GC--A-A--TAGT.....G-CT-TGCA--A--G--A--C--AAGG-AT.....		1514
O_MVP5180	-A--A--G--T--TCA--G--GCAGC-GC--A-A--TAGT.....G-CT-TGCA--A--G--A--C--AAGG-AT.....		1986
O_CM.ANT70	-A--A--G--T--TCA--G--GCAGC-GC--A-A--TAGT.....G-CT-TGCA--A--G--A--C--AAGG-AT.....		2010
HIV2_A_ST	-----A-----CT-A-AGG-G-CC-TGGCAC--CC--ATC.....CCATTTCG-GCAGCC--ACAG.....		1702
HIV2_A_BEN	-----A-----CC-A-AAG-G-CT-TGGGAC--AGCC-TATC.....CCATTTCG-GCAGCC--ACA.....		2257
HIV2_A_ALI	-----A-----CT-A-AGG-G-CC-TG-CAC--C-TATC.....CCATTTCGCGGAGCC--ACA.....		2250
HIV2_B_D205	-----C-----CT-A-AAG-G-CCCT--CAC-T--AC--ATA.....CCGTTTGCTGCCG-T--ACA-AAAGCAGGG-A.....		2230
HIV2_B_EHO	-----G--TT-A-AAG--CCCTG-CAC-TT-C--AAT.....CCGTTTGCCGCGCT--ACC-A-AGCAGGG.....		2228
STMM83293	-----A-----CT--AAG-G--CTTCC-AC--ACC-ACTG.....CCCTTCGC-GCAGCC--ACAGCAGGGA.....		1872
SMM251	-----A-----CC--AAG-G-CCCTGCAC--TGC-AATC.....CCCTTCGC-GCAGCC--AGA-GGGACCA.....		2201
SMM9	-----A--G--T--AAG-T-C-CTG-CGCA--GGC-ACTC.....CCATTTCGRCAG-C--ACAGAAAGGACAA.....		1698
SIVCPZANT	-----A-----T--GCTTCT-CT-AT--G--CAGGGAACC.....GCAG-CT-TC-----GGA-AT-GAGG.....		1374
SIVCPZGAB	-A-----C--TCAATG--TCAG--CA--GG-GAGCA.....GATG-TT-CT-C--AG--AC-AGG-GC-GG--C.....		2014
SIVCPZUS	-A-----C--T--A-G-A--CC-CT-A-G-.....G--T-TC-A--A--A--CGC-G--G-C.....		1995
SIVSTAN1	---TA-----CAA--A-GCA-GGAGTT.....	---AC--G--AG--CACCCGAG-AGGTAGGGCCAGAGGC.....		2089
SIVGRI677	---TA---ATG---TA-T-G-CAG.....	---AT---G-C---AGTG---ACC-CAG-AA---GGGCCCC.....		1884
SIVVER9063	---A---ATG---AAC-TGC-GA-TCAA.....	---AC---G---CA---GGCAGCAG-AGGAGT-AGGAGACAG.....		2129
SIVVER155	---A---ATG---GAACCTGC-GAGTCAG.....	---AC---G---A---CAG--AGG-GGA--GGAAG-CCA.....		2112
SIVVERTYO	---A---ATG---GACC-TGC-AAATCAA.....	---AC---G---CAG--AGG-CCA-AA-GA--.....		1610
SIVVERAGM3	---A---ATG---AAT-TGC-AAGCCAG.....	---AC-----ACAG--GG-CAG--AGGAAG-CCA.....		1624
SIVLHOEST	---A-----GCTTC--CC-T-C-GCA-CAAGTAAGA.....CAA-AC--G-----T--C--CCGCTT-GAA-TGCGCAAGGCAGGTTTGTGAGAACAGGAGGAGGAGGTCCT.....		2245
SIVSUN	---A-----GC-TCT-CTTTT-GGCAGCAAGTAGGA.....AGACAAGCT--G---CA-AA-TTACCACCC-GAA-CAGCCAGGGGAGGTTTGTAGAAATAGGAGGAGGAGGACCT.....		2238
SIVMNDGB1	---A---ATG---AAC--GGT-GGACA-T-A.....CAA-ATT-TG---CAGA-AGGGCC-CAA-GAGG-CCAGTTAGACAACCTACTTGA.....		1598
SIVSAB1C	---G---T---CAGCG-CCTTTC-GCA-CAA--AGTT.....GGGAAC--CT-TG---CA--GGCAAGACCC-GGGTCCCTTAGGGGGAAGAGGTAGACCTCTG.....		2252
SIVSYK173	---A--G--C--GTAATG-CCCAGC-G--TA.....	---AT--G--A--AG-GCCATCAAAG--GATC-ATGATTAATGCTACAAATTGGGCACATGCAAAA.....		2068
GAG	M_A_E_A_M_V_M_A_Q_Q_S_V.....	N_M_V_Q_G_P_S_K_G_R_S_M_I_K_C_Y_N_C_G_Q_I_G_H_M_Q_K.....		GAG

B_FR.HXB2RAGAAAGATTGTTAAGTGTTC AATTGGCAAAGAAGGGCACACAGCCAGAAATTGCAGGGCCCTAGGAAAAAG.....GGCTGTTGGAAATGTGGAAGGAAGGCACCAAA	2055
GAGR K I V K C F N C G K E G H T A R N C R A P R K K.....G C W K C G K E G H Q	GAG
A_UG.U455-GA-GAA-----C-----A--CT---A---T-----C-G-A-----	1504
C_ET.ETH2220-A-GAGCAA--A--C-----G-----CT-----A-----	1426
D_ZR.84ZR085-A-----C-----T-A---T-----	1575
FL_BE.VI850-GAG--A---T-----A---T-----A-----GA-----	1392
G_SE.SE6165-GA-C-A-C-A-----G--A--TCT-----T-----T--	1461
H_CF.90CF056-AT-----A-C---C-----G--A--T-----GA-----T-G	1405
J_SE.SE92809-A-GA-----C-----C-A---T---A-A---C-----A-----	1375
CRF01_AE_TH.CM240-A-GAA-----C---C---T-G---A---CT-----A---C-----T-----C-G-----T--	1614
CRF02_AG_NG.IBNG-GA-CAA-A-----C-----A---CT-----A---G-----G--A-----	1580
CRF03_AB_RU.KAL153-A-GAA-----C-----A---CT-----C-----C-T--	2100
CRF04_cpx_CY.94CY03-GA-C-A-----C-----G--A--TCT-----C---G-----G---T--	1424
N_CM.YBF30-G-ACCCA--A-----G--C---TTTG--A---C-T-A-----A-G-GGA.....G--C---GC-A-----T--	1622
O_MVP5180-A-GG-CCCA-A-A-C-----A---G--A--T-T--A-A---C-TC-A-A---A---G-----T-C-----C---T--	2094
O_CM.ANT70-A-GGA-C-A-A-A-----C---A---G--A--T-T--A---TC-A-A-----T-C-----C---T--	2118
HIV2_A_ST-G-CAA-----C-GG---C-A-G-----T-G-A---C-A---C-A-A---A-G-C-A-----C---C---ATC	1810
HIV2_A_BEN-GCAA--G--A--GG-C---A-G-G--A--T-G-A---C-G---C-A-A---A-G-C-----C---G---C---CC---ATC	2365
HIV2_A_ALI-G-CAA-----C-GG---A-G-----T-G-A---C-A---C-A-A-C-A-C-----C---C---CC---T-TT	2358
HIV2_B_D205-GG-CA--G-CA-C-GG-C---C-G-A-----GC-A-----A-G-C-----A-C---AAC---ATC	2338
HIV2_B_EHO-AG-G-CA--G-CA-C-GG-C---C---G-CG--A--T-----GC-G---A-----A-GGC-----A-C---GC---C---ATC	2336
STMM83293C---G-CA--A-A-C-GG---G---C---T---A-A-C-G-T-AA-G---G-A-G-C-A-----T---C---AC-G-C-T-G	1980
SMM251-CCAA-----GG---G--G--A---T-T--A--GC-A---A---A-G-C-----A-C---AATG-AC-TGTT	2309
SMM9R---AA--W---RG-C--R-A-G---A---T-T--A---C-A-T---A---A-G-C-R-----C---G---A-C---TGTT	1806
SIVCPZANTGGA-A--GACC-C-C-A---T-C-C-T---G--C-T--T--A---A---A-GG-A-----C---G---C---G	1485
SIVCPZGAB-A-GA-AAA-A-A-C-T---G---T---CT---T---T-AA-A-A-A-G-A-----CGG---GC-A--G-----	2122
SIVCPZUSG-G-GA-AAA-C-A---T-----G--A--TCT-----T-A-A-G--A-----G-C---G---GC-----T--	2103
SIVSTAN1-GGCCACCCCGC-C---A---AC-ATT---T-T-CAA-A-G---TCCAAGAG-AG-ACC---C---AAA-CCTA-----G---CC---TT-	2197
SIVGRI677C--GG-CCGC-A-A-C-T---TTT--A--T-TGCAA--GG-A---A--A--AC-G-TC.....AAA-C-TT--G--C--AATT-C-TATG	1992
SIVVER9063-CCCCAC-A-A---A---C-A---TTC-A---TGCAA--GC-A---CA-AA-A-A-G-TA.....AAA-CCTT-----G---CT---TG	2237
SIVVER155-CCCCCGCCA--A---C---A---TTT-C---TGCAAG--GC-G--TCT-AG-A-A---TA.....AAA-CCTT-----CC--G---TT-	2220
SIVVERTYO-CCCCAC-A-GA---AT-----A---TTT-C--T-TGCAA---C-A--TCC-AA-A---CA.....AAA-CTA--G---ATTG-----T-	1718
SIVVERAGM3-CCCCA-A---A---A---TTT-C--T-TGCAA---C-A--CCT-AA-A-A--G-T-----A-A-C-T-----G--ACC--G--TTT-	1732
SIVLHOEST-G-ACCC-C-CT-C---A---GCC--T-----A---TG---ACAG-C-AC-GG-A.....T-----C---GTA--AG--TAG-	2353
SIVSUN-G-CCCA-G-CT-T---T---CC--A--TCT--A-ATC-G---A-AG-A-A---GGACCACCCGGCTC---C-----ATG--TA--	2355
SIVMNDGB1-G-AACC-TC-C---C---AAT-----TGT--A--GTTT-T-A---A-GG-A-----G-C---T-----GCAATG-AT--T-G	1703
SIVSAB1C-ACCCA-A-A---A---A---GCCT-T---CTG-T---TT-T-A---A---G-C-A-----C-----C-GCCC-AC-T-G	2360
SIVSYK173	AGACTGCAAAAAGCC-CTT-AA-C---A-C-T---A---AC--A---CT--A--GGCC-T---CAA-C-A-G-----	2153
GAGD C K K P L K A K C F N C G K T G H L A R A C R Q P K R.....	GAG

stem-loop after TTTTT slip potentiates slippage
Gag-Pol -1 ribosomal slip site
Gag p7 end \ / Gag p1 start

Gag p1 end \ / Gag p6 start

B_FR.HXB2R	ATGAAAGATTGTACT.....GAGAGACAGGCTAA.TTTTTTAGGGAAGAT..CTGGCC.....TTCTTACAGGGGAAGCCAGGGAATTTCTTCAGAGCAGACCAGA.....	2150
GAG	M_K_D_C_T.....E_R_Q_A_\$ _F_L_G_K_#_#_W_#.....#_S_Y_K_G_R_P_G_N_F_L_Q_S_R_P_#.....	GAG
POL	F_F_R_E_D..L_A.....F_L_Q_G_K_A_R_E_F_S_S_E_Q_T_R.....	POL
A_UG.U455	-----C-C-----.....-A--..T-----A-----G-----C-----.....	1599
C_ET.ETH2220	-----C..C.....GAC...T.....A.....C.....C.....	1521
D_ZR.84ZR085	-----C..C.....A-----A-----T-----C-----G-----C-----.....	1670
FI_BE.VI850	-----C..C.....A-----A-----T-----A-----G-----C-----A-----C-----.....	1487
G_SE.SE6165	-----C-C-G.....A-----A-----T-----A-----G-----C-----A-----GA-----.....	1556
H_CF.90CF056	-----C-C-A.....A-----A-----T-----A-----G-----A-----A-----C-----G-----.....	1500
J_SE.SE92809	-----C..C.....A-----A-----T-----A-----G-----A-----G-----C-----C-----.....	1470
CRF01_AE_TH.CM240	-----C..C.....A-----A-----T-----A-----A-----G-----G-----C-----C-----.....	1709
CRF02_AG_NG.IBNG	-----C.....A-----A-----T-----A-----G-----G-----A-----C-----G-----.....	1675
CRF03_AB_RU.KAL153	-----C..C.....GA...T.....AG--A-G-----A-----C-----.....	2195
CRF04_cpX_CY.94CY03	-----C..C.....GA...G.....AG--A-G-----A-----C-----A-----G-----.....	1519
N_CM.YBF30	-----AAAATGAA.....-GA-----G.....T-TCC.....-T--A-G-A--A--C--C-CC--CA-C-A--AGGAAAGA..	1729
O_MVP5180	-----C-AAAATGGA.....-A-----TA.....-C-GGGGGGCACG-----C--A-G-G--AACA-GTGC.....	2192
O_CM.ANT70	-----C-GAAAT.....-GA-A--A-----C-ATA.....-C-GGGGGGCACG-----C--A-G-G--ACC-G--C.....	2216
HIV2_A_ST	---GC-A-A-CC-A.....A-----GGG.....TT-GGCCCA--GG.....A--AAGCCC-G...C--C-C-GT-GC-CA-ATCCCGCA..G	1903
HIV2_A_BEN	---GC-A-C-CC-G.....A-----AGG.....TT-GGCCAC--GG.....A--AAGCCT-G...C-C-C-CCGT-C-CA-G-CCCTCA..G	2458
HIV2_A_ALI	---GC-A-C-CC-A.....A-----GG.....CTTGGCCCA--GG.....A--AAGCCT-G...C-C-C-CCGT-C-C-GTCCCGCA..G	2451
HIV2_B_D205	---TC-A-A-CC-A.....A-----GGG.....TTAGGACC--GG.....A--AAGCCT-G...C-C-C-CCAT-C-CA-GTGCCTCA..G	2431
HIV2_B_EHO	---TC-A-A-CC-A.....A-----GGG.....TTGGCCCA--GG.....A--AAGCCT-G...C-C-C-CCGT-C..CAGG--CCCCA..G	2426
STMM83293	---GCCA-A-CC-A.....A-----TGGG.....CTTTGGCCCA--GG.....A--AAGCCC-A...C-C-C-CCAT-GC-CA-AT-CCTCA..G	2073
SMM251	---GCCA-A-CC-A.....C-----GGG.....CCTTGGTCCA--GG.....A--AAGCCC-G...C--C-CCAT-GCTCA-GTGC-TCA..G	2402
SMM9	---GCTA-A-CC-C.....A-----GGG.....CTTGGCCCA--GG.....A--AAGCCC-G...C--C-CCAT-GC-CAGATGCCTCA..G	1899
SIVCPZANT	C-T--A-C--C-AGCAACAAATACA-GA-A-GTA...--ACCGAC-CCCA.....G-GG-GGGG-T-C-A-----C--G-G--AGGAGGA--T.....	1592
SIVCPZGAB	-----C-C-AGGA.....TG-----AGG...T----TTCCCGCAGCGGGAGCCAGGCAACTTTGTGCA--CAGAACAG-ACCA-CGG-CCC.....	2229
SIVCPZUS	-----C-C-AGCAGGA.....A-C-----ACA...T--T-CCCATCGTGGAGCGGGGATCAAAG--A-----C-C--GG-A-A--GAAG.....	2216
SIVSTAN1	GCA-----CCGCAGT.....G-----C-----GA--GCCAACCTG.....GGGA-CAAAGCC--G...C-----T-AG--CAAG-CGG--CCGT...	2292
SIVGRI677	GCA-----C-AGAAT.....G-----A-----T-TGGCCAT--GG.....AGGAGC-AA-CCAAG...A-----G-G--ATA--GG--CACAGTT	2091
SIVVER9063	GCA-----C-C-GG.....G-----TG-----T-TGGACGG--A-----GGGG-CAAA-CC-AG...A-C-----CGTGC--CT-A.....T	2324
SIVVER155	GCA-----C-C-GG.....G-----TG-----T-TGGCCGG--AT.....GGGG-CAAA-CCAAG...A-----CCGCAGC--CT-T.....T	2307
SIVVERTYO	GCA-----C-C-GG.....G-----TG-----T-TGGACGG--AT.....GGGGGCAAA-CC-AG...A-----CCGCCGCT--CT-T.....T	1805
SIVVERAGM3	GCA-----C-GA.....G-----TA-----T-TGGCCGG--AT.....GGGAGC-AA-CCAG...A-----CCGCCGCT--CT-T.....T	1819
SIVLHOEST	T-TGC-C-G-CC-A.....AAACCAA-AG-A-TG-----AT-TGGTCC--AG.....GCAA-CCT---C---A-C--TACTGG--GG--CTGC..A	2455
SIVSUN	CAAGCTC-A-CC-ACAA.....GA-A-----TG-----T-TGGCCC--AA.....G--GCC-----C---CAGTC-T.....	2441
SIVMNDGB1	--AAGTCC-A-CC--AAGCCAGCTCAGC--CAGAG-T-----T-TGGCCCT--GG.....TCCTCC-AA-G-----A--CGGCACAAGAGGTGAC.....	1807
SIVSAB1C	-----CAA.....A-----T-----TTTGGCCC--GG.....GA--G-AA-CA-----C-C-T--C-TCCATCAG.....	2449
SIVSYK173	...--TC-AG-CC-----CCA-TTGC-----AGGA...--GGGT.....C-G-A-GCCC--CC-C-C-CAGT--G--CGA.....	2245
	frameshift	
GAG	.._N_Q_G_P_..._P_V_A_Q_A_#_F_L_G_K_G_.._W_G_V_..._#_S_R_R_P_P_A_N_F_P_V_R_S_#_..._	GAG
POL	.._G_G_#_..._L_Q_E_A_P_S_Q_L_P_S_E_E_R_..._	POL

B_FR.HXB2RGCCAACAGCCCCACCAGAAGAGAGCTTCAGGTCCTGGGTAGAGACAAC.....	2198
GAG# P T A P P E E S F R S G V E T #.....	GAG
POLA N S P T R R E L Q V W G R D N.....	POL
A_UG.U455-C--A-T--TG--ATG---A-A--T.....	1644
C_ET.ETH2220-GAGTCT-AGACCAGACCCAAC--CCC--C--ACCAGA.....	1575
D_ZR.84ZR085-C---AGCT-CGG-T-TG---AGGAGATA-C-.....	1718
FI_BE.VI850-C-----G-----TCA-A-AG---T.....	1535
G_SE.SE6165A-----T-C--A--C-G--TC--A-AG---T-G-.....	1604
H_CF.90CF056A-----G-----TC--A-AG---TG.....	1548
J_SE.SE92809-C-----C-G-CTC--A-AG---T.....	1515
CRF01_AE_TH.CM240-C--A-A-GGG--ATG---A---T.....	1757
CRF02_AG_NG.IBNGA-----T-----C-----TG--ATG---A---T.....	1720
CRF03_AB_RU.KAL153-T-----C--A-A--TG--ATG---A---T.....	2243
CRF04_cpx_CY.94CY03-G-C-CG--AT--AGA-AGGAAA-AG-----	1567
N_CM.YBF30-C-----G--CT-----T-ATG--T-CA--AG-----	1771
O_MVP5180C-----T-----	2207
O_CM.ANT70C-----T-----	2231
HIV2_A_ST	GGGCTGACA-----A-C-GAT--CCCAG-AGA-GACCTAC-----AGTACATGCAGCAAGG.....	2006
HIV2_A_BEN	GGGCTGATA-----A-T-G-C---TCCAGCAGC-GAACT-T-G---G-TATATGCAGCAAGG.....	2561
HIV2_A_ALI	GGATTAACA-----A-C---C---CCAGCAGC-GACCT-C-G---CAGTATATGCAGCAGGG.....	2554
HIV2_B_D205	GGAGTGACA--T-T-A-C-GATGA-CCCAGCAGA-GGCAT-AC-CCTCGGGGGCGCACCATC.....	2534
HIV2_B_EHO	GGGATAGT---T-T-G-C-GATGA-CCCAGCATTCCGGCAT-AC-CCTCAGGGTGCAGTCCATC.....	2529
STMM83293	GGGCTGAC---T---T---CCCAGCTGC-GA-CT-C-GAG.....	2123
SMM251	GGGCTGAC---T-T-C---G---CCCAGCTGT-GA-CT-C-A-.....	2452
SMM9	GGGCTGACA--T-T-Y---G---TCCAGCTGT-GA-CTAC-GA-.....	1949
SIVCPZANTAGTGGAGC-AA-G-TCC-CC.....	1613
SIVCPZGABA-G-T-AGAGTTAT-GGT-CCA.....	2253
SIVCPZUS-C---T---AT---GAT---G---A-CAA-A.....	2255
SIVSTAN1T-G-G-T-C-GATGCCTGCGCA-G-A-TCCCCAC--GTT-GC-AGCTGCGGGAGC.....	2387
SIVGRI677	GGTCTGGAA-----C---ATG-A-CAGC.....	2161
SIVVER9063	GGAGTGGAA--G-GC-G-C--TCC-CC--CACC.....	2394
SIVVER155	GGGGCGGAA--GT-G-C--TCC-CC--A-AA-TCTA-ACC.....	2386
SIVVERTYO	GGAGCGGAA--G-GT-G-T--TCC-CC--GG--CCA-CCC.....	1884
SIVVERAGM3	GGGGTGGAA--T-G-C--TCC-CC--TCC.....	1889
SIVLHOEST	GGGAGAATT--T---T-C---ATG---TGCTCCAA-AAAA-C---G-G-ACTGGA.....	2517
SIVSUNTT-T-T-G-TCT---GATC-G-CA-TG---AACAGA-TG--.....	2518
SIVMNDGB1T-----T-----TT---GAGAAACCTCTGCA-AA-ACTCTG-GCACTTATCAGAAATTAGGGAGAGGGCTCAGGCAGAAGAT.....	1894
SIVSAB1CA-----T-C---ATG--AC-GGATTACAGCA--CCG--GA--ATTGGTATGCAGACAGACCTCCGACCAGAGGACCAGGCCAGACGATCCAGCAACAGCCCTGTTAAAGCAGTATG.....	2571
SIVSYK173-T-T-T-C-TCTG--GA-A--GAAGA---CC-TG-CTC--ATGGTCAGCACAGATGAGCCAAACAAGCACAAGCGAA.....	2349
GAG# P S A P P L E D I E D G P W L T W S A Q M S Q Q A Q A #.....	GAG
POLA I C S P S G G H R R W A M A H M V S T D E P T S T S E.....	POL

	Pol TF \/ Pol p10 protease	Gag p6 end <-	
B_FR.HXB2RAACTCCCCTCAGAAGCAGGAGCCGATAGACAAGGAACT...GTATCCTTTAACTTCCTCAGGTCACCTCTTTGGCAACGACCCCTCGTCAACAATAAGATAGGGGGG		2303
POLN S P S E A G A D R O G T . . . V S F N F P Q V T L W Q R P L V T I K I G G		POL
GAG# T P P Q K Q E P I D K E # # Y P L T S L R S L F G N D P S S Q \$		GAG
A_UG.U455G-CT----GC--A-----T-A----G----AGACTCC...---GT-----AA-----T-----G-----A---A		1749
C_ET.ETH2220GAGCTTCAGGTTTCGAGGAAG-AACACTTCTCC-AA-CAG--GCT-A--GACAG-G-AG-C---AA-----A--T--T-----A-----A		1695
D_ZR.84ZR085C-CT--AGA-AC--G-AC--AAAGAC-AG-CA-GGAAC--C---G-----AA-----G-T-----A-----A		1825
FI_BE.VI850CC-CT-T--GA--C--G--AAAGAC-GGG-ACTGTACCC...---C---G-----AA-----T-A-----A-----A		1640
G_SE.SE6165CC-CT--GA--C--G--AT-AA-GA-A-GG--CT-TA.....TC-C-----AA-----G--A-----G--A-----A		1700
H_CF.90CF056CC-CT-T--GA--C--G--T--AG-----C.....C---G-----A-----G-----T--T--G--A-----A--A		1650
J_SE.SE92809CC-CT--GA-AC--G--CC-AA-GACA-GG-ACTGTA.....C-----AA-----G-----T-----GA-----A		1617
CRF01_AE_TH.CM240GGGGGAAGAGAT--CT--TTA-C-----A--A-----ATCCTCC-----GT-----AA-----T-----G-----A-----A--A		1877
CRF02_AG_NG.IBNGC-----T--CAC-----A--GG-----GG--ATATCC-----AA-----T-A--T--G--GA-----A--A		1828
CRF03_AB_RU.KAL153CC-CT--TGA-AC--G-AC--AA-GACAGGG-AC-G-A...TCC---C-T-----AA-----G-----G-----GA-----A--A		2348
CRF04_cpx_CY.94CY03CT--T--T-----A--GG-----A-----AA-----G-----T-----AC-----A		1669
N_CM.YBF30GAAGAGCACACAGGGGAAGGAG-TG-AGGAGA-CC-G--AG-A-AGAGA--TCTCTGTACCCACC...---A-----A-----G-A-----G-----AAA		1897
O_MVP5180AATGGAGGAGGCAGTGAAGGAA-AAGAGA-TC-AGTC-AA-GGG--TC---GAGCT--C-A--TG-C-----AA-C-----G-CA--AA-A-----GC--G-T--T		2333
O_CM.ANT70GATGGAGGAGGAAGTGAAGGGA-AGGAGA-TC-G-AC--AAAGGG-G-CC-A-CGAGCTC-----G-TG-C-----AA-C-----G-CA--AA-A--T--GC-GAG-T--C		2357
HIV2_A_STAAGAAGTGACAGAGGA...CTTCCTG-AG-TCG---A--A--A-ACC-TG--GA--GACGACAG-GGAC--GCTGCA---AT--T-----A--A--AG-A-----GC-C-TG-T-A--C		2133
HIV2_A_BENAAGAGGTGACGGAGGA...CTTGCTG-A-TCG--C--AGA--A-ACCTC--GA--GGAGACAG-GGAC--GCTGCA---AT--T-----A--A--AG-A-----GCCT-C--C-A--AT		2688
HIV2_A_ALIAAGAGGTGACGGAGGA...CTT-CTG-A-TCG--C--GGA--A-ACC-C-----A--GG-GACAG-GGAC--GCTGCA---AT--T-----A--A--AG-A-----GCCT-C--T-A--C		2681
HIV2_B_D205AAAGTTACATGCAGATGGGGAG-CAA-AGAGAG--GC-GA--AGACCCCT-----GG-GACAG-GGA--GCTGCA---AT--T-----AG-A--AG-A-----A-GC-TGT--C-A--T		2664
HIV2_B_EHOAGAACTACATGCAACTAGGGGAG-AG-AGAAGG--CAGA--AGACCCCT-----GG-GACAG-GGA--GCTGCA---AT--T-----AG-A--AG-A-----A-GC-CT--T-A--T		2659
STMM83293AAAGTTACATGCAGCTGGGCAAG-AG-AGAGAG--GCAG-A--A-ACCTC-----GG-GACAG-GGA--GCTGCA---AT--T-----AG-A--AG-A-----GCCCT-T--T-A--T		2252
SMM251GAACTACATGCAGTTGGGCAAGCAG-AGAGAG-A-GCAGA--AA-CCTT-----GG-GACAG-GGA--GCTGCA---AT--T-----AGGA--AG-A-----TGCTC-T--T-AA--A		2581
SMM9GAATTACATGAAGGTGGGAG-AG-AGAGGG--CAGA--AGACCTT-----GG-AAC-G-GGA--GCTGCA---AT--T-----AG-A--AG-A-----TGCCT-C--T-AA-AA		2078
SIVCPZANTCATAGAGATCTATCAGGAGGAGCACAAGGACTC--A--GT-TC-AG-GGG---GGAACTACC--C-CGTA-----G-AA-C-----A--AA-GA-GGA-G-TCTC--CA--A		1742
SIVCPZGABGGAGGA-GAGAAGAGC--G-A--A-AGA-GG--G-GCAGT--ATATCCA--AACCT--...---AA-----G-----A-C--G--AG--AA--		2364
SIVCPZUSGAGA-AGTGACAC--G-A-A-A-GG-A-GG-AA-GGAGCCC-T-AGC-----AA-----G-----C--G-----AG--CA--		2363
SIVSTAN1CCAAAAAGGG.....AG-T--ATT-AGA-AGC-G-----GAAGGAATGGAGGACTA---T-G-T-C-----AGGA---AA-GAGT-AG--CT-C--CA--C		2493
SIVGRI677CAGAGAAGGGACAGCGCTGAG-GAGGAGAGAG-AC--ACAAG-AAACAGA-GG--A-GA...AGTGGAGGATGT---T-G-C-C-----AGGA---AA-GAAACG-G-C-TC---AA-A		2288
SIVVER9063CAGAGAAGGGAAAAACAGTTGAGGGAA-AAGGGA---GAACACCA--A-C-A-TCCA--CTG...-GCGGAAGG-TA---TT-G-AC-C-----AG-A---AA-AAAA--G-C-TT--T-AA--A		2521
SIVVER155CAGAGAAGGGAAAAACAATGAG--A--AGAACAGA--C-CCCCAG--ACA-TCCA--TTG...-A-CGAGGG-TA---TT-G-AC-C-----AG-A---AA-AAGG--CTGT-TA---A--A		2513
SIVVERTYOCAGAGAAGGGAAAAACAACACTGAGGGAG-AAAAGAG--T-CACC-G-A--GA-TCC---TTG...-ACCAGGGG-TA---TT-G-AC-C-----AG-A---AA-AAAA--G-GT-T---AA--		2011
SIVVERAGM3CAGACAAGGGGAAGCAGTTGAGGGAA-AAAGGA-A--A-CACCAG-AG-GA-TCCC--TTG...-ACAGAGGG-TA---TT-G-AC-C-----AG-A---AA-AAAA--G-TT-C---AA--		2016
SIVLHOESTGACTTACAGGACTTTAGG-CAA-AG-TGA-A-G---C---A-CA--T-CCAC-GAAATGTGTGGACGAGC-C-G---G-AT-----CCAG-----AA-A-AA-GT---TT-G-CAAT---		2643
SIVMNDGB1TGGAGACCTACAGACTGTTAGG-CAGGG--TCAGAGCA--C--AA--GGA-GG-AAGGGGGAA-G--AGGAGC-A-G-T-G-ATATG-----CCAG-A---T-A-AG-AGG-CTTC--TAAT---		2648
SIVSAB1CAG--GG--A-GAGAGA-G-GGA---TCA-----G-TA-T-----CAAG-A---AA-A-AAGAGG-CTCAG-G-AT--T		1978
SIVSYK173CTGTTTCAGGGGAAACGGCAGAA-CAG-AGTGG--A--C--CTC---ACA-C-G-GCCCTACGA-G-GG--AC-GC-----T-----AG-A---AA-GAAA--TG-GT-C---A--A		2701
GAGAGAAACCCCCACCAACAGGGA-GTG-T-T-G-C---G--AGCAGTGGGA-GG---GACCAA-AG--TC-ATC-C-T---C-T-----AG-A---AA-GA-AGA-G--G-TG---AA--T		2479
POLK K P P T N R E V L S P K E S S G K E E T K S L Y P S L S S L F G E D Q \$		GAG
E T P H Q Q G S A L A E G E Q W E G G D Q E S L S L S Q L S L W R R P M I E V D V E G		POL

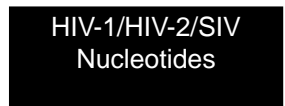
B_FR.HXB2R	CAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTG CCAGGAAGATGGAAACCAAAAATGATAGGGGAATGGAGGTTTATCAAAGTAA	2421
POL	<u>Q L K E A L L D T G A D D T V L E E M S L P G R W K P K M I G G I G G F I K V</u>	POL
A_UG.U455	--G-G-TA-----C-----C-A-A-----A-----A-----	1867
C_ET.ETH2220	--G-----G-----C-----C-----A-A-----A-----A-----T-----	1813
D_ZR.84ZR085	--G-----A-----C-----A-A-----AG-----A-----	1943
FL_BE.VI850	--GA-----G-----C-----A-A-----A-----A-----	1758
G_SE.SE6165	--G-----TA-----C-----C-----C-----A-A-A-----A-----	1818
H_CF.90CF056	--GT-----G-----C-----G-----A-A-----G-----A-----	1768
J_SE.SE92809	--G-G-G-----C-----AGAC-----C-----A-----	1735
CRF01_AE_TH.CM240	--G-G-A-----T-----A-A-----A-----A-----G-----	1995
CRF02_AG_NG.IBNG	--G-----TA-----C-----C-----C-----A-A-A-----A-----	1946
CRF03_AB_RU.KAL153	--G-----A-----C-----A-A-----A-----A-----G-----G-----	2466
CRF04_cpx_CY.94CY03	--GA-----G-----G-----T-----A-----A-----A-----G-----C-----	1787
N_CM.YBF30	G--G--GA-----T-----T-----A-----GC-ACAA-A-----GAG--A-----A-----A-----G-----	2015
O_MVP5180	--T--TGT-G--T-C-G-----G-----A-TA-C-ACAA-A-----GA-----C-----T-A-C-----A-----	2451
O_CM.ANT70	--C--TGT--T-T-GC-G-----C-----C--A-CA-C-ACAA-----GA--A-----T-A-----A-----	2475
HIV2_A_ST	--G-C-GTA--T-T-GC-C-----G--T--C--CT--A--G--C--GCG-AGAG-A-----GGGA-C-AT-AT-GT--G-AG-----A-G-A-C-A--TACC--	2251
HIV2_A_BEN	--G-CGTA--TCT-C-----C-----G--T--CT--A--G--C--G--AGAA-A-----GGG-AC-AT-AC-CT--AG-----A-G-A--A--CACC--	2806
HIV2_A_ALI	--G-C-GT--T-T-C-----C-----G--T--C--CT--A--G--C--G--AGAG-A-----GGGA-C-AC-AT-CC--AG-----A-G-A-C-A--TACC--	2799
HIV2_B_D205	--GTC-GTA--TAT-C-----C-----TT--C--CT--A--G--C--GG-AGAA-A-----GGTA-C-AT-AC-CC--AG--A-G-A--G-C-A--TACC--	2782
HIV2_B_EHO	--TC-GTA--TAT-C-----C-----T-----CT--A--G--C--GG-AGAA-A-----GGCA-C-AT-AC-CC--AG--T-G-A--A--A--TACC--	2777
STMM83293	--G-CTGTA--TA-----G--C-----G--T--C--CT--A--G--G--CG-GG-AGAGC-A-----GGGCT-CA-AC-CC--GG-AG--T--A--A--A--TACT--	2370
SMM251	--G-CTGTA--TAT--G-----G--T-----T-TA-TG-AC-G-AGAG-A-----GGTCC-CAT-AT-CC--AG--A--A-----T-TACT--	2699
SMM9	--G-CC-TA--TAT-----G--T-----T-A-TG--C--GG-AGAA-----GGTCC--AT-AT-CC--T--AG--A--A-----C-C-T--CACC--	2196
SIVCPZANT	--AA-TGTC-G-----T-----C-----G-----G-----TCA-----A-----AC--G-----CA--T--T-----G-----TTCC--C	1860
SIVCPZGAB	--TGT-----T-GC-----T-----A-----GAG-ACAA-A-----A-----CTT-----A-----C-----	2482
SIVCPZUS	--GA--GTA-----T-----C-----T-----C-----CA-C-ACAAA-T-----GA--G-C-----G-----G-----A-----	2481
SIVSTAN1	--G-C-GTAC--T--G-----T-----T-----CA--A--A--G--GAAGA-A-C-----CATTAGA--C-C--CGG--T--G-AG--A--T--A--GGGA--TC--	2617
SIVGRI677	ACG-C-GT-C--CT-G-----C-----TA--A-TC-----A-GAC--CACTTCCCCCA-ATAA-CC--CGTT-C--GG-AG--A--T--A--AGGG--TC--T--C--	2418
SIVVER9063	GTC-CTGTAA--TC-G-----G-----C-----CA-CA-A-----CAGACA-A-----CAACTAA--CCT--G-----G-A--A-----AGGAC--T--T--	2645
SIVVER155	ACTGCCGTTA-G-AT-----G-----C-----TA--A--A--G--T-CAGA--A-----CAATTAAGG--TC--G-----G-----AG--A-----G--AGGT--A--C--	2637
SIVVERTYO	GTC-CC-TTA-G-A-GC--C-----G-----C-----CA--A--TA-----ATGA--A-----CAATTAT--TCC--G--C-----T-----C--A--AGCC--T--T--	2135
SIVVERAGM3	GTC-CC-TCAG--AT-----G--G-----C-----CA--TA--A-----GCAGA--A-----CAATTAT--C-----A-----C-----GGGAC--T--	2140
SIVLHOEST	--C-GT-TC--A--G--G--C--T--G--T-----C--A--T--TTC--ATTCAG-A-----AGAATAGA--CCC-AT-CT--G-G-AG--C--A--G--GCAA--T-GG--	2767
SIVSUN	--CTGT-ACT--T--G--C--G-----C-----A--G--GTCT--GATTC-G-----TCCATAGA--TCAC--G--C--GGG-AG-G--T--T--A--G--ACAA--T-G--	2772
SIVMNDGB1	GTCAC-TAAG-----C-----T-----CA-C-TA-T--GA-A-A-A-----AAATTA--AT--C-G-----T-----A--T--AACT--A-G--	2102
SIVSAB1C	--GAA-GTCACT-----C-----CT-----A--C--G--GT--AGAA-A-----GG--AC-AT-----G--A--T--A-----A--G--T--T--	2819
SIVSYK173	G-C--GGT-C--ATGT-G-----G-----TA--A-TAG--GAAGA-A-T-----CAATTA-ACCA-CC--TC--AG--A--T--A-----AAAC--T-C--T--	2603
POL	<u>D L V Q M L V D T G A D D T I I R E E D I Q L H Q P W S P K I V G G L G G N I T V</u>	POL

B_FR.HXB2R	GACAGTATGATCAGATACTCATAGAAATCTGTGGACATAAAGCTATA.....GGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGTTG	2536
POL	R_Q_Y_D_O_I_L_I_E_I_C_G_H_K_A_I.....G_T_V_L_V_G_P_T_P_V_N_I_I_G_R_N_L_L_T_Q_I_G_C	POL
A_UG.U455	-----T-----T-----A-A-GA-----G-----G-A-----	1982
C_ET.ETH2220	-----A--A-----T-----A-A-G-----C-----C-----CA-----C-----AC-	1928
D_ZR.84ZR085	-----C-----C-----G-----C-----C-----T-----C-----C-----	2058
FL_BE.VI850	A-----A-C-----T-----C-G-----G-----G-----A-----A-----	1873
G_SE.SE6165	---A---AG---CT---TA---A-A-G-----G-GA-----A-----G-----A-----	1933
H_CF.90CF056	-----G-AG-GC-----A-A-G-----T-----G-A-A-----A-----	1883
J_SE.SE92809	-----A-CG-G-CG-----TGAG--A-A-G-----G--A-----G-CA-----C-----	1850
CRF01_AE_TH.CM240	AG--A-----T-----A-A-G-----C-----A-----	2110
CRF02_AG_NG.IBNG	---A---A---T-----T-----A-A-G-C-----C-----A-----	2061
CRF03_AB_RU.KAL153	---A---A---T-----T-----A-A-G-----G-----C-----A-----C-----	2581
CRF04_cpx_CY.94CY03	---A---A---CT-----T-----A-A-G-C-----C-----G-----C-----CA-----C-----	1902
N_CM.YBF30	---A---A-T--ACAG--C--ACAG--AGA--AG-T.....A-----T-T-T-A-----T-A-C-----C--	2130
O_MVP5180	A-G---A-CA-TG-GACAG---G-ACAA--A-GG--TACAG.....-A-----G-G-----T-----T-T-TC-G---CA-A---AGGAT-A--A--	2566
O_CM.ANT70	A-G-A---A-TG-GACAG---AGAA--AGGG-G-TACAG.....-A-----G-G-----T-----T-T-T-----A-A---AGGAT-A--	2590
HIV2_A_ST	A-G-A---A-AA-TG-GAA---AG-G-A-TAAATA-A-G-TA-G-----CC-CA-A-GAC--TGA-C-AA--TT-C---CA-TC--AGCCT-A--CAT	2366
HIV2_A_BEN	A-G-A---CA-AA-TG-GAA---A-G-ACTAAATA-A-G-TA-G-----CC-CA-A-GAC--GA-C-AA--CT-C---A-TC--AGCCT-A--CAT	2921
HIV2_A_ALI	A-G-A---AG-TG-GAA---A-G-ACTAAATA-A-G-TA-A-----CC-CA-A-GAC--TGAC-C-AA--T-TT-C---CA-T---AGCCT-A--CAT	2914
HIV2_B_D205	A-G-A---CA-AG-TG-GAA---G-AGTG--A-A-G--TA-GG.....CA-TA-A-GAC--GA-C-AA-A--TT-C---A-T--A-A-ACCT-G--CAT	2897
HIV2_B_EHO	ATG-A--CA-AA-TG-GAA---G-AGTA--A-A-G--TA-G-----CA-----A-GAC--GGAC-C-AA-A--TT-C---A-T--A-A-AGCT-A--CAT	2892
STMM83293	AGG-A-T-A-GA-TG-AA---GG-A-TA-CA-G--ATA-A-----A-TA-TA-GACT--GAC-C-CA-T--TT-C-G-C-A---AGCTT-G--GAT	2485
SMM251	A-G-A--CA-AA-TG-AAA---G-T-TA-CA-A-GGAT--A-----G-A-CA-GAC--GGAC-T-GA-T--TT-T-G---T-C-A--AGCTC-G--GAT	2814
SMM9	A-G-A--A-AG-TG-AAA---A-G-A-TA-CA-AGT-AT-AG.....-A--A-TA-GAC--GA-C-AA-T--TT-C---C-A--AGCT--G--CAT	2311
SIVCPZANT	A-----A-CA-AG--C---C---TG-A-ACAGG-C-TAC-----CA---C-GT-G---A-AC-A-A--T-A-T---G-T-ATG--TTT-A--A--	1975
SIVCPZGAB	A---A-T---A-TG--A---G-AGAA--GAGA--TAG-----G-----A-----A-T-----A-T---A-AT-G-----	2597
SIVCPZUS	A---A-----CG-CAAT-----AGAG--AGA---ACAG.....T---T---G---A---A-A-T-T---C-G---A-T--A--A---A--	2596
SIVSTAN1	A---A-----CA-T--TT-AG-GA-T-AGA--A-AG--A---G-----AT---C-A-T---A---AA-A-T-TG-A-----A-TC-AG-A-CIT-A--AGC	2732
SIVGRI677	A-G-A---C-GGG-G---AAG-C-T-GGAG-ATA-A-TCATC-CC-----CT--A-TC-A---AG---AA--T-T-A-----A-T-AG-----GCA--CAT	2533
SIVVER9063	A-G---AG-A-TG-GAAG--A---AGAA-ATA-A-TTTTG-G-----C---TC-TT--G-A-C--A--T-----T-AC-AG-A-CAGCA--AGC	2760
SIVVER155	A-G-A---A-TG-GAAG--C--T-GGAA-ACA-G-T-TTA-G-----A---CC-CA---G-A-T-CA--T-C-A-----CT-T-AG-C---GCA--AGC	2752
SIVVERTYO	A-G-A---A-CG-C-GGGAAG--A---AGAA-ATA-A-TTTTG-G-----A--A--GT--G-A-T-CA-T-T---A-T-----T-C--G-C-C-GCA-TGCC	2250
SIVVERAGM3	A-G---AG-G-T-GGGAAG--AG-T-GGAA-ACA-A-TTTTG-G-----G-CA--GA--AGC-T-CA-A-----A-A-AG-A-CAGCA--AGC	2255
SIVLHOEST	A-G-A---AGAG-TG-CT-T-----AGCA--A-A-CTA-AG-----G-C---T-----C-A-AG-T-TG-A-G-G---A-T-A---GCA--G--GGC	2882
SIVSUN	A-G-A---CAG-TG-CT-TG-----TA-G-ATA-A-C---C-----A-----G-G-G-A---A-AG-T-C-A-G-----A-TC-TT-AGCCT-A--GG-	2887
SIVMNDGB1	A-----A-TG--TATG-----AA-A--GA-GGG-A-AT-T.....G-----G-----T-AA-AG-T---A-G---CA-AA--GAAA-AT-A--AG-	2217
SIVSAB1C	A-GCA--CC-CA-TCAGGAGG--A---AGAA-ATA-A-CTTG--A-----C--CA-T-G---GAA--A-T-T-T-A-----G-C--AG-A--AT-A--AGT	2934
SIVSYK173	-G--A---AGAA-T--TAGAT-CACTG-CACTG-TGT-AA--CCTC--GA-A-AGAAAGCAAGTAGAG--G-TC-T---G-T---G-A-A-T--T-A---G---A-AC-T--CA-AT-G--AGT	2733
POL	R_Q_Y_R_N_I_R_F_T_V_V_K_P_S_G_K_R_K_Q_V_E_G_T_L_L_V_G_P_T_P_V_N_I_L_G_R_N_I_L_T_K_L_G_V	POL

Pol p10 protease end \ Pol p51 and p66 reverse transcriptase start		
B_FR.HXB2R	CACTTTAAATTTCCCATTTAGCCCTATTGAGACTGTA.....CCAGTAAATTAAGCCAGGAATGGATGGCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAAGCATTAGTAGAAATT	2660
POL	<u> T L N F P I S P I E T V . . . P V K L K P G M D G P K V K Q W P L T E E K I K A L V E I </u>	POL
A_UG.U455	T-----A-----T-----A-----.....-----A-----A-----G-----G-----AC-----	2106
C_ET.ETH2220	--A---C---A---T---C---A---.....-----C-----G-----G-----AC---C---	2052
D_ZR.84ZR085	-----C---A---T---A---.....-----G-----G-----AC-----	2182
FI_BE.VI850	T-----AG---T-----A-----.....-G-----G-----G-----AC---A	1997
G_SE.SE6165	-----A---T-----A-----.....-----GG-----AC-----	2057
H_CF.90CF056	---C-----A---T-----A-----.....-----G-----ACG-----	2007
J_SE.SE92809	T-----A---T-----A-----.....-----A-----G-----AC---C---	1974
CRF01_AE_TH.CM240	T-----C---A---T-----C-----.....-----C-----A-----G-----AC-----	2234
CRF02_AG_NG.IBNG	T-----A---T-----A---G-----.....-----G-----G-----AC---C---	2185
CRF03_AB_RU.KAL153	T-----A---A---T-----A-----.....-----C-----G-----A-----G-----AC---C---	2705
CRF04_cpx_CY.94CY03	T-----A---T-----A-----.....-----G-----G-----C---AC---G---A	2026
N_CM.YBF30	T-----A---A---T-----A-----.....-----A-----G---A-----T-----AC-----G---G---AG-----	2254
O_MVP5180	T-AC-----C---T---A---T---C---A---CCC---A---G-----.....-GC---A-----A-----A-----CC-AT-TAG---G---G---C---ACT---C---A	2690
O_CM.ANT70	T-AC---C---C---T---A---C---A---CCC---A---G-----.....-C---A-----A-----A-----CC-AT-TA-----G---C---GAC---C---A	2714
HIV2_A_ST	GT-A-----C---A---AG---CGC---AAG---A---AC---AA-----.....AA---A---TGC---G-----A-----A---C---G---G---C---A---A---G---G---C---AA---G---C	2490
HIV2_A_BEN	GT-A-----A---AG---GC---AAG---A---C---AA-----.....AA---C---G---G---AA---A---GGC---G---CC---A---A---G---G---C---AA---G---C	3045
HIV2_A_ALI	GT-A-----CC---A---AG---GC---AAG---A---C---AA-----.....GAG---G-----AA---C---G---T---A---G---C---A---A---G---G---C---AA---C	3038
HIV2_B_D205	G-----C---AG---GGCAAAGG---A---AC---A-----.....AA---TG---G---A---T---AA---G---GA---C---G---TC---AT---CAGG---G---CT---CC---CAA---C	3021
HIV2_B_EHO	G---C-----C---AG---AGCAAGG---A---AC---A-----.....AA---CC---G---T---A---AA---G---A---C---G---CC---AT---CA---G---CT---CC---CAA---C	3016
STMM83293	GT-A-----AG---AGCTAAAG---A---AC---A-----.....AAG---G---C---A---AA---A---A---A---AT---A---G---G---G---C---AA---G---C	2609
SMM251	GT-C---C---AGCTAAGG---A---C-----.....AA---C---CC---A---T---A---T---G---G---G---AT---A---G---GTT---AG---C	2938
SMM9	GT-C---C---AGCTAAGG---G---RC---A-----.....AA---C---C---A---AA---A---T---A---G---G---GC---AT---GA---G---TT---AG---C	2435
SIVCPZANT	T-----A---TAAAG---A---A---G-----.....C---T---AGA---AA---A---G---A---GC---CT---A---G---G---G---CC---AA-----	2099
SIVCPZGAB	T---C---GTG---C---A---TT---A---A---C-----.....C---G---A-----T---G---A---G---C---AT---C-----T---G---C---AC-----	2721
SIVCPZUS	T-----T---C---T---A---A-----.....G---A-----T---G---A---G-----C---AT---C-----T---G---C---AC-----	2720
SIVSTAN1	A-GA---TAA---GGGACAGCTAAA---GAGA---A---TACCCATAACCAAG---C---G---A---T---G---TCA---A---G---T---A---CC---AT-----CC---GAC---A	2862
SIVGRI677	G-AA---GT---A---GGGAG---CTAT---G---C---T---AGGAACAAA---C---C---AGA---G---AA---A---T---T---G---C---AT---AG---TG---T---AC---A	2663
SIVVER9063	T-GAC---GTGA---GGGACAATTGT---AGAAC---A---T---CCCATCACC---C---C---G---AGAT---GCAAGG---A---C---TA---A---TC---AT---A---G---G---G---C---AC---G---A	2890
SIVVER155	--AA---GTGA---GGGGCAATTGT---GCAGACA---TACC---ATCACC---G---CGC-----GA---GGCCAGA---A---CG---T---G---G---C---CT---TA-----T---CC---GCA-----	2882
SIVVERTYO	-CGG---GTAA---GGGACAATTAT---AGAAA---A---T---CCTGTGACA---T---C---G---G---GA---GGCTCGG---A---CTGT---A---G---TC---CT---TA---G---G---TG---T---CAG---A	2380
SIVVERAGM3	--AA---GTAA---GGGTCAACTGT---AGAAC---A---T---CCCATTACC---T---G-----AGA---GGCTAGA---A---TTCT---A---CC---CT---CA-----C---CAG---A	2385
SIVLHOEST	--AA---G---T---AG---TCAGTTAAG---GA---A---T---CCTATTACAAAG---TCT---A---T---GTGT---A---G---A---G---G---TC---AT---A---G---G---GCC---CA---CT---	3012
SIVSUN	A-GA---GT---AG---C---CTTAG---GAAA---TCCC---ATTACAAA---T---G---AGA---GGCC---C---A---G-----CC---T---AGG---G---G---GCC---CAGA---A	3017
SIVMNDGB1	A-AA---TA---GG---ACAATTGT---GA---A---TACC---ATAACAAA---G-----A---G---A---C---G---A---A---T---A---GTA---G---GTT---GTC---TCAGA---A	2347
SIVSAB1C	T---C-----GA---ACAA---AGAG---A---AC---A-----.....AA---CC---C-----CAA---G---GGA---A---GG-----T---T---A-----G---G---C---AAG---CC---	3058
SIVSYK173	--AA---GTAA---GGTTCAA---CAG---AT---A---C---A---C-----.....AA---GTC---G---A---AC---AA---ATTG---CCGGC---A-----AT---G---T---G---T---G---G---CC---CAAG---C---A	2857
POL	<u> K L V M V Q T A L E P V . . . K V S L K P D K E L P R L K Q W P L S V E K L E A L K A I </u>	POL

B_FR.HXB2R	TGTACAGAGATGGAAAAGGAAGGAAAAATTTCAAAAAATGGGCCTGAAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTAAATGGAGAAAATTAGTAGATTCAGAGAACTTA	2790
POL	_C_T_E_M_E_K_E_G_K_I_S_K_I_G_P_E_N_P_Y_N_T_P_V_F_A_I_K_K_K_D_S_T_K_W_R_K_L_V_D_F_R_E_L	POL
A_UG.U455	---AT-----A-----T-----G---C-----G-----C-----	2236
C_ET.ETH2220	--GA--A--GC--A-----G-----C--T-----A-G-G-----G-----G--C--	2182
D_ZR.84ZR085	---T-----A-----G-----A-----G-----G-----G-----	2312
FL_BE.VI850	--CT--A-----A-----A-----G-----G-----G-----A-----	2127
G_SE.SE6165	--A--A--G--A-----C--A-----G-----G-----G-----G--C--	2187
H_CF.90CF056	--G--A--G--A-----C--G--A-----G-----GC--A-----A-G-G--T-----G-----C--	2137
J_SE.SE92809	--G--A--G--A-----C--G--G-----T--C-----A-----G-----C--C-----G-----T-----C--	2104
CRF01_AE_TH.CM240	--A-----G--A-----C-----A-----T-----G-----C--C-----G-----G--C--	2364
CRF02_AG_NG.IBNG	---AG-----A-----A-----G-----T-----G-----C-----G-----C-----	2315
CRF03_AB_RU.KAL153	---C-----C--G-----A-----T-----A-----C-----C-----C-----G-----	2835
CRF04_cpx_CY.94CY03	---A-----A-----T--G--A-----G-----T--C-----A--T-----T--A--G-----T--C-----G--T--A--	2156
N_CM.YBF30	--CA--A--C--A-----C--G--A--A-----T--T--A--TA--T-----T--A--G-----T--C-----G-----C-----T--A--	2384
O_MVP5180	--CAG--A--C--A-----G--A--A-----T--T--A--TA--C-----T--A--G-----TG-----G-----G-----C-----T--A--	2820
O_CM.ANT70	--GAGA--A--GA--G--CC--GC--AGAGG--GGCACCT--AACT--T--T--C--CAC--A--C-----G--AA--AC-----TGC--A-----T--A--	2844
HIV2_A_ST	--GA--A--A-----A--G--CC--GC--AGA--G--GGCACCT--AACT--T--T--C--CAC--A--T-----G--AG--AC-----G--TGC--GA-----T--A--	2620
HIV2_A_BEN	--GA--A--A-----GA--CC--T--AGAGG--GGCACCT--AACT--C--T--T--C--CAC--A-----G--G--AA--AC-----TGC--A-----T--A--	3175
HIV2_A_ALI	--GA--A--A-----G--AC--T--AGA--G--GCACCC--ACT--C--A--CACC-----G--T--AG--AC-----TGC--A-----C--T--GT--A--	3168
HIV2_B_D205	--GA--A--A-----A--G--AC--GT--AGA--G--GGCGCCT--ACT--T--G--CACC--C-----A--G-----AA--AC-----G--TGC--A-----A--	3151
HIV2_B_EHO	--GA--A--A-----G--T--C--GC--AGA--G--GC--CCT--AAC-----T-----C--CACT--T-----G--AA--AC--G-----G--TGC--GA-----A--	3146
STMM83293	--GA--A-----T--TC--GT--GGAGG--GC--CCC--GACC-----C--C--CAC-----T--AG--AC-----TGC--GA-----T--G--A--	2739
SMM251	--GA--A-----G--A--T--CC--GT--AGAGGR--GC--CCT--AACC-----T--C--C--CAC-----T-----R--GR-----G--A--A-----G--TGC--A-----T--A--	3068
SMM9	--GATA--T--A--GCA--AAT--G-----T--G--A-----A--T--C-----A--A--T-----A--A--G-----C--T--A-----GC-----T-----T--A--	2565
SIVCPZANT	--CA--A-----A--G-----A--A-----A-----T-----A--TA--T-----A--C--A--G-----G-----C-----C-----T--A--	2229
SIVCPZGAB	--CCA--A-----A--A--G--A-----G--G--A-----A-----T-----A--A--T-----T--C--A--G-----T--C-----C-----T--A--	2851
SIVCPZUS	AC--CAG--CTC--A-----A-----T--GG-----AGGA--G-----T-----C--G--TGT--CCG-----G-----AA--AC--G-----G--TG--A-----T--CT--C--	2850
SIVSTAN1	--C--A--C--A-----G-----G--A-----T--A--T--GG--A--AGGA-----T--T-----A-----G-----G-----T--AA--AC-----TGC--T-----G--A--	2992
SIVGRI677	--GA--C--T--A-----A-----CAGC--G--A--AGGA-----G--T--T-----C--A--C--TGT--C--GA--G--G-----AA--C-----G--TG-----C-----T--A--	2793
SIVVER9063	--C--A--ACAT--A--GG--A-----A--T--AAGC--GGG--A--GGA--C-----G-----CTGT--G-----AAT--AC--G-----TGC--G-----G-----C--	3020
SIVVER155	--T--CC--AT--A--GC-----A-----CAGT--G--G--A--AGGA-----G-----C-----A-----TG-----G-----AAT--CC--G-----G--TGC-----C--T--GT--A--	3012
SIVVERTYO	--GACC--AT--A--G--A-----A-----AGC--G--A--AGGA--G-----G-----C-----G-----TG-----AGT--AC-----TG-----T--G--A--	2510
SIVVERAGM3	--GATAG--T--A-----A-----A-----TCC--G--G--AT--A--GG--C-----C-----A-----A-----T--C-----G-----AA--A--G-----G-----A--T-----A--G--C--	2515
SIVLHOEST	--GATAGAT-----GCA-----A--G-----AGC--GGGCA--AC--A--GG--C--T--T--C--A--A--T-----T-----G-----AA--A--G-----G-----A-----T-----T--A--	3142
SIVSUN	--GATAGAT--A--GG-----A-----AGT--GGG--A--AT--A--G-----AAT-----A--TA--C-----G--G--T--AA--A--G-----A-----C--T-----T--A--	3147
SIVMNDGB1	--GA-----CT--A-----AC--AC--TT--GGA--G-----A-----A-----T--C--C-----A-----G-----G-----T--AA--C--G-----T--C--CA--G--C-----GC--GT--A--	2477
SIVSAB1C	GTAGA---C---CTC---C---AC--T--GGAG---GC--A--T---ACC---T--T---C-----TA--T--G-----G---AA--AG-----G--TGC--CA--T---T---A--T--A--	3188
SIVSYK173	_V_E_D_M_L_K_A_G_Q_L_E_K_A_S_P_T_N_P_Y_N_T_P_V_F_V_I_R_K_K_D_K_K_K_W_R_M_L_I_D_F_R_K_L	2987
POL	POL	POL

B_FR.HXB2R	ATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAAGAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCTTAGATGAAGACTT	2920
POL	N_K_R_T_Q_D_F_W_E_V_Q_L_G_I_P_H_P_A_G_L_K_K_K_K_S_V_T_V_L_D_V_G_D_A_Y_F_S_V_P_L_D_E_D_F	POL
A_UG.U455	---A-----G-----C-G--A-A-G--TC-----A-----G-C-----T-----AG---	2366
C_ET.ETH2220	---A-----T-----G-----C-A-----G-----A-----G-----C-----T-----GT---	2312
D_ZR.84ZR085	-----G-----G-----T-----A-----G-----G-----A-----C-----A-----TG---	2442
FL_BE.VI850	---A-----T-----G-----T-----T-----G-----G-----A-----G-----T---	2257
G_SE.SE6165	---A-----G-C-----T-----T-----A-----G-----G-----C-----T-----A-----T---	2317
H_CF.90CF056	---A-----G-----C-A-----T-----C-----G-----A-----G-----C-----T-----A-----T---	2267
J_SE.SE92809	---A-----G-----T-----G-----A-----T-----C-----A-----G-----C-----T-----T-----AG---	2234
CRF01_AE_TH.CM240	---A-----G-----T-----G-----A-----T-----C-----A-----G-----C-----T-----T-----AG---	2494
CRF02_AG_NG.IBNG	-----T-----G-C-G-----T-----G-A-----T-----G-A-----G-----A-----A-----	2445
CRF03_AB_RU.KAL153	-----C-T-----T-----T-----G-----A-----G-----A-----G-----C-----A-----	2965
CRF04_cpx_CY.94CY03	---A-----G-----G-----G-----G-C-A-----G-----T-----G-----C-----G-----	2286
N_CM.YBF30	---A-G-C---T-T---G-GC---T---A---A---A---GC---G---TT---A-A---T---TG---G-CA---T---	2514
O_MVP5180	---A---A---T---G-G---T-T---A-GG-T---GC-A-GGC---T-T---CT-A---A-A---T---C---TGC---T---CC---G---	2950
O_CM.ANT70	---A---G---G---G-A-GC---T-C---G-GG-T---G-GC-A-GC---T-T---CT-A---A-A---T---C---TG---CCC---T---	2974
HIV2_A_ST	-C---GT---ACA---A-C-G---T---C-A---AC-GCC---G---CG-A-T---T-C-A---A-G---T-C---CA-A-AC-C---G---T---	2750
HIV2_A_BEN	---GTG---T---ACA---A---GC---T---C-G---AC-GCC---A---G---AGGA-CT-TA---T-A---A-G---C---CA-A-AC-C---T---	3305
HIV2_A_ALI	-C---GT---T---ACA---GA---G---G-T---A---A---GCC---G---AG-A-C---T---G---A---G---T-C---CA-A-AC-GC---GAG---	3298
HIV2_B_D205	-C---GT---C-G---ACA---CA-C---G-T-T-C-C-G---TG---GC-G-A---GGAG-A---A-A---A---C-C-CAGTA-C-AC---CC-A---	3278
HIV2_B_EHO	-C---GT---C---A---TACA---G---C---GC-G-T-T-T-C-A---AC-GGC-TCA---G---AG-A---A---A---A---C-C-CAGT---C-AC---CC---	3276
STMM83293	---AGTG---A---T---TACA---A---GC---T---C-T---AC-GC---GG-G-AG-A-C---T---A---G---C---CA-A-C---G-G---	2869
SMM251	---G-GTC---G---TACA---C---C-T---AC-GC---A-GG-AGGA-T---A-A---C---CA-A-TC---A---A---	3198
SMM9	-C---GTC---T---TACA---A---G---G---C-TR---AC-GC---G-G-AGGA-C---GT---A---C---CA-A-T---A---A---	2695
SIVCPZANT	---A---A---T---T---GA-A---T-T---A-C-A---GC-A---G---G---T---A---A---C---C-CA-A---C-G---	2359
SIVCPZGAB	---A---A---T---G---G---C---T-C-A---A---G---G---G---T---A---A---A---C-C-C-TTG---C-G---A---T---	2981
SIVCPZUS	---A-G-A---G---G---C---C-A---A---G---G---C---A---A---A---G---C---TTG---AC-G---A-G-A---	2980
SIVSTAN1	---AGC---A-G---TT---G-A---T-T-C---A---A---G---GA---G---CAGA-C---A---A-A---A---C---A-AGTA-A---A---G---CC---A---	3122
SIVGRI677	-C---AGCT---T-TT---G---GC-G---T-T-C-A---G---CC-TC-G-A---G---GCA-A-C---A-A---CA-A-G---C---A---A---A---TGCA-G---A---	2923
SIVVER9063	---AGT---A---TT---G---G---T-C-G---GC-G-A-TG---CAGA---T-A---A-A-G---T---A---CA-A-A---CC---G---	3150
SIVVER155	-C---AGCT---A---TT---C---T---C---A---G---A-TG---GCA-A---CA-TA-A---G---A-AGCA-A-AC-G---CCT---G---	3142
SIVVERTYO	---GC---C---T---TT---G---G---C---C-A---A---G---TG-G-CAGA---TT-A---A---A---C-C---A---CA-A-A---G---CC-A-T---	2640
SIVVERAGM3	-C---AGC---A---T---T-TC---A---G---C---T---AT---CG---TG-CGGA-A---T-A---CA-A-G---C---A---A---A---CCC---G---	2645
SIVLHOEST	-G-ATT---TCAT---GC-A---T-T---C-A---TA-C---G---ATGC---CAGA-T---G-A---CA-A-A---C---CAGTA-C---T---CC-A-T-A---	3272
SIVSUN	---GGT-TG---TCAT---GT-A---G---G-T-T---A---A---G---TGT---CAGA-T---G-A---G---C---C---A---A---A---G---CC---T-A---	3277
SIVMNDGB1	-C---TT---A-G-T---TCAT---T-A-G---T---T-C-A---AA---TG---AG-A---C---A---A-A---G---C---AGTA-A-TC-G---CC---T-A---	2607
SIVSAB1C	---A---T---TCA---C---G---G---C-C-A---C---C-GC-A-GGG-GCAGA-T---GT---CA-A-A---C---AGCTGC---T---CC---G---	3318
SIVSYK173	---G-GCC---A---T---T-TC---G-A---G---C-C-A-G---AC---CG-C---.C---A-TA-A---T-AAAA---AC---C---G---CA---G---	3114
POL	N_E_A_T_Q_D_F_F_E_V_Q_L_G_I_P_H_P_G_G_L_K_R_Q_K.L_T_I_I_D_L_K_D_A_Y_Y_S_V_P_L_D_K_E_F	POL



B_FR.HXB2R POL	CAGGAAGTATACTGCATTACCATACCTAGTATAAACAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATGACAAAAATC R K Y T A F T I P S I N N E T P G I R Y Q Y N V L P Q G W K G S P A I F Q S S M T K I	3050 POL
A_UG.U455	T--A-----G-C-----AG-C-G-----G-----T-----G-----	2496
C_ET.ETH2220	--A-A-----C-----C-----A-----A-T-----C-C-----C-----G-----C-CC----	2442
D_ZR.84ZR085	T-----C-----C-----C-----A-----A-----A-----A-----T-----T-----	2572
FL_BE.VI850	--A-----C-----C-----G-C-----T-A-----G-C-----A-----A-----T-----	2387
G_SE.SE6165	T--A-----C-T-----T-----G-----T-----T-----T-----G-----	2447
H_CF.90CF056	--A-----C-----C-----C-----T-----T-----G-----G-----G-----	2397
J_SE.SE92809	--A-----C-T-----C-----C-----A-----C-----C-----T-----GT-----	2364
CRF01_AE_TH.CM240	T--A-----C-----C-----A-C-----C-----G-----G-----G-----	2624
CRF02_AG_NG.IBNG	T--A-----T-----G-----T-----C-----G-----G-----T-----GGCA-----	2575
CRF03_AB_RU.KAL153	--A-----C-----C-----C-----A-----C-----T-----T-----	3095
CRF04_cpx_CY.94CY03	--C-----C-----CC-----A-----C-----T-----T-----	2416
N_CM.YBF30	T--A-----A-T-----T-----C-----T-----G-----A-C-----G-----T-----T-G-----CA-----T	2644
O_MVP5180	T--A--A--C--C--C--T--T--G--G--C--AG--A--C--C--G--A--G--T--T--C--T--G--TCA-----G--T	3080
O_CM.ANT70	T--A--A--T--C--T--T--G--G--T--C--A--A--C--C--G--A--T--T--T--TCA-----T	3104
HIV2_A_ST	T--AC-----TC-----ATCA-----CTGA-----A-AA-----CATA--T--A--CTCA-----G-----T--T--GTAC--CA--GGC--GG--	2880
HIV2_A_BEN	T--C-----C-----AGCAG-----AT--GA-----A-AA-----ATA--T--A--CT--G-----A-----G-----T--T--TAC--CA--GGC--G--	3435
HIV2_A_ALI	T--AC-----TC-----ATCAG-----CAGA-----A-AA-----ATA--T--A--CT--A--G-----G-----T--T--CAC--CA--G--C--G--	3428
HIV2_B_D205	--AC-----A-----C-----T--G--ATCAG-----CAGAG-----A--AG-----CATT--T--A--T--A-----A--G-----G--G--C--AGT--T--GT--TACTC-----G--GG--A	3408
HIV2_B_EHO	--AC--A-----A-----TT--G--AGCAG-----T-----CAGA-----A--AG-----TT-----A--C--A-----G-----C-----T-----GTAC--C-----G--GG--A	3406
STMM83293	T--C-----C-----T--C-----T-----ATCAG-----CAGA-----A--AA-----CATC-----G--T--G-----G-----G-----T--T--TA--CA--G--T--A	2999
SMM251	T--C-----C-----TT-----ATCAG-----T-----CAGAG-----A--AAC-----CATT--T--G--T--G-----G-----G-----C-----C-----TAC--CT--G--C--TG--G	3328
SMM9	--C--A--C-----C-----T-----ATCGG-----CAGA-----A--AA-----CATC--T--G--AT--A--T--A--G-----G-----T--T--T--RCA--CTR--G--TG--A	2825
SIVCPZANT	T--A--A-----A--T--C--A--T--A--CG--G-----C-----A--A-----TGT--T-----T--A-----A--C-----T--T--T--GCA-----C--G--	2489
SIVCPZGAB	--A-----A-----A--T--C-----T-----C-----AG-----A--T-----TT--G-----A-----G-----T--T--C-----C--T-----T	3111
SIVCPZUS	T--A--A--C--A-----C-----G-----C-----A-----C-----T-----T--A-----A-----G-----T-----G--C--T-----T	3110
SIVSTAN1	--CCC-----A--T-----C--ACAAG-----T-----TAGGG--T--A-----G-----A--T-----TGT--A--T--A-----CA--C-----T--GCAC--CAGCAGCG--G--	3252
SIVGRI677	--A--A-----A-----C--CTCAG-----T-----ACAGGG-----A--G-----T-----TGT--G--T-----T--TA-----T-----G--A--CGGCAG-----C--T	3053
SIVVER9063	T--A--A--C--A-----A--C-----TCAG-----T-----C--AGG-----TG-----G-----A--T-----TGC-----A--C-----G--T--A--T--T--T--G--A--CAGCAG-----T	3280
SIVVER155	T--A--A--C--A--T--C-----C--CGG-----G-----C--A-----A--TT-----TGC--A--G-----C-----G-----C--GA-----T-----AC--CAGCAT-----T	3272
SIVVERTYO	T-----A-----T-----T--T-----CAG--G--T--C--GG--C-----G-----A--T-----CTGT--C--G--A--G-----T--TA-----T-----A--CAGCAG--TCC--T	2770
SIVVERAGM3	T--A-----C--T-----T--ATCAG-----T--C--AGGG-----T--C-----T--CTGT-----A-----G-----C--A--T--T--T--G--AC--CAGCAG--TTCC--T	2775
SIVLHOEST	--A-----A-----T-----TC--C--C--T--C--AGAG-----A--AA-----G-----A--G-----G--GAGC--TTGT-----T--GG--A--CAG--AG--GG--C--T	3402
SIVSUN	T--A-----A-----C--A-----GTAG-----T--C--G-----A--AA-----GT--T-----C--A-----C-----AGT--TTGC--C--T--GG--A--CGG--AG--G--GCC--T	3407
SIVMNDGB1	T--ACCC-----C-----GG-----ATCAG--T--T--C--AG-----A--AA-----CAT--T-----T-----T--A--G-----G--AGT--TGT--C--T--G--G--CAG--AG--TC--C--G	2737
SIVSAB1C	TCA--A-----A-----C-----ATCAG--C--T--AG--GAG-----C--A-----CAT--T--G-----C-----G-----C-----CA--T--T--G--CA--CAGCC--AC-----T	3448
SIVSYK173	T--CCA--C--G-----AG--G--A--C-----T--CCT--C--G--AGAA-----A--TT--CA--T--A-----T-----CA--T--C-----C--CA--A--ATC--G--A	3244
POL	R P Y T A F T V P S I N N A S P G E R Y Q F T V L P Q G W K G S P T I F Q S T I N Q I	POL

B_FR.HXB2R	TTAGAGCCTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGACTGACTTAGAAATAGGGCAGCATAGAACAAAATAGAGGAGCTGAGACAACATCTGTTGA	3180
POL	L E P F R K Q N P D I V I Y Q Y M D D L Y V G S D L E I G O H R T K I E E L R Q H L L L	POL
A_UG.U455	-----C-----TC---C-----C-----G-----T-----A-----G-----AT-A--GCT----A----	2626
C_ET.ETH2220	-----C-----GGCC-CC-C---A-----T-----C-----A-----G-CCC---A--T-A--G-----A--A-	2572
D_ZR.84ZR085	-----C-----G-----AG-----T-----A-----G-----A-----A-----A-----G-----	2702
FL_BE.VI850	-----C-----TGA---C-----C-----C-----G-----T-----A-----G-----AC---	2517
G_SE.SE6165	-----C-----GC-A-T---A-G-G---C-----A-----G-----T-----A-----G-----A-----	2577
H_CF.90CF056	---C---C-----G-----T-A-G---T-C-----A-----G-----A-----G-----T-A--GCT--T----	2527
J_SE.SE92809	---A---A-----G-AG---C---A-----C-G-----C-----G-----AA---G---A-----GG---A-----	2494
CRF01_AE_TH.CM240	---A---C-----T-A-----A-G-----A-----C-----T-----T-----C-----A-----GCT--A-----	2754
CRF02_AG_NG.IBNG	---C-----C-A-----G-----G-----C-----A-----A-----T-----G-----T-----GG---AC---	2705
CRF03_AB_RU.KAL153	---C-----C-----G-----G-----G-----G-----G-----A-----G-----C-----	3225
CRF04_cpx_CY.94CY03	---C-----TTCA---C---A---C---A---C---T-----G-----G-----A-----G-----A-----G-G---A-----	2546
N_CM.YBF30	C---A---A---C---G-GA---C---G---A---C---T---C---G---CC-C---G---C---CA---A---GAGGC-G---A---C---T---G---T---	2774
O_MVP5180	C---T---A---AGC---C---AG---AA---T---G---A---C---A---A---T---CC-T-G-CAG-A---AG-GGG-C---ATT---T---GG---T---A-ATC	3210
O_CM.ANT70	C---T---A---C---G-G-C---C---AT---AA---T---G---G---CC-A---A---T---CCCC-GACAG-A---A---GG-T---ATT---T---G---C---A-ATC	3234
HIV2_A_ST	---A---A---C---GC---C---G---T---CA---C---AT---G---A---C---TGA---CCAGC---AGGAC-GATTTAG-A---GAC-G-GTGG-TCT-C---A-AGG---T---AAATG	3010
HIV2_A_BEN	---A---C---GC---C---TG-CA---C---ATC-G---A---C---TAA---CTAG---AGGAC-GTTTGTAG---GAC-A-GTGG-CCT-C---A-A-G---T---AAATG	3565
HIV2_A_ALI	---A---C---GGC---C---AG-TG-CA---C---AT---A---C---TAA---CTAG---AGGAC-GATTTAG-A---GAC-AGTGG-CCT-C---T-A-AGG---TC---AAATG	3558
HIV2_B_D205	---T---GGCC---CAGC---TG-CA---AAT---G---CA-CCT-A---C-AG---AG-AGTGATCT-G---CGAC-GGGT-G-GTCCC-A---A-A-G-GTAT---AAATG	3538
HIV2_B_EHO	C---C---C---GCC---CAAT---TG-CAC---AATC-G---CA-TCTC-G-C-AG---AGGAGCGATCT-G---GAC-GGGT-G-GTCTC-A---A-A-G-G-TAT---AAAT-	3536
STMM83293	---A---C---G-G-GCC---C---TG-AC-C-GATC---A---C---TGA---CCAG---AG-AC-GATCTAG---GAC-GGGT-G-TTTAC---T---AGG---TC---AAAT-	3129
SMM251	C---A---C---G-G-GGC---TG-GACCT-AGTC---G---T---CA-C-TAA---CTAG---AGGAC-GACCT-G-A---GAC-GGGT-G-TTTAC---A-AGG---TCT---AAAT-	3458
SMM9	---A---GC---TG-GACCC-GATC---CA-C-TAA---CTAG---TAG-AC-GATTTAG-A---GAC-GGGT-G-TTTAC-AT-A-AGG---T---AAC-	2955
SIVCPZANT	---CA---A---GG-TA-GT---CAG---AA---T---C---TA-G---TACTGCA---GA-TG---AA---T---T---T---ACA-G	2619
SIVCPZGAB	C---A---C---C---G-A-G---T---TAC---T---C---G---CC-A---G-G---TC-T---T-AT---A---A---GG-G---A---A---A---T---C-T-	3241
SIVCPZUS	C---T---C---G---TG-TA-A---T---T---C---C---G-A-TC-A-CT---AAA---GGA---GG---ACT---C---T---T---C-T-	3240
SIVSTAN1	C---A---AGAAA-AG---A-GGT---A---TT---AA---AAT---G---CA---GGA---A---T---GCC-GAGAAAAG-A---GATGA---GG---A---G-A---A-G-T---C-AT	3382
SIVGRI677	---GAGA-C-A---GG---C-C---T---GGT---AA---TGTC---C---GGT-G-CG-A---CAT---TGAGACTAGA---ATCA-C-GG---CATAG-A---A-GATG---C-AG	3183
SIVVER9063	---AAAA-A-A-G-GG---TTAAA-CCAT-GACA---AGTG-G---GG---G---C-AGA---TGA-TACACC---GAC-AGTT-G---AC---T---GTCA-G---AAAA-	3410
SIVVER155	C---AGAAA-A-AG---G---TTAAA-C-GC-GACG---TGTC-G---CC-C-GG---AC-AGA---GGGTCCAA---GATCAGCT-G---C-AACA---T---GA-TAGAT---CAAG	3402
SIVVERTYO	---G---GAGA-A-A---G-A-CTTG---CAC---ACC---TGTA---A---GG---T---C-AGA-A-TGA-CACACC---GAC-A-TT-G---AC-T-A---AC-A-AT-ACAAG	2900
SIVVERAGM3	C---AGAAA-A-A---GG-GTTAAA-CC-C---ACC---TGTC---CC-A-GG---G---C-GGA---TGA-TACAC---GATCGGTTGG---AC-A-A---ATGA-AT-AAATG	2905
SIVLHOEST	C-CTCAGAG---G---TT---GA-C---T---C---T---A-TCA---A-TAG---G-G-AAAGGA---GATCAGGC-G-A-A-A-C---G---T---TA---	3532
SIVSUN	---AAA---G---GT---AG-TCAAT-G---C---ACTCA---T-AT---G-A-AA-A---GA-GA-T-G-GA-AC-AT---G-G-TA---CA---G	3537
SIVMNDGB1	C-G---GTA---GA-CC---ACAG---CAGT-A---T---G-A---ATAC-GC---AAG---GAG-A-GCT---TA-AT-A---GGCTTTAT-AA---	2867
SIVSAB1C	C-GC---GAA---GC-A-G-C---TG---A---A---A---TAA---CCAG---TAGGCC---AG-CAG-A---CT-GT---TGG---C---C---T-A---G-CT---AGAA-	3578
SIVSYK173	C-C---A---GA-GT---T---TC-CAC-C-AATA-G---C-A-TGA-T-CA---TAG-TC-GAGAA-GCA---CAGGAG-T-G---C-AC-AA-AGTGAC-GCA---A---A-	3374
POL	L Q P F R K K Y S D L T L I Q Y M D D L L I G T D R S E K A H Q E I V Q Q I V T A L L	POL

B_FR.HXB2R	GGTGGGGACTTACCACACCAGACAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGC.....TG	3304
POL	R_W_G_L_T_T_P_D_K_K_H_Q_K_E_P_P_F_L_W_M_G_Y_E_L_H_P_D_K_W_T_V_Q_P_I_V_L_P_E_K_D_S.....W	POL
A_UG.U455	-C-----T-C-TT--C-----G-----T-----G-----T-----C-----T-----CA-----.....	2750
C_ET.ETH2220	A-----T-----G-----T-----G-----T-----C-----T-----CA-----G--T-----	2696
D_ZR.84ZR085	-----G-----T-----T-----G-----T-----ACA-----G-----A-----	2826
FL_BE.VI850	-A-----T-----T-----GC-----G-----CAAT-----A-C-G-----	2641
G_SE.SE6165	AA---GT-----T-----A---G-----C-----G---A---CA-----C-G-A---	2701
H_CF.90CF056	AA---T-----C---T-----A---A---C-----A---G---AAA-----	2651
J_SE.SE92809	A---T-----T-----T-----A---G---T-----C---A---CAA-----AGA-----	2618
CRF01_AE_TH.CM240	-C---T---T---G---G---T---A---A---C---G---C---A---AA-----	2878
CRF02_AG_NG.IBNG	AA---T-----G-----T-----A-----C---G---AA---A-----	2829
CRF03_AB_RU.KAL153	-----T-----T-----T-----	3349
CRF04_cpx_CY.94CY03	-A---T---C-----C---T-----G-----C-----G---CAA-C-G---G---T---	2670
N_CM.YBF30	A---CT---G---C---T-----G---C---G---C---A-----A---C-----C---A---AA-T-A---G---TGTA-----	2898
O_MVP5180	A-----T---T---C---T---G---G---C---TT-A---A---G---C---A---C---G---C---CAAT---T---C---AGTG-----	3334
O_CM.ANT70	A-----T---C---T---C---T---G---A---G---C---T---A---G---G---A---C---G---T---C---CAAT---TA---C---G---TG TG-----	3358
HIV2_A_ST	-CCT---T---T---C---TG-G---GTTT---A---C---A---AA---C---GTGG---AAC---AGC-G---AAGA---CAAT---CC---G---AGTA-----	3134
HIV2_A_BEN	-CCTA---GT---T---T---TG-G---GTTT---A---C---T---AA---C---G---ATGG---AAC---AGC-G---AAAC---CAA---CC-G---TA-----	3689
HIV2_A_ALI	-CCTA---T---T---C---TG-G---GTTT---A---C---A---AAA---C---G---GTGG---AAC---AGC-G---AAAA---CAAT---CC-G---AGTA-----	3682
HIV2_B_D205	ACAT---T---CT---T---C---AG---GTTT---A---C---G---AAA---G---G---TGG---AA-A---G---A---C---G---AAAA---CAA---AGTT-----	3662
HIV2_B_EHO	ACAT---T---CT---T---T---AG---GTTT---A---C---G---AAA---G---G---TGG---AA-A---G---A---C---G---AAAA---CA---A---GTT-----	3660
STMM83293	AT-TA---T---T---C---T---AG-G---TT---G---T---T---AA---A---GT-ATGG---AACA---G---A---C---G---AAAA---AAT-A---TC-G-G---GTT-----	3253
SMM251	-CATA---GT---CT---T---C---AG-G---TTC---A---T---C---T---AA---G---C---T---GTGG---GACA---AGT-G---AAAG---A---T---C---G---G---C-----	3582
SMM9	-TATA---T---T---C---AG-G---GTTT---A---T---C---AG---A---T---GTGG---CACC---A---C---G---AAAA---A---T---GC---G---A---C-----	3079
SIVCPZANT	TC-----AGAG---T---T---G---G---A---T---AA---A---GT-A---A---C---T---AAA---AA---A---GCC---TGAT-----	2743
SIVCPZGAB	AA---GT---C---A---C---A---G---G---A---C---TT-A---A---G---C---A---C---C---A---C---TCAAT---A---GGTA-----	3365
SIVCPZUS	CT---T---C---T---C---T---T---G---G---G---G---A---T---A---C---A---G---G---C---CA-T-A---C---A---TT-----	3364
SIVSTAN1	TAA-A---CT---CGAG---G---T---G---GGTG---AG---ATGA---A---CA-GT-GT-C---TC-C-C---GTC-CA-CAAA---A---A---A---GC---ACA-----	3506
SIVGRI677	AAAAA---T---AGAA---C---G---GTC---A---G---G---GGGAA---G---A---T---G---GA---CA-TA-CAAA---AAT-A---CCCTT---AG-AGAA-----	3310
SIVVER9063	CA---T---AGAA---C---G---GTTG---A---GA-G---A---ATGAA---A---CA---T---ATGG---AC-A---CA-C---A---TAG---AACT-AGA---G---ACAA-----	3534
SIVVER155	AA---T---AGAA---C---G---GTTG---A---G---C---TGAG---A---A---T---ATGG---C---AGT---AAG---AAT-AGAGA-G---ACAA-----	3526
SIVVERTYO	CC---CT---AGAA---C---A---G---GATG---A---A---T---ATGAG---A---CA---TTGG---C---GA-C---AGCAGA---CAA---GAG---TGAA-----	3024
SIVVERAGM3	CC---T---AGAA---G---GTA---A---A---T---ATGAG---A---CA---T---GTGG---AC-C---G---CAGA---AGCAGC---AAT-AGA---C---AGAA-----	3029
SIVLHOEST	CA---AACT-AGAG---A---G---GTTT---AGC---G---A---C---AT-A---T---G---GG---GA-A---G---AAG---TAAAT-A---G---TG---TCTAACAAAAAC	3662
SIVSUN	AA---AAT-AGAA---G---T---GG---G---A---C---ATAAG---G---CAT-T-A---C---A---C---G---GAGA---G---AAG---AAAT-A---CCCTCT---AGAAGAG...CC	3664
SIVMNDGB1	CA---AACT-AGAA---T---A---G---T---C---T---A---G---GT-A---C---A---G---AGA---G---AAAGG-TCAA---A---TT---CAGAACAG...CC	2994
SIVSAB1C	CT---GT---AA---C---T---A---G---GTT---A---G---T---A---A---C---G---T---GT---AA-G---CA-C-G---GAG---CACCC---A---GG---AGAA...CC	3702
SIVSYK173	AAGT---T---AGGTC---A---AG---TGG---G---CC---TA---A---G---AG---T---A---CACT---T---A---C---G---CAGT-G---AAA---AA---C---TA-T-T---TGA-GAA...AT	3501
POL	K_V_G_F_K_V_P_K_E_K_W_Q_D_Q_Y_P_M_Q_W_L_G_Y_T_L_H_P_D_K_W_Q_L_Q_K_I_E_L_P_N_I_D_D_E...I	POL

B_FR.HXB2R	GACTGTCAATGACATACAGAAGTTAGTGGGAAATTGAATTGGGCAAGTCAGATTACCCAGGGATTAAAGTAAGCAATTATGTAAACTCCTTAGAGGA...ACCAAAGCACTAACAGAAAGTAATACCA	3431
POL	T V N D I Q K L V G K L N W A S Q I Y P G I K V R Q L C K L L R G . . T K A L T E V I P	POL
A_UG.U455	-----T-----A-----A-A-C-A-----TG-----A---C-G-----T--C-G---...G-----TA--G--A-C	2877
C_ET.ETH2220	-----T-----A-----A--A-C-----C-----T-----A-----A-----C-G-----G---...G-----CA--G--A--	2823
D_ZR.84ZR085	-----T-----A-----A--A-----C-----T-----A-----A-----G-----G---...G-----G-----	2953
FL_BE.VI850	-----T-----A-----A--A-C-A-----T-----C-----G-----G---...G-----CA--G--G--	2768
G_SE.SE6165	-----T-----A-----A--A-C-A-----T-----C-----C-----G-----G---...G-----CA--G--T--	2828
H_CF.90CF056	-----T-----A-----A--A-C-A-----T-----AAT-----A--C-----G-----G---...G-----T-----CA--G--	2778
J_SE.SE92809	-----T-----A-----A--A-C-A-----T-----A-----A-----C-----G-----G---...G-----CA--G--	2745
CRF01_AE_TH.CM240	-----T-----A-----A--A-C-A-----A-----TG-----G-----A-----C-G-----C-G---...G-----CA--G--	3005
CRF02_AG_NG.IBNG	-----T-----A-----A--A-C-A-----TG-----A-----A--A-----C-G-----G-----C-G---...G-----TA--G--G--	2956
CRF03_AB_RU.KAL153	-----C-----C-----A-----A-----TG-----A-----A--A-----C-G-----G-----C-G---...G-----	3476
CRF04_cpx_CY.94CY03	-----C--T--C-----A-----A-C-A-----T-----T-----A-----A-----G-----G---...G-T-----C-----CA--G--	2797
N_CM.YBF30	-----T-----A-----A--A-G-A-----C-----T-----A--C-G-----AA--GC--C-----T-AA--C-----TT-G-----G-CAAC	3025
O_MVP5180	---A--A--T-----A--A--A--A-----A--C--T--A--A--G-----AAG--G--C--GT-AA--C-----T--T--G-----G--G--T	3461
O_CM.ANT70	---A--A--T-----A--AC--A--A--A--GC--A-----GCA--A--C--T--A--A--G-----G--AG--G-----GT-AA--C-----C-----GT--T-----G--T	3485
HIV2_A_ST	---A--A--C--A--AC--G-----TGTC--A-----GCA--A--C-----A--GACC--AA--C-----GGT-AA--C-----GGA-AA--TGA--C-----AGG--AG	3261
HIV2_A_BEN	---A--A--C--A--C-----AGTC--A-----GGCA--A--C--TT--A--A--ACC--AA--C-----G--AA-----GGA-AA--TGA--C-----A--G--G--AG	3816
HIV2_A_ALI	---A--A--C--A--ACAT-----TGTC--A-----GCA--A--C-----A--A--GACC--AA--C-----G--AA-----GGA-AA--TGA--C-----G--G--AG	3809
HIV2_B_D205	---A--G--CA--T--A--AC--G--A--AGT--A--C-----GC--AC--C--TT--T--A-----GAC-----CA-----AA-----G-----GGA--AG--TGA--C-----A--G--AG	3789
HIV2_B_EHO	---A--A--T-----A-----A-----AGT--A-----GC--AC--T--G-----GACC-----TA-----AA--A--G-----GGA--AG--TGA--C-----AGG--A--	3787
STMM83293	---A--G--T--C--A--AC--G--A--GTC--A--C-----GCA--A--T-----A-----ACT--AA--TC--G--C--T--AA--C-----GGT--AG--TG-----G--T--G--AGG--AG	3380
SMM251	---A--G--T-----A-----A-----AGT--A-----GC--A-----T-----T--A-----ACC--AA--TC--C--GGT--AA-----GGA--AA--TGA--T-----G--A--G--T--AG	3709
SMM9	---A--A--A--T-----A--A-----A-----AGTGC--A-----GCA--A--T-----A--A-----GACT--AA--TC--T--C--G--T--AA--C-----GGA--AA--TGA--TT-----AGG--T--AG	3206
SIVCPZANT	---A--T-----C-----A-----A-----A-----A-----A-----A-----C-----A--C-----ACT--A--G--G-----G--A--C-----GT--GAGT-----TAG--G--A--	2870
SIVCPZGAB	---A--T-----T-----A-----A-----A-----A-----A-----A-----A-----A-----A-----A-----G--GA--A-----A--GAA--G-----T--G--T--T	3492
SIVCPZUS	---A--T-----T-----A-----A-----A-----C-----C-----A-----T-----A--A-----A-----A-----G-----C--T--AA--A--A-----G--T-----T-----TG--AAT	3491
SIVSTAN1	---A--G--T-----T-----A--A-----A-----AGTG--A-----T-----AT--A--T-----A-----GAC--AAGCCC--C--T--A--AC-----G--TC--ATT-----CT--G--G--A--	3633
SIVGRI677	---A--A--CA--A-----G-----A-----AGTTC--A-----A-----T-----A-----ACC--AA--TACC--GCCA--GT--G-----G-----AA--GAAC--CCT-----A--G--T--	3437
SIVVER9063	---T-----T--C--A--AC-----A-----C-----GC--C-----C-----T--A--G--AC--AA--GA-----C--GGT--AA--C--G-----AG--AAT--TT--G--A--G--A--	3661
SIVVER155	---A--G--TC--T-----A--G--A-----A-----GCA--AT--A--T-----AT--G--G--AC--AAA--TA--C-----G--A-----AG--AAT--TT--C--GG--GA--	3653
SIVVERTYO	---A--T-----C-----T-----C-----A-----GCA--AT--G--T-----TC-----GAC--A--A-----C--GT--AA--C-----GGA--AG--AAT--GTT--GC--G--G--T	3151
SIVVERAGM3	---A--T-----A-----GAC-----C-----A-----GCA--C-----AC--C--G--ACT--AAA--C--G-----T--AA--C-----GGA--AA--GAAC--CT--AC--G--A--C	3156
SIVLHOEST	T--A--A--C--A-----AC--G-----AGT--C--T-----GC--AT--G--TGAT--T-----GGAC--AAG--C--C-----T--AA--A--G-----GGAGTA--C--CT--GGA--A--C--AAC	3789
SIVSUN	A--A--G-----A-----AA--T--A-----AGTTA--C-----TTCA--AT--A--TGA-----C--C--G--AC--AAG-----AA-----GGG--AT--C--T--G-----GAA--G--AAG	3791
SIVMNDGB1	A--A--A-----A-----A--G--A--T-----A-----TGCA--T--A--T-----C-----AC--AA--C--G--C--G--AA--A-----GGAGGACT--AACA-----GAA--G--CA--	3121
SIVSAB1C	---A--G--T-----T-----A--A-----A--A-----G--A-----A-----A-----GACT--AA--T--G--G-----AA-----GGAG--T--G--C--TT--G-----GA--G--C--A--	3829
SIVSYK173	C--A--G--C--GT-----A--GA--A--AGT--C--T-----C-----TT-----A-----GAC--AAG--GC--C-----GTGTA--C-----GGA--T--CA--T-----C--CA--	3628
POL	T V N Q L Q K L I G V L N W A S Q I Y S G I K T K E L C K C I R . G T K Q L T E V L T	POL

B_FR.HXB2R	CTAACAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATT.....CTAAAAGAACCAGTACATGGAGTGTATTATGACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGCCAG	3552
POL	L T E E A E L E L A E N R E I L K E P V H G V Y Y D P S K D L I A E I Q K Q G Q	POL
A_UG.U455	--G-T-----AT---T-----G---G-----.....-----C-T-G-----A-----C-G-----A-A-----	2998
C_ET.ETH2220	---T-----AT---T-----G---G-A-----.....-----G-----A-T-----T-----A-----A-T-----	2944
D_ZR.84ZR085	-----AT-----G-----G-----.....-----G---A-G-----T-----T-----A-----A-----	3074
FL_BE.VI850	--G-T-C---G---T---T---A---T---G-----.....-----G-----G-A-----A-----A-A---AG-C-----	2889
G_SE.SE6165	--G---C---AA-G---GT---G---G-A-----.....-----G-----T---C-----A-----G---A-A---A---T-----	2949
H_CF.90CF056	--G---A---G---AT-G---T-----G-----G-----.....-----G-G---A---A---T---G---A-----G---A-----	2899
J_SE.SE92809	T-G-TAG---AT-G---A-----A-----.....-----G-A-----G-A---T---G---A-----G-G---A-A---TG-----	2866
CRF01_AE_TH.CM240	--G-T-----AT---GT---G---G-----.....-----ACC-T-G---A-----G-----G---A-A---G-----	3126
CRF02_AG_NG.IBNG	--G-T-G---AT---T---G---G-A-----.....-----T---G-A-----A-----G---T-----A-A-----	3077
CRF03_AB_RU.KAL153	-----C-----G-----G-----.....-----G-----G-----.....G-----A-----	3597
CRF04_cpx_CY.94CY03	-----TAC---G---T---T-A---G---G-----.....-----G-CA-----G-CA-----A-----	2918
N_CM.YBF30	T-T-----AT---A---A---G---A-----.....T-----CC-G---C-----GG---A---G---T---A---A-A-----	3146
O_MVP5180	T---GTA---G---A---T-A-A-----A-AG-----.....-----G-----A---CC-G---TGAC---GTGG-TTAGT---T---AG-----	3582
O_CM.ANT70	T---GTAG---G---G---T-A-AG-----A-GG-----.....T---C---G---G-A-C---C-A---TGAT---G---TC---TGG-TTA-T---A---AGGG-----	3606
HIV2_A_ST	TGG-----TT---AGCG---A-A---A-ATC-C-----.....T---GCC-GGA-CA-G-A---TGC---CC-AGAGGA---G---GC---GA---AC-G-C-A---AG-TCAAG-CA	3382
HIV2_A_BEN	TGG-----CT---AGC---G-A-A---A-ATT-C-----.....T-G-GCC-GGA-CA-G-A---TAT---CC-AGA-GA---A---GAG---AC---C-A---AAGCCAAGG-C	3937
HIV2_A_ALI	TGG-----CT---AGC---A-AG-G---A-ATT-C-----.....T---GTC-GGA-CA-G-G---GCAC---CC-AGA-GA---G---G---GA---AC-G-C-A---AG-TCAAG-CA	3930
HIV2_B_D205	TGG-----CT---AGC---G---ACAG---G---T-A-ATC-C-----.....T---G---C-GGA-CA-G-A---TCC---C---CA-GGA-AGGGT-CCGC---GA---AC-G---AA-CCTAGC-A	3910
HIV2_B_EHO	TGG-T---TTG---GC---T-CCAG---A-ATC-C-----.....-----G---C---GAGCAGG-A---TCC---CA-AGA-GGGGT-CCT---GA---AC-G---AA-TCTAGC-A	3908
STMM83293	TGG-----ATG---GC---TAT---T-A-ATT-----.....-----C-GTC-GGAGCAGG-A---GCAA---C---CCGAGA-GAC---GCCTC---GAG---AC-G---GT---AG-TCA-G-CA	3501
SMM251	TGG-T---GATG---AGC---TAT-AG---T-AGATA-----.....-----C-GTC-GGA-CA-G-A---TGT---CC-AGA-GGC---GCCA---GA---CACGG---ATA---AGTCA-G-CA	3830
SMM9	TGG-T---GATG---GC---TAT-A---AGATA-----.....-----CC-GTC---GA-CC-G-G---TGT---C---C-GGAGGG---GCCAC---GAG---CC-G---ATA---AGTCA-G-TA	3327
SIVCPZANT	A-G---TAGG---AT---T-A-A---T-A-C-A---NNNNNNNNN---C-GC-GAA-A---G-G---TAC---C---G---TGGTCT-CCAC---A---ACC-----A-A---TC---	3000
SIVCPZGAB	--C---CC---AT---T-A---T-G---A-----.....G---GCAC-----G-A-C---T---GAC---GC-T-----A---CA-CT-----	3613
SIVCPZUS	T-T---C-T-G---AA-G---GT-A-A-----A-----.....-----G-----G-C---GA---A---G-----G---A---A---AAGGA-----	3612
SIVSTAN1	TGG-G---GC-G---G---GT-CAG-C---TCA---C-----.....T-----GAGCA---G---CT-C---AGT---TCT---G-G---CACT---TACC---T---ATG-T-----	3754
SIVGRI677	TGG-----G---GC---TATAAGA-C---TCA-G-A-----.....G-GC-G---A---CA-G-A---ACA-C---TCTC---A-----AC-G-T---A---A---AG-----	3558
SIVVER9063	TGG-T---G-G-AGC---GTAT---G---A---A-----.....T-G---AC-GA-CA-G-A---ACT---CA---GG---GCCTC-T-G---C-G-C---ATTA-AAG-T-----	3782
SIVVER155	TGG-CCC---G---AGC---GTAC-A---AG---C-----.....-----AC-GAGCA-G-A---TACT---CA---GA---CC-C-T-GG---C-G---ATTA-AAG-T-----	3774
SIVVERTYO	TGG---CCT---G---AGCT---TAT-----TG-----.....T---AC-GA-CAGG-A---ACC---CA-A---GG---T-CCTA-T-GG---C-G---ATT---AAGG-----	3272
SIVVERAGM3	TGG---G---AGC---TAT-----A---C-----.....T---ACGGA-CAGG-A---GACC---C---CA-A---GG---G-CC-A-C-G---C-G-G---A---TA-AAGG-----	3277
SIVLHOEST	TGG---G---CTT-GA---TAT-GGC-G---A---G-A-----.....T-----AAGA-G-G---CC---GA---G---AC-T---T-TCAGGG---C-AA-CAAAA-G-----	3910
SIVSUN	A-G-----TAG-GA---TACCAGAGT---TCAG---G-A-----.....T---C-G---GT---TCA-CAGT---C---A---GAT---A---TTGCAG-G-G---A---GTAAA-----	3912
SIVMNDGB1	A-G-----AGA-G---TAT-A-C---T-A-----C-----.....T-GGCT---GA-CA-G-A---GTCT---T---TAAT---G---A---TAT-T-AG-T-T-----AACCAACAGG-----	3242
SIVSAB1C	TGG-----G---A---T-A-A---T---C---C-----.....T---G-C---AAGCAG-G---CAA-----TG-TCTTCT---GG---TA-GG-T-TT---A-T---AG-T-----	3950
SIVSYK173	--T-----C---T---GC---T-A-A---G---T---C-A-C-----.....T---G---GA-CAGGCA-GTCT---TAA---CCT---GA---C-C---CACT---A-TA---TCCC-----	3749
POL	L T E A A E A E L E E N R Q I L K E E Q A G S Y Y D P K K P L E A H I T K L G S	POL

B_FR.HXB2R	GCCAATGGACATATCAAATTTATCAAGAGCCA...TTTAAAAATCTGAAAACAGGAAAATATGCAAGAATGAGGGGTGCCCCACACTAATGATGTA AAAACAATTACAGAGGCAGTGCAAAAAATAACCAC	3679
POL	G_Q_W_T_Y_Q_I_Y_Q_E_P...F_K_N_L_K_T_G_K_Y_A_R_M_R_G_A_H_T_N_D_V_K_Q_L_T_E_A_V_Q_K_I_T_T	POL
A_UG.U455	A-----TC-----T-----A-TG-----G-GT----	3125
C_ET.ETH2220	A-----T-----T-----C-----A-----C-----G-G-T-----A-GAG-AC-----G-G-----C-T-----G--G--CT	3071
D_ZR.84ZR085	-T-----G-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----T	3201
FL_BE.VI850	-G-----GA-C-----A-----G-----A-G-----TCG-----A-----A-----G-----G-TTT	3016
G_SE.SE6165	A-----AC-----G-----G-----A-G-G-TC-----G-----G-----A-T-----G-----	3076
H_CF.90CF056	-----G-----G-----A-----AAC-----A-----A-----G-----T-T--	3026
J_SE.SE92809	A-----G-----C-----G-----A-G-A-----G-----G-----A-TG-----G--TT	2993
CRF01_AE_TH.CM240	A-----A-----C-----G-----C-----GAG-TC-----T-----G-----A-TG-----G--G--	3253
CRF02_AG_NG.IBNG	A-----T-A-----A-A-TC-----G-----T-----G-G-T-T	3204
CRF03_AB_RU.KAL153	-----C-----C-----G-----G-----G-----G-----G-----G-----	3724
CRF04_cpx_CY.94CY03	-T-----A-----CA-----G-G-----CC-ATC-----T-G-----A-----G--G--T	3045
N_CM.YBF30	-T-G-----G-----G--TT-----CA-----T-A-----G-----A-----ATC-----T-----A-----G-----GTT-A-TG--AAGG--G-GG-A--	3273
O_MVP5180	-G-----T-C--GG-A-----G-TGA...CA-G-C-T-----T-GCAA-A-CCT-----A-----A-----G-----GG--A-T--C-G--GG-GT-TCA	3709
O_CM.ANT70	AG-----T-C-G-A-----G-AGA...CA-G-C-C-----G-----A-T-GCAA-A-CCT-----A-----A-----G-----G-----A-T-A-C--G--GG-GT-TCA	3733
HIV2_A_ST	AT-G-----A-G-AC-C-G-GA...GGA-T-A-GT-----G-----AGG-A-AAAA-A-----C-C-GA-C-G-TCC-G-C-A-T--T-----GGA-A	3506
HIV2_A_BEN	AT-----CA--AC-C-G-A...GAG-TC-A-GT-----G-----AG-A-AAAA-A-----T-C--GG-C-G-TT-C-G-C--T--T-G-----GGA-A	4061
HIV2_A_ALI	AT-----A--AC-C-G-GA...GAA-T-A-GTG-A--G-----A-----AAAA-A-----T-C-C-GG-C-G-TTG--G-C--T--T-----GGA-A	4054
HIV2_B_D205	AT-G-----CA--C--G-GA...AA-GTC-A-GT-----G-----AGG-T-AAAAA-G-----C-C-GG--G-T-C-GG--C-T-T--T-G-----GG-A	4034
HIV2_B_EHO	AT-G-----CA-G--C--G-GA...GA--TC-A-GT-----G-----AGG-T-AAAAA-T-----C--GA--G--T--GG-TC-T-T--C-----GGA-A	4032
STMM83293	AT-----T-C--A-G-AC-C-----A...GAC-TAT--GT--C-G-T-----A--A-AAAAA-A-----T-A--GA-T-G-TT--GG-C-T-TGA-T-G--G--GGG-A	3625
SMM251	AT-G--T-T-A--C-C-----A...GAC-TA--GT-----T-----AG--A-A-AA-A-A-----T-C--GA-T-G-T--G-C-T-T-A-A--G-----GGA-A	3954
SMM9	AT-----T-----A-----C-----GAC-TA-A-GT--C--T-----AG-T-AAAA-A-A-----T-A--GA-C-GCTT--G-C-T-T-----G-----GGA-A	3451
SIVCPZANT	-A-----C-----A-----A-TGA...GGA-CTGT-A-G-----T-----GCCT-CA-AA-T-----G-T-GG-----G-T-GA-T-----GGATT	3127
SIVCPZGAB	-G-----T-----G-A-T-G-A-----CA--G--T-----G-----C--CAA--TCA-A-----A-----CA-C-GG--C-GG-G-A-----A-----TG-T--	3740
SIVCPZUS	-T-G-----C-----T-----A-GG...CA--G--A-G-----C-----CAA--ATCA-A-----T-----A-C-G--GC-GTT--A-TG-----G-T--	3739
SIVSTAN1	-G-----GG-----T-CA-A-G-----GGA-GTA-----A-----AG-GC-AAAA-A-----T-T-A-----GT-T-G-GT--GG--GAATG-C--GGG-GTG--A	3878
SIVGRI677	-G-----C--T-CACC--A...GGGGC-GTAT-A-GGTG--G-----C-AGCAA--A-AAA-T-----C-----GGACTC-G-C-CCTT-C-----CTGT-A	3682
SIVVER9063	-G-----GT-----GT--A-A-G-A...GGA--TCT--GT--G-----G-----A-CAA-A-AAAC--T-A-----AC-TCGCAC--GG-C-GCCT--A-----CTGT-A	3906
SIVVER155	-G-----T--C--T-CA-G-G-A...GGA--TCT-A-GGT--G--G-TC-C-A-CA-AA-C-A-T-----C-----GT-CGCTGT-C-G--GA-T--A--G-----GGG-A	3898
SIVVERTYO	-A-G--GT--C--T-CA-A-----A...GGAC-GTCT--GT-----CA-C-AGCAA-A-AAAC--T-A-----AC-TCGCAC--G-T-GTTT--G--G--TTG--A	3396
SIVVERAGM3	-T-----GT--C--T-CA-G-----GGAC-GTAT-A-GT--T-----CA--AGCA--AAAAA-T--T-C-----GT-CCGTGT--GG--GATT--A-----C-TTGT-A	3401
SIVLHOEST	-GAT-ATT-T-TC-GTGGAGA--GA...AA--C-TCT-A-G-G-T--G-GG--CA--CA--A-CA--A-----A-----CCCC--C-GA--GT--A--TA-T-G--G--GGA-A	4034
SIVSUN	-AAT-TA-C-T--TGG-TG--G-GA...AAAC-GTAT-A-GGGT--G-----CA--AG-GAG--CA--A--TGAG--CC-TGTC--G--G--CT--CC-A-----G--GGA-G	4036
SIVMNDGB1	-AG-TATAT--T--TGGA-G--GA...AA--GGT-T-A-G-G--G-----GG-A-CA--AAACA--A--T-G-----CC-C-TGA--GG--GT--TAC--G--GG--GGA-G	3366
SIVSAB1C	-----GG--G--A-C--GCCAGA...AA--TA--GGTT--C-----A--C-AAACA--A-----C-----AT--G-ATG--G--GCCT--A--G--G--GGA-A	4077
SIVSYK173	AG-----GG--CATG--A-G-----A-AAAGGCCC-CCC-A-T-----GACA--A--CCTTT-CA--A--T-G-----CTATC--TCTC-TG-TC-ATTGT-AA-T-----GGG-T	3879
POL	Q_Q_W_G_Y_M_I_K_Q_E_Q_K_G_P_P_L_I_T_G_K_T_A_K_G_F_A_A_H_S_N_D_Y_Q_S_L_A_Q_L_L_N_K_I_G_I	POL

B_FR.HXB2R	AGAAAGCATAGTAATATGGGGAAAGACT...CCTAAATTTAAACTGCCCATACAAAAGGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTCTGAGTGGGAGTTTGTTAATACC	3806
POL	E S I V I W G K T . . . P K F K L P I O K E T W E T W W T E Y W Q A T W I P E W E F V N T	POL
A_UG.U455	-----TC...-----G--A-----G-----TG-----G--T-----A-----C-----	3252
C_ET.ETH2220	G-----G-----G--A-----C--G--A-----G-----C-----A-----C-----	3198
D_ZR.84ZR085	-----G-----G-----G--A-----C-----T-----C-----A-----A-----C-----	3328
FL_BE.VI850	-----G-----GATCT-----G-----A-----TG--A--G-----T-----T-----A-----C-----	3146
G_SE.SE6165	--G-----C-----C-----A-----G--A-----G--A-----T-----C-----G-----C-----	3203
H_CF.90CF056	-----A-T-----G--A--T-----A-----G--C-----A-----A-----C-----	3153
J_SE.SE92809	G--GC-----A-----G--A-----GA-----G-----C C-----G-----C-----	3120
CRF01_AE_TH.CM240	-----C-----G--A-----GA-----TG-----G--T-----A-----C-----	3380
CRF02_AG_NG.IBNG	G-----C-----G--A-----GA--G-----TG-----G--T-----A-----C-----	3331
CRF03_AB_RU.KAL153	T-----A-----A-----A-----A-----A-----A-----A-----C-----	3851
CRF04_cpx_CY.94CY03	G--T-----G--G--T--A-----C-----A-----G-----C-----A-----A-----C-----	3172
N_CM.YBF30	---T-----T-----G--T--A--AG-----GTG---GG-----C TC-----A--T-----A-----C--C--T	3400
O_MVP5180	---GCT---T-----G--ATTA...---C--GG---AG--TACT--GA---T---T---G--A-----G-----A-----C--GC--A	3836
O_CM.ANT70	---TCT---A--T--C-----ATTG...---G---AG--CACT--GA---T---G--C-----G--A-----A-----A-----C--GC--A	3860
HIV2_A_ST	---GCAC---C--T-----CGA--TA...---A---C--C--A--AG--G--GA--T--C---CAG---GATA--C--C---TG--A---C--A--C---C--CA--ATC---	3633
HIV2_A_BEN	---GGCAC---C--T-----CG--TA...---A---C--C---AG--GG--G--GA--G--C---GCAG---GATA--C--C---TG--A---C--A---C---C---ATC---	4188
HIV2_A_ALI	---GCAC--G--C--T-----CGA--TA...---A--G---C--C--A--AG--G--GA---C---CAG---GAT--C--C---TG--A---C--A--C---C---ATC---	4181
HIV2_B_D205	---GC--C---C--C-----G--TA...---AGTG--CC--T---AG--G--GA--G---CCAG---T--C---TA---C--A---C---C---CTCG---	4161
HIV2_B_EHO	G---GCAT--G--C--C-----G--TA...---A--TG--CC--T--A--AG--G--GA--G---TCAG---T--C---TA---C--A--A---T---CTCA---	4159
STMM83293	---GCA---G--C-----C--TC...---A---CC--CT--AG--G--GA--GGTC---CA---G--T---G--TA--A---A--A---T---ATCC--A	3752
SMM251	G---GCA---G--C-----C--GTC...---A---CC--CT--A--AG--TG--G--G---TGT---CAG---C---G--TA---A--G---T---A--CTCA--G	4081
SMM9	---GCA---A--T-----C--GTG...---A--G---C--TT--AG--G--RA---TT---CA---T---TA---A--A--A---C---GTCA--A	3578
SIVCPZANT	---T---T--C-----G--GTA...---A---C--T--A---ACT--G---T---TG--C---T--C---G---A--A--A---C---A	3254
SIVCPZGAB	T-----T-----T-----A--A...---GGT--A--AG---G--A---GT---G---G--A--C---G--A---C---A--A--C---	3867
SIVCPZUS	T--G-----T--C--T-----GTA...---A---T--TAG--G--G---GTC---T--A---G---A--T---A---C---	3866
SIVSTAN1	G---GCG---CACCT--T-----ATTA...---C--TCA--GG--G--A--AG--G--GA---GT---CA---GT--C---T---A---C--A--A---ATA--A--CA--CTCC--A	4005
SIVGRI677	G---GCAC--TACC--T-----GACT...---ACG--G--AC---C--AG--G--C---A---T--TG---CAG--C---TAT---A--A--A---GC--A	3809
SIVVER9063	---GGCAC--G--C--T-----G--ACTA...---A--TT--G---C--AG--G--G---GT---CA---G--C--C---TG--G---A---A---G--G--GTA	4033
SIVVER155	---GGC--C-----T-----C--ATTA...---C--CT--G---C--AG--GG--G--G---C---CA---G--C---T--GT---A--C--A---C---C--G--GTT	4025
SIVVERTYO	---GCTC---T--T-----G--TATTA...---AGTTC--AG---C--G--G--GA---GGT---GCA---G--G--T--C---G--TA--G---C--A---T---C--GC---	3523
SIVVERAGM3	---TCTT---T-----G--TTG...---AGTCC--G---C--A--G--G---GT---CA---G--T--T--C---G--TA--GT---A--C---A---C--G---	3528
SIVLHOEST	---C-----C-----CTTTGTG...---A--A--C--G--C--AG--GACT--GA---GTC---GCAC---GC--CC--C---G--TT--A---A--C--TA--A--CA--TC---	4161
SIVSUN	---T-----T-----TTTTGTG...---A--GA--AC--GG--A---GA--G--T---AGTCAG---G--C--C---GTGT--A---C---TA--A---C--GC--T	4163
SIVMNDGB1	-----T-----C-----TTTTGTA...---A--A--GC--GA--A---CTAC--G--G--T---GAT---CAT---GTGT--A---A--A--AGTA--A---A--C--GC--A	3493
SIVSAB1C	G-----C--C-----C--TA...---TTA--GG---AG--G--GA---GCTC---CA---T--T--C---G--T---C--G--A---AA--G--C--G---T	4204
SIVSYK173	CC--TCAC--TGGTAT-----GTA...---A--C---CC--T--C--AG--A---GA---GGA---A---C---A---G--A--A--AGTCA--A--CA--A--G---	4006
POL	Q S L W Y W G K V . . . P T F H L P V K R E E W E K W W T D Y W Q A T W V P E V K F I S T	POL

	Pol p66 reverse transcriptase continue (RT is dimer of p51 and p66)			
	Pol p51 reverse transcriptase end \ Pol p51 Rnase H			
B_FR.HXB2R	CCTCCCTTAGTGA	AATTATGGTACCAGT	TAGAGAAAAGAACCCATAGTAGGAGCAGAAAACCTTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAAGCAGGATATGTTACTAATAGAGGAA	
POL	P	P	L V K L W Y Q L E K E P I V G A E T F Y V D G A A N R E T K L G K A G Y V T N R G	
A_UG.U455	----	C----	A-----C-----G-A-----T-----A-----GC-----G-----C--G-C-----	3382
C_ET.ETH2220	----	C-----	C-----C-----T-----T-----A-----T-----A-----A-----G-----G-----G-----	3328
D_ZR.84ZR085	----	T-----	A-----G-----A-----T-----T-----AA-----T-----AA-----G-----C-----	3458
FI_BE.VI850	----	TC-----	C-----A-C--G-----C-----C-----T-----T-----A-----C--AA-----G--A-----	3276
G_SE.SE6165	----	TC-----	A-G-----T-G-----A-C-----CC-----T-A-----T-----A-----G-----G--C-A-----	3333
H_CF.90CF056	----	ATC-----	A-----T-----A-C--G-----C-----T-A-----A-----T-----A-----C-----G-----C-----	3283
J_SE.SE92809	----	C-----	TC-----A-G-----A-----A-G-----A-G-----T-----T-----T-----A-----AC-----G-----G--C-A-----	3250
CRF01_AE_TH.CM240	----	TC-----	A-----A-----A-----C-----G-----T-----GT-----GC-----G-----C--G-C-----	3510
CRF02_AG_NG.IBNG	----	TC-----	A-----A-----A-----C-----T-----A-----T-----A-----GA-----G-----C--G-C-----	3461
CRF03_AB_RU.KAL153	----	-----	A-----T-----A-----A-----C-----C-----G-----C-----G-C-----	3981
CRF04_cpx_CY.94CY03	----	TC-----	A-----A-C--C-----C-----T-----T-----A-----A-----A-----CAG-----G-----C-----	3302
N_CM.YBF30	----	C-T-----	T-----A-C--G--A--CAGT--G-----T-----A-----T-----A-----A-----G-----T-T--G--AG--G-----	3530
O_MVP5180	----	C-A--GA--C-----	A--C-----T--T--G-----T-----A-----T-----A-----A-----C-----G--G-----AG--ACA-----	3966
O_CM.ANT70	----	C-A--GA--C-----	AG--C-----A--GT--T--TA--G--G-----AT-----A-----T-----A-----A-----C-----G-----AG--ACA-----G-----	3990
HIV2_A_ST	----	G--AC--G--C--G-----	GTA--TTA--CC--G--T-----T-----C-----C-----CAC-----AT--CTGC--T--A--C--AT--A--G--GA-----CA--A--AG-----	3763
HIV2_A_BEN	----	A--AC--G--C--GG-----	ACA--TTA--CC--TAGG--T--T--CC--C-----G-----CAC-----AT--TGC--T--AC--T--A--GAG-----A--AG-----	4318
HIV2_A_ALI	----	A--A--A--A--G--C-----	GCC--A--CC--TC-----C--C--A--GAG-----A--CAC-----C--CTG--T--C--AT--A--GA-----A--A--AG-----G-----	4311
HIV2_B_D205	----	A--A--A--A--GG-----	GCC--TA--CC--G--TC-----C--C--A-----T-----T--A--CAC-----AT--CTG--AA--CCT--A--GA--G-----C--AG--C--G-----	4291
HIV2_B_EHO	----	T-----	A--G-----GTC--TTA--T-----TA-----T--G--A--A-----A--T-----G--CT--TG-----C--AT-----GA--G-----C--CA--A--G-----	3882
STMM83293	----	A--AC-----	A--G-----GTC--T--A--TC--T--G--C--T-----AG--A-----AT--AC-----AT--TG--T--AAC--T--A--GA--G-----A--C--AG-----G--C-----	4211
SMM251	----	-----	C--G--C--GTT--T--A--CC--TA-----G--T--CAG-----A--T-----A--AT--CTG--T--C--T--A--G--GA-----C--G--AG-----C-----	3708
SMM9	----	C--A--A--T--G--C--C-----	TA--TC--TT--GC--C--T--TCC--AG--T-----T-----T-----T--AA--CT--C--G--G--G-----C--G--AG--C-----	3384
SIVCPZANT	----	-----	C-----AGT-----C-----T--CC--ACCA--T--T--AT-----A-----A--T-----A--A--ACT--G-----A--AG--C--A--G-----	3997
SIVCPZGAB	----	-----	C--C--A--C--T-----TA-----C--GG-----AG--G-----A--T--G-----A-----CC--AC--A--A--C--GAA-----G-----G--G--T-----	3996
SIVCPZUS	----	G-----	A--C-----TA--C--G--TA-----T-----CC--AA-----A--GGTG--AT--TGC-----TGC--T--AA--TT--C-----G-----C--TCC--A-----C--C--G-----	4135
SIVSTAN1	----	A--T--C-----	A--C--G-----TTCC--TA-----A--CAA-----A--TGTT--AT-----G-----ATC--AA--T--C-----T--G-----C--GT--AG--G-----	3939
SIVGRI677	----	A--TC-----	G-----ACTC--GACC--G-----T-----CC-----AG--TGTA--AT-----CTG--T--AA--CT--A--GA-----CA--A--AC--ATAT-----	4163
SIVVER9063	----	G-----	A--CT--G-----TACAC--GACT--G-----C--CCG--AG--TGT--A-----A--CTG--T--AC--T--G--GAG-----C--CA--A--CC--ACA-----C-----	4155
SIVVER155	----	A--T--GC--C-----	ACA--ACA-----C--CAAG--AG--CGTT--A-----A--TGC--AA--TT--A--GA-----CA--CT--AC--ATAC-----	3650
SIVVERTYO	----	A--C-----	A-----TACCC--GACA-----C--AAG--A--TGT--A-----G-----A--TTG--T--AA--TT--A--GGGA-----A--C--AC--ATAT--G-----	3658
SIVLHOEST	----	G--AA-----	AAC--GAG-----ATT--GG--GC-----A--G--T--C--A-----A--AG--A--TAGGA--AC-----A--A--AC--ATC--G-----	4291
SIVSUN	----	-----	AAG--G--A--C--GAG-----TACTC--CC--C-----AG--GCC-----C--A-----G-----A--A--CTAGAA--GAG--G-----A--A--CC--GG-----	4290
SIVMNDGB1	----	-----	ATG-----AA--GGGA-----TAGC--GTCCCC-----TC--AG--G--T-----A--AT-----T-----A-----C--G-----A--G-----A--AG-----TT-----	3623
SIVSAB1C	----	-----	AA--GA--T--G-----G-----A--A-----TA-----T-----CCC-----A--C--GT--AT-----T-----AA--TT--C-----GA-----CT--A--AG-----GG-----	4334
SIVSYK173	----	G--A--G-----	T--GG--GG--AC--TA--C--G--TACC-----CC--A-----T-----C-----A-----T--A--C--GC--GAC-----T-----AG--A--G--GAT--G-----	4136
POL	P	P	L V R W Y Y N L V P E P I P E A V T F Y V D G A A N R D S K T G N A G Y V A S D G	POL

B_FR.HXB2R	GACAAAAAGTTGTCCACCTAACTGACACACAAATCAGAAGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGACTCACAAATATGCATTAGGAATCAT	4066
POL	R_Q_K_V_V_T_L_T_D_T_T_N_Q_K_T_E_L_Q_A_I_Y_L_A_L_Q_D_S_G_L_E_V_N_I_V_T_D_S_Q_Y_A_L_G_I_I	POL
A_UG.U455	-----G-----TT-----G-----A-----A-----T-----CC-----C-----A-----C-----T-----G-----G-----	3512
C_ET.ETH2220	-G--G--A--TT-T-----A-----A-----G--CC-G--G--A-----A--C-----G-----G-----G-----	3458
D_ZR.84ZR085	---G-----C--T-T-----A-----A-----A-----A-----T-----T-----C-----C-----	3588
FL_BE.VI850	A-----T-----G--G--C-----G-----T-----A-----C-----G-----	3406
G_SE.SE6165	AG-----A--A--T-----A-----C--A--G--A-----G-----C--G-----A--C--AA--C-----G-----G-----C-----	3463
H_CF.90CF056	AG-----T-----G--A-----A-----A-----A-----A-----A-----A-----G-----G-----G-----T-----G-----C-----	3413
J_SE.SE92809	---G--A--T-----A-----AC--C--C-----T-----A--G--A--GC-----T-----T-----G--G--T--	3380
CRF01_AE_TH.CM240	---G--A--TT-----G-----A-----A-----A-----T--G--CC--T--C-----A--C-----T-----T-----G-----	3640
CRF02_AG_NG.IBNG	---G--TT-----G-----A-----A-----T-----C--T--C--A-----A--C-----T-----G-----	3591
CRF03_AB_RU.KAL153	---G-----T-----C-----	4111
CRF04_cpx_CY.94CY03	---G-----T-----A-----A-----CT-----A-----C-----G-----A-----	3432
N_CM.YBF30	---G-----G--T-TA-TG-A---C--C---A--G-----T--CCT-A-G-C--A--A--G--A--CGG-T-----C--T-----T--G-----TA-G---A--	3660
O_MVP5180	A--G--CA-AA-A-AGT--GAA--G--C---A--G--A--ATG--TG-A-TAA--C---CAAGGAGC-----T--G--C---	4096
O_CM.ANT70	A--G--GA-AA-A-AAT--GA--G--C--C---A--AG--A--ATG--G--A-TA--C--A---CAA-GA-AC-----T-----T-----T-----G--CG---	4120
HIV2_A_ST	--G--C--G--GAGGCTAT--GAGC-A--C--C---C--AG--A--G--CT--GCGA-G--AG-AACA--C--A--TCC-A-G-CC--TA--GT-----T--A--G---AG--	3893
HIV2_A_BEN	A--G--C---AAAAGTAT--GAAC-A--T--C---C--G--A--A--G--TCT--CGGA-G--AC--GCA--C--A--CCC-A-G--T--T--CA--GT--T--G---T--A--G--G--AG--	4448
HIV2_A_ALI	--G--C--G--GAAGGTA--GAAC-A--T--C---C--AG--A--A--G--CT--CGCG--G--AG-AACA--C--A--TCC-A--CC--T--TA--GT--T--G---T--A--G---AG--	4441
HIV2_B_D205	A--G--T--G--AAAGTGT--GAAC-G-----C--AC--AG--A--AC--TG--T--GCAT--A--AACA--C--A--A--CC--C--T--CA--GT--T-----TCA-G---A--	4421
HIV2_B_EHO	AGG-T-----AAAC-AT--GAAC-A-----C--AG--A--C--TG--T--GCA--AC--A--C--A--CC--C--G--C--T--CA--GT--T-----TCA-G---AG--	4419
STMM83293	A--A--C---AAAAG--TT--GAAC-G--T--T---AC--AG--A--A--G--T--GCCA--G--C--GC--C--A--TCC--A--C--T--T--GGT-----TGA-G---	4012
SMM251	A--G--C---AAAAGTGT--GAAC-G--T--T---AC--AG--A--A--G--T--CTCA--G--A--ACA--C--A--GCC--A--GAC--T--TA--GT--T-----TTA-G---A--	4341
SMM9	--GGC--GACAAAAGTGT--GGAAC-G--T--C---AC--AG--A--GG--CT-----C--AGCA---A--CC--A--C--T--TA--GT--T--C-----TTA-G--C--AG--	3838
SIVCPZANT	--AGT--GG--AAAACA---CAAA--G--C--C---AC--AG--A--A--G--G--CT--A--G--C--AG-----AAC--GGCCC--C--T--C-----T-----G--G--T--	3514
SIVCPZGAB	A-----A--CA--T--G--T--GAAA--T--C--T---C--AG--A--A--G--TT--GCT--G--C--AG-----A--ATCA--C--G--T-----G--T-----TG--G--T--	4127
SIVCPZUS	-----A--A--TT--GAAA--T--T--T---A--G--A--C--CAC--GG--A--CT---A--AA--A---A--AGAATAC--T--TG--C--T--G--C-----TC---	4126
SIVSTAN1	A--TC--C--G--CA--AGAAT--GAG--T--C--G--C--AC--AG--A--A-----G--CCT-----A--AA--AC--AGTAAGACCA--T--T-----G--T-----TCA-G---C--	4265
SIVGRI677	A--AGT--G--A--AGGGAAT--GAAA---C--T--C--AC--AG--A--A--AC---G--A--GA--G--A--G---CAGT--GA--A--T-----T-----T-----T--A--GAAC---T--	4069
SIVVER9063	A--G--G--A--AG--AAT--GAGC--G--T-----C--C--AG--A--A--GGT--C--AA--AA--G--C--G---AGT--GA--A-----T-----CA--G---AT--	4293
SIVVER155	A---G---ACAACAG--GAAA---T---C---AC--AG--AC--GAC--C--AA--AA--G--C--G---AGC--CCCTA--C--T-----T-----GA--G--C--AT--	4285
SIVVERTYO	A---G--G--A--AA--AT--GAAA---T--C---C--AG--AA--A--AC--T--AA--AA--G---G--A--CAGT--GCCTA--T--G-----T-----A--G---TT--	3780
SIVVERAGM3	A---G--G--G--AA--AAT--GAAA--T-----C---C--AG--A--A--ATG--C--AA--AA--G--AC--AG--A--AGT--GCCTA--T-----T-----A--G---AT--	3788
SIVLHOEST	A--G--G---AAGGAGT-----C---C--AG--A--AC--G--G--G--CTAA--G--A--A--A--AGCAATAGTA-----T-----T-----T-----T--A--GAA---AT--	4421
SIVSUN	--TCC--G---AAG--AAT--GGAAA---C--C---C--AG--A--AC--G--G--T--AA--AA--G--C--G---AGTA--GAGCTC--T--T-----G-----C--A--C--GC--	4420
SIVMNDGB1	TT---GG--A--AAGAATATCTAA--T--C--C---C---A--A-----T--G--AA--A---C--AG--A---AGT--AGTT--T-----T-----T-----T--G---T--AC--	3753
SIVSAB1C	AT---G--G--AG--AT--GAGA--T--T--C--C---AG--A-----G--G--C--CTGT--C--AAGA--C--T---AGCA-----T--A-----TA--G--T--T--	4464
SIVSYK173	C---G--G---ACAGTAT--GGAGC--A-----C--AC--G--A--A--G--G--GCT--GCTCA--G--A-----C--AAA--GACA-----T-----GT---T-----CT--T--AT--C--TC--	4266
POL	T_Q_R_V_Q_Y_L_E_Q_T_T_N_Q_Q_A_E_L_E_G_L_L_M_A_L_Q_D_S_K_D_K_V_N_I_V_V_D_S_Q_Y_S_Y_G_I_L	POL

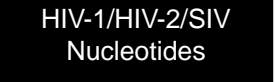
B_FR.HXB2R	TCAAGCACAAACCAGATCAAAGTGAATCAGAGTTAGTCAATCAAATAATAGACAGTTAATAAAAAAGGAAAAGGTTCTATCTGGCATGGGTACCAGCACACAAAGGAATTGGAGGAAATGAACAAGTAGAT	4196
POL	Q A Q P D Q S E S E L V N Q I I E Q L I K K E K V Y L A W V P A H K G I G G N E Q V D	POL
A_UG.U455	---G-----CAGG-----AA-----A-C---G-----A---C---T-----G-----G-----G-----	3642
C_ET.ETH2220	C-TG-----A-G-----A-----A-----GC-----G-----C---T-----T-----T-----	3588
D_ZR.84ZR085	-----A-G-----A-----G-----A-----C-----C-----T-----T-----	3718
FL_BE.VI850	---G-G---A-G-----AA-----A---C---C-----G-----C---T-----G-----	3536
G_SE.SE6165	-----AGG---G---A-----A---C-----G-----C---T-----G-----C-----	3593
H_CF.90CF056	-----C---A-G-----T-----G-A-----G-----C---T-----G-----	3543
J_SE.SE92809	A-----A-G-----A-----G-----G-----G-----C---T-G-----G-----	3510
CRF01_AE_TH.CM240	---G-----CAGG-----AG-----C-----G-C-----G-A-C---T-----G-G-----	3770
CRF02_AG_NG.IBNG	---G-----CAGG-----A-C-----C-A-C---T-----G-----G-----	3721
CRF03_AB_RU.KAL153	-----A-G-----G-----G-----G-----C-----C-----T-----T-----	4241
CRF04_cpx_CY.94CY03	-----AG-----T-----T-----CGG---C-----C---T-----G-----	3562
N_CM.YBF30	---TT---G---A-----A-G-G-C-----AG-C-C---GA-T---CT-T---T---T---G-G---C	3790
O_MVP5180	ATCCT-C---ACA-G---C-CCCTA--TC-G-G---G-AC-C---CGA-G---TA---T-T-T---C-A---A-A---	4226
O_CM.ANT70	CTCCT-C---TACA-G---C-CCCTA--TC-G---G-AC-C---C---G---TA---T-T-C-T---C-A---A-A---	4250
HIV2_A_ST	AGC-GC---ACAG-GTCA--GAGTA-AA---A---C---AG-AA-G---GCAA---G-T---C---C-T---C-A---CC-GG-G---	4023
HIV2_A_BEN	AGC-GC-G---ACAG-GTCA--AATAGAA--G-C-G-C---AG-AA-G---G---GCAA---G-T---C---C-T---C-A---CC-GG---C	4578
HIV2_A_ALI	AGC-GC-G---ACAG--TCA--AATAGAA--A---C---AG-AA-G---GCAA---G-T-G---C---C---C-A-G---C-GG---C	4571
HIV2_B_D205	AGCT---G---ACAG---CA---CCAA--AGCAA--T---AG-AA-G-C---A-GGCA-A---G-A-G---T---G---C-G-T---C-GG---C	4551
HIV2_B_EHO	AGCT---G---ACAG---CA---CC-A--A-GAG---T---AG-AA-G-C---AA-A---G-A-G---T---G---C-G-T---C-GG---C	4549
STMM83293	AAC-G---G---CACAG--TCG--GAGCA--A---G---AG-AA-G-T---GCAA-T---G-A---T---A---C-GG-G---	4142
SMM251	AAC-G-TGC--TACAG--TCA--GAGCAG-C---T-C---AG-AA-G-T---TC-G-AA-T---G-A---T---A---CC-G-A---C	4471
SMM9	AGC-GT---CAC-G--TCA--AGTAGA--A-C-G---A-AA-G-T---GCAA-T---G-A---T---RR-A---A---C-G---	3968
SIVCPZANT	G---GTACC---CCCC---GG-AG---CC-A-A---GA---C-A-T-C-CT-C---T---C-A---G---	3644
SIVCPZGAB	---GT---G---C---A---G---AG---T---A---AA---C-CT-C---T---A---G---	4257
SIVCPZUS	C-CT-T-G---C-----A-G-----G---AG-A---G-A---AAGT--A-CT-----T-----A---	4256
SIVSTAN1	GAC-----TACA---C-C-T-C---A-----T---AG-AA-G-T---T-A---TGCTA---T-AA-C-----T---G-A-G---C-G-A---	4395
SIVGRI677	GAC---TGT---C-GG---A-C---CCC---GG-A-G---C-AGCCC---G---AGGC---CT-ACA---T-T-G-G-A---C-AC-G-A---	4199
SIVVER9063	GAC---G---CACA---C-CCCA---AG-A-G-C---CATTAA-GG---G---AAGGCAA-A-C-ACA---G-T---T-A-G---G-A---C	4423
SIVVER155	GAC---G---CACA-G---C-CCAC---AG-A---CA---A-GG-C-G-A---GCCA---CA---T-T-T---T-A-G-C---G-A---C	4415
SIVVERTYO	GAC-----CACA---T---CCA---AG-G---T---CCTTAA-G---C---C-AA-A---T-CAG-----T---A---GG-GA---	3910
SIVVERAGM3	AACT-C---CACA-G---C-CCC--A-AG-A---T---CA-TAA-GG-C---AC-TC--A-A-CT--CA-----G---A---C---G-GA---	3918
SIVLHOEST	GTC-CA-AG---ACAG---CA---CACCCCA--A-AG-C-C---A---A-GC-AGC-G-A---TC-A-T---A-G---G-T-T---T---C-G-G---	4551
SIVSUN	CAGTAAGAG--CACAG---CA-T---C---A-GG-G-CG---TCA---GGC---CC---A-C-A-GT---G-----A---C-G-A---	4550
SIVMNDGB1	AGC-AGCAG--TAC-G---CA--TCACCCCA--A-GG---ATTAA-G-A-GG--A---AA-T---T-AAGT--C-----G-A---G---A---	3883
SIVSAB1C	AGC-GGG---TACAG--TCA--TAATA-CA--AC-A---T---G-AC-----G---GGCA-G-A-A---T-C-T---G-A-T---C---GG-A-T---	4594
SIVSYK173	-ATGA--TGC--CAC-A-C-CA--CACCCAA---GG-G-G---C-AG-GCC---GCCA-A---G-A---C---C-T---A---GCT-T---	4396
POL	<u> M T C P T N T E H P I V E Q I I Q E A I K K E A I Y V T W V P A H K G I G G N E A V D</u>	POL

		Pol p66 RT and Pol p15 RNase H end \ / Pol p31 integrase start					
B_FR.HXB2R	AAATTAGTCAGTGTGGAATCAGG.....AAAGTACTATTTTAGATGGAATAGATAAGGCCCAAGATGAACATGAGAAATATCACAGTAATGGAGAGCAATGGCTAGTGATTTTAACCTGCCACCTG			4320			
POL	K L V S A G I R K V L F L D G I D K A Q D E H E K Y H S N W R A M A S D F N L P P _			POL			
A_UG.U455	-----T------G-G-----G-----	T-----G-C-----A-----	T-C-C-----	T-----	3766	
C_ET.ETH2220	-----A-AG------G-G--C-----	T-----A-G-----A-----	C-----	A--A-----TA-C-----C-	3712	
D_ZR.84ZR085	-----AG------G-----G-----	T-----G-----	C-----		3842	
FI_BE.VI850	-----G------A--G-----G-----	A-----A-----A-----	AC-----G-----	T-----A	3660	
G_SE.SE6165	-----AG------C-----C-----	A-----A-G-----G-----	AC-C-----C-----	TT-----A	3717	
H_CF.90CF056	-----T-T-G-A------G--C-----G-----	A-T-----A-----A-GG-----	T-AC-----G-----	T-A-----A	3667	
J_SE.SE92809	-----T------G-G--C-----	G-----A-T-----A-T-----	T-C-----	T-----	3634	
CRF01_AE_TH.CM240	-----T-A------G-G-----G-----	A-----T-----A-----A-G-----	C-----A-----	TT-----A	3894	
CRF02_AG_NG.IBNG	-----AA------G-----C-----	A-----A-G-----A-G-----	C-----AG-----	T-----A	3845	
CRF03_AB_RU.KAL153	-----G------G-----A-----	A-----A-----	G-----		4365	
CRF04_cpx_CY.94CY03	-----CAA------G-G-----	T-----A-----	A--C-----	T-----T-A-	3686	
N_CM.YBF30	-----T-CT-A------A-T--CC-----	T-----A-A-----A-T-----	C-G-----C-----	A-----C-----T-A-C--CA	3914	
O_MVP5180	-----A-CAAA-AC-T-A------G-C-G--CC-G-A-----	C-----A-----A-T-----	A-----T-----	C--GGAT-A-----AA	4350	
O_CM.ANT70	-----A-CAAG-AT-T-A------G-C-G--CC-A-----	CC-----A-----A-T-----	A-----T-----	A--C-A-----A--GGA--A--A-	4374	
HIV2_A_ST	C-C-----A-CAG-C-----C-----T-----	CC-----GAA-----ACCC-T-G-G-----	A-----T-C-----	GTA-A--A-C-AT-CCA-A-A--GGA--	CAAAT	
HIV2_A_BEN	C-T-----A-CAA-C-----C-----T-----	CC-G-AAAG-----GCCC-T-G-----	A-----T-C-T-ATA-A--A-C-AA-CCA-A-A--	GGAA-A--C-T-C-	4702	
HIV2_A_ALI	C-T-----A-CAG-C-----C-----T-G--C-----	AAA-----GCCA-T-G-A-----	A-G-----T-C-----	GT-A--A-C-AT-CCA-A-A--GGAT-A--CAA-C	4695	
HIV2_B_D205	C-CC-----A-CAA-----C-G-CT-G--CC-----	AAA-----ACCA--G-A-G-----	A-----TG-C-----	GTA-A--A-C--TCCA-A-A--CGGA-T--	AAT	
HIV2_B_EHO	C-CC-----G-CCAA-----C-A-C-----C-----	AAA-----ACCA-T-----	A-----T-A-----	GTA-A--A-C-A-TCCA-A-A--GGGA-T--	AAT	
STMM83293	C-T-G-A-CAG-----C-G-C-----CC-----	AAA-----ACCA-T-----	A-G-----A-----	T-C-----GT-A--AGC-A-TCTTCA-G--	GG-A-A--AGGC	
SMM251	C-CC-----T-CAG-G-T-A-----C-T-C--C--G-AAAG-----	GCCA-A-----A-----	T-----C-T-----	GTA-A--A-T--TATTCA-A--GGAT-A--	CAGAA	
SMM9	C-TC-----T-CAA-A-----C-----C-----	AAA-----ACCA-G-G-A-G-----	A-----C-T-C-----	GTA-A-AA-T--TATTCA-A--	RGTT-A--TAGGC	
SIVCPZANT	-----CAA-G-----C-C-C--C-G-A-C-----	A-----T-----G-C-----	T-----T-C-----	T-T-A--AGA--A-AC-T-T-C--	A	
SIVCPZGAB	-----A-----G--CC-----	C-G-----A-----	A-GG-----T-----	A--T-----TT-A--	CA	
SIVCPZUS	-----C-----T-A-G-----T-T--CC-----	C-----A-A-G-A-C-----	C-----T-----	C--T-----	A	
SIVSTAN1	-----GC-----A-CCAG-G-T-----C-G--C-----CA-----	G-AG-A--G-ATC-----	T-----T-C-----	CA-TTT--AAGAGA--CA--CGTA-A--	G-AC	
SIVGRI677	-----A-CAAA-----A-A-----C-GA-C--C--C-----	A-----AG-A-A-----	C-----CA-G-C--T-AC-----	AGT--TACAG--A--GGAT-A--	TAA-A	
SIVVER9063	-----A-CAA-----A-A-----G--T-----	A-----GAA-----	AG-A-T-----A-----	C-G-----T-----	A-AAT-A--GGACACA--GGAT-A--	AGA
SIVVER155	-----A-CAAG--G-T-A------G-A--T-G--CA-T-GCA-G-----	AG-A-A-----A-----	T-GG-----C-----	AATC-A--AGACACA--GGAT--	AAA	
SIVVERTYO	-----G-AAA-C-T-A------G--TT--C-----	AAA-----AG-A-T-----	AA-G-----A-G-----	T-A-----A-AACC-A--AGA-ACA-A--GGG--	T--AAA	
SIVVERAGM3	-----C-----A-CAA-G-G-----A-TT-----	AAA-----AG-A--G-----	A-GG-C--T-A-----	GAAC-T-A--AGACAC--A--GGG--	A--AAA	
SIVLHOEST	C-CC-----A-AAA-T-A-----ACAGAAAC-G--CA-G--C-----	AAA-----ACCT-AGT-----	A-----AGC--TC--T-AC-----	GCA-AG-ATC-A--AAGAAA-A-----	A--C--CA	
SIVSUN	C-C-----T-CCAG-----AAAGAGAC-G--A-G--A-----	AAA-----ACCA-TGT-----	A-----G--G-TC--T-AC-----	GCAGCCT-TT--CAGGAGATG--G--A--	C-TA-	
SIVMNDGB1	-----GC-----A-T-A-----CT--CC-C-AAAT-----	ACCA-A-G-A-----	A-----T-C-----	GAAGC-CA-T-AAGAGAGA-A--CC--T-A--	G--C	
SIVSAB1C	-----C-----T-CAA-----A-A-----C-----	CC-G-CA--T-AG-A-A-----	A-G-----T-----	C--TGCC-----GAGT--CAACAA--A--	GG-T-A--TG--A	
SIVSYK173	-----C-----AAA-----A-A-----A-C-----	CC--AA--TCCC-A--A-G-A-T-----	A-GG-C--TCA--AT-GAGTATC--TAGGCAG--G--	C-T--A--C-GAC	4520	
POL	K L V S K g I R K I L F L E R I P Q A Q E D H E R Y H S N M E Y L R Q E F H L P R _					POL	

B_FR.HXB2R	TAGTAGCAAAAGAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCAACTAGATTGTACACATTTAGAAGGAAAAGTTATCCT	4450
POL	V_V_A_K_E_I_V_A_S_C_D_K_C_Q_L_K_G_E_A_M_H_G_Q_V_D_C_S_P_G_I_W_Q_L_D_C_T_H_L_E_G_K_V_I_L	POL
A_UG.U455	-G---G--G-----A-----A-----G-----G-----T-----C-----C-----A-----	3896
C_ET.ETH2220	---C-C-----TT-----G-----A-----A-T-----G-----T-----C-----G---A-C-----	3842
D_ZR.84ZR085	-G-----T-----A-----C-----G-----T-----C-----	3972
FL_BE.VI850	-----T-----G-----G-----C-----G-----T-----C-----C---G-----	3790
G_SE.SE6165	-----G-----T-----G-----C-----T-----C-----G---A-----	3847
H_CF.90CF056	-----T-----G-----C-----T-----C-----G---C-----T---	3797
J_SE.SE92809	---G---G-----T-----G-----G-----T-----C-----C-----	3764
CRF01_AE_TH.CM240	-----G-----A-----A-----A-----G-----T-----G-----G-----T-----C-----C-----	4024
CRF02_AG_NG.IBNG	-----G-----A-----G-----G-----G-----G-----T-----C-----G---A-----T---	3975
CRF03_AB_RU.KAL153	-G-----T-----AT-----C-----C-----	4495
CRF04_cpx_CY.94CY03	-G-----G-----T-----A-----G-----G-----G-----GT-----T-----A-----	3816
N_CM.YBF30	---G---G-----C-----C-----G-----G-----G-CA-T-----G-G-----T-----C-----G---A-C-----	4044
O_MVP5180	---C---G---CA-T--T--T--CC---C-TA---G---A-C---T---AC-C---AG---A-G---C---A-C--AA	4480
O_CM.ANT70	-G--G--C--G---CA-T--T---CC---TA---G---A-T---T-G---C---A-G---A-G---C---A-C--AA	4504
HIV2_A_ST	---G---G-C---AA-CA--CACCC---AG---G-G-T-A---G---A-TGCAGAA-TT--C-CT--A-G-C-C---C---A-C--TA	4277
HIV2_A_BEN	---G-C-G---AA-TCA--CCC---C-A-AG---G---A---G---A-TGCAGAAAT--CG-T---A-G-C-AC---C---A-C--TA	4832
HIV2_A_ALI	-G--G--G-C-G---AA-CA--CCC---AG-G---G-T-A---G---GA-TGCAGAA-T--C-CT--A-G-C-C---C---C--A	4825
HIV2_B_D205	---G---C-G---AATTC---C-A-A---G---T-T---G---A-TGCAGAC-T--G-C---GA-G-C---A---AA	4805
HIV2_B_EHO	---G---G-C---AATTC---C-A-A---G---T-T---G---A-T-CAGAA-T--G-C---A-G-C---G---G---AA	4803
STMM83293	---G---GC---G-AT-CA---C-G-C-C-AG---T-A---G---A-TGCAGAGTT--G-CT--GA-G-C---C---C-GA-A--TA	4396
SMM251	---G-C-G-C-G---A-C---T-AG---T-A---G-G---A-T-CAGA-T--G-CT--A-G-C---C---A-AG--A	4725
SMM9	---C---A-CA---C-C-G---Y-A---G---A-TGCRGRA-T--G-CT--A-G-C---CC---G-C---A-A--TA	4222
SIVCPZANT	-T-G-T---TA---ACAG---CG-G---G---G---C---G-T---G---C---C---T---A---A	3898
SIVCPZGAB	---G---CAT---G-C-G---G---G---G---C---G-T---G---C---CC---C---G---A	4511
SIVCPZUS	-G-C-T-G-G---G---C---G---A---G---G---T-C---T-C---G---G---A-C--TA	4510
SIVSTAN1	-G-----TGCT--CCCC---AA-TCG---C-G-A---G---TGCC--GT--GG-C---A-G-C---G---A-C--TA	4649
SIVGRI677	---G---G---GGCA--CCC---C-AA--G---C-T-A---G---GCCTCCATT-A--CT--GA-G-C-C-C---AA	4453
SIVVER9063	---T-C---G---A-TG--CCA--TGCCA-AA---G-C-A-A---G---GCCTCC--GG---A-G-C-C-G---G---A-AG--A	4677
SIVVER155	---T---T---A-TG--CCCC---AG---G-C-A-A---G---TGC-TCA---G-G---GA-G-C-C---A---A-AG-GA	4669
SIVVERTYO	---G---G---TG--CCA---AA---G---C-AG---G---G---TGCCTCA--T--C---GA-G---T---C---AA---AG-A	4164
SIVVERAGM3	-T-G---TG---CCA---A---G---C-AG-C---G---G---TGCCTCG--GG---A-G-C---C---G-A--A	4172
SIVLHOEST	-G---C---C---T---CAATGA---CA-C---AAAG---AACA---G---TGTCTCAGTG--T-T---G---C---T-C---C-G--C--TA	4811
SIVSUN	-G---GC-G---CAATGAA---CAC---A-AG---AACA---G---TGCCTCAGTG--C-C---GA---C-C-CA-G---A---C--AA	4810
SIVMNDGB1	---C---C-G-T---GCAA--T---CAG---G---CTGT-ATC-T---GC---AA-G---GAC---TGC-TCA-TT---G-C---GA---C---C-G---AATC--A---TA	4137
SIVSAB1C	---G---G---GGCA--CC---C-A---G---GT-TG-A---G---TGC---G-G---A-G---CC---C---A-C--A	4848
SIVSYK173	AG-----C---TA---CAACAA---CCA---AAAT-G---C-A-AA---G---TGTAGA-AT-TAT-AT---GA-G-C---T---GAG---C---A---TG	4650
POL	Q_V_A_K_A_I_I_Q_Q_C_P_K_C_Q_N_R_G_E_P_K_H_G_Q_V_D_V_D_I_Y_N_W_Q_M_D_C_T_H_E_E_G_K_V_I_C	POL

B_FR.HXB2R	GGTAGCAGTTCATGTAGCCAGTGGATATATAGAAGCAGAAGTTATCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTTAAAAATTAGCAGGAAGATGGCCAGTAAAAACAATACATACTGACAAT	4580
POL	V A V H V A S G Y I E A E V I P A E T G Q E T A Y F L L K L A G R W P V K T I H T D N	POL
A_UG.U455	A-----C-----C-C-----C-----A---G-----C--A-AC-----GT-----C-A-----	4026
C_ET.ETH2220	-----C-----C-C--G--G-----A--A-----C--AC-----G-----C-GGGT-----A-T---	3972
D_ZR.84ZR085	-----C-----C-----C-----G-----C-----C-----GT-G-G-----A---	4102
FL_BE.VI850	-----C-----T--AG-C-----A-----C-C-CA-A--G-----T-----A-----	3920
G_SE.SE6165	A-----C-----C-----C-----AA-----C-C-G-G--C--A-C-----C-GT-----A-T---	3977
H_CF.90CF056	-----C-----C-----C-----A---G-----T--A-A-----C-----GT-----A-----	3927
J_SE.SE92809	-----C-C-G-----C-----A---G-----C--GC--C-----GT-----C-A-----C	3894
CRF01_AE_TH.CM240	A-----C-----C-----A---G-----C--A-A-----G--GT-----C-A-----	4154
CRF02_AG_NG.IBNG	A-----C-----C-----A---G-----C--G-C-----T-----A-----	4105
CRF03_AB_RU.KAL153	A-----G-T-----C-----C-----C-CA-AC-----G--TG-----G-A-----C	4625
CRF04_cpx_CY.94CY03	T-G-G-G-C-----G-CT-----T--G-----A-----G--T-----C--T-----	3946
N_CM.YBF30	A-T-T-C-----A--AC-T-----G-G-A-----A-----T-C--C-G-----C-----T-C--GT-----A-----	4174
O_MVP5180	A-T-T-C-----G-A--G-TC-----G-A-----A--A-----T-C-C-G-----C-G-T-C-----T-T--GT-----A-----C	4610
O_CM.ANT70	A-----A--T-A-----T-----C-C--CAG--T--AAG-C--G--CTC-C-AC--C-G-CA-T-G--A--C-CATT-G-C-A-----	4634
HIV2_A_ST	A-----A--T-A-----TC-----C-C--CAG--T--AAG-C-G--CTC-C-A--C-G-CA-T-G--A--CGCACT-G-C-A-----	4407
HIV2_A_BEN	AA-----G--T-----TC-----G-C-C-GCAG--T--AAGAC-----CTC-C-A--C-G-TA-T--A--C-CACT-G-C-A-T-G--	4962
HIV2_A_ALI	A-G-----C-----G-T-----G-A-A-CCA-G--AAGAC-G--TCTC-C-AC--G-G-CA-C--TA-C-C-CACC--C-A-----	4955
HIV2_B_D205	A-G-----A-----T-----TC-----A-A-CCA-----AAGAC-G--TCTC-C-G--GC-G-CA-C--TA-C-C-CACC-G-C-A-----	4935
HIV2_B_EHO	A-----A-----T-----G-T-----G-----A-C-CA-----AGAC-G--TG--G-G-G-A-C--T-C-C-CACC-G-C-G-T---	4933
STMM83293	A-T-----A-----T-----TC-----A-----CA-G--AAGAC-G--CTA--G--G--C-----TA-T-C-CATC--C-A-T---	4526
SMM251	A-T-----A-----G-GT-----TC-----G-----A-C-GCRG--AAGAC-----CTG--GC-----CA-C--CA-C-C-CATC-G-----T---	4855
SMM9	A-----G-----T-T-----C-TC-----T-----A-GG--AT--G--A-AAGT-----C-C-G-----CA-C-----C-----	4352
SIVCPZANT	A-G-----C-----C-----C-----C-----T-G-----A-A-----T-C-G-----T-----T-T-C-A-T---	4028
SIVCPZGAB	A-G-----A-----G-----CC-----G-C--T-----A-A-G--C-----A-C-----T-----GTG--T-----T---	4641
SIVCPZUS	A-----C-----G-----C-TCT-----C-AG-TAGG--AA-A-----A-A-C-G--A-CATTA-T-G--TA-T-C-A-T-----	4640
SIVSTAN1	A-----A-----TC-----G-G-C--AG--T-----G-G-C--AG--T--A--G--C--GC-G--C-GTT--C-----G--CATC--C--T---	4779
SIVGRI677	A-----G-----G-----C--TC-----C-A-TAGG--AAGA-G--CA-A-C--C--A--TTGA-T-----A--GTCCAGT--C-A-T---	4583
SIVVER9063	A---G-C-----G-T-----G--C-TAGG--AA-A-G--A-G-CT-G--A-AT-----CA-C-CTCACC-C--A-T---	4807
SIVVER155	A--T-G-C-----TC-----C-A-TAGG--AA-A--G--A-G--A--A--CTGA-T-----TA--C-CAGT--C-A-----	4799
SIVVERTYO	A-----C-----TC-----A--TAG--A-A-----A-A--T-A-----GA--CT-A-T-----CA--CCCA-C-G-----A-----	4294
SIVVERAGM3	AAAT-----C-----G--TCA--C-TC--G-TT-----A--AT--G--AA-AAC--GT--C--C-G--TGCA-T-----T-----CA-----	4302
SIVLHOEST	A-----G-----G--TTCA--G--G-T-----AC-A-CAAT--GCA--A--AC--CAC--GG--G--GC-CTGT-CC-TG-----CAG-----A-T---	4941
SIVSUN	A-----G-----TTCA--C-TC--GA-G-----AA-----T--AA-AA-G--G-G--G-----C-CA-----TA-T-GT-A-C--C-A-T---	4940
SIVMNDGB1	A-G--T-----T-T-----T-----A-C--CAG--AA-A-CT--C-C-C-GC-G-G--CA-C-----CA-C-C-CAGC-G--A-T---	4267
SIVSAB1C	T-----AA-CAC--T-----A-CA-GA-CT-AAA-AGG--G--G-T-----TG--CA-GC-A--CA-T-----A--CAG--C--T---	4978
SIVSYK173	POL	4780
POL	V A V N T A S G Y I E T K I L K R E T G D E T A L F L M Q I A S R W P I K Q I H T D N	POL

B_FR.HXB2R	GGCAGCAATTCACCGGTGCTACGGTTAGGGCCGCTGTTGGTGGGCGGGAATCAAGCAGGAATTTGGAAATCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAGAAAA	4710
POL	<u>G S N F T G A T V R A A C W W A G I K Q E F G I P Y N P O S Q G V V E S M N K E L K K</u>	POL
A_UG.U455	-----A-C--G-A--AA--A-T-----AAAT--C-A-----G-C-----G---C---C-G-----	4156
C_ET.ETH2220	---T-C---A-AA-G-A--AA--A-----A-T--TC-A-----G---C---C-G-----	4102
D_ZR.84ZR085	-----A--A--A--A-----C-----A-T-----G---C---C-G-----	4232
FL_BE.VI850	-----A--CG--A--AT-----A-T--C-A-----C-----A---G---A-G---	4050
G_SE.SE6165	-----A--G-A--A-A--A-----AAAT--CA-----C-----G---G-----	4107
H_CF.90CF056	-----GA--G--A--A-----A-AT--C-A-----G-----G-----G---	4057
J_SE.SE92809	-----C--A--G-G-T--G-A--A-----A-AT--C-----C-----G---G-----	4024
CRF01_AE_TH.CM240	---T-----A-C--G-A--AA--A-----CAATG--C-A-----G-C-----G-----	4284
CRF02_AG_NG.IBNG	-----A--G-A--A-A--G-A-----AAATG--CA-A-----C-----G-----	4235
CRF03_AB_RU.KAL153	-----A--A--G--A--T-----A--G-----C-----C-----C---	4755
CRF04_cpx_CY.94CY03	---CC-----A--G--A--A-----A-AT--C-----C-----G---G-----	4076
N_CM.YBF30	---ATC-----TA--C-T--A-AA--A-----AAAT--A-----G-A-----T--G--C--G-C-----	4304
O_MVP5180	---ACCT---T-AA--AG-CA-G-AA--T--A-----A-A--C-AC-A--T--G--G-A--A--T--A-----G-C-----ATCT-	4740
O_CM.ANT70	---GCCT---T-AA--A-A--TA-G-A--T--A-----CAAC-AC-A--T--G-----A-A--T--A-----G-C-----G-----ATC-	4764
HIV2_A_ST	---TGC--C---TTCACAGGAA--A-A-ATG-TGGCA--ATA--T-AG-A--ATCC--C--G-A--T-----A--C-----G--G-A--C-CC-CC--A--TC	4537
HIV2_A_BEN	---CC--C---TTCACAGGAA--G-A-ATG-TGGCA--TA--T-AG-A--ATCC--G-A--T-----C--A--C--G-----G-A--C-CC-CC--TC	5092
HIV2_A_ALI	---TGT--C---TCACAGGAA--A-A-ATG-TAGCA--TT--T-AG--ATCC--G-A--T-----A--C-----G-A--C-CC-CC--A--TC	5085
HIV2_B_D205	---TGC--C---TCAC-A-GT--A-A-ATG-TAGCC--TA--G-AG-A--AACT--G-A--T-----T--C--A-----G--G-A--CC-TC-CC-G-A--TC	5065
HIV2_B_EHO	---TGC--C---TTCACAAGAT--G-AAATG--AGCC--ATA--G-AG-A--AAC--C--G-G--T--AG-----G-----G-A--CC-TC-TC-G-A--TC	5063
STMM83293	---GCT--C---TCGCAAGAA--A-A-ATG-TAGCC--AG-A--AACC--G-A--T-----A--G--C-----G--G-C--C-TC-T-----CCC	4656
SMM251	---TGCT--C--TG--TCGCAAGAA--A-A-ATG-TTGCA--A--G--AG--CACC--GG-A--A-----A--G-----G-----G-A--C-CC-CC-G-A--TC	4985
SMM9	---TGC--C---ATCACAAGAA--G-AAATG--TTGC--A--R--TG-A--ACC--R-GG-A--T--T-----A--G--C--R-G--G--G--C-TC-TC--A--CCC	4482
SIVCPZANT	---AGCT---AA--AG-A--A-AA--G--A-----TAAT--C-A-----A--A-----A--G-----G-----C-----G--G--C---	4158
SIVCPZGAB	---GCCA--T--AA--G-A--C-A--T-----A-AC-----A--A-----A--G-----G-----CT-A-----GC-----	4771
SIVCPZUS	---ATCT--C--T--AA--AG--A-----A--A--T--C-----A--C--AC-A--A--G-----A--A-----G-----C-----	4770
SIVSTAN1	---GCCA--T--GTTA--CAGGA--ACAAA--TATT--GACA-G-AG--CAC-ACA--G--C-----T--C--A--TCA--AGT--AG--G-C-----AG--	4909
SIVGRI677	---CCA--C--T--TC-CAG-AT--GGCA--G-TG--C-----GTAAT--AG--CACCAC--A--T--T--C--A--GTCA--G--TAGT--AGC--C-G-C-GC--G--	4713
SIVVER9063	---ACCA--T--TTCCCAAGA--AGC--AAT-----GAAAG--AG--CACCACA--G-G-A-----A--GTCA--G--TC--AGC--C--C--C--AG--	4937
SIVVER155	---ACCA--T--TTC-CAGGAA--AGCT--TATG--C-----GAAAGG-AG-A--CAC-ACG--GG-A--A--T-----A--GTCC--G--TCTA--AG--C--G--AG-G--	4929
SIVVERTYO	---GCCT--C--T--TCCCAAGAA--GGCA--AATA-----GAAA--TG-A--TAC-ACA--T--A--A--T--C-----TC--TC-A--AGC--C--GC--AG-G--	4424
SIVVERAGM3	---ACC--T--GTC-CAAGAA--AGCA--AATG-----GAAA--AG-A--CACCACA--TG-A--T--C--T--TCA--CTCTA--GAG--C--G--AG--	4432
SIVLHOEST	---TCCA--TGTA--AAGGA--ACA--A-TAACC--ATA--G--AG-A--TAC-AC--G-A--T--T--C--T--G-----G--T--G--A--AA--C--TT--A--GT--	5071
SIVSUN	---GCCA--T--A--AAAGAT--GGA--A-TA-----TTA--C--AC--TAC-ACA--G-A--T--C--T--G-----G--G--A--A--C--G--T--AC--	5070
SIVMNDGB1	---GCCT--C--T--TA--CAGGAA--AGAAA--ATG-----TTA--G--AG-A--CAC-----A--T--C--A-----G--G-----G--AA--AA--GT-TC--AG--T	4397
SIVSAB1C	---T-C--C--T--TA--CCAGCAA--AGCA--TATT--C-----GAAA--AG--CAC--GG-A-----C-----G-----G--G--AG--G--	5108
SIVSYK173	---ACCA--C--TGT-A--AC-A--T-C-AA--A--A-----TGT--C--AG-A--CACTACA--T--A-----C--A-----C--G--A-----G--C-AA--GGT-T--AG--G	4910
POL	<u>G P N F V S D K F K A A C W W C G I E H T T G I P Y N P O S Q G I V E S K N R Y L K E</u>	POL



B_FR.HXB2R	TTATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAAATGGCAGTATTCATCCACAATTTTAAAGAAAAGGGGGGATTGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAAT	4840
POL	I_I_G_Q_V_R_D_Q_A_E_H_L_K_T_A_V_Q_M_A_V_F_I_H_N_F_K_R_K_G_G_I_G_G_Y_S_A_G_E_R_I_V_D_I_I	POL
A_UG.U455	-C-----G-----G--A-----C-----T-----A-----A-----	4286
C_ET.ETH2220	-C-----G-----A--A-----G--C-----T-----G-----A-----T-----	4232
D_ZR.84ZR085	-----A-----A-----G-----G-----A-----A-----	4362
FL_BE.VI850	-C-----A-----A-----A-----A-----A-----A-----	4180
G_SE.SE6165	-C--C--G----C--G----A-----G-----T-----A-----A-----	4237
H_CF.90CF056	-C-----G-----C--A--A-----C-----T-----A-----A-----	4187
J_SE.SE92809	-C-----A-----A-----C-----A-----A-----A-----	4154
CRF01_AE_TH.CM240	-C-----G-----G--A-----C-----A-----T-----A-----	4414
CRF02_AG_NG.IBNG	-----C-----A-----C-----G-----T-----A-----T-----	4365
CRF03_AB_RU.KAL153	C-----A-----A-----A-----A-----A-----A-----	4885
CRF04_cpx_CY.94CY03	-C-----G-----C--G----A-----C-----G-----T-----A-----	4206
N_CM.YBF30	-----AA-C-----A--A-----A-----G-----G--T-----C-----A-----	4434
O_MVP5180	---CAG---G--G--C--A--A--G--T--A--A-----C--TG--T-----C-----G--C--A-----C--	4870
O_CM.ANT70	---CAG---G--G--C--A--A--CT--A--GA-----TG--T-----C-----A-----G--A-----T--	4894
HIV2_A_ST	AG--A-CAGAA-T---G---AA-CACAG-AG-A--AT---T---TCATTG-ATG-----G-GG-A-A-A--ATATG-CCC--CA---C-A-CA-T--GG-	4667
HIV2_A_BEN	AG--A-TAGAA-T---A---AA-TACAA-AG-A--AT---TG---TCATTG-ATG-----GG-A-A-A--ATATG-CCC--CA---C-A-CA--G--	5222
HIV2_A_ALI	AG--A-CAGAA-T---G---AA-TACAG-GG-A--AT---T---TCATTG-ATG-----GG-A-A-A--ATATG-CCC--CA---C-CA-CA-T--G--	5215
HIV2_B_D205	AA---ACAGAC-C---C--A--A--T-TCAA-AG---TT---T---ACTCA-TG-ATG-----GG-A-A-A--ATATG-CCC-T-CA---C--TA--G--	5195
HIV2_B_EHO	AG---ACAGAA-T---A--T-TCAA-AG---TT--GTT---ACTCA-TG-ATG-----GG-A-A-A--ATATG-CCC-T-CA---CA--G--	5193
STMM83293	AG---ACAGAA-T---A--AA-CACAG-AG-A--T-T---TC---TCA-TG-ATG-----G-A--GG-A-AC-A--ATATG-CCC--CA---T--TA-T--G--	4786
SMM251	AA---ATAGAA-C-G-A-A--AA-TTCAG-AG-A--CAT---TT---TCATTG-ATG-----GG-A-A-A--ATATG-C-C--CA---T-A-TA--G--	5115
SMM9	AA---ATAGAA-T---A---AA-TTCAA-AG-A--TAT---T---A-TCATTG-AT-----GG-A-A-A--ATATG-C-C--CA---T--CA-T--G--	4612
SIVCPZANT	-----AA-T---C--A--A---AT-A-----GT-----GCA--T-----CAC-T--AC-G---T-----C--	4288
SIVCPZGAB	-A-----T--G---A--A-----T-A--A-----G-----G-----T-----C-----A-----	4901
SIVCPZUS	-C-----AA-C--G---A--A--G---T-A-----T---G-----T-----A-----CA-----C--	4900
SIVSTAN1	C---A-A-AA-T---AG-CTG-C-TT-T-GG-A--G---C-TC---CTGCCA--T-----G-A-A-A--TATG-CCC--CA--G--GC-A-A--G--	5039
SIVGRI677	-C--CTCT--AA-----G-TTG--GAGAT-GG-----G-----TACGCAT--G-A-A-A--TAT-TC-AGT-CA---T-G--TA-T--GC-	4843
SIVVER9063	-A--T--A-AA-----G-TTG-C-GT--ACAG-A-----TT---TGCCA--T-----G-A-A-A--TA-CAC--CA---C-A-TA-T--G--	5067
SIVVER155	-A--T--A-AA-----G-CTG-C--T--ACAG-A-----TT---CTGCCA--T---CTA-CA--T-CA-G--C-A--A-T--G--	5059
SIVVERTYO	-A--T--GA-AA-----G-TTGCC--T--ACAG-G-----TG---TTGCA-AC-T-----G-A-A-A--AC-G-C-T--CA-G--C-A-TA-T--G--	4554
SIVVERAGM3	-A--T--GA-AA-----G-CTG-C--T-CACAG-A-----TT---TGCCA--T-----G-A-A-A--TA-CAC-G-CA-G--T-A-CA-T--G--	4562
SIVLHOEST	-----A-AGA-----AG-T--AC-G--A--A--A-----T-----CCA--T--T-----CA-GG-A-AC-A--CTT-CAC--CA-G--GT-TA-TA-T--G--	5201
SIVSUN	-----CA-CAGAA-T---AG-T--AC-GG-AT-G-A--T-----TC---TC-GCA--T--T-----CA-GG-A-AC-A--AATG-CA--CA-G--GT-TA-CA--G--	5200
SIVMNDGB1	-G--T-AGA-AA-----AG-TTGCA--G-AT-A--A-----GCC---CAC--T--T-----CA-GG-A-AC-A--ATG-CA--A-G-----A-T--G--	4527
SIVSAB1C	-C-----AA-T---G--T--A---AGAT-GG-A-----CAT-----GCA--T-----TC---CA---C-A-CA-T--G--	5238
SIVSYK173	C---CTC---AA-T---CG-C-TCAC---T-GC-A-----GGC---ACC---TTA-----G-A-A-A--ATATCAC--A---GTACA-TA-T--GC-	5040
POL	A_I_S_Q_I_R_D_D_V_T_H_L_Q_T_A_V_A_M_A_T_F_I_L_N_F_K_R_K_G_G_I_G_G_I_S_P_G_E_R_Y_I_N_M_L	POL

B_FR.HXB2R	AGCAACAGACATACAACTAAAGAATTACAAAACAAATTACAAAAATTCAAAAATTTTCGGGTTTATTACAGGGACAGCAGAAATCCACTTTGGAAAGGACCAGCAAAGCTCCTCTGGAAAGGTGAAGGG	4970
POL	A T D I O T K E L Q K Q I T K I O N F R V Y Y R D S R N P L W K G P A K L L W K G E G	POL
A_UG.U455	-----T-----G-C-CA-----A-A-----	4416
C_ET.ETH2220	---T---G---C-C---C---TT-----A-----G-C-TA-----C-A-A-----	4362
D_ZR.84ZR085	---T---T---G---G---C---C-----T-----G---A-----T-----	4492
FL_BE.VI850	-T-----G-----G-----C-----C-G-----A-----	4310
G_SE.SE6165	---T---T---C-----C-----A-----G-C-G-----C-A-A-----	4367
H_CF.90CF056	-----G-----T---C---A-----G-C-A-----A-----	4317
J_SE.SE92809	-----G-----G-C-A-----A-A-C-----	4284
CRF01_AE_TH.CM240	-----G-C-A-----A-A-----	4544
CRF02_AG_NG.IBNG	---T---T---C-----G---T-----G-C-CA-----A-A-----	4495
CRF03_AB_RU.KAL153	-----T-----A-----G---A-----A-A-----	5015
CRF04_cpx_CY.94CY03	---T---T---C-----G-A-A-----G-----A-A-----	4336
N_CM.YBF30	---G-A-C-A-T---C---TT---G-----A-----G-CA-----C-A-T-G-----A-----	4564
O_MVP5180	---T-C-A---A-C---C---TT---CA-C---C---A-T---G-C-TA-----G-C-A---G-----G-----	5000
O_CM.ANT70	---T-C-A---A-C---TT---N---C-A---C---A---G-C-TA-C-----G---G---G-A-----	5024
HIV2_A_ST	CA-TG---ACAGG---TAC-TTCC-C---GC-A---A-T---T-A---C---T---A-AG---G-AG-G---T-GGG-A-A-G---G-G-C-A	4797
HIV2_A_BEN	TA-C---ACA-G---TAC-TTCC-C---G-A---A-T---TT-A---C-A---C---A-AG---G-AG-G---T-GTG-A-A-G---G-A---A	5352
HIV2_A_ALI	CT-C---ACA-G---TAC-TTCC-C---C-A---A-TTG---T-A---C-C---C---A-AG---G-AG-G---T-GGG---A-G---G-C-A	5345
HIV2_B_D205	-A-C---GCA-G---TAC-GTTC-C---GC-A---A-TT---T---C-A---C---A-AG---G-A-C---G---T-GTG-A---AT-G---G-A	5325
HIV2_B_EHO	-A-T---ACA-G---TAC-TTCC-C---C-A---A-TT---T-C---C---C---A-AG---G-A-C---G---T-GTG-T---AT-G---G-A	5323
STMM83293	CA-C---ACA-G---TAC-TTTCA---TC-A---A-T---T-CA---C---C---A-AG---G-A-A---G---C-GTG---AT-G---G-A	4916
SMM251	CA-T---ACA-G---TAC-TTTCA---TC-A---ACT---T-A---C---C---A-AG---G-A-G---G---C-GTG---AT-G---G-A	5245
SMM9	TA-C---ACA-G---TAC-TTCCA---TC-A---A-T---T-A---C---C---A-AG---G-C-AG-G---C-GTG---AT-G---G-A	4742
SIVCPZANT	-----G-A-CTC---T---TT---C-A---C---T---G-C-TG-G---C-A-T-G-----	4418
SIVCPZGAB	-----A-GC---T---TT---G---A---C---A---G-C-A---G---CCT-A-G-----	5031
SIVCPZUS	---T---AC---AG-CTT---TT---G---C---A---G---C-A---T-G-----	5030
SIVSTAN1	-A-C---C-AC-G---TAC-C-CA---C---CAACA---TC---AA---C-C---A-AG-A---G-G---G-G---C-CT---G-A---A	5169
SIVGRI677	-A---C-AC-G---CTA-TACTC---C---CCA---TTG---AA---C-C---A-AG-T---G---G-G---GCGA---A---A---C	4973
SIVVER9063	-A---C-AT-G---CTAC-AC---CCA---CA---TT---A-A-C-C---A-AG-G---G---G---G---C-AT-AA---G---T	5197
SIVVER155	-A---C-AT-G---TC-CACTC---CCA---CCA---TTG---A-A-C-C---A-AG---G---G-G---G---T---TCGC---GA---A---C	5189
SIVVERTYO	-A---C-AT-G---TAC-C-T---CCA---CA---TT---A-A-C-C---A-AG-G---G-C-TG-G---G---C-AT-AA---G---A	4684
SIVVERAGM3	TA-T---C-AT-G---TTAC-C-CC---CCA---CA---TT---A-A-C-C---A-AG-G---G---TG-C---G---C-T-AA-T---G---T	4692
SIVLHOEST	TAATG---AC-G---AC-T-TC---TT---A-T---TT---A---AA---C-C---AC-AG-A---G---T-AG---GC-A---T---G---A---T	5331
SIVSUN	TAATG---TC-G---AC-T-C---G---AT---A-T---TT---A---AA---C-C---A-AG-A---G---T-AG---T---T---A---T---G---A---A	5330
SIVMNDGB1	CAAT---AT-G---TA-C-T-TCA---T---T---A-TT---AA---C-T---A-AG-A---G---AG-G---T-GT-TC---TT-G---A---A	4657
SIVSAB1C	-CAT---AC-G---C---AC-C-T---C-A---T---C---C---A---G---G-C-TG-G---T-G-T---A---A---A	5368
SIVSYK173	-TAT---AC---TTAC-C---AATAC-C-TC-...C---T---TCG---A---C---C-CC-G...G-G---GAACGAG---G---TCGT-A-A-T---T	5164
POL	Y T E L Q L Q Q N T T S . P K F S N F R V Y Y R Q . G K N E W K G P A R L L W K G E G	POL

B_FR.HXB2R	GCAGTAGTAATA.....	4982
POL	_A_V_V_I_.....	POL
A_UG.U455	-----	4428
C_ET.ETH2220	-----	4374
D_ZR.84ZR085	-----	4504
FL_BE.VI850	-----C-----	4322
G_SE.SE6165	-----	4379
H_CF.90CF056	-----	4329
J_SE.SE92809	-----	4296
CRF01_AE_TH.CM240	-----	4556
CRF02_AG_NG.IBNG	-----	4507
CRF03_AB_RU.KAL153	-----G-----	5027
CRF04_cpx_CY.94CY03	-----	4348
N_CM.YBF30	-----G---T-----	4576
O_MVP5180	-----C-----	5012
O_CM.ANT70	-----C-----	5036
HIV2_A_ST	-----CA-G-C-----	4809
HIV2_A_BEN	-----CA-G-C-----	5364
HIV2_A_ALI	-----CA-G-T-----	5357
HIV2_B_D205	-----CA-C-----	5337
HIV2_B_EHO	-----CA-C-----	5335
STMM83293	-----CA-CG-----	4928
SMM251	-----CA-CT-----	5257
SMM9	-----CA-CC-----	4754
SIVCPZANT	-----CANNNNNNNNNNNNNNNNNNNGGGGATCCTCTAGAGTCCATCTTGGAAACGCCACATGGCTCCCGTTTCATCAGTCGTGGGGTCATTGGCGGTGCGATCACTAGGAAGGTCAAAGTTA	4548
SIVCPZGAB	-----G-C-----	5043
SIVCPZUS	-----C-----	5042
SIVSTAN1	-----G-C-----	5181
SIVGRI677	-----G-G---T-----	4985
SIVVER9063	-----G-C-C-----	5209
SIVVER155	-----G---TC-C-----	5201
SIVVERTYO	-----G-CC-C-----	4696
SIVVERAGM3	-----G-C-C-----	4704
SIVLHOEST	-----T---GG-----	5343
SIVSUN	-----C-C-----	5342
SIVMNDGB1	-----TT-----	4669
SIVSAB1C	-----T-----	5380
SIVSYK173	-----GG-G-----	5176
POL	_A_V_V_V_.....	POL

B_FR.HXB2RCAAGATAATAGT...GA	4996
POLQ D N S...D	POL
A_UG.U455C-----	4442
C_ET.ETH2220	4388
D_ZR.84ZR085C-----	4518
FL_BE.VI850C-----	4336
G_SE.SE6165C---AC....	4393
H_CF.90CF056C-----	4343
J_SE.SE92809G-C-----	4310
CRF01_AE_TH.CM240C-----	4570
CRF02_AG_NG.IBNGG-C-----	4521
CRF03_AB_RU.KAL153G-C---AC....	5041
CRF04_cpx_CY.94CY03G-C-C-----	4362
N_CM.YBF30CG-G-----	4590
O_MVP5180AG-A-----	5026
O_CM.ANT70GG-A-----	5050
HIV2_A_STA-G-TAGGGGC-..	4823
HIV2_A_BENA-G-TAGGG-CA..	5378
HIV2_A_ALIA-G-TAGGG-CA..	5371
HIV2_B_D205A-G-TAGGG-CA..	5351
HIV2_B_EHOA-G-TAGGG-CA..	5349
STMM83293A-G-TAGGG-CA..	4942
SMM251A-G-TAGGG-CA..	5271
SMM9A-G-TAGGG-CA..	4768
SIVCPZANT	GCGGCGACCTCCTCACGAAGGCTCCGCGAAAAGTCACGGAAGCTGGCATAGCCCAGCTTGCGCACAAAGCGCGTCACCGACGCGTCGAGGTCCCCGTCGCAGCACTGATCTCGTCCA-GTTCTCCAAG-C	4678
SIVCPZGABG--C-AG-G....	5057
SIVCPZUSA-G--A--CGAG....	5056
SIVSTAN1A---AGG-GAG....	5195
SIVGRI677A---GGGGGAA....	4999
SIVVER9063A---AGGAGAA....	5223
SIVVER155A-G--AGG-GAA....	5215
SIVVERTYOA-G--CGGA....	4710
SIVVERAGM3A--GAGG-GTG....	4718
SIVLHOESTA---AGGGGAG...A	5357
SIVSUNA--C-GGG-GAG...A	5356
SIVMNDGB1A--T--C-AGAA....	4683
SIVSAB1CAC-GG--....	5394
SIVSYK173ACAG-GGAGGGA--	5193
POLQ T E E G D	POL

```

B_FR.HXB2R      CATAAAAGTAGTGCCAAGAAGAAAAGCAAAGATCATTAGGGATT..|-> Vif      Pol p31 integrase end <-|
POL             ATGGAA.....AACAGATGGCAGGTGATGATTGTGTGGCAAGTAGACAGGATGAGGATTAGAACATGAAAAAGTTTAGTAAAA 5118
VIF             I_K_V_V_P_R_R_K_A_K_I_I_R_D..#_Q_#_#_Q_M_A_G_D_D_C_V_A_S_R_Q_D_E_D_$_      POL
               M_E_.._N_R_W_Q_V_M_I_V_W_Q_V_D_R_M_R_I_R_T_W_K_S_L_V_K      VIR

A_UG.U455      T----G---A-----A-----A-----A-----A-----G-----A-----A-----C-----G----- 4564
C_ET.ETH2220   ----G---A---G-G---A-----A-----A-----C-----G-----A-----A-----T-----G----- 4510
D_ZR.84ZR085   ---G---A-----A-----A-----A-----A-----C-----G-----A-----A-----T-----G----- 4640
F1_BE.VI850    A---GA---A---G-G---A-----A-----A-----T-----G-----G-----AC-----AC-----C-----G----- 4458
G_SE.SE6165    A---G---A-----A-----A-----A-----A-----G-----G-----G-----AC-----C-----G----- 4515
H_CF.90CF056   A---G---A-----G-G---A-----A-----A-----A-----G-----G-----G-----AC-----C-----G----- 4465
J_SE.SE92809   A---G---A-----A-----A-----A-----A-----G-----G-----G-----AC-----C-----G----- 4432
CRF01_AE_TH.CM240 T-----A-----A-----A-----A-----A-----G-----G-----G-----AC-----C-----G----- 4692
CRF02_AG_NG.IBNG T---G---A-----A-----A-----A-----A-----G-----G-----G-----AC-----C-----G----- 4643
CRF03_AB_RU.KAL153 T-----A-----A-----A-----A-----A-----G-----G-----G-----AC-----C-----G----- 5163
CRF04_cpx_CY.94CY03 T-C-----A-----A-----T-----C-----C-----A-----G-----G-----A-----A-----C-----C-----G----- 4484
N_CM.YBF30     T-----C---C-T-G---A-A---A-----A-----G-----G-----G-----AA-C-G-A---T-C-----C-----G----- 4712
O_MVP5180     --T---G-A---G---G---A-A-C-A---A-----A-----AC---A-A---A---ACA-AAG-G-A-G---C-CC-G-G--- 5148
O_CM.ANT70     --T-G---A-----G-G---A-A-C-A-G---A-----A-----AC---A-A---G---ACA-AAG-G-A-G-G---C-CC-G-G--- 5172
HIV2_A_ST      ---A-A-A---G---T---C-AA-C---GGAAGGC--G---ATA-C-G-TCCAAC--AGG-GCCAG--G--T...GAG-GG--C-T--CC-TA-C-G 4948
HIV2_A_BEN     ---A-----G-G-C---T-C---GGAAGAC-GG-AC---ATA-AG-CCCCAC--AGG-GCCAG--G--T...GAG-A---C-TGCC--T-C-G 5503
HIV2_A_ALI     ---A-----G-G-C---C-A-C---GGAAGGC--G--C---ATA--G-CCCCACT--AGG-GCCAG--G--T...GAG-AG--C-T--CC-T-C-G 5496
HIV2_B_D205    A-C-----A-C-G---A-T-A---C---GGAGGAA--GGAT--ATT-AG-CCGACA---AGGA-CCAG-C-G-CTA...GAG-G--C-C--C-GA-T-G 5476
HIV2_B_EHO     A-C-----A-A-C-----A-A-C---GGAGGAA--G-AT--ATT-CAG-CCGAC---AGGA-CCAT-C-G-CTA...GAG-GG--C-C--C-A-T--- 5474
STMM83293     --T-----A---G---G-T-A---C-AA---GGAGGAA--G--G---ATA--G-TCCCAC--AGGA-CCAG--G-CT...GAG-GG--C-T--CC-CA--- 5067
SMM251        --T-G---A-C---G-T-A-T-C-AA---GGAGGAA--G--G---ATA-CAG-TCCCACA---AGGA-CCGGA--G-CTA...GAG-GG--C-T--CC-CA--- 5396
SMM9          G-C-G---A---G-G---T-A-T-C-AA-C---GGAGGAA--G-AT--ATA--G-TCCCAC--AGGA-CCGG--G-CTA...GAG-AG--C-T--CC-CA-C--- 4893
SIVCPZANT      ACATG-G-A-AA-ATCTTC-C-GT-CTG-GC-ATGACT-CATCG..G-C-GCGTCATCCTGTG--A-A.....AT-TC-----TCC-----G--- 4788
SIVCPZGAB      AC---G---A---G---A-----A-----A-----A-----A-----A-----AG--T---T-----T----- 5179
SIVCPZUS       GG-T-----A-C-----A-T-A-AA-C---C---A-A---G-----A-----AG--AGG-T---T---C-G--- 5178
SIVSTAN1       GT-G-----A---G---G---A-T-C-AA-A---GAGAGAA--ACTG--GT-ACA-GAC-AACA--AGG--TAGTG-AGA-CAT--AGAC-A---GG-CA---G... 5320
SIVVER155      --C-G---C-C-G---G-T---T-C-AA---GAGAGAA--ACA---ATA---G-G-A---A---AGG--GTGAGA--G-CAA--A--C-G---GGG-GA---G... 5124
SIVVER9063     GC-G---A-T-C---A---C-AA---A-CCAAGAA--AGA---GT-A---G-G-AACA--A-G---CTCAG--GACCT--A-CC-A---C--G-AA---C... 5348
SIVVER155      AC-G---G---T-G---G---A---A-AA-C---AGCCAAGAA--ACAT--GT-A---GAC-CACC--A-G---GC-GGA-GAAGT--C-CC-A---C-GG-GA---T... 5340
SIVVERTYO      -C---G---T-A---G---T-A---A---A-CCCAAC--AGAG--GTAA--G-G-GAC--A-G---CCAG--GATCT--A-CT-A---C-GG-AA---C... 4835
SIVVERAGM3     AT---A-TAC---G---A---A---A-CCAAGAA--AGA---GT-A---GAG-AACT--A-G---GCCGA-GA-CT--A-CT-A---C-G-GA---G... 4843
SIVLHOEST      ---CTTCTC--C-C---AC-AG-A-AA---G...AGGG--CA-A-ATA---GAG-A-CC--AT-AC-ACT--ATCCTTA--AGA--AG-TTTT-GACA---C-GG 5482
SIVSUN         ---CCTG---A-T-G---AC-AG-A-AA---G...G...G---AATCCTCC-CA---AG-TGGTGGGATGAA...G---G---GA-AG-AG-C-G 5463
SIVMNDGB1      G---GA---A-T---GTGT--A-A-A-AA---G...AGAGTGA-A-AATA--C-G-T-AACT--A---GTCT--TCA---A-AGA--AG--C-CT-G---G 4808
SIVSAB1C       AT-G---AC-A-C---C-A-T-A-AA---G...G...G---GCAT--ATA--C-G-CCCC-C---A-G---ACGG-CGGACA-CAGGAG-GG---CT--CC---G--- 5516
SIVSYK173      ---TTTT-C---A---G---A-A---CA-CC---G...G...G---AGA---ATA--G-TCCCAC--AG-A-GAC-CCA-GACA---GAT-G--T-C-GCA-A--A--- 5315
POL            I_F_A_V_P_R_R_K_A_K_I_I_T_D..#_Q_#_#_R_M_D_S_G_S_H_V_E_N_D_P_K_T_D_$_      POL
VIF            M_E_.._K_E_W_I_V_V_P_T_W_R_M_T_P_R_Q_I_D_R_L_Q_H_I_I_K      VIF

```


B_FR.HXB2R	CACCATATGTATGTTTCAGGAAAGCTAGGGGATGGTTTTAT...AGACATCACTAT.....GAAAGCCCTCATC	5185
VIF	H H M Y V S G K A R G W F Y . R H H Y E S P H	VIF
A_UG.U455	--T-C-----C-CAA-----CAA--T-----...-----TAGA---T	4631
C_ET.ETH2220	-----C-A---A--G---AT---G-----...-----T-----C---AGA---T	4577
D_ZR.84ZR085	T-----C-A---AA-----AA-----C-----...-----C-----C---	4707
FL_BE.VI850	T-T-----AA-----C-AA-----C-----...-----T-----C---AGG---T	4525
G_SE.SE6165	--T-----C-AA-----A-C-----...-----C-----C---AGG---T	4582
H_CF.90CF056	T-----C-A---A-----A-----...-----T-T-----A-----	4532
J_SE.SE92809	T-----A---AA-----AAA---C-G-----...-G-----C---AA---T	4499
CRF01_AE_TH.CM240	--T-----A-C---AA-----AAAAG-----...-----T-----C---AG---T	4759
CRF02_AG_NG.IBNG	--T-----C-AA-----A-T-----...-----T-----TAGG---T	4710
CRF03_AB_RU.KAL153	-----A-A---AA-----G-----A-----...-A-----AGGA---T	5230
CRF04_cpx_CY.94CY03	--T-----AA-----AA-----A-C-----...-----TAGG---C---	4551
N_CM.YBF30	--T-----G-AAA-G-A-AA-----...-----T-----CA-A---C---	4779
O_MVP5180	T-----AA-CA-G-TAA--G-CGC-AAC--CG-----...-G-----T-----GAATCCA	5209
O_CM.ANT70	T-----A--CAGG--TA-A--GA-CGA-AAC--GG-C-----...-----T-----GAATCCA	5233
HIV2_A_ST	T-T-TA-AA--CAGAA---AG-TCTAGA-AAGGT-GC--GTTCCC-C-TA-G-----...-TGG-ATGGGCGT	5018
HIV2_A_BEN	T--TG-AA--CAGAA--AAAG-CCTAGA--AGGT-CGC--GTTCCC-C--A-G-----...-TGG-ATGGGCGT	5573
HIV2_A_ALI	T--TA-A--CAGAA--AAAG-TCTAGAAAAGGT-GC--GTTCCC-C-TA-G-----...-TGG-ATGGGCGT	5566
HIV2_B_D205	T-T-T-A--AGAA---AG-GTTGCAACAGGTC-C--GTCCT-C--A-G-----...-T-G-ATGGGC-T	5546
HIV2_B_EHO	T--TG-AA--AGGA--AAAG-CTTGCAACAGGTC-C--GTTCCCT-C-TA-G-----...-T-G-ATGGGC-T	5544
STMM83293	--TTTA-A---AACA-TAAAG-GCTA--CAAGGCA-GC--GTGCC-C-TA-G-----...-TTG-GTGGGCAT	5137
SMM251	T-T-TG-AA---AAAA-TAAAG-TCTACAAAAGGTT-GC--GTGCC--TTTA-G-----...-TCG-ATGGGCAT	5466
SMM9	----TG-AA---AACA-TAAAG-CCTACA-AAGGCT-G--GTACCC--TA-A-----...-TTG-ATGGGCAT	4963
SIVCPZANT	-----C-A-GG-AGA-TAAAGT-CT--AACC---AAG---...-----T-----ATGA---C---	4855
SIVCPZGAB	T-T-----T--AGG--AAA-G-A-A-T-----C-----...-----T-----TCA---A---	5246
SIVCPZUS	T-T-C-----AGA--AA-----AA-----C-C...-G-----T-----G--A-A---	5245
SIVSTAN1	..A-ATACC-CA-GAGGAACCGTTACA--AT--ACA--...CTTATG--C-----...CAATGTG-ATGGGC-T	5387
SIVGRI677	..AC-TACA-GA-CAGGAAT--CAATT-CCT--GAA-C...-----T-GG-----...CAG-TGCAATGG--GT	5191
SIVVER9063	..A-ATAT-GGA-GA-TAAA-GGAACCTTAA---CAA-C...-TGATG--C-----...CAAAT-ATTGGGC-T	5415
SIVVER155	..AGATAT-GGA-GAATAAA-GGAA-CT-AA---GAA-C...-A-ATG-T--...CAAATC-CTTGGGCAT	5407
SIVVERTYO	..AGATAT-GGA-GAG-CAA-G-AA-TT-AA---AAC---...TATATG--T-C-----...CAAAT-CATGGGC-T	4902
SIVVERAGM3	..AGGTAC-GGA-GAGGACT-G-AAATTA-AC---AAA---...C-ATG-----C-----...CAAAT-CATGGGCAT	4910
SIVLHOEST	TTAA-C-CAG-G-CAAA--AA--TTGGGTA-C-TACA-GGAACAG-T.....-T-G-TTGGG-AT	5543
SIVSUN	T-TA-G--AGT-AGAAT--TATGGCT--TA-ATA-AA--GC-GTAGAGA-ATTT-TAGATTTAAGAAGATGCATAGGGAAACAAAAGATGATTGGGTATCTATGTATGGAACAGGTAAGT-ATGGG-GT	5593
SIVMNDGB1	AGA--G--GCATGGG-CACTGC-AA--AT-AGGAAGGA-G-TGGT-G-TGT-TCC.....CATTTTATG-CTTATAA-G-AT	4887
SIVSAB1C	T-----C--C---G-CAAAC-GTG-GTTCAT---CGG---...-CC-C---TAC-----...AAAAT---ATGGA--T	5586
SIVSYK173	ACA-----A---AAAAGCAAAG--TTGGA-AA-GCCACA---...-A-----C-----...CAAAT-GAATGG--AT	5385
VIF	T H K Y K S K E L E K A T Y . K H H Y Q I E W Q	VIF

B_FR.HXB2R	CAAGAATAAGTTCAGAAGTACACATCCCCTAGGGGATGCTAGA.....TTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGACTGGGCAT.....TTGGGTCAGGG	5291
VIF	P_R_I_S_S_E_V_H_I_P_L_G_D_A_R.....L_V_I_T_T_Y_W_G_L_H_T_G_E_R_D_W_H.....L_G_Q_G	VIF
A_UG.U455	-----G-----A-----A-G-G-----A-----C-----T--	4737
C_ET.ETH2220	--AGG-----T-----G-----AA-A-----T-----A-----T-----T--	4683
D_ZR.84ZR085	-A-----T-----A-----C-----G-----A-----C-----C-----	4813
FL_BE.VI850	---G-----A-A-T-A-----A-----G-----CC-----A-----C-----	4631
G_SE.SE6165	---G-----A-----C-----C-A-G-----G-----A-----A-----T--	4688
H_CF.90CF056	--G-----T-----A-A-G-----A-C-----C-----C-----A-----A-----A-C--	4638
J_SE.SE92809	--A-----A-----G-----T-----A-----G-----G-----T-----C-----	4605
CRF01_AE_TH.CM240	--AGG-----C-----T-----A-----G-----A-----G-----A-----AG-----C-----T--	4865
CRF02_AG_NG.IBNG	--A-G-T-----G-----A-----G-----G-----C-----C-----C-----T--	4816
CRF03_AB_RU.KAL153	-----T-----A-----A-----A-----A-----C-----C-----	5336
CRF04_cpx_CY.94CY03	--A-G-----T-----G-----A-----G-----G-----GC-----G-----CA-----C-----T--	4657
N_CM.YBF30	--A-----T-----G-----TC-G-A-----A-----G-C-GTC-T-----G-----AAC-----CAGTCT-----C-A-A--T--	4885
O_MVP5180	GG-ATCC--AAGTCAGTTCGGCGG-GTATA-TCCA-TA-AGA-GCTGATA-A-GG-C-C-----AT-AATGC-----G-----GGAA-----C-----A-T--	5321
O_CM.ANT70	G--ATCC--AGTTAGTTC-AGTG-ATATA-TCCA-TA-GGGTGGCCCATG-A-G-----C-----AT-ATGC-----G-----TGAA-----A-T--	5345
HIV2_A_ST	GGT-G-CTT-CAGCAGG--ATAT-----T-AAA-GA-AA-TCAT...C--AG--CAGG--C--AAC-A...--CC--A-GA--TCTCCTCCTATTGAGTA.....	5128
HIV2_A_BEN	GGT-G-CTT-CAGCAGG--ATAT-----CAA-GAAAA-TCAT...C-A-A--CAGG-----AAC-A...--CC--A-GA--TCTCCTCATGAGTA.....	5683
HIV2_A_ALI	GGT-G-CTT-CAGCAGG--ATAT-----T-CAA-GAAGA-TCAT...C-A-AG--CAGG-----AAC-A...--CC--A-GA--TCTCCTTATGAGTA.....	5676
HIV2_B_D205	GGT-G-CTT-CAGTAG-A-ATAT-T-C--AAC-A-GAGC-TGG...C-A-A-G-CCA-GG-----AAC-A...--CCC--G-GA-TCTTGAGCTCCTATGCTGTA.....	5656
HIV2_B_EHO	GGT-G-CTT-CAGTAG--ATAT-T-C--GAAA-A-GAGC-CAT...C-A-A-G-CCA-GG--C--AAC-----CCC--G-GA-TCTTGAGTTCCCTATGCTGTA.....	5654
STMM83293	GGT-G-CTT-CAGTAG--GATTT-T-CT-GCAA-GA-AAGC-CAC...C-A-A-G-TCA-GGG--AA-T-A...--CC--A-GA--TCAGTGAATATGCTGTG.....	5247
SMM251	GGT-G-CCT-CAGCAG--AT-T--C--CA--A-GA-CCAT...C-A-A-G-CA-GGG--CA-T...--CC--GG--TCAGTACTTATGAGTAGT.....	5576
SMM9	GGT-G-CTT-CAGCAG--AT-T--GT--A--AGGCTCAT...A-G-CA-GG--AA-T...--CC--GRA-GG--TCAGTACTTATGAGTAGT.....	5073
SIVCPZANT	--AG-A-G-AGA-----T-A-----CCAACATTGGATAAAAAA...G-G--GTG-TC--AT-A-ATGT-----C-CCCA-----C-----A-A-T--	4967
SIVCPZGAB	--A-G-TGC--T--A-----A-TT-CA-A--TA-TCTAAA...AA-G-G--C-----CCT-AAGCC-----G-C-----A-C-T--	5355
SIVCPZUS	-C--G--A-----T-----TA--ATA-A-A-A...A-A-G-T-T--G-----GT-A-A--CA-C-----C-----A-G-T--	5351
SIVSTAN1	GGTAC-C-T--AGTAG-T-C-T--CT--A-GA-AAG-AAA...A-A-G-GATTGT-C--CAC-A...--GCC--CAG-GA--TCTCTACCTATGAGTAGGAA-A.....	5503
SIVGRI677	TTT-G-CCTACAGCC-GT-CATT--CT--TCAA-A-A-GATTAC...A-A-A-G-AT-TT--CACAA-C...--CCC--GA--TCTCAAGTCATGGAGTAGGG-A.....	5307
SIVVER9063	GGTAT-CT-TG-GCAG-TACAT--A--ATT-AAG-GAA...A-ACAG--GACTGC-----CA-T-A...--CC--G--TCTCTACTACGCCATAGGAA-A.....	5531
SIVVER155	GGTAC-CT-TGAGCAG-TATGTA--A-C--CCCA-GAAG-G-GAA...A-CCATG-GGAT-TC-----CA-T-A...G-TCC-A-CA-GA--TCTCAACTTATGAGTAGGAA-A.....	5523
SIVVERTYO	GGTAC-CC-TGAGTAG-TATGTA--A--A--AA-ACA-G-GGAA...A-ATGTG--GACCT--CA-T-A...--CC--GCA-GA--TATCCACATATGAGTAGGTA-A.....	5018
SIVVERAGM3	GGTAC-C--TGAGTAG-TATG-G-A-C--C-ACA-G-AGT...A-ACATG--GATCT--CA...--CC--AG-GA--TATCAACATATGCTGAGGGGA-A.....	5026
SIVLHOEST	TCTAT-CCTA-CATA-T-TGT-A-T--A-TCA-GA-CTTTA...G-A-C-G...-T--GCACA-T...-G--C-G-GA--ATCTACAATGGGCAGTT.....	5650
SIVSUN	GGTAC-C-TACAATA-A-TATT-T--G--ACTT--GG-C-GTA...G-A--GG...-TC--G-ACAC-C...--CC--C-A-GG-T-AT-AACCAGTGGGGATGT.....	5700
SIVMNDGB1	GGTAT-CTT-CAGTA--GTG--TAT-A--AATAGG-AC-T-AGA...--AA--G-T-G-GC-----CA-T...CA-AT--GGT-GA--CTTAAGTACTTATGAGTAAGCA-A-AAGCA-T	5011
SIVSAB1C	GGTACTCCTA-CA--TGGGT--A-C--GAA--GAGCTCTA...A-AAA-G-T-T-AT-----CA-T-A...--CCC--G-A-GA--TTAGAAACTTATGCAACA.....	5696
SIVSYK173	GGTAC-CTTA-GCC--TGGACA-A--G--A--GA-CTATA...G-A-C-CG...TTC--CACAA-T...G-CCC-----GA--TTACATATGAGGGGAATT.....	5492
VIF	W_Y_T_Y_C_Q_W_T_I_P_V_G_D_G_T_I...W_I_T...F_Y_H_N_L...A_P_E_R_G_W_L_H_M_Q_G_I.....	VIF

B_FR.HXB2R	AGTCTCCATAGAAATGGAGGAAAAAGAGATATAGCACACAAGTAGACCCCTGAACTAGCAGACCAACTAATTCATCTGTATTACTTTGACTGTTTTTCAGACTCTGCTATAAGAAAGGCCATTATTAGGACAC	5421
VIF	V S I E W R K K R Y S T Q V D P E L A D Q L I H L Y Y F D C F S D S A I R K A L L G H	VIF
A_UG.U455	G-----CTG-A-----T-----C-----T-----C-C--T-----A---C---G-GA---A-----A	4867
C_ET.ETH2220	-----ATTG-GA-C--A-----GC-G-----C-----A-C-T-----T-----G-A---C---G-A---A-----T-T	4813
D_ZR.84ZR085	-----G-----GC-G-----A-----T-----T-----G-----C---A---A-----T	4943
FL_BE.VI850	-----C-GGG-AG---G---A-----GC-G-G-----G-C--A-A--T-----T-----A---C---G-A---A-----T	4761
G_SE.SE6165	G---G---C-G-GA---A-----G---G---T-G-----C-T-----T-----C---G-A---A-----A	4818
H_CF.90CF056	-----ACTG-A-G-----G---GC-G-----A-C-T-----T-----A---C---G-A---A-----GT	4768
J_SE.SE92809	-----C-G-GA-G-C-A-----GC-G-----A-G-----A-G-----T-----C---G-A---A-----A	4735
CRF01_AE_TH.CM240	-----C-G-GA-A-----A-T-----C-----A-G-----AC-A-T-G-----C---G-A---A-----A	4995
CRF02_AG_NG.IBNG	G-----A-C-G-A-----A-----CT-----G-----C-----T-A-T-----C---A---A-----G-A	4946
CRF03_AB_RU.KAL153	-C-----A-G-----A-C-----T-----T-----A-----T---A-----T	5466
CRF04_cpx_CY.94CY03	-----CTC-GA-----G-T-----C-G-----A-C-----T-----A---C---G-A---A-----T	4787
N_CM.YBF30	--A-----ACT--GA-A--C-AG-----T-T-----A-G-----A-G-----A-TC-----T-----A-C---C---GC-A--GG-C---GAGA	5015
O_MVP5180	G--TAGT-----CAAT-C--GAG---AA---GA-T-T-----AC-----AGGA-G-A-----C-T-CACA---A---A-A-A-C-G---A-TC---G-G	5451
O_CM.ANT70	G--TAGT-----C-TT-T-----AG---AA---GA-T-----AC-----AGGA-G-A-----C-T-AC-----A-CA-A-A-AG-C-G---A-CC---G-G	5475
HIV2_A_ST	...AGAC--ACT--TAT-C-G-A-A--TCT-G--G-T-TAC--A-CTGT--G--TCC--A--AGCAC--T-CTCT-C--A-G-CAGGC-AAG---GA--A-CAG--GG-A	5254
HIV2_A_BEN	...AGAT--ACC--TAT-C-G-A-AG-TCT-G--G-T-TAC--A-CTGT--ATC--A--AGCAC--T-CTCT-C--A-G-CAGG-AAG---GA--A-CAG--GG-A	5809
HIV2_A_ALI	...AGA--ACT--TAT-C-G-A-AG-TCT-G--G-T-TAC--A-CTGT--TCC--A--GGCAC--T-CTCT-C--A-G-CAGG-AAG---GA--A-CAG--GG-A	5802
HIV2_B_D205	...AGAC--AC--TATG-G-G--AC-T-TAT--G-T-ACA--TG-G--G--C-G--GG-C-T-CTCT-C--C-AA-AAG--G-GA--A-CAGG--G-A	5782
HIV2_B_EHO	...AGAC--AC--TATG-G-G--C-T-TAT--G-T-ACT--TG-----G-T-C-G--GG-C-T-CTCT-CG--A-CTAA-AAG--G-GA--A-CAGG--G-A	5780
STMM83293	...AGG--ACT--TAT-CC-GA-AT-TCT-GT--G-T-ACA-A-CTGT--T-G--C-G--GGCAC--T-CCCT-C-A-CGGGA-AAG-G--GA--A-CAGG--G-A	5373
SMM251	...AGG--ACC--TACTC--G-AC-T-T-G--G-T-ACA-A-CTAT--ATTT--C-G--AGCAC--T-CCCT-C-A-CGGGA-AAG-G--G--A-CAGG--G-A	5702
SMM9	...AGA--ACC--TACTCC-GA-AT-TCT-G--G-T-C-ACA-A-TTAT--ACTT--C-G--GGCAC--T-CCCT-C--GGGA-AAG-C--G--A-CAGG--G-A	5199
SIVCPZANT	T--GAG-----ATGTGGA-AG--T-----A--AC--T-T--A-G-A-C-AA--T-CCC-----TCAA-AG-C--C-A-AA-CC--G-A	5097
SIVCPZGAB	---A-A--TC-----CT-GGATCT--GT-----G-G-T-CTTCAC--TAG-T-G-C-CTCTC-A-T-----T-C-G-AA--C-C--GA--TA-----G-G	5485
SIVCPZUS	---A-A--AC--GGA-CT-C-AG--A--T-----GA-G--A-G-----T-----T-----A-C---C---A-TG-T---GAGA	5481
SIVSTAN1	.AGT-TTGAGA-T--CAA-C...-CG-C-AA-G--ACT--TG-G-G--TA-G--TGCC-C--CCCT-C-A--TAGA-C--CCA-C--A-CAG--GG-A	5629
SIVGRI677	.TC--ATTACC-TCAA-A-GG-...-AG-G--T--A-G-AC--AG-A-G-A-C-A--T-A--A--TAGA-C--CCA-C--TA-CAG--GG-G	5430
SIVVER9063	.CAA-A-T-G-G-CGAGTAGG-...-AT-TC--A--G-T-----A-C-AC--TGG-A-G-A-C-AC-C-T--TCA--A--AAGG-C--TCA--A--TA--AGG--G-G	5657
SIVVER155	.CAA-ATG-TAGCCTAGTA--TG-T-A-----A-G-T-----T-CA-TAC--TCCA-G-A--TGTC--T-AC-----A--TAGA-C--CCA-C--AC--AGG-A--	5652
SIVVERTYO	.CAA-ATG-AGCAATTTAG--TCT-A-----A-G-T-----CTAC--TAGTA--A--GGTC-C-T--A-T-----AA--AAGA-C--CCA-C-A--TC-GAGG--C--	5147
SIVVERAGM3	.CAG-ATC--AGCAAT--G-TCCTT-G--G-G-T-G--T--C-AC--TAGC--G-A--ACCC--T-ACT-----A--AAGG-C--C--G-A--C--G--G	5155
SIVLHOEST	...GTCT-T-GAGAC--CAA-G-A-----CAT-----A-G-C-CTATA-A-T-A--C-CAT-C-AAGA-AC--T--G--AC-TC-G--G-A	5776
SIVSUN	...AGT--G-----TTT-T-C-T-CAA-G-A--T-CTTG-G--T--GA-G-C-CT-TA-G-T-----T--GGA-CTCTAGGTG--C--GA--TA-G--GG-A	5826
SIVMNDGB1	---TAGACCGCC-CCCTTTG-G--AGAG-GGT-T--G-GA-ACT-A--GG-----T-T--A--T-AC--TT-A-----C-CATG--AG--AG-T-TG--A--A-CAGG--G-A	5141
SIVSAB1C	...GGA--G--AC--T--GG-GA--GGTTT--G-GC-T-T--TGGAC--G--TA-T-----TG--CC-T-CCC-----A--TAGA--G-CCA-C-A--A-CAGG--G-A	5822
SIVSYK173	...AGG--TC--ACCAATGG--CCAG-GG-A--TG-CT-ACA-A-C-G-----GC-T-A--AAT-T--T-CCCT-C--A--CAAGA--GG-C-ACC--AG-CAG--GG-A	5618
VIF	...R I Q Y Q W N Q W N T D L T P A V A D R L I H N F Y F P C F T A R A V N Q A V R G E	VIF

B_FR.HXB2R	ATAGTTAGCCCTAGGTGTGAATATCAAGCAGGACATAAC.....AAGGTAGGACTCTACAATACTTGGCACTAGCAGCATTAAATA.....ACACCAAAAAAGATAAAGC	5521
VIF	I V S P R C E Y Q A G H N.....K V G S L Q Y L A L A A L I.....T P K K I K	VIF
A_UG.U455	-----A-----T-----T-GAA-----G-----C-----C-G-GC-----	4967
C_ET.ETH2220	-G-----C-----T-----G-----G-----G-----AG-----GC-----	4913
D_ZR.84ZR085	-----T-----CT--G-T-A-----A-----G-----G-----G-----	5043
FL_BE.VI850	-G-A-----A-C-----T-----C-----T-----A-----G-T--G-G--C-----	4861
G_SE.SE6165	-----T-----C-----T-----T-GAA-T--G-----T-----G-GT-G-----	4918
H_CF.90CF056	G-----A-----A-C--C-----A-----C-----A-----T-----G-G-----G-----	4868
J_SE.SE92809	-----C--C-----TC-----A-----G-----G--G-----	4835
CRF01_AE_TH.CM240	G-----A-G-----C-T-----C-----T-----GAA-----C-----G-----G-----	5095
CRF02_AG_NG.IBNG	-----A-----T-----G-----G-----T-----GAAT--T--G-----G-----C--C-----	5046
CRF03_AB_RU.KAL153	-G-----T-----G-----G-----G-----G-----	5566
CRF04_cpx_CY.94CY03	-G-----T-----T-----CT-----C-----T-C-----C-----	4887
N_CM.YBF30	CC--ATTA-----C-----G--C--A-----C-----CA-C-----TC-A-----A-----C-GGG-G-----GG-G-----G-----	5115
O_MVP5180	-G--GCTGA-C-A-----C-TG-----GT-----C-----GA-A-----T--A-CT-GAA--G--G-G-----A-GT--GA-AT-----	5551
O_CM.ANT70	-G--ACTGA-C-A-----C-CTA-----GT-----C-----GA-A-----CTGC-A--T--AG--G--G-----A-G--G--GC-GG-----	5575
HIV2_A_ST	-AGC-ATTAT-CT-C--CA-C--C-CCCA--CC---GTAC.....C---CCG--A-C-G-TTC--CT--TG-T-G-GCA.....CA-AATGGC-G.....	5350
HIV2_A_BEN	-AGT-ATTGT-CT-C--CA-C---CCCA--CT---AGCA.....C---CC--A-T---C-A-C---T-T-G-CA.....CA-AATG-C-GA.....	5905
VIF	-----V V V Q-----Q N D R-----	VIF
VFX	-----S Y #-----K M T #-----	
HIV2_A_ALI	-AGT-ATTGT-CT-C--CA-T---CCCAG-CC---GTCA.....C---CCG--A-C---TTC--CT--TG-T-G-GCA.....CA-AATGGC--A.....	5898
HIV2_B_D205	-AGA-ATTGT-CTAC--CA-C---C-T---CT--CG-AGGG.....C---CC-AGCT--G-TTC-A-C---AGG-TCG--CAG.....GA-GG---TGG-....	5881
HIV2_B_EHO	-AGA-ATTGT-CCAC--CA-C--C-C-T---CT---CAGGC.....C---CC-AG-T--G-TTC-A-C---AG--TTG--CA.....GA-GG---G-TGG-....	5879
STMM83293	-AGT-ACTGT--T-C--CAGG-TCACGAA--CT---GAAC.....C---CC-AGCT--G---A-TT-AA--T-G-GGA.....CATGTC-G-----	5466
SMM251	CA-C-GCTGT--T-C--CA-G-TC-CGAG--CT---GGTAC.....C---CC-AGC--G---A---AA--T-G--GC.....GATGTC-G-----	5795
SMM9	-A-T-GCTGT--T-C--CA-G-TC-CGAA--CT---GAAT.....C---CC-AGC--G--TC-A-----A--T-G--GT.....CATGTCRSR.....	5292
SIVCPZANT	-G-A-ATTGA-ATAC--CC-C--CA-GAAG-----GT.....C---GA-A-G-G-T-A--TT-TTGTAAGA-CC--GAATTCAGAGGATAC--C--GGGTCC--GAA	5209
SIVCPZGAB	C-T--AGCG--C-T-----CA--A-----C-GG.....C-----CT--G-T-----C--AA--T-----T--GAG-G--G-CAT-GA-	5585
SIVCPZUS	CC--A-TT--C-AA-----C-CTA-----G-C-T.....C---C--T-G--T-----TT--AAG--C--G-----GG--AG-GC--A-G--GA-	5581
SIVSTAN1	-GCT--TTGTGG...CAC--CA-G-A-----CGTAGCAGAAAACCATTGGGGTC--C-G-C-G-G-TT--A--C-GA--T--AT-C.....GACTTCCT--GA-ATGGCA	5744
SIVGRI677	-AGTA--CGTGG...CAC--TCA-G-A-----A.....GGTC--CA--A-G--CTT-----TT--AT-C.....-ATGGC-TC-G--AG-GA-	5529
SIVVER9063	-G-T-CGT-TTC...C--TC-CC-AG-----AACCACA.....GGA--CCC-A-G-G---A-C--TCTG--CCATCAG.....-ATGGC-TC-G-GAG-G-	5762
SIVVER155	-GGT-C-T-TTC...C--T--C--G-----ACTAACA.....GGTC--CCC-CT-G---T-A--T--CT--CCATCA.....-ATGGCCTC-G--AG-GA-	5757
SIVVERTYO	-G-T--GT-TTC...C-G-T--C--A--G---AAGCAC.....GGAC--CC--T-G-G--C-A--T--GCTC--CATCA.....-ATGGCCTC-G-GAG-GA-	5252
SIVVERAGM3	-GGT-C-C-TTC...C-G-TC-CC-AG---C--GAAAACA.....GGAC--CCC--T-G-----A--T--CCTT--CACCA.....-ATGGCCTC-G-CAG-GA-	5260
SIVLHOEST	-G-A-A-CTATGT-T--CC-G...CTTT-----AGGT.....TGT--TCTC--T-G-G-TTA-CTGCT-GAG-CA-C--CA.....CATGTCC-GGCA.....	5869
SIVSUN	-A-A-ACTA-A-GA--CAG-A-----T--CT--C--AGGC.....CT--TCTC--T-G-G-TTC-CTGCT-GAG--TTC-CCAT.....GGC-AGC-GG-----	5922
SIVMNDGB1	GA--GTTAAAAGTT--AG--T--C--T--C---AGCA.....CAAGGT--TCTC--T-G-G-TTC-CTGCT-GAG--TCA-CTAT.....GGG--G--G-----	5240
SIVSAB1C	-AGTA-CTTTGG...A-C-C--G-TG-----C-GCCAACA.....GGGC-A--CC--AT--G--T--A---TAGG-T...TAT.....--AATGGCCTC.....	5918
SIVSYK173	C--T-G-CAT-ACAC--CTGGACC-C.....CTGAC.....C---CCC--A--T--G--T--A--T--CA--TG...TAC.....CT-AA-G-TGGC.....	5705
VIF	L L T S H C W T P.....H T D.....Q V P S L Q Y L A L Q V...Y.....L K D G.....	VIF

B_FR.HXB2R	CACCTTTGCCTAGTGTACGAAACTGACAGAGGATAG.ATGGAACAAGCCCCAGAAGACCAAGGGCCACAGAGGGAGCCACACAAT.....	5606
VIF	P_P_L_P_S_V_T_K_L_T_E_D#.W_N_K_P_Q_K_T_K_G_H_R_G_S_H_T#.....	VIF
VPR	M_E_Q_A_P_E_D_Q_G_P_Q_R_E_P_H_N.....	VPR
A_UG.U455	-----A--T-A-----G-----GT--GC-----	5052
C_ET.ETH2220	---C---GT--T-AGT---A-----G---G---A--T-----	4998
D_ZR.84ZR085	-----C---A--G--A---A-----G-C-----G-----T-----	5128
Fl_BE.VI850	---GC---CCA---AGT---C---G-----G-----T-----	4946
G_SE.SE6165	-----C---G--T-AG--A---G-----A--A--CT-----	5003
H_CF.90CF056	-----GA--G--AGT-----G-----	4953
J_SE.SE92809	-----CA---AGT-----C-----A---T--C---	4920
CRF01_AE_TH.CM240	---C---A..G--A-A--A---G-----A--A--CT-----	5174
CRF02_AG_NG.IBNG	-----G--GT-AG--A--C---AG-----G-----GTC-----	5131
CRF03_AB_RU.KAL153	-----A---AGTG-----G---A-----T-----	5651
CRF04_cpx_CY.94CY03	-----A---AGTG-----G---G---A--A--A-T---	4972
N_CM.YBF30	---C--A---G--T--G--A---A---G--A---TG---A--CT-T---	5200
O_MVP5180	---CC-A--C---CCA--G-T-A---A---TG--A-T--G--GCT---TT---	5636
O_CM.ANT70	---CC-A--C---CCA--T-A---A---A--T--GA-T--G--GCT-AA---CTT---	5660
HIV2_A_STCCCCAG-GAG-CAAT--CACCAGG-A..ACAGTGGC-AAGA-ACTAT-GG--A-G-CTTCGA-TG--T-G--GGACGGTAGAAGCCATAAACAGAGAGGCAGTGAACCACCTGCCCCG	5467
HIV2_A_BENCCCCAG-GA--GGGT--C-CCAGG-A..ACAGTGG--AAGAGACCATTGG--A-G-CTTCGA-TG--T-G-G-GGACCATAGAAGCCTTAAACAGGGAGGCAGTGAACCATCTGCCCCG	6022
VIFP_Q_R_K_G_T_G_R#.#Q_W_R_R_D_H_W_R_G_L_R_V_A_R_E_D_H_R_S_L_K_Q_G_G_S_E_P_S_A_P	VIF
VPX#_P_R_E_R_V_P_P_G#_T_S_G_E_E_T_I_G_E_A_F_E_W_L_E_R_T_I_E_A_L_N_R_E_A_V_N_H_L_P	VPX
HIV2_A_ALICCCCAG-GA--CAGT--CACCAGG-A..ACAGTGGC-AAGAGACTAT-GG--A-G-CTTCGA-TG--T-G--GGACAGTAGAGGCCTTAAACAGAGAGGGCGGTGAATCACCTGCCCCG	6015
HIV2_B_D205TCCCAGGGAG-GAGTG-CACCAGG-A..ACAGCG-C-AAGA-AC--TAGG--AAG--TTCGCTTG--T-G-A-GAACAAATAACAGAGCTCAACAGGGTAGCGGTCAACCATTTGCCCCG	5998
HIV2_B_EHOTCCCAGGGAG-GAGT--CACCAGG-A..ACAGCGGC-AAGA-AC--TAGG--A-G--TTCGAATG--T-G-G-CAACATTAGAACATCTCAACAGAGTAGCAGTCAATCACTTGCCCG	5996
STMM83293TCCCAG-GAG-GAAT--C-CCCG-A..ACAGTGG--AAGAG-C-ATAG-G-AAG--TTCGA-TG--T---GAACGGTAGAGGACATAAACCCAGAGAGGCAGTAAACCATCTACCCG	5583
SMM251TCCCAGGGAG-GAATC-CACCTGG-A..ACAGTGG--AAGAGAC-ATAGG--A-G-CTTCGAATG--T-A---GAACAGTAGAGGAGATAAACAGAGAGGCAGTAAACCATCTACCCG	5912
SMM9TCCCAGGGAG-GRATC-CACCTGG-A..ACAGTGG--AAGA-ACRRTAGGC-A-G-YTTCGACTG--T---GAACASTAGAAGAAATAAACAGGGCAGCAGTGAATCATTGCCCRAG	5409
SIVCPZANT	G--AG--C-----C-GT-T-TCT-A---A-----GC-G--TG-G-----A--A--ATG---	5294
SIVCPZGAB	-----A-----CG-C---T-A-----A-----G---G---T---CA--A--A--TT--C-G.....	5670
SIVCPZUS	---C--A-----G--T---T-----A-----TG-----GCC--A--A--T-----	5666
SIVSTAN1	G-AGGAA-AGATTCCAGGGA--GAA--CCG-ATGGT..A--ATTTGG.....	5792
SIVGRI677TCCAAG-GA-CCT-T--CAG-ATGGC..---ATCTGG.....	5566
SIVVER9063TCCCAG-GA---A-G-CAG-AGA---A--GATCTGG.....	5799
SIVVER155TCCCAG-GAGG-GA---CAG-AGG-C..-A--ATTTGG.....	5794
SIVVERTYOTCCAAG-GAGGCAA---CAG-AGA---A--ATTTGG.....	5289
SIVVERAGM3TCCCAG-GA-GCAA---C-G--GA-C..-A--AT-TGG.....	5297
SIVLHOEST	5869
SIVSUN	5922
SIVMNDGB1	5240
SIVSAB1C	5918
SIVSYK173	5705
VIF	VIF

B_FR.HXB2R	5606
VIF	-----	VIF
VPR	-----	VPR
A_UG.U455	5052
C_ET.ETH2220	4998
D_ZR.84ZR085	5128
FI_BE.VI850	4946
G_SE.SE6165	5003
H_CF.90CF056	4953
J_SE.SE92809	4920
CRF01_AE_TH.CM240	5174
CRF02_AG_NG.IBNG	5131
CRF03_AB_RU.KAL153	5651
CRF04_cpx_CY.94CY03	4972
N_CM.YBF30	5200
O_MVP5180	5636
O_CM.ANT70	5660
HIV2_A_ST	AGAGCTTATTTTCCAGGTGTGGCAAAGGTCCTGGAGATACTGGCATGATGAACAAGGAATGTCAATAAGTTACACAAAGTATAGATATTTGTGCCTAATGCAGAAAGCTATGTTACATACATTCTAAGAGA	5597
	Vif end in HIV-2 and SMM sequences <-	
HIV2_A_BEN	AGAGCTCATTTCAGGTGTGGCAAAGGTCCTGGAGATACTGGCATGATGAACAAGGGATGTCAGCAAGCTACACAAAGTATAGATATTTGTGCCTAATGCAAAAAGCTATATTTACACATTTCAAGAGA	6152
VIF	R A H F P G V A K V L E I L A S	VIF
VPX	R E L I F Q V W Q R S W R Y W H D E Q G M S A S Y T K Y R Y L C L M Q K A I F T H F K R	VPX
HIV2_A_ALI	GGAGCTCATTTCAGGTGTGGCAAAGGTCCTGGAGATACTGGCATGATGAACAAGGAATGTCAAAAGTTACACAAAATATAGATACTTGTGCTTAATGCAGAAAAGCTATGTACACACATTTTCATGAAA	6145
HIV2_B_D205	AGAACTTATTTCCAGGTCTGGCAGAGGTCCTGGGCATACTGGCGTGAGGAACAGGGCATGTCAATTAGCTATACCAAATATAGATACTTGTGCTAATGCAGAAAAGCAATGTTTGTGCACTATACAAAAG	6128
HIV2_B_EHO	GGAACTTATTTCCAGGTCTGGCAGAAGTCTTGGGCATATGGCGTGAGGAACAGGGCATGTCAATTAGTTATACCAAATATAGGTATTTGTGCTAATGCAGAAAAGCAATGTTTATACACTTTGCAAAG	6126
STMM83293	GGAACTGATTTCCAGGTCTGGCAGAAGTCTTGGGAATACTGGCATGATGAGCAAGGGATGCCCGGAGCTATGTAAAATATAGATACTTGTGCTTGAATACAGAAAAGCTTTGTTTATGCATAGCAAGAGG	5713
SMM251	GGAGCTGATTTTCAGGTTTGGCAAAGGTCCTGGGAATACTGGCATGATGAACAAGGGATGTCAAAAGCTATGTAAAATACAGATACTTGTGTTAATGCAGAAAAGGCTTTATTTATGCATTGCAAGAAA	6042
SMM9	GGAGCTAATTTTCAGGTCTGGCAGAAGGTCCTGGGAATACTGGCATGATGAATGGGGATGTCAGTCAGCTACACTAAATRCAGATACTTGTGCTTGAATACAGAAAAGCTCTGTTTATGCATTGCAAGAAA	5539
SIVCPZANT	5294
SIVCPZGAB	5670
SIVCPZUS	5666
SIVSTAN1	5792
SIVGRI677	5566
SIVVER9063	5799
SIVVER155	5794
SIVVERTYO	5289
SIVVERAGM3	5297
SIVLHOEST	5869
SIVSUN	5922
SIVMNDGB1	5240
SIVSAB1C	5918
SIVSYK173	5705
VIF	-----	VIF
VPR	-----	VPR

B_FR.HXB2R	5606
VIF	-----	VIF
VPR	-----	VPR
A_UG.U455	5052
C_ET.ETH2220	4998
D_ZR.84ZR085	5128
FI_BE.VI850	4946
G_SE.SE6165	5003
H_CF.90CF056	4953
J_SE.SE92809	4920
CRF01_AE_TH.CM240	5174
CRF02_AG_NG.IBNG	5131
CRF03_AB_RU.KAL153	5651
CRF04_cpx_CY.94CY03	4972
N_CM.YBF30	5200
O_MVP5180	5636
O_CM.ANT70	5660
HIV2_A_ST	GGGTGCACTTGCCTGGGGGAGGACATGGGCCGGGAGGATGGAGATCAGGACCTCCCCCTCCTCCCCCTCCAGGTCTAGT...CTAATGACTGAAGCACCAACAGAGTCTCCCCGGAGGATAGGACCC	5723
	Vpx end in HIV-2 <-	
HIV2_A_BEN	GGGTGCACTTGTGGGGGAGGAC.ATGGGCCGGGAAGGATTGGAAGACCAAGGACCTCCCCCTCCTCCCCCTCCAGGTCTAGT.CTAATGACTGAAGCACCAACAGAGTTTCCCCCAGAAGATGGGACCC	6281
VPX	_____G C T C W G E # M G R E G L E D O G P P P P P P P G L # \$_____	VPX
HIV2_A_ALI	GGGTGCACTTGCCTGGGGGAGGACATGGGCCAGGAGGATGGAGATCAGGACCTCCCCCTCCTCCCCCTCCAGGTCTAGT...CTAATGACTGAAGCACCAACAGAGTTTCCCCGGCGGAATGGGAC	6271
HIV2_B_D205	GGCTGTAGGTGCCTGCAGGAGGGCCATGGGCCAGGGGGATGGAGATCAGGACCTCCTCCTCCTCCTCCCCAGGCTGGC...CTAATGGCAGAAGCAGCCCCAGAGATCCCTCCAGAGAACGAGAACC	6254
HIV2_B_EHO	GGTTGTGGATGCTGCAGGGAAGGCCATGGGCCAGGGGGATGGAGATCAGGACCTCCTCCTCCTCCTCCCCAGGCTTAGC...CTAATGGCAGAAGCAGTCCAGAGATTCCTCCAGAGGATAAAAACC	6252
STMM83293	GGATGCAGATGCTTGGGGGAAGGACATGGGGCAGGAGGATGGAGACCAGGCCCTCCTCCTCCTCCCCCTCCAGGCTTAGCATA...GAATGACACACAGACCTCCAGAGGATGAAGGCC	5829
SMM251	GGCTGTAGATGCTTAGGGGAAGGACACGGGGCAGGAGGATGGAGACCAGGACCTCCTCCTCCTCCCCCTCCAGGACTAGCATA...AATGGAGAAAGACCTCCAGAAGATGAAGGCC	6157
SMM9	GGGTGTAGGTGCTTAGGAGGAGACATGGGGCAGGGGGATGGAGACCAGGCCCTCCTCCTCCTCCCCCTCCAGGACTAGCATA...AAATGACAGAAAGACCTCCAGAAGATGAAGCCC	5655
SIVCPZANT	5294
SIVCPZGAB	5670
SIVCPZUS	5666
SIVSTAN1GATC	5796
SIVGRI677GATC	5570
SIVVER9063GATC	5803
SIVVER155GATC	5798
SIVVERTYOGATC	5293
SIVVERAGM3GATT	5301
SIVLHOESTGAGGCAGCAAAGAAGTCCAGAGAGACCACCAGAGGACTTTGGAC	5913
SIVSUNAGAGCCAGTAGAACAGCCACCAGAAGACGAGGGGC	5957
SIVMNDGB1AGATGAGCAAGTATCAGAAGATCAAGGAC	5269
SIVSAB1CAGGAGGTGGCTCCCACCAGTAGGAGGGGATCCTCCCAAGGATCCCCCAAGAATCCCAGA...GAAGAGATACCAGGATGGCTAGAACATGGGATT	6013
SIVSYK173GGAGGCTTTTCTCAATCCC.TCCCAGCATGTGCAAGGAACACCATGGTTCCTTCATTCACAAGAAATG	5770
VIF	_____G G F L Q S # P A C A R N T M V L H S K K C_____	VIF
VPR	_____# E A F F N P . S Q H V Q G T P W F F I P R N_____	VPR

```

Vif end iN HIV-1 <-|
B_FR.HXB2R .....GAATGGACTAGAGCTTTAGAGGAGCTTAAGAATGAAGCTGTAGACATTTTCTAGGATTGGCTCCATGGCTTAGGGCAACAT.....AT 5695
VIF .....#_N_G_H_$_..... VIF
VPR .....E_W_T_L_E_L_L_E_E_L_K_N_E_A_V_R_H_F_P_R_I_W_L_H_G_L_G_Q_H.....I VPR

A_UG.U455 .....A-G-T---T-G---A-----C-----A-----CAG-----A---A----- 5141
C_ET.ETH2220 .....-G-----C---C-G-----C---C-----ACCA-----AA-----A---T----- 5087
D_ZR.84ZR085 .....-T-----G-----C-----C-----A---T---A-----A---T----- 5217
FL_BE.VI850 .....-G-----AA-----A-----A-----G-----G-----CCA---T---C-----C----- 5035
G_SE.SE6165 .....-G-T---A-G---A-----A-----A-----C-----T---A---G----- 5092
H_CF.90CF056 .....-G-----A-----A-----A-----G-----G-----AG-A-----CAA---A---G----- 5042
J_SE.SE92809 .....-G-----G-----A-----A-----C-----CCA---T---A---T---A---GT----- 5009
CRF01_AE_TH.CM240 .....-T---A---G-----A-----A-----C-----CC---A-----A---GT---C----- 5263
CRF02_AG_NG.IBNG .....-G-----A-G---A-----AC-----C-----CCG-----A-----A----- 5220
CRF03_AB_RU.KAL153 .....-G-----G-----G-----G-----C-----G-G---T---A-----A---T---C----- 5740
CRF04_cpx_CY.94CY03 .....-G-----G-----A-----A-----G-----G-----ACCC-----C-----A---G----- 5061
N_CM.YBF30 .....-G-----AT-A---A---AT-A---A-----GC-C-----A-----A-----G-----C----- 5289
O_MVP5180 .....-G-----CC-G---A-----G---AG-A---A-----A-----C-----CC---T---A---A---C---GT-----GT---C----- 5725
O_CM.ANT70 .....-G-----C-----A-----A---AGCA---A---A-----C-----CC-----A---C---G---A---T---C----- 5749
HIV2_A_ST CACCGAGGGAGCCAGGGGAT--G--GT-A---AACCC-GAGA--A-A---AT-A---T-A-AG-C---GACCCCTCGC-T---AAT-ACTC-T-CA-CT----- 5832
HIV2_A_BEN CACCGAGGGACTTAGGGAGT--C--GT-A---AAC-C-GAG--AA-A---G-A---CT-A---G---GA-CCCCG-T---AAT-CTC-T-CT-CT----- 6390
HIV2_A_ALI CCCACCAGGGAGCCAGGGAT--G--GT-A---AG-CC-GAGA--AA-A---AG-A---T-A---G---GACCCCTCGCAT--AAT-CTC-T---CGGCT----- 6380
HIV2_B_D205 CACAAAAGAGAACCCTGGGAA--G--T-GGG---A-CC-G---AA-A---C-A---CT-A-AG---GA-CCTCGC-T---AAC-CGC-T-TA-CTT----- 6363
HIV2_B_EHO CACAAAAGAGAACCCTGGGAAC-G--GT-G-G---CG-CC-G---AA-A---AC-A---CT-A-AG---GACCCCGC-TAT-AAC-CAC-T-CA-TTT----- 6361
STMM83293 CACAGAGAGAACCATGGGAT--G--GT-G-G---AG-C-G---AA-A---AC-A---CC-GC-G---GA-CCTCGC-T---AAGC-CGC-T-TA-CT----- 5938
SMM251 CACAAAAGGAACCATGGGAT--GT-G-G---G-C-G---A-AT-G-AG-A---T-A-A---GA-CCTCGC-T---AAC-CAC-T-TA-T----- 6266
SMM9 CACAGAGGAACCCTGGGAT--GTRG---AG-C-G---A-KA-AG-A---CC-G-AC-K---GA-CCTCGC-T---AATY-CGC-T-TA-CT----- 5764
SIVCPZANT .....-G---TTGT---AAC-----A-AA-A---A---G---A---A---C---ACA-CC-ACAT-A----- 5365
SIVCPZGAB .....-G-T---ACC-G---A-AT-A---A---A---A---C---ACCA---T-G---CAA-----A---TTC----- 5759
SIVCPZUS .....-G---G-T---TG-C---A-A-A---AG-A---G---T-GC-C---ACCC-----A---A-----A---GT-C----- 5755
SIVSTAN1 TCAGCAGGGAGCCGTGGGAC---ATT-AG---CA-GG---CA-A---C---TC-G---ACAAC-C---GGG-AGAAGCTC---TTCCAAG--T--A-CTTC.....TGTCAGGAAGAAGG 5917
SIVGRI677 TAGACAGGGAGCCATGGGAC---CT-A---CA-GC-CAG--T-A---CG-A---CAGA-G---C---GGA-T---ACAT---AATCC-AG--T--A-TT-C.....TGTTAGAGGAGGG 5691
SIVVER9063 TAAACAGAGAACCATGGGAC---CT-AGG---CA-G---A-T---C---TC-G---CAAAT---GGC-AGAGCTC---TTCCAAG-TT--A-TT-----TGTCAGGAGGAAGG 5924
SIVVER155 TCAGCAGGGAGCCGTGGGAC---CTCAG---TA-GG---A-AA-C---C---CAAAGT---GGC-AGAAGCTC---AT-CCAAG--T--A-TT-----TGTCAGGAGGAAGG 5919
SIVVERTYO TAAGCAGGGAGCCGTGGGACA---CT-AG---TA-G---C-A---T---C---CC-G---GAGGCTG---GGG-AGAAGCTC---GTTCCAAG--T--A-CT-C.....TGTCAGGAGGAAGG 5414
SIVVERAGM3 TGAGCAGGGAGCCGTGGGAC---CT-AG---CA-GC---A-TA-C---TC-G---CAAG-TG---GGG-CGAGCTC---GTTCCAAG--T--A-CT-----TGTCAGGAGGAGGG 5422
SIVLHOEST CACCAAGGGAGCCATGGGGTC-G---TT-GC---CACCA-G---AA-A---AG-G---G---AAGAC-C---C---AGAAG-AAT-T-GGCCCAAG-CTC---T-----TG 6022
SIVSUN CACCGAGAGAGCCATTTGCAC---CTGGC---TACCA-G---A-AA-C---G-G-C---G-AA---G---C---AGCAGCAATAG-TGCCCAAG-CTT---T-C-----TG 6066
SIVMNDGB1 CTCCCAGAGAGCCATACAATC-G---CT-GC---TAC-A-G---AA-A---G-A---AAGA-AG-C---C---CTC---ATC-AA---CAG-TCAG--T-----TG 5378
SIVSAB1C TGCCCAGAGAGCCGTTCCGAC---CTC-G---CA-G---C---C---C---TTCA-G---CCAGT-C---C---C---A---A---CTC---TTTC---TC---TT---TGGA.....ATTGTGGAAGAACAGC 6137
SIVSYK173 TAGAGTTGACCCCAAACGTGATCAATGTCACT-TAAAGGC---ACT-G-AGT---CA---ATCC-AG---A---CCTCAAGAAA-AT---AG-TT--A-T---ATCCTTAAACGAAGAGCCAGGG-C 5900
VIF .....R_V_D_P_K_R_D_Q_C_H_C_K_G_R_T_G_S_D_R_S_I_Q_A_F_Y_S_S_R_N_I_W_S_L_E_S_I_L_K_R_R_G_R_D_ VIF
VPR .....V_E_L_T_P_N_V_I_N_V_T_V_K_A_E_L_V_V_T_E_A_S_K_H_F_T_P_Q_E_I_Y_G_V_W_N_Q_S_L_N_E_E_A_G_T VPR

```


	frameshift	insert in HXB2	HXB2 Vpr	premature end (without frameshift)	
B_FR.HXB2R	CTATGAAACTTATGGGATACTTGGGCAGGAGTGGGAAGCCATAATAAGAATTCTGCAACAACCTGCTGTTTATCCATT	.TTCAGAAATGGGTGTCGACATAGCAGAATAGCGGTTACTCGACAGAGGAGA			5823
VPR	Y E T Y G D T W A G V E A I I R I L Q Q L L F I H # F R I G C R H S R I G V T R Q R R	# # Q N W V S T \$ Q N R R Y S T E E			VPR
VPR					VPR
A_UG.U455	---A-C-C-----AG---T---T-----T-----T-A-----T---...-----C-A-----A--T-C-GG---A---				5268
C_ET.ETH2220	---A-C-----C---T-G---A---TT-----C-----AA-----A-C-AG-----A--TTA-----A---				5214
D_ZR.84ZR085	---A-C-----C-----C-----T-----T-----A-----T-----A-----C-----A-----C-----A---				5344
FL_BE.VI850	---A-C-C-----C-----AG---T---T-----G-AT---A-----A-----C---T-----A--GT-C---A---				5162
G_SE.SE6165	---A-C-----A-----T-----T-----CG---A-----A-----A-----C-A---A-----A--A-C-G---A--G				5219
H_CF.90CF056	---A-C-C-----A-----T-----TT-----CG---A-----A-----T-----C-A-----AA-----A---				5169
J_SE.SE92809	---AGC-----A-----A-----C-----A-----AT---T-----C---C-AT-----A--TA-C---A---				5136
CRF01_AE_TH.CM240	---A-C-A-----A-----G-T---T-----GT---A-----G-T---A-----A-----A--TG-C-GG---A---				5390
CRF02_AG_NG.IBNG	---A-----A-----A-G-TA---T-----T-----AT---G-T---A-----A-----A--T-A-GG---A---				5347
CRF03_AB_RU.KAL153	---A-----A-----A-----T-----T-----T-----A-----T-----G-----A-----A--T-A-G---				5867
CRF04_cpx_CY.94CY03	---A-C-----A-----C---A-G-T---T-----T-----A-----T-----C-A---T-----A--CT-A---				5188
N_CM.YBF30	---A-C-A-----A-C-C---AG-G-A-G-A-T-C-G-A-A---T-A-----T-----AT-G---C-C-AG-C---GA-C---CT-A---				5416
O_MVP5180	T---G-----A-C-----A-----TATG---A-T-----CT-A-----A-----C-----AT---A-C-A-----T-----AA--CTC-C-TCT-AC-C-				5852
O_CM.ANT70	T---G-----C-----T-----TATG---A-T-----CT-A-----T-----AT---A-C-A-----T-----AA--AC-C-AGAG-A---				5876
HIV2_A_ST	---CT-GAC---A-C---CCTT-A---C-CCAG--GGC-C--T-G-C-A---G-GCC--CC-CT-G-C---GCA-A--CG-C-GCTCA--G--T---T---...				5941
HIV2_A_BEN	-C--A-T-GAC---A-C---CCTT-A---C-CCAG--AGC-C--T-A-CC--A---G-GCC--C-CG-G-C---GCG-A--AAC-GCTCA---T---...				6499
HIV2_A_ALI	---ACT-GAC---A-C---CCTT-A---CCAGG--AGC-C--T-ATGCC--A---G-GCC--C-C-G-C---GCA-A--G-C-G-TCA--G-T---...				6489
HIV2_B_D205	---CAGT-GGC---A---CCTT---CA-G--AGC-C--T-A---C--G-GCC--C-CC-G-C---GCC-T---A--CTCA--G--T---A---...				6472
HIV2_B_EHO	---CA-T-GGC---AA---CCTT-A---CA-G--AGC-C--T-A-C-C--G-GGGCA--C-CC--C---GGG-T-C-A--CTCA--G--T---A---...				6470
STMM83293	---T-GAC---A-C---CCTT-A---CA-G--AGC-C--T-AG-C-C--G-GCC--C-C-G-C---GGC-T-C-C-CTC--G--T---...				6047
SMM251	---A-T-GAC---A-C---CCTT-AG---CA-G--AAC-C--T---C--G-GC--C-C-G---T---GGC-A--CAAC--CTC---C---...				6375
SMM9	---T-GAC---A-C---CCTT-A---CA-G--AGC-C--T---C--G-GG-GCA--C-C---GG--A-CA--CTC---C---...				5873
SIVCPZANT-ACT---AA--TGGGTATAT-C-AATT-C-GAGATTC-T-G-GAGGAGTG--GG-TT--A-A--T-TCCTCAA-A-GC-CT--T-AC---TAT---CAC--ATG-G---ATAGC---T-				5488
SIVCPZGAB	T---C-----A---A---T-----A-G---C-----CT-A---CT-----T---C-----C-AG-----A-CCTC-C---A-A---				5886
SIVCPZUS	---CAGT---A-C---AG---A---A---T-----G--CT-A-----TT-----AT-G---C---AAT-----G---GA--T-CTGGCA---				5882
SIVSTAN1	GG-AAGG-A-GGG-CACCA-TGATA-A-A-G-CTT-TAGATATTAT--GT-GG-A--GA-GCCTT-A--G-G---TC--TG---C--CA-GAGA--A-CTCCTTTT-AGC-CTACG-AGAA--G				6044
SIVGRI677	AAGGAG-CA-A--ACCCATGGAAT-AGAT--GCT-CAAGTACTAT----G-T--A-GTCTA----G-A-----TG---T---A--AGG--AG--CCTTTTTC--TACG-AGA--G				5818
SIVVER9063	AG-GAG-CGGGA-CCCA-TGAC--AGA-G-CCT-TAGATACTAT--GT-GG-T--GA--GCA--C--G-T---C-TTG---T---TAGA--ACAGCCCTT--AAC--TACG--GAA--G				6051
SIVVER155	GG-GAG-CAGGAA-ACCC-TAGC--A-A-G-CAT-TAAGTATTATC-CT-AG-T--GA--GCT--C--G-G---C-GTG---A---CAGG--ACA-CCCTTT-AGC-ATACG--GA--G				6046
SIVVERTYO	AG-AAG-CA-GG-ACTCCC-TGAT--A-A-G-CCT-CAATATTAT--GC-AG-A---A-GGCT--C--G-G---TC--TG---CA-GAGA--GCAGCCCTT--AAC-ATACG--GA--G				5541
SIVVERAGM3	AG-AAGG-A-CGCACTCCC-TGCTA-A-A-G-CTT-TAAATATTAT-A-T-GG-----A-GCT--C--G-G---C-GTG---A-C--CAGA--ACA-CCCTTT--AAC-ATACG-AGAA--G				5549
SIVLHOEST	TGTA--C-G-ACA---TCAGAAGA--AG-CCTGCATGAAAT-C--C-CCC-AG-CA-TAG-GCTT---C-T-C---CTCCCAACT---CCAGCTATGC-GC--C---AGA-CAG				6135
SIVSUN	TGCG--CT--ACA---TCAGAAAA--A---GTGTATGAGAGC---C-CTT-G--CA-TAG-GCCT---C-T---CTGCCAAT---CTTC-TA--TC-G-AT--AGAAGTGG-TCA-GA-				6188
SIVMNDGB1	TGTGC---ACACA---AG-GAGGAA--AG-CCTGT--GAAAT-T--T-CCT-AA--A-TAG-GCCA-T-GGG---C---C-AGCTCAA---GATGGA-C-TTC-GG-AA---G-AGACCAC				5497
SIVSAB1C	-AT---TCA-GGGCAAACC-GGCT--A---CTG-T-TAAATATTGT-----A-GCCTT-A--G-G---ATGAAAG-T-GAT-T--AAG-CA-AG-CTCACCTT.....AGAAGTGG-TCA-GA-				6249
SIVSYK173	TG--TCCC-GACAAT--C-TGGGAAAG-ACCA--TTG-A---GG---GCA--CA-T-TTA---C---GAA---TGCTGCA--A---CT--AC-G-CC-G-TAT.....				6009
VPR	D S P T M A W E R T M L D M V R A L N L M L F E H . F A A G C P Q R T R Y				VPR
TAT					TAT

HIV-1/HIV-2/SIV
Nucleotides

HIV-1 Vpr normal end <-		
-> Tat start		
B_FR.HXB2R	GCAAGAA.....ATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAACTGCTTGTACCAATTGCTATGTAAAAAGTGTTCATTGCCAAGTTT	5939
TAT	_A_R_#_..._M_E_P_V_D_P_R_L_E_P_W_K_H_P_G_S_Q_P_K_T_A_C_T_N_C_Y_C_K_K_C_C_F_H_C_Q_V_	TAT
A_UG.U455	-GC-G-.....AC-----A-C-G-----C-----G--C-T-----GT-----GG-----C----	5384
C_ET.ETH2220	-----AC--A-----C-----G-----ATC-A-T-----A--A-A-----T-T----	5330
D_ZR.84ZR085	-----T-----ACA--T-----C-----GG-----A-----A-----A-----	5460
F1_BE.VI850	-T-----CT--T-----T-----C-----C-----A--T-----CGA-----TGG----	5278
G_SE.SE6165	-T--GG.....C-G-----AC-----T--G-G-----C-C-----A--A-T-T-----GT-----GG-----	5335
H_CF.90CF056	-T-----C-----A-----C-----C-----A-----T-----C-----A-----A-G-----	5285
J_SE.SE92809	-GG-G-.....ACAG-----T-----T-----T-----A-----A-----T-----G-----	5252
CRF01_AE_TH.CM240	-GC-G-.....A-TG-----AC-----T-----G-----C-----G--T-C-----A-----GG-----C-A-----	5506
CRF02_AG_NG.IBNG	-GC-G-.....TG-----C-----C-C-G-----C-----G--G-T-C-----T--C-----G-----T-G-----	5462
CRF03_AB_RU.KAL153	-----C-----C-----C-----C-----C-----TC-----	5983
CRF04_cpx_CY.94CY03	AG-G-C-GGGG.....GAC-----C-----G-----C-----A-----A--G-T-TC-----GG-----	5308
N_CM.YBF30	AGG.....A-----T-----C-A-----A-----C-AT-----GA-----A--C--TT-TA----	5529
O_MVP5180	AG-G--GAGGAAGAAGAA...-T-----GAGA-GCCC--T--C-T-C-T-G-CA--CC--CC--AT-----C--GA-C--A-----T-T----	5980
O_CM.ANT70	-G----GAA.....T-----GAGG-GCCC--T--C-C-C-T-----CC-G-TCC--AT-----C--GA-C--A-----T-T----	5995
HIV2_A_STCAGC--AG-GGAC--AAT-CTTT-T-G-T-TACCAA-CC-TAG-GG-ATGCGA--A--A--T--C-----C-----AC-----GA-G-----	6039
HIV2_A_BENTA-A-AAG--GA--A-T-CTTGCC--G-TGCACC-A-CC-TAG-GC-ATGCAC--A--A-----GCGA--CA-T--C-----GC-G-----	6597
HIV2_A_ALICA-A-AAG-GGA--AAT-CTCTCT--G-T-TACC-A-CC-TAG--ATGCAA--A-CA-----GG-A-C--AC-----GC-G-----	6587
HIV2_B_D205CA--AGG-GGAG--AAT-CTCTCT--A-T-TACC-C-CC-T--GG-ATGCGA--AT-CA--C-----G-A-C--AC-----GC-----	6570
HIV2_B_EHOCA-C-GGG-GGAG--AAT-CTCTCT--G-T-TACC-C-CT-T--GG-ATGCA--AT-CA--C-----A-----C-AC-----GC-----	6568
STMM83293CA-C-AGGAGGAG--AAT-CTTTG--A-T-TACC-C-CA-TAG-GGAGTG--A--G--T--C-----A-----A--C-----GC-G-----	6145
SMM251CA-C-TGG-GGAG--AAT-CTCTCT--A-T-TACC-C-CT-T-G-GG-GTGC-A--A-CA-----AC-----GT-----	6473
SMM9CAG--TGGAGGAR--AAT-CTCTCT--A-T-TACC-C-CKGTAG-GG-GTG--A--A--G--Y--C-----G-A-----A-----T-GCA-----	5971
SIVCPZANT	-G-C.....C-C-----CG--GA-AC-CCT--A--TTA--C--TCCTGC-ACA--AGC--GC-A-C-AT--C-T-C-----TGC-----T-AC-----CTC-C-----	5601
SIVCPZGAB	AG-TCC.....T--A-----GAC--G-----A-C-----C-----A-G--A-T-----AT-----C--GCT--T-A--C--TAT-TA----	6002
SIVCPZUS	A-TCCTCAAGGGAGAAGGA...-T-----ACA-----G-----C-----G-----C-AT-----C-----TGTTG-----	6010
SIVSTAN1	AG--.....ATGGAGTC---GGG-AC-GGATGGC-GAGAGC-T-CTCC--G-CTTGCAT--G-CTTTA-C-C-ATG-AC-A--A--T--C-----GA-C--A--C--T-----	6163
SIVGRI677	AG--.....CAAG-AG--GGAGCCC-ACCCC-T-CTCC--G-CTTGCAT--G-CATTAC--C-ATG-AC-AT--A-----C-----A-C-T-A--C--TG-GC-----	5931
SIVVER9063	AG-G.-TGGACAAGGGGGAGGAA---G-ACC-TG-TCCAC-AG--CTTGATT-GA--ATATAAG-AG-C--TA-C---TG-A-A-AT--G---T--C-----A--C--T--GC---	6179
SIVVER155	AG--.-TGGACAAGGGGGAGGAA--C-AG-AC-TGT-CCAC-AG--CTTGATT--A--ATACA-G-AA-CC-T-G-G--ATG-ACA-AT--A--T--C-----A--C--T-A-----T-C-----	6174
SIVVERTYO	AG-G.-TGGACAAGGGGGAGGCA---A-A-C-TGT-CCAC-AG--CTTGA-TG-AG-CTATCAG-AG-CT-TGC-G---TG-AAA-AT--A--T-T--C-----A-----T-AC--C--T--GC-----	5669
SIVVERAGM3	AGGG.-TGGACAAGGGGGAGGAC--AG-GC-CGTACCAC-AG--CTTGATTG-A--A-TCAA-GCA-CC-TG--GCGGTG-ACA-A--G-----TGT--C--T-A--C--T--GC-----	5677
SIVLHOEST	-A-CAGCCCATCCAACAGAAGCA-C-C-A--ACCTAATG-AGGAGACATA-AAACG--GCAGTACT-AAGCC-TGC--G-GTG-CAG-AT-CC--T-GC-----A-----TATTC-----	6265
SIVSUN	ATGTC--CCCAAGGCCACCAACAA--C-A--ACC-GGGC-AGGGGACATTAGAAG--GCCTACAAG-C-A-CTTGG-GG-CTG-GA--A--A--GG--GG-GA--C-----C--T--C-----	6318
SIVMNDGB1	AACT-CCCCCTCAGGATTCAGGCC-AGAG-AGATAGA-TATAAGC-A-TGCTAG-AG-GTATTATCAG-CTTTGC--G--TG-GAG-AT--A--GG--C--G--A--C-----TATGC-----	5627
SIVSAB1C	..GC-T-TGGACCAGGAGCAGGA--CC-GCCCC--GGTCT-GGAG-----TTCA-G--G-G-TGCATC-G-C--TCC-GG-CTG-GA--AT-C-----TC-C--GTT-C-----TATTC-C-----	6377
SIVSYK173GCCAG-CAC--AGG--CCC--ATCCT-CTTAGAAG-A-CATTC-T--AG-AAGGA--ACC--AC-A---AT--A---T--C-----C-----A-----GC---	6123
Vpr end in SIVSYKES <-		
VPRA_R_H_R_G_Y_P_H_P_S_\$_	VPR
TAT#_Q_T_Q_R_V_P_P_S_F_L_E_G_T_F_L_E_K_G_P_P_T_P_C_N_K_C_F_C_K_N_C_C_Y_H_C_Q_L_	TAT

	-> Rev start	Tat and Rev exon end \ / intron	
B_FR.HXB2R	GTTTCATAACAAAAGCCTTAGGCATCTCCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGAGCTCATCAG.....AACAGTCAGACTCATCAAGCTTCTCTATCAAAGCAGTAA.....		6049
TAT	C_F_I_T_K_A_L_G_I_S_Y_G_R_K_K_R_R_Q_R_R_A_H_Q.....N_S_Q_T_H_Q_A_S_L_S_K_Q		TAT
REV	M_A_G_R_S_G_D_S_D_E_E_L_I_#.....#_T_V_R_L_I_K_L_L_Y_Q_S_S		REV
A_UG.U455	-C--TC-G-A-----G-----T-----A-----A--CC-----G--C--C--A.....GG---A--GA-----A--CT-A--C-----		5494
C_ET.ETH2220	-C--TC-G-.-.-.-G-----T-----C-----C-----A.....G---A--GA-----AA--CT-A-----..GTACCAAATAAT		5451
D_ZR.84ZR085	-C-----G-----G-----T-----C-----C-----C-----G-----G-----A--C--A--C-----		5570
FI_BE.VI850	-C--T-C--G--G-G-----T-----A-----A--CC-----G--C--C--A.....G---A--GA-----A--CT-A--C-----		5388
G_SE.SE6165	-C--TC-G-AC--G-----G-----T-----A-----A--CC-----G--C--C--A.....G---A--GA-----A--CT-A--C-----		5445
H_CF.90CF056	-C--TT--AG--G-----A--T-----C-----C-----A.....G---A--GA-----A--CT-A--C-----		5395
J_SE.SE92809	-C--TC--CAG--G-G-----T-----C-----C-----A.....G---A--GA-----A--CT-A--C-----		5362
CRF01_AE_TH.CM240	-C--TC-G-A-----G-----T-----A-----A--CC-----G--C--C--A.....G---A--GA-----A--CT-A--C-----		5616
CRF02_AG_NG.IBNG	-C--TC-G-AC--G-G-----T-----C-----C-----A.....G---A--GA-----A--CT-A--C-----		5572
CRF03_AB_RU.KAL153	-C--TC-G-A-----G-----T-----A-----A--CC-----G--C--C--A.....G---A--GA-----A--CT-A--C-----		6093
CRF04_cpx_CY.94CY03	-C--TC-G-A-----G-----T-----A-----A--CC-----G--C--C--A.....G---A--GA-----A--CT-A--C-----		5418
N_CM.YBF30	-C--C--AG--G-----A-----T-----C-----C-----A.....G---A--GA-----A--CT-A--C-----		5639
O_MVP5180	-C--AG--G-GT--G-A-----C-----A-----A--CC-----G--C--C--A.....G---A--GA-----A--CT-A--C-----		6087
O_CM.ANT70	-G--G--G-GT--G-A-----C-----A-----A--CC-----G--C--C--A.....G---A--GA-----A--CT-A--C-----		6099
HIV2_A_ST	---TT--AC--G-GGC-C--G--A--GG---AAC-A--GGCA-A--AGAA--ACTCCGAAGA-A.....C-A--G---TCGT---GC--G-CA--G-----		6143
HIV2_A_BEN	---TCG-A--G-GGC-C--A--A--A--AGC-A--GGCA-AC--AGAA--GACTCC-AGGA-A.....C-A--G---TCGT---GC--G-CA--G-----		6701
HIV2_A_ALI	---TT--AT--G-GAC-C--G--A--GG---A-C-A--GGCA-C--AGAA--TCTCC-AAGA-A.....T-A--G---TCGT---GC--G-CA--G-----		6691
HIV2_B_D205	---TC-T-A--G-GTC-T--G--A--GT---A-C-CTC--GAA--TCTGC-AA--G--A-G.....AC-GC-C---GC-C--G-CA--G-----		6668
HIV2_B_EHO	-C--TC-T-A--G-G-C-G--G--A--GT---AAC-CTC--GAAA---TCTTC-AA--G--A-G.....AC-AC-T---GC-C--CG--G-----		6666
STMM83293	-TG-T-----GTC-T-A--AA---AGC-ATCA-GAA-A--GTTAAGAA--AG--A-G.....T--C-AT---GC---CA-----		6243
SMM251	---TC-T-A--G-GA--G--G--A--GT---AGCA-TCACGAAA---AGAA--ACTCCGAAAA-G.....GC-A--G--A--AC-T---GC---CA-----		6577
SMM9	---TC-Y-A--G-G--G--A--A--G---AGCA-C-ACGAA-A--ACT-CGAAG-AGA-A-G.....G--A--AC-TT---GC---G-CA-----		6069
SIVCPZANT	-C--C--A-----G--G--T-----GAA-AGCT-GAA-GAAC---CGAACAACTGCTGAA-G---G-A-A-A-----A-C-G-----T-----		5717
SIVCPZGAB	-C--C--A-----G--G--T-----A-CCACAA---C-CA-GGC--CT...GCAGG---A-A-A-A-----A--A-C-----		6115
SIVCPZUS	---TCG-AG--G-G-----A--T-----G-G-GA--CA-C-A-AAC-GCT...AAG-G---GC-GA-A-----A-C---CGC-----		6125
SIVSTAN1	-C--TC-GCAG--G-G-----A-----CATGTCTCT-G-ATC--AGA-CTAAG-AGAA---TCA-A-----GAA-CT-G---C-----		6264
SIVGRI677	-C--C--GCAG--G-G-----AG-GCGT---CATGTCTCT-G-AAA--AGAAA-AC---CA---A.....GA-A-----ACC-A--CGGC-----		6029
SIVVER9063	---TC-GCA-----G-C---TG-TA---CATGCCCTT-G-ACC---AGAAAAGAA--T-T--G.....C--A---GCCTA---A-CAGG-C-----		6277
SIVVER155	-C--T--CGG---GAC---T--TA---CATGCCCTT-G-ACC---AGAAAAGAAAG-TT---TCG.....G--G---GCAT-C--G--C-GC-----		6272
SIVVERTYO	-C--TC-GCA-----G-C---TG-TA---CATGCCCTT-G-ACC---AGAAAAGAAAG-TTCG-TCG.....CT-A--TTG---C--T--CAGC-C-----		5767
SIVVERAGM3	-C--TT--CA--G-GA---TG-TA---CATGCCCTT-G-ATC---AGAAAAGAAAG-TT---CG.....CT-G---GCTT-C--GA-CA---A--G-----		5775
SIVLHOEST	-C--TCAGCAG--GG-----GC-GCA---TATGTCC-C-G-AAAC-GAGA---GTT.....ACCTCG-TA-TTGAGAC---GCAGGA--CTGTGG-CAG---		6366
SIVSUN	-C--TC-GCAG--GG-----G--CA---TATGTCT-C-G-G-AC-----AT-A-----TCA-TATCT-AGG--A-GCAAGAGGCTGT-CGAAG---		6416
SIVMNDGB1	-C--TCA--AG--G-GT---A--AAGG---CCATGTCT-C-G-AAAC-TGTA-C-G--CTAA.....AAG--G-TA-CTGG-AG-GGT-AAGAAGCTAT-CGAAG---		5734
SIVSAB1C	-C--CAT-A--G--C-----TCG---TATGTCCCT-G--CA--AGAGTCTCT.....AAG--G-T-TC---A-TCAAGT-TCTCT-C-CA-C--G-----		6475
SIVSYK173	---C-TCA---G-GA--G-A---A-T---ACCA-G-AA-C--GCAGCT--C.....GT-T-TCTGAGG--G-T...AGCCG-C-T-CAGG---		6218
	Rev start except in macaques and manglebeys		
TAT	C_F_L_Q_K_G_L_G_I_T_Y_A_R_P_R_K_R_A_A_R.....S_I_S_E_D_D..S_A_P_T_G		TAT
REV	M_P_D_Q_G_S_E_Q_L_#.....#_V_F_L_R_M_#_#_A_H_L_Q_G		RAV

	(ACG start codon in HXB2, ATG in others)	
B_FR.HXB2RGTAGTACATGTAACGCAACCTATACCAATAGTAGCAATAGTAGCATTAGTAGTAGCAATAATAATAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATAT...AGGAAAATATTAAGA	6166
VPU	..._V_V_H_V_T_Q_P_I_P_I_V_A_I_V_A_L_V_V_A_I_I_I_A_I_V_V_W_S_I_V_I_I_E_Y.._R_K_I_L_R_	VPU
A_UG.U455GTAAT---CTT-----T-AC---T-GGA---CTGG-----AC--GGC-GA-----C---CT-----A-----A-T-----GGT-----A-AA-AA---T-GC---AG	5617
C_ET.ETH2220	AGATGTAATG--T-ATTTAC--G-AA--GTAGATTAT-G-A--T-----CA-----C---C-----A-----C-TAT-----...-T-G-----AG	5578
D_ZR.84ZR085T-----T---T---A---T-----T-----T-----C-----C-----A-----T-----C-T-G---AA---G	5687
F1_BE.VI850	GTAGTGTTAAA----T-----T-TC-TA-T-GTT-GC-A--GG---C-----A-----C-----C-----A-T-----TAT-----...-A---C-GG---G	5515
G_SE.SE6165	GTAGAAAAAAT----T-----T---GT-AT--GT---T-----G-----G---A-----T-C---GC---C-----A-----T-T-----...-G---AG---A	5572
H_CF.90CF056TTATC-TAAATGTA--TT-GG-T--G-----G---GC-----A---T-T-C---CG--A-----A-----TAT-----...-AA---T-GG---G	5509
J_SE.SE92809	GTAATCAAGTA--C-T-----T-AT---T-G-A---C-----TA-----TT-CC-T---G-A---A-T---GTAT-----...-A---C-G-----	5488
CRF01_AE_TH.CM240GTAATA--T-----T-AC---T-GGA---TAGT-----G-C-GA-----GC---CT---T-GGA---A---AA-T---GCT---GT---...-A---C---G	5737
CRF02_AG_NG.IBNG	GTAGTAATAAT--C-T---G-T-----T-A---ACT-----G-C-----T-C---GC---C-----AG-----TAT-----...-AGG-A---	5699
CRF03_AB_RU.KAL153T-----T---T-CT-G---C-----T-GG-----G-----T-----...-----AGG-A---	6210
CRF04_cpx_CY.94CY03GTATTA---T-----T-TT-TTCTGGGA--CTGG-----G-C-G---GC---T---T---A---A-TT---T-T-----...-A---T-GAG---G	5539
N_CM.YBF30GTAAA-C-TGT-TA--G-TGTC-T-G-G-T-CA---G---G-C---G-T-GC-----G-A-A---C---G-AT-C-TAT-G-----...-A-----AA-TTG	5750
O_MVP5180	...GTAAAGC-GA-G--CA-GA-A-C-TGC--G-CT--A---TT--AGT--T--GTGTC-TAT--ATG--C-TAT-TGGT-GT-TAACCTT-G-A-TTATT--TGC-AAGA-AAC--GATAG--G	6213
O_CM.ANT70	...GTAAAGC--A-G--CAT-G-G-C-TGC--G--A---ATT--TAGT--T--GC-GT-TAT--ATG---TTT-TGG-GGT-TATCTT-G-AA-TATT---C-AAAGGAAC--GACAG--AG	6225
HIV2_A_ST	6143
HIV2_A_BEN	6701
HIV2_A_ALI	6691
HIV2_B_D205	6668
HIV2_B_EHO	6666
STMM83293	6243
SMM251	6577
SMM9	6069
SIVCPZANT---TCT---T-ACTAA---TTTGTAGTAT--TT-TC-T-C--TAGTA---T-C--TGG--TATCTGTA-ACCTATACT-TATAAGC-TTATA--ATATAT-A-C-GCAGCA--T-	5834
SIVCPZGABGTATATGACTCT-TT-GTCGGCTT-G-TC-CAT-T-----GC---A-T-C-TGG-AC--TTGTAT-TGG-GATACATTAT-A-TGGGGATATAG-AGGTAT-AA-G-CATAGGCTT	6233
SIVCPZUS	GCCCTTTTGT-CTA-G-T-AATTG-TTGAA--GGTT--A-----T---G-AT-GA--GTA-TTT-G--G-G--ATCTGG-GAC-AGTTG-T-G-T-GTGG-G-C--ATAAAA-TA---GA-AAT-CT	6255
SIVSTAN1	6264
SIVGRI677	6029
SIVVER9063	6277
SIVVER155	6272
SIVVERTYO	5767
SIVVERAGM3	5775
SIVLHOEST	6366
SIVSUN	6416
SIVMNDGB1	5734
SIVSAB1C	6475
SIVSYK173	6218

```

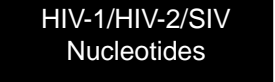
B_FR.HXB2R      CAAAGAAAAATAGACAGGTTAATTGATAGACTAATAGAAAAGCAGAGACAGTGGCA. |-> Env gp160 start in HIV-1
VPU             .ATGAGAGTGAAGGAGAAATATCAGCACTTGTGGAGATGGGGGTGGAGATGGGGCACCATGCTCTTGGGA 6294
ENV             _Q_R_K_I_D_R_L_I_D_R_L_I_E_R_A_E_D_S_G_#. _E_S_E_A_E_I_S_A_L_V_E_M_G_V_E_M_G_H_H_A_P_W_D_VPU
               _M_R_V_K_E_K_Y_Q_H_L_W_R_W_G_W_R_W_G_T_M_L_L_G_ENV

A_UG.U455      ---A-----T-AA-C---A--G-----T-G--T-C-GAG-A-T...-ATCCTT-CCT-----A-T--A--T-G---T 5742
C_ET.ETH2220   ---G---T-----A-A---ACT-GG-----..---A---T-G--T-C-GAG-A-T...-TCA-CAAT-----T-----A-T-T-AGG-T-T--- 5703
D_ZR.84ZR085   -----T-----A-G-----A-----..---G--T-A-GAG-A---..-ATCA-CCCTT-----A-----T----- 5812
FI_BE.VI850    -----A-T-AA--TA-A-A---A-G-----..---G--TGC-GAG-A-T...-CAGCACTT-G---A-----CTTT-AT---G--- 5640
G_SE.SE6165    --G-A--G---GA-A--C---A--G-----A-----..---C--G--T-C-GAG-A-T...-TA-CACTT---A-----G--TT-A-----T 5697
H_CF.90CF056   --AG-----C-----A--AG-----..---C-----T-----C-C-GAG-A-T...-ATCC-A-CTTA-----T-AA--T-G---T 5634
J_SE.SE92809   ---G-----T-A-----A-G-----T-----..---C-----T-----C-C-GA-A-T...-CAGACCTT---G-----CT---A-T---A 5613
CRF01_AE_TH.CM240
CRF02_AG_NG.IBNG
CRF03_AB_RU.KAL153
CRF04_cpx_CY.94CY03
N_CM.YBF30     --G-AG-----C-----A-G-----T-----..---T-G--T-C-GA-A-T...-ATCC-CTCTTA-----T--G-ATA-AT--T--- 5824
O_MVP5180     G-GCAGG---CTTGAAG-T-AAGG--A--AG---TCAGG-T-----A-T...---A--A-T-A---GA-CA-CAGGAA--CAT-GA-CTTA-AC-T-.....GC-A-G-CTT 6326
O_CM.ANT70     G-G--G---CTTGA-AGGT-AAGA--A--G---TTAGG--T-T---A-T...---A--CA-T---GAGGA-CA-GAA--TAT-GACCTT--ACTT-.....GC-A-G-CTT 6341
HIV2_A_ST     .....GTAA-AT-T-TG-TA--AATCA-CTATTTGTGCCAG-T-GCTAG 6188
HIV2_A_BEN     .....GTA--GA-CCTG-TA--AATCAGCT-TTTGTTG-CA-TT-ACTA- 6746
HIV2_A_ALI     .....GTA--AT-TCTA-TA-AAATCAGCT-CCTGTT-CTA--T-ACTAG 6736
HIV2_B_D205    .....GTA--GCATAT-TTA-CA-CC-CCT-CCT-TTGC-----GCTT- 6713
HIV2_B_EHO     .....GTA--GCACATGTTAATAATTACCTACTTGT-CA-----GCTT- 6711
STMM83293     .....GTA--GCCT-CCCT--AAATCA-CT-CCT-T-GCTA--T-GCTAT 6288
SMM251        .....GTA--G-AT-TCTT---AATCAGCT-CCT-T-GCCA--T-GCTTT 6622
SMM9          .....GTA--G-AT-TCTT---AATCAGCT-CCT-T-GC-----TCGCTAT 6114
SIVCPZANT     G-T...-T-AGAGA-ATCA--GAATC-T-GA-G-GTT--TAG-AG-CTT--ATAG...-AGTGCAAT--A---GATGAG-A-GCC-ATACATATTATT--G-TCT-.....GCTT-GCTA 5950
SIVCPZGAB     G-G-C-G---T--G--AC-T-ACCTA-T-T-G-G-----T-----A...---A--A-T---GA-GA--AGAGAC---ACA-CTTA-CC-T-ATTACA-T---AACAA-CATTT 6361
SIVCPZUS      ---CA-G-G--TC-A-ATC-TT-A--A---A-C-G-AT---A-----T-----A...---A--T---GA-GA--AGAC-C---.....CTA-GTTATTG-C TTC 6362
SIVSTAN1     .....GTA--G-ACC--TAA--G-A-A-G---TATTAT-AG-AA--TT-G 6309
SIVGRI677     .....GTA--G-----T-CTTATA-A-ATACTA-TA--AGCAA-A---- 6074
SIVVER9063    .....GTA--AAACCTCTTTTAG-T-T-GTTCTT-TAGCTA--A-A--A- 6322
SIVVER155     .....GTA--ACA-AG-TCTTAG-A-TT-TTATAGTAT-AGGAA-A--- 6317
SIVVERTYO     .....GTA--A--TATACAATAATA-CC-TA--A-TA--AG-GA-A--A- 5812
SIVVERAGM3    .....GTA--AA-CTGACATTACT--T-G--ATACTAT-AA-AGGGATAG 5820
SIVLHOEST     .....G-A--TGCCAT-TCCAG--TA--A-TTC----TT-GCTTC 6408
SIVSUN        .....GTAA--TGC--T-CCCTGA-ATATTG-TAGGTT--TC-TT-C 6458
SIVMNDGB1     .....-T-TGAA-T-TCCAG-TCTTTTCTT--A--ATG-TTAG 5773
SIVSAB1C      .....-AA-T-TGAA-CTTCTT-C-GTACTGCTTTG---AAG---T 6517
SIVSYK173     .....G-AAAA-TGGCA-CT-TT--ACTTA--TAG-TTGT--TTT- 6261
ENV           _M_A_A_F_R_T_Y_I_V_C_L_F_ ENV

```

	Vpu end <-	
B_FR.HXB2R	TGTTGATGATCTGT...AGTGCTACAGAAAAATTGTGGGTCACAGTCTATTATGGGGTACCTGTGTGGAAGGAAGCA...ACCACCCTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGT	6418
VPU	V D D L # #	VPU
ENV	M L M I C . S A T E K L W V T V Y Y G V P V W K E A . T T T L F C A S D A K A Y D T E V	ENV
A_UG.U455	--A-A-A-T---...A--A..C-C-----G---C-----A-T---.GTT---C-----G---A--	5863
C_ET.ETH2220	---A---T---...A-GA-TG-GG-C-----A---A-T-T...-G-C-T-----A---A--	5827
D_ZR.84ZR085	-----GACCTAT---G---T-C-----T-----A---A---...-T-G-----T---A-A---C	5939
F1_BE.VI850	-A--A-A---...A---G---C-C-----A---A---...-T-G-----C-----A-G---C	5764
G_SE.SE6165	--G---A-T---...---CT-A-T-T-----T-----G-A-T---.GAT---C-----C-----AG-G-T---AAG	5821
H_CF.90CF056	---C---A-T---...---G-C---C-----T---A-----A---G---...AA---T---C-----G-----AAA	5758
J_SE.SE92809	---A---T---...-AG-A-A---G-C-----T-----A---T---...-AG---T---C---T-----AG---T---AAA	5737
CRF01_AE_TH.CM240	--G---A-T---...---CT---C-C-----T---T-----G-----GA-T---.GAT---C-----C-G---C---A--	5986
CRF02_AG_NG.IBNG	--AA---A-T---...A---...---C-----G---C-----ACG---.GAG---C-----A---A--	5945
CRF03_AB_RU.KAL153	---C-----...---T-A-----T-----A-----AG-AG---	6450
CRF04_cpx_CY.94CY03	--G---A---...---CT-A-C-C-----T-----G-A-C---.GAG---C-----A-----G-A---A--	5788
N_CM.YBF30	--AAGCT---AAGGTA-T-GGT-T---C-CAT---A---G-C-----A-A---GA---.GAG-A---T-C---T-----CC-AG---C	6005
O_MVP5180	--C-A-CCCA---TTG---TA-GTA-C---AT-C---T---C-----A-G-A-G---.G-AC-AGTA---C---T-----CCT-ACAAGC-T-ACA	6456
O_CM.ANT70	--A-A-CCCA---.TTGAGCCTTAG-C-GC-A-AT-CA---GC---G---A-G-A-T---.AC-AGTA---C---T-----CCT-ACAAGC-T-AAA	6465
HIV2_A_ST	CTAGTGCTTG--TAATA.....TATTGCGTCCAA-AT-G-T-T-T-C---C-G-C---GAA-T---.T--TTC-C-C-----CA--TA-AAA	6288
HIV2_A_BEN	CAAGTGCTTG--TAGTA.....TATTGT-GCCA-AT-G-T-T-T-C---CA---C-C---AA-T---.T-T-TTC-CT-----CA--TA-AAA	6846
HIV2_A_ALI	CTAGTGCTTG--TAGTA.....TATTGT---CAA-AT-G-T-T-T---C-G-A-CA---AA-T---.T---TTC-C-C-----CA--CA-AAA	6836
HIV2_B_D205	-AGGT--C-GTG-GTTT.....GT-TGT---CAA-AT-T-T-T-C---CA---C-CA---G-A-C---.AGTTC-C-CA-----CA--CACAAA	6813
HIV2_B_EHO	-AAGT--CTATG-GTAT.....TG-GC---GAAC-TT---T---TC---TA---C-CA---AA-T---.T-A-TTC-C-C-----C---CAGAAA	6811
STMM83293	-AAGTGCTTG--TGACC.....TATTGC-CTCA-AT---G-A-T---T---A-CA---G-A-T-G...-G-TTC-C-C-C-----CA--CA-AA	6388
SMM251	-AAGTG-CTATG-GATC.....TATTGT-CTCAA-AT---T---T---A-CT---G-A-T-G...-A-TTC-C-C-C-----CA--CA-AA	6722
SMM9	YAAGTGCTTCAG-GATT.....TATTGTGTTCAA-AT-A---A-TC---T---A-CA---G-A-T-G...-AGTTC-C-C-C-----CA--CAGAAA	6214
SIVCPZANT	-CCA-T-T-AGAG...-AG-GG-GA-TG-GAC-AT-A---A-TC---A-C---C---GAA-T-G...-AC-T-----CA-A---CTCCATGACAAG-	6074
SIVCPZGAB	--C-A-CCCA---.TTG---CTCTG-G-A---A---A---A---T---C-T-T-T...GA-C-GGTA-C-----C---C---G---C-AG-----C	6482
SIVCPZUS	-C-C--GCT-GATAATCCA-GATTGCT-GC-A---CT---A---A---A---A---GA-T-T...GAG-A-CT---C---T---A-G---CA-GCAG---C	6489
SIVSTAN1	GA--A-GCT-AATAGGACTGTTATAT-GG-C-CA-ATA-T---G-T---TA-C-A-A---AA-CAGC...T-AGTGCAGCT-C---.ATG--CCTAA	6415
SIVGRI677	-AAGT--AGGAATA.....-GT-CC--AT-G---G-T---AA-C-A-A---AA-TT---.AGTTCAGGC---C---.ATG--GCCCA	6168
SIVVER9063	-AAGTTCCT-A-TAAGT.....-G-CAA---A-A---G-TC---A---A---A-CAGC...T-AGTACAAGCC-C---.ATG--TCCTAC	6419
SIVVER155	-AGGA--AGGGATAAGT.....-A-C-GCA---A-A---G-TC---A---A---A-CAGC...T-AGT-CAAGCT-C---.ATG--CCTAC	6417
SIVVERTYO	-AGG--AG-G-TAAGT.....-GCAA---A-A---T---AA---A-A---AA-CAGC...T-GGTGCAGGCT-C-C...ATG--TCCAC	5906
SIVVERAGM3	GAG-AG--C-TAA-ACA.....AGGC---CAA---A-T---A---A---A-CAGC...T-AGTACAGGCT-C-C...ATG--CCCAC	5917
SIVLHOEST	-TGGC--A-A-GGGA.....-CAA-AT-A---A-T---G-AAAT---G-T-TAATGTTT-AGTGC-CT-GA---.C-T---CCAA	6505
SIVSUN	---AGGT-GTATAGCA.....-T-CA-AT-A---A-TC---ACT---AA---G-ACC---.GTAGTGC-AT-A-A---.C-T---CTAA	6552
SIVMNDGB1	C-C-AC-AG-AG--ATT-TA.....-G---CAA-AT-G---TC---A---A---A---T...-AA-ACA-T-GA---.C---TAA	5872
SIVSAB1C	GT-G--GCT-AGTGTGGTTA.....GTCCA-AT-T---G-T---TA-C-A-A---AA-TAGT...T-AGT-CAAGCC-----A-G-CCCCAA	6614
SIVSYK173	GC--A--A-GT-TAGGATT-ATGGA-A-C-C-CAA-AT-A---A-T---A---ACAT---G-A-T---.TAGT-AC---C-----ACC-CTCCA	6367
ENV	S L I S L G F M E K Q Q Y V T V F Y G I P H W E D A . Y A P L F C T T S H	ENV

B_FR.HXB2R	ACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTA.....AATGTGACAGAAAATTTAACATGTGGAAAAATGACATGGTAGAACAGATGCAT	6539
ENV	H N V W A T H A C V P T D P N P O E V V L V N V T E N F N M W K N D M V E Q M H	ENV
A_UG.U455	G-----C---T-----A--AT---C.....-G-A-----A-----T-----	5984
C_ET.ETH2220	-----C---GT---TT-----C-----G-----T-G-GT---A-----G-G-----	5948
D_ZR.84ZR085	-----A-C---T-----A--A-C---AA-----C-----CA-----G-C-----	6060
FL_BE.VI850	-----C-----T-----T-TC---AA-----A-----G-T-----A-----A-----	5885
G_SE.SE6165	C-----C---T-----A--A-TA---A-----A-----T-----A-----G---A-----	5942
H_CF.90CF056	G-----C---T-----A-----GA-G---CA---AG-----A-----G-GC---T-----G---A-----G-G-----	5879
J_SE.SE92809	-----C---T-----T-----T-A-G-----GA-GAATC---CC-----A-----C-----C-----G	5858
CRF01_AE_TH.CM240	G--C-----C-----T-----A-----A-----CACC---A-----A-----A-----G-----G	6107
CRF02_AG_NG.IBNG	-----C-----T-----T-----T-----A-----CAT---A-----A-----G-----A-----G-----	6066
CRF03_AB_RU.KAL153	-----C-----T-----T-----G-----G-----A-----CC---AA-----A-----G-----A-----	6571
CRF04_cpx_CY.94CY03	-----A-C---T-----T-----CTC---A-----A-----G---C-----A-----	5909
N_CM.YBF30	T--C--CA-C-----A--A---T--T--T-----T-----GC---ACCC.....-A--T---A-----T-----G---A-A---C---C-A---A	6126
O_MVP5180	G-----A-----AT---A---C--T--T-----T-----T---TCC-C-A-GC.....-T---C---G-T---A-----T-----G---C---A---	6577
O_CM.ANT70	G-----A-----AT---A---T--T-----CT---T--T---TATCC---CAC.....-T---C---T---A-----T-----A-----G	6586
HIV2_A_ST	TAGAG--AC---GA---CATACAG--CT-G--AGACA-TGATG--TAT--G--A---CT--A-----GGCC--CG--GCA.....A-T-ATAC--T-ACAGAA--A	6400
HIV2_A_BEN	TAGAG-CAC---GG---CATACAG--CT-G--AGACA-TGATG--TTAT--G--A---A-T--A-----GGC---G-TGCA---A-T-ATAC--TGACAGAA--A	6958
HIV2_A_ALI	TAGAG--AC---GA---CATACAG--CT--AGACA-TGATG--TTAT--G--A---CT--A-----GGC---CG-TGCA---T-ATAC--T-ACAGAA--A	6948
HIV2_B_D205	CAGAG-CACC---GA---TGTACAG--C--C--AGACA-TGGTG--TACT--GA-CAGGC-A-----CA-A---GGC---G-TGCA---T-ATAC--TGACACAA--G	6925
HIV2_B_EHO	CAGAG--ACC---GA---TGTACAA--CC-C--AGATA-TGATG--TATACC--A-CCA--A-----A-A---GGC---G-TGCA---T-ATAC--TGACAGAT--G	6923
STMM83293	TAGGG--AC---GA---AC-CAA--C---AGACA-TGGTG--TTACTC--T-G-C-A-A-----C---GGC---G-TGCC.....ATAC--TCACAGAA--A	6500
SMM251	TAGGG--AC---GA---AC-CAG--CC---AGATA-TGGTG--TTATTC--T-G-CCC-T-----T---GC---G-TGCT---G-ATAC--TCACAGAA--G	6834
SMM9	TAGGG--AC---GA---CACAAA--CT-G--AGATA-TGGTG--TTACTC--T-G-C-A-T-----C---GGC---G-TGCT---T-ATAC--TCACAGAA--A	6326
SIVCPZANT	G--C---A---A---TACCAGT---G--A-T---T--AG-T--TATT-T---TAGGC-C-----ACCT--TCTGG---TGCT-AT---T-T---G---AGT---ACA	6192
SIVCPZGAB	T-----A-----G--A---T-----T-----GT--T--G-----T-TC-TCC-----A-T---TCA---TGCT-AT---A-T---G---C---A---	6603
SIVCPZUS	C--C--CA-----G--A---C--T--T-----T-----CAT---CC-----T---G---G-----G---A-T---C-----A---G	6610
SIVSTAN1	TACC--CC-----A---AC-AAT--A---AGAT--T-ATG--TTATACT--CAG--A-----CT-T--G--A---G-AGCA---GG-CAGG-ATCC--TT-GT-GCA--G	6533
SIVGRI677	TACC--A-G---A---CACAA--CA---AGAT--T-ATG--TAATAC--G--GCCTC-A-----CA-T---GC---CG-GGCT---T-ATCCGCTGGTA-AA--A	6280
SIVVER9063	TACC-GGC-A---T---AC-AAT--CA---AGAT--T-ATG--TTATACT--G--GCC--A-----CA-T--T--GCC--G-AGCC--GC-G--AGG-ATCC--TT-GT-GCA--G	6537
SIVVER155	TAC--GTT-G---A---TAC-AAT--CA---AGAT--T-ATG--TATAC---CC-C---A-A---GCCA--G-AGCA--GC-G-CAGA-ATCCCTT-GTAGCA--A	6535
SIVVERTYO	CACA-GCT-A---T---TAC-AAT--CA---AGAT--T-ATG--TATAC---CCTC-A-----A--C--T---CCA--G-GGCA---GGTG--AGA-ACCC--TT-ATAGCA--A	6024
SIVVERAGM3	CACC-GAC-A---A---TAC-AA--CGA---AGAT--T-ATG--TACAC--G--CC--A-----CA-C--T---CCA--G-AGCA--GCTG--CAGA-ACCCCTT-GTAGCA--A	6035
SIVLHOEST	TAC--GCC-C---T---GACAT--T--CT--G--AGATTTG--AATCTTATGC--G--CCTA-C-----TAT--A--TT---C---CA--ACAGT---A---ATCAG--TTATACA--A	6626
SIVSUN	TA--GCT-G---TG--TAC-T-A--CT--G--AGATTTA-AA-CTTATGC-----GCCCA-AAC.....GGGT--AGA--G-----CAGA-G-A-T--GCA-T-ATCA-AT-GT-CA--A	6673
SIVMNDGB1	TTCA-G-C-C---TA---CAC-AAT--CA---TT--TTG--AG--TTATG-T--G---A--A-TCCT.....G--A--A-AG-----CAGGACTT-T--GG--A-ATCAGAT--GTTTAT--A	5993
SIVSAB1C	TACG--CC-A---T---TAC-AAT--CA--C---GAT--TGAACCAGA--GG-AC-A---C-GAA---CCCATACCT--A-C-----A---G-TGCT---CGG-ACCCCTT-GTAGGC--A	6744
SIVSYK173	TA-AGGA-GA---T---TA-GAAT---G--GT---CAGATC-AAT-G---T-AGG--CAAC.....T-ACAGG---T-C---CCAGCA---T-GCAGTCACA-GAT-AG-CA--A	6485
ENV	K G G W A T K N C V P S A D Q I E V R V N I G G E Y F P A W N S S H M I R Q Q	ENV

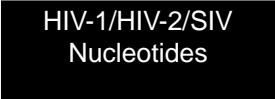


B_FR.HXB2R	GAGGAT.....ATAATCAGTTTATGGGATCAAAGCCCTAAAGCCATGTGTAAAAATTAACCCCACTCTGTGTTAGTTTAAAGTGCCT.....GATTGAAGAATGATACTAATACCAATA	6648
ENV	<u>E D I I S L W D Q S L K P C V K L T P L C V S L K C T D L K N D T N T N</u>	ENV
A_UG.U455	--A-----C-----C-----G-----T-----C---CG--G-T---CA-AACATCACCATTAACA-CACC--T--CA-C-C---T--CAG	6108
C_ET.ETH2220	C-----G-----G-----G-----G-----C-----C---T-A-GCTATCAAAAACAATAC-AA-GTT-CAA---A-G--T-----	6072
D_ZR.84ZR085	-----T-----T-----A-----G-----C-----T-----GATGCATCGAGAACAACAGCACTG-T--A-C-GC-CACTGCCA-	6184
FL_BE.VI850	ACA-----G-----G-----G-----G-----C---T---AATGCCACCAATAACAG-CA-G-A--GCCAGGGGCA.....	6002
G_SE.SE6165	-----G-----GC-----T-----CC-----C---T---GATGTAAACCAACAAA-GCAAC--A-GAA--A--GC--TG---	6066
H_CF.90CF056	ACA-C.....T-G-A-----C-----C---T---AATGTCAGAAAACAATACC-CT--C-GCAC--GC-G--TGG-GG	6003
J_SE.SE92809	--A-----G-----G-----T-----GA---T---A---C---T---T-G-AACATCACTAGTAACAG-AAT-CA-C-AG--AC-G--GTGTA-	5982
CRF01_AE_TH.CM240	-----G-----T-----C-----G-----T---C-----T---CAATGCTAAATTTGACCA--GGC-GT-GCA-A--C---GT-TC--	6231
CRF02_AG_NG.IBNG	--A-----T---C-----C-----T-----G-----T---C-----C---G-T-TCA-AAATTTCAACAACAGCT-CAGC--C-G-AGC-ACCTA--T-G-G	6190
CRF03_AB_RU.KAL153	-----G-----G-----G-----G-----C-----T-----G-----GGT--C--GT-CC-	6680
CRF04_cpx_CY.94CY03	-----A-G-G-----A-----C---GC---T-T-----C---T-CA--T-T-AATGCAACTACTACTA--AGT-CC---GC--GTG-TT--A.	6032
N_CM.YBF30	-----T-----C-G---A-G--T-----C---T-----T-A-----A-C-A-GCTT-T-ACGATAGCTATGGGGAG--AAG---C---ACA-A--TG--A-CA-	6250
O_MVP5180	--A-C.....C-T---G---A-G--TT-----T---AG--A-G-TTCT-A-----ACAAA-G-C---TGTAGATCTGCAAAACAATA-AACAGGCCCTATTA-A-G-G--A-TA.	6700
O_CM.ANT70	--A-C.....T-T---C-G--T-T-A-T-----TC-A-G-TTTC-G-----ACAAA-GG--T---AAACATAGCTGGAACAACAATG-A.....	6691
HIV2_A.ST	-CA-TAGAAGATG-CTGG--C--TT--GAC-TCAA--A-----C-C---A-CT-A-----AGCAA-GCGT-T-ACAGCACAACTGCAAAAA-CACA-CCTCCACAC-A-CA-----CC-	6530
HIV2_A.BEN	-CA-TAGAAGATG-CTGGCA-C--TT--GAC-TCAA--A-----C-GC---A-T-A-----GGCAA-G-T-T-GCAGGGTTCAAGGGAATACCAC--CCCCGA--C-C-GG--TCG-	7088
HIV2_A.ALI	-CA-TAGAAGATG-CTGG--AC--TT--GAC-TCAA--A-----C-G---A-TT-A--A-AGCAA-G---T-GCAACATAAGCACAGAGAGCACA-CC-CATCCC-G-GCC-AGGG-	7078
HIV2_B.D205	-CA-TAGATGATG-GTGG--AC-C-TT--AACCTC-A--A-----C---C-----G---GGCAA-G-C-T-G-AAAACCGAAAACAACCCAGG--TGCCAG--G--C---CC-	7055
HIV2_B.EHO	-CAACAAGGATG-GTGG--C-C-TT--GACCTCAA--A-----T-----G-----G-CAA-G---T-A-AAGACGTGGAGCTCA-C-AGC--AG-GACC---CGT--TCCG	7053
STMM83293	-CAATAGAGGATG-CTGG--ACC-T-TT--AAC-TC-A-T--A--T-----G-T-----A-A-GCAA-G-G---T-A-AAGAATGAAACAGACA-A-G-GGATTGACAGGG--A--AGTA-	6630
SMM251	-CAATAGAGGACG--TGGCAAC-C-TT--GACCTCAA-----T-----T-----T-A-CA--C-A-G-GA---A-AAAAGTGAGACAGATAGA-G-GGATTGACA-AAATCAT-A-CA-	6964
SMM9	-CAATAGAGGATG--TGG-ACC-C-TT--AAC-TC-A-T-----C-----C---A-A-A-CAA-G-G---T-A-AAAAGTGAGMCAAGMCAAG-A-GGGTTTAAACAGGAGSAC-AGCAC	6456
SIVCPZANT	--A-----GN-ACAA--TCC-A-----AT-----C---A-TA-G--A-A-AAA-G-T-T-T-AGGATACAATGGAAACACC-ACA-CACCAAG--A-CA--A-G--	6316
SIVCPZGAB	--A-C.....T-----C-C-----G-T-C-----T-----G-G---TT-A--A-CC-GC-----G-AAGGCTAACTTTAGCC-GGCA--A--CCTA--A--CCAG-CAT	6727
SIVCPZUS	-A-----T-T-CC-----C-G--T--A-T--A--G-----T-G-----C-A-G-CT--TCT-AAACCCGATAGTAATAG-AGTGTGTAA-----C-----	6727
SIVSTAN1	-C--GAGTAAC--CATTTGC-C-TT--AAGC-CT-G-----C---G-A-TA-G-CA-A-AGA-G--T-T---AAATTAACAAGTACC-CCCCT-CATCAAG--CCC---ATCA-	6663
SIVGRI677	-CA--GAGTAAT--CATCTAC-C-TT--A-----CGA-G-G--T-----T-GC-CT---CA-A-A--AAA-GTCC--TGTAGAGCTGAATGGTACA-CCAC--CA--G-CC--C--TGCA-	6410
SIVVER9063	-CA-GAAGTAAT--TCATTTG--G-TT--A--CAT-G-A-----C-----T-T-T-A--A-A-AAA-G-GC--TGTAGAGTTGAACTCAAGTAGAAGA--CCCAAC-C-A-C-T-AGC-T	6667
SIVVER155	-CA-GAAGTAAC--TCA-CTGC-G-TT--A--G-CAT-----C-----GC--T-A-T-A--A--C-AAA-G--T-TGTAGAGTTAAAAGGCTTCC-CAACCTCT-CCCCAG-A-CCT-T-C--	6665
SIVVERTYO	-CA-CCAGTAAC--CCATTTAC-C-TT--G-----CTA-G-A--T-----G-G--T-A-----A-CA-C-AGA-G--T-TGTAGAGTTAAAATCCACAAGAGAA-G-GCGACA--ACC--A-CG-	6154
SIVVERAGM3	-CA-GAAGTAAT--CA-CTGC--TT--G--G-CT--G-----T-A--TT-G-CA--AAA-GTCC--TGTAAGATTTGAACTCCTCT--GCCT-CC-CCAC-C---A-GT-CC-	6165
SIVLHOEST	-CCTGGTCTGCT--G-ATGC-A-GGTA--CAGT-TTA-G--A-----A-C-AT--TAT--G--AA-GC--TGGGGAAGTTACAAAAACACCAACA-CA-CACCA-AA-C---T-CAC	6756
SIVSUN	-CATGGCAAGCT--G-C-TC-A-GGTA--GCC-TTA-G-----T-----GA-C-AT-CTAT-----C-AAA-G---TC-ACCAGAAACCCGACAACCTCCC-CCTCC--CTCA--AGTC-AA-G--	6803
SIVMNDGB1	-CATGGCATGCT--GGGA--A-G-TA--ACC-TA-T-----GA-T-A--TAT-----AGA-GC-A--TCAGGAAACAGAAAATGATATCAGCA-CA-CA-C--AGCC--TA-C--	6123
SIVSAB1C	-CA--GAGTAAC--CA-CT--G-TT--AAGT-CTT-----T-GC-CT-A--A-G-CA-A-AAA-G--C---TA-CGCTTAGAAGGAGGA-CAGCA-CA-CAACAT-ACCAT-A-CAT	6874
SIVSYK173	ATATTAGAGGAC--GTCAGC--TCTTG--GC-AAT-GA--C-----G-GC--G-A-CA-G--A--GA-GTT--T--ATTGGACAACAGCCCT-C-ACA-GT-CCCC--CA-GTCCCC	6615
ENV	<u>I L E D M S A L F L Q A N R P C V K L A P M C I R M L C T L D N S P A T S T P T T S P</u>	ENV

B_FR.HXB2R	GTAGTAGCGGGAGAATGATAATGGAGAAA.....GG	6679
ENV	S_S_S_G_R_M_I_M_E_K.....G	ENV
A_UG.U455	A-G-AGTGA--GA-GAA.....	6125
C_ET.ETH2220	--GCC-AT.....	6082
D_ZR.84ZR085	CGGTG-AGCCA.....	6197
FL_BE.VI850	6002
G_SE.SE6165	A--C-CT-AA-CC-AC-AT-GCACTGTGGATAACCCA.....	6106
H_CF.90CF056	CAG-AG-G-AA.....	6014
J_SE.SE92809	----CCT-AT-TC.....	5996
CRF01_AE_TH.CM240	AC-TA-TA--A-AT--A-C.....	6253
CRF02_AG_NG.IBNG	AC-TG-AT--GA.....	6204
CRF03_AB_RU.KAL153	A--C---TA-C-T--AA-G----A.....	6706
CRF04_cpx_CY.94CY03	6034
N_CM.YBF30	-AGAACCA-AC-T-GGATAC-AA.....	6273
O_MVP5180	6702
O_CM.ANT70	6693
HIV2_A_ST	CA-CAGCAAAC-C--CA---GGA----TTCTTCATGCATACGCACAGACAACCTGCACAGGGTTGGGAGAAGAAGAGATGGTC.....	6613
HIV2_A_BEN	--TCC-CAACTCG-GACC-CCACATCCGCGAGCCTCCATAATAAATGAAACTTCTAACTGCATAGAA.....	7185
HIV2_A_ALI	-C-CACT-AAACCCC-----AT---GGGATCCATGCATAAAGGCAGACAACCTCCCCAGGGGACTAGGGGATGAAGAGATG.....	7161
HIV2_B_D205	C--AGCCTACT-CC-CCTCTCGT-G-CTGAAAACGATTAACGAAAACAGACCCATGCATAAAAAAT.....	7146
HIV2_B_EHO	CGTCCCT-A-ATCTTCT-CTCA-ACCCT-CTCAACGAAGATAGCAAATGTATTCAAAATGAC.....	7138
STMM83293	CA-CAGTGACACC--CAGC-GCA-CAGC-GCAACAAGCCAGAGTTAGTAAATGAAACAAGCTCCTGTGTAAGT.....	6733
SMM251	CA-TA-CAACAGC-GCACC--CATCAGC-CCAGTATCAGAAAAAATAGACATGGTCAATGAGACTAGTTCTTGTATAGCTCAG.....	7073
SMM9	CA-CA-CAACACA--CATC--CAACACC-CCATCACCAATAATAGCAAAGGTTGTAAATGACAGTGATCCTTGTATAAGAAGT.....	6565
SIVCPZANT	CAGTA-CACCA.....	6356
SIVCPZGAB	C-TC-CCGCT.....	6740
SIVCPZUS	6729
SIVSTAN1	-C--C-CTAC-GATCCCTGTCCAA-T-CCGACGAAAGTAGCTGTAAACGCCACCTTAGTTACAAATAGC.....	6742
SIVGRI677	C--CA-CAAT--CT-CCCCCTGTC---TTGCAGTACAGAGCAGATAGAAGGAGAAATGGCAGAG.....	6477
SIVVER9063	CA-CA-CAAAAGCTCCA-A-CA-GTG-TCCCTGCATTAAATCTACCAATAATAATGTT.....	6776
SIVVER155	CGGCAG-AACC-A-C-ACCCTGT-TT-G-AATAAAACAGACTCC.....	6759
SIVVERTYO	CGCGC-AATCT-CCGGCC--CCCTGTGT-GGCCCGACGTCAGGTGAA.....	6251
SIVVERAGM3	CGGCCTCAACA-CC-AT--C-CA-CCTC-ACAACCCTTTGCCGTGTCCAGAACAAGACAAGTACTGTGTGTAATCATGTAATGAAACAATC.....	6283
SIVLHOEST	AG-TGCCCTT-TTTT--C-ACGAAC--GT-ACAGTTAAGAACCCTGGGGAATGAAACAAGATTAGAGGAAGATCTTAATTGCACAAGAGGGCTTAATGAGACCACAGAGAGGAAT.....	6869
SIVSUN	-CT--GATTATT-G-C--CC-CCAC-GCTAAGACTACAACCTCAGACTACAAGTAGTACAGCTCCACTGTAGTACAACACTCTCCAATGCCTCTAGATTGGAATTGTACTGACACAGAGAACATAGCG-A	6933
SIVMNDGB1	CACC--CTACT-C-TCT-C-G-T-CA-GTACTACAGAGATTTACTTTAGATGTAGATAAAAAATAATACAGAAGAAAAGGTAGAGAGGAAT.....	6212
SIVSAB1C	CA-CAGCAC-CCCCGA-G--G--AGTGTGGGCTTTAATGACTCAGTGTAGTAACAAGAGATGGAAGAAA.....	6944
SIVSYK173	C--C--CACCTCC--ATGA--C-TG-TGGGGAGACAACAGTACAGAGCCAAGA.....	6668
ENV	P_T_T_P_P_N_E_T_W_W_G_D_N_S_T_E_P_R.....	ENV

B_FR.HXB2R	AGAGATAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAGAAAAGAA...TATGCATTTTTTTATAAACTTGATATAATACCAATAGATAATGATACT.....	6788
ENV	E I K N C S F N I S T S I R G K V O K E . Y A F F Y K L D I I P I D N D T	ENV
A_UG.U455	...-G-----T-----G-C---GAAC-----A---AAA-----T...-T-C-----G-----G---A---CA---AAC-GA-AAC.....	6233
C_ET.ETH2220	T--A-G-----T--T-----A-C---GAAC-----A---AAAAG--G-C...-C-----C-----G-----G---C-TA---C-G-T--...	6191
D_ZR.84ZR085	--A-G-----G-----A-C---GTAG--CG-A---AG-A--C--T...C---C-----G-----G---A---GGG-AAAAATGAAATCAAT...	6321
FL_BE.VI850	...-GC-----T-----G-C---GAAG-----A---AA--T---TT...-C--GC-----GG-----G-----GC---A-C-A-AGC.....	6110
G_SE.SE6165	--A-----G-----G-A-C---GAA-----A---AAAA-----...-C-G---C---G-----G---TA-----CA.....	6212
H_CF.90CF056	...C--C--T-----G-A-CT--GTAC-----A---CA-----T...C---C-C---G-----G---G---T---CA--G-.....	6119
J_SE.SE92809	...-G-C-----G-----A-CT--GAA-----AA---AAG-A-C--G...-C--GC---C---G---AA--G-G---T---G-A-C-A.....	6101
CRF01_AE_TH.CM240	T--AG--G-----TA---T-----G-C---GAAC--C-A---AA---AGAG--G-TC...C---C-C---G-----G---A---T--AG--A-G-AGACTAGT.....	6368
CRF02_AG_NG.IBNG	...-G-----G-----A-C---GAAG-----A---AAAA---GATG...C---C---G-----G---G---A---TA--G-AA--A-GGT.....	6312
CRF03_AB_RU.KAL153	...-G-----C-----G-A-C---GA-C-G---A---A-A-----...-C-C-----G-----G---A-----A-----	6808
CRF04_cpx_CY.94CY03	--GA-----G-----G-A-C---GAA-----G-A---AA-A-----...-GC---C---G-A-----G-G---TA--GC-AGAGTGCCAAATTAATGGTAGTAAT	6161
N_CM.YBF30	.C-A-G-----T---A-----GCA-C---TGAGC--C-A---AAA-A-C--TT...-CT-TC-G-----GT-GAA--G-G---CA-GCC.....	6375
O_MVP5180	T----G-G---T-TAG--T--G-A-CT--GT-C-C-C-AC-AAA-G-C-A--...C-G-TC-A-C---GT-TCA--C-G-GTAAGG-TA--G-CTCA-A-GCAGTAAATGGA.....	6823
O_CM.ANT70	CCTT--G--G--G--TGAG--T--G-A-C---TGTT--C-A-AC-AAA-G---A--...C-G-TC-A-C---GT-TCA--T-G--GGA-C-GA--G-GACA-GCAGCACAAATAAG.....	6814
HIV2_A.STG---TCAG-----G-CAGG-TTAGAG--G-A---AAAA-ACT-T-T...A--A-ACA-GG-CTC-AAA--G-G-CTGTGA-TCA---C-CAAGAAAGAG.....	6721
HIV2_A.BEN	G-----G-TGC-A--TGAG-----G-AGGGTTAGA-CA-A---AAAAG--GGT-T...A-G-ACACA-GG--TT-GAA--G-GG-TTGTGACA-C-CAACA.....	7291
HIV2_A.ALIGTC--T--TCGG-----C-G-CAGG-TTACAG--A---CCAA-AC-GT-T...A--A-ACA-GG-CTC-AAA--G-GG-TTGTGA-CCATT-A-C--CACCACAAAC.....	7272
HIV2_B.D205	G--A--TG-C-A--TAA--T-G--G--CGGG-CTA-G---A-G-C-AA-AC--T-T...A-A-ACACC-GG-CTC-GAA--T--GAGTGT-ATA--CCAGG-AGTATACC.....	7261
HIV2_B.EHO	--A--G-T-G---TCAA---A--G-CAGG-TTA-A---A-G--TCAA-AC--T-T...A-G-ACACC-GG---AA--T--G-GTGTGA-A-GGGGACA-GGAGTAAT.....	7253
STMM83293	-TCAC--GT--GGT--TAAA-----G-CAGG-CTA-A---A---AAA-AGAG-GT-C...A--A-ACG-GG-CTCTAGC--C---CTGTGA-C-A---...GTCACTGCC.....	6845
SMM251	GC-A--G-T--G---TAAA-----CC--G-CAGGGTTA-A---AC--ACAA--G-GT-C...A--A-AC--GG-CTCTACA--T-GG-TTGTGA-C-AGGGA--...AGCACTGAT.....	7188
SMM9	GCCC--GGT--GT--TAAA--T--C--G-CAGGGTTA-A---AC-AAA-AGAG--T-T...A--A-ACA-GG--TC-AGA--T--G-GTGTGA-C-A---AGC-A-GGAAATGAA.....	6683
SIVCPZANT	GA--C--C--G-A--TAAC--T--CAG-----G-AT-T-A-A---AAA--A---ATG...A-A-CA-A-----GGA--C-T--GAAGTGC-GG-CA-C-A-GAGACTAACTGC.....	6477
SIVCPZGAB	C--A--G-----T--A--T--G-G-CA---GAAT-----A---AA-A-AC-G-TC...-T-C-A-----GTGGAG--G-GG-AACC--GG-----G.....	6846
SIVCPZUS	TATA--G-G-----T-----A-CT--TGAAT-----AC-AAA-A-AC--TG...-T-C-A-----GT-GA--C-GCT-AT--CA-----	6829
SIVSTAN1	GA-T-GTTCT-TA---C-TGC--GGCAGG-TAT-GG---A-GTAAAA---T-T...A--AGTAC--GG--G-T-AG--GT-GG--TGTGAGA-GG-AA-C-ACACCACA.....	6857
SIVGRI677	--CAGC-TCC-----A---TGCA--TGCAGG-TATCA---A-GTAAAA---TT-T...AGCATGACC-GG--G-T-AG--GT--G-CTGC-ATA---AACAGGAAGTGAAAAG.....	6595
SIVVER9063	--CAGCTTC--T--TA-A--TGC--GGCAGGGTA--T---A-C--AA-A-A--TT-T...CA-T-G-A-GGA--G-TGCA--G--CTATTGT-AGA-C-GCTC--GCTCTAACTCT.....	6894
SIVVER155	G-CAGCGTC-----A-C-TGC--GGCTGGGTA--T--G-ACC-AAA-A---TT-C...CA-T-G-A-GGA--G-TGCA--A--CTATTGT-AGCG--G-ACA...TCGCATAAT.....	6874
SIVVERTYO	GCCCGCCTCT--T--TA-A--GCA--GGCTGGCTATG-----A-C--AA-A-A--TT-TTAT-C--TGG-G-GGA--G-TGCA--A--CTATTGC-A-A---GAC...AATAGC.....	6366
SIVVERAGM3	GCCTGTCTCT--T--TA-A--TGCA--GGCAGGGTATG-----A-C--AAAA---GT-T...CA-TGG-G-GGA--G-TGCA--A--C--GTGT-AGA-GGG...AACAAAT.....	6392
SIVLHOESTGC-G-A---CAG-AT--G-A-CAGGGTTATGC---ACTGCAG-AC-G--AT-AAAC-AAGC---AGA...GA---G---CGTGTCT--GAG-GAGGGAGAACAGA.....	6977
SIVSUN	GAGC-AT--GTA--TAAA--AT--G-G-CAGGGCTTTGT---A-TGTAAAAC-G-G-TG...G-ACA-AAC---AGGG-TACA--AG-T-CTTGC-AT-G-----C.....	7039
SIVMNDGB1C-TGTA--TAGG-AT--C--A-CAGG-CTATGC--G-A-TC-AA-G-AG--AT...GTAA--AA---AGAGGGGA---G-G-A-TGTGA-A-----	6308
SIVSAB1C	-C--GCC-TG-----T--C--TGC--GGCAGG-TAT-G--G-A-GTAAA-A-A--TT-C...CAA--G-A-GGG--G-T-AA--GG-GG-TTGTGA---GGGAAGGGAAAAATCAAAC.....	7062
SIVSYK173TTC--T-----T--CT-G-CAGGGG-GT-C-AG-A---AAAA---C--T-T...AGA--T--C--C---GGA--CT-G--GAA-GAG--AGCA---G-AGC.....	6773
ENV	... F N C S F N L T G G F K D K K Q Q Y . R A F F Y K D D L M K E E G N S S	ENV

B_FR.HXB2RACCAGCTATAAGTTGACAAGTTGTAACACCTCAGTCATTACACAGGCTGTCCAAAGGTATCCTTTGAGCCAATCCCATACATTATTGTGCCCGGCTGGTTTTCGATTCTAA	6903
ENVT_S_Y_K_L_T_S_C_N_T_S_V_I_T_Q_A_C_P_K_V_S_F_E_P_I_P_I_H_Y_C_A_P_A_G_F_A_I_L	ENV
A_UG.U455-AT--T--GAC-A-T-A-----T-----AC-----T-----C-----G-----A-----	6348
C_ET.ETH2220--TGAT--GA--A-T-A-----T-----AC--A--A-----C--T--G--C-----T-----T-----A-----A-----	6306
D_ZR.84ZR085GATACTTATGGT-CT--G--A-T-A-----T-----C-----A-T-----A-----A-----A-----	6445
FL_BE.VI850-GTGAA--G-C-A-T-A-----T-----AC-----T-----T-GG--T-----C--T-----T--A-----A--A-----	6225
G_SE.SE6165-T-AT--G-C-A--CA-----TGT--AC--C-A-----T-----A-T-----T-----T--AA-----T-----	6327
H_CF.90CF056-TCAG--G-C-A-T-A-----T-----T-----C-----G-----A--T-----T-----T--A-----C-----	6234
J_SE.SE92809AA-AT--TTC-A-T-AC--T-----A--A--T-----G--T--C--A-----T-----A-----A-----	6216
CRF01_AE_TH.CM240-GTGAG--G--A-T-A-----T--T-----AG--T-----A-----T-----T-----A-T-A-----A-----T-----	6483
CRF02_AG_NG.IBNG-GTCAG--G--A-T-A-----T-----C-----T-----T-----A-----A-----A-----	6427
CRF03_AB_RU.KAL153G--T-----T-----G-----A-----A-----A-----	6920
CRF04_cpx_CY.94CY03	AGGAATAATAGTACAGAAGAG--T--A-T-A-----G--AC--A--T--C-----G--T-----A-----A-----A-----T-----	6291
N_CM.YBF30TATAAT-AA-CA--G-C-A-T-A-----T--A--CTG-G--A--T--ACT-----A-----C--A--AC-A-C--C--A-G--	6496
O_MVP5180-A-CA--T--A--T-A-----T--A--A-T-C-AG-----G-----AGT-----C-----C-----T--AA-A-A-A--C--CT-T--	6938
O_CM.ANT70ACAAACAGC-AA-TG--CA--A--T-A-----T--A--AC--C-G-A-----T--A--A--C-----T--A--A--A--T--CT-T--	6938
HIV2_A_STAA-CA-G-T-CA--ACCAC--C-----A-----C--G--T-A--GAC--CAC-AT-GG--TA-T--GAGGT-TAGA--A--AC-G-----CC-G--	6836
HIV2_A_BENGCTGG--CA-G-T-CA--G-CA--C-----A--A--C-A-G--T-A--GAT--CAC-AT-GG--TG-T--GAGGT-TAGA-C--A--AC-G--C-----CC-AT--	7409
HIV2_A_ALICAG--G-G-T-CA--ACCA--C-----A-----C--G--T-A--GAT--CAC-AT-GG--TG-T--AAGGT-TAGA-C--A--AC--AC--CC-A--	7390
HIV2_B_D205G--A-GCT-TA-A-G--CC--C-----AA--A-T--CCA-G--T-A--GAC--ACAT-AT-GG--CAGCT--AAGGT--TAGG-----T--CC-G--G--TTTC-A--	7376
HIV2_B_EHOGAA-G--AA-G-T-TA-A-A--CC-----T--A--A--T--CCA-G-AT-A--GAC--ACAT-AT-GG--TAGTT--AAGAT--TAGG-----T--CC-A--A--TT-G--	7371
STMM83293GAAGAG-G--A-GCT-CA--G-CA--C--T--AGT--A--CA-G-AT--GAC--ACAT-AT-GG--TG-CT--AAGGT--TAGG-----A--AC--A--TT-G--C--	6966
SMM251AATGAA-G--A-GCT-CA--ATCAC-----T--T--T--CCA-G-AT--GAC--ACAT-AT-GG--TA-T--AGAT--TAGG-----A--TC-A--A--TT-G--T--	7309
SMM9ACTGAC-GT-AA-GCT-TA--ACCA-----AGT--T--CA-G-AT--GAC--CAT-AT-GG--TG-T--AGAT--YAGA--A--C-A--A--TT-G--T--	6804
SIVCPZANTT-C--ATGGCAC--C-----A--ACT--C--AT--GA--TCTA-T--A-----T-----T--A--A--A--T--AT-G--	6586
SIVCPZGABAAC-A--CA--G-A-A-TT-A--C--T--TA--C--A--A--A--AAC-----T--A--C-----A--A--A--G--A--T--	6964
SIVCPZUSAAT-AT-C--GAC--T-A--C-----A-C-CT-C-----T--T--ACC-----A--C--C-----A--AC-A-C--C--C--	6947
SIVSTAN1GGCAG-GAG-T-G-T-CA--TTCAC--C--GA--T--A--A-A-G-A--T--GAG--AACT-AT-GG--TA-CT--AAGAT--AGA--C--A--A--A--C--T--CT--	6978
SIVGRI677GGAAGT-AGGAT-G-T-CA--T-CA-----TGAT--G--A--A-G-A--T--GAT--AAC--AT-GG--TA-TT--AAGAG--AGA--C--A--A--A--G--A--TT-G--	6716
SIVVER9063AGCACA-AAGAA-GCT-TA--TTCAC-----TGAT--T--A--A-G-A--GA--AAC--AT-GG--TGAGC--AAGCC--AGA--C-----A--A--A--TC-GT--	7015
SIVVER155GGGACA-AAGAG-GCT-TA--TCCAC-----TGAT--T--A--AGG-A--T--GAT--AC--AT-GG--TGA-T--AAGAC--AGA-----T--A--A--A--AC--TT-G--T--	6995
SIVVERTYOACTAGC-AAGAG-G-T-CA--TTCAC-----TGA--T--A--A-G-A--A--GAC--AAC--AT-GG--T-AGT--GAGGT--AGG-----T--A--A--A--TT-G--	6487
SIVVERAGM3TCTAAC-GAGAA-G-T-TA--TTCA-----TGAT--T--A--A-G-A--GAT--AAC--AT-GG--TGAGT--AAGAT--AGG--C-----A--G--TT-AT--	6513
SIVLHOESTC--G-T-CA--TCA-----TGAT--A--A--A--A--AAT--GGGTAA-GC-AAATGC-TATT--TAGGCT-----T--A--A--C--A--AT-C--AT-G--	7089
SIVSUNCA-G-T-TA--CA--C--TGAT--A-T-----A--A--C-AT--GTATTA-GC-AAATGC-TA-T-CAGGCT-----A--T--G--G--A--AT-T--AT--	7151
SIVMNDGB1CT-GCT-TA--ATCA-----TGAG--T--A--G--A--A--AG-----GGAC--TT-AT--GATGT--TTTAGG-----TG--TC-A--A--A--TC--GT--	6417
SIVSAB1CGCAACACATACTGTGG-A-G-T-CA--TCCA-----T-----A--C--A--G--T--GAT--AACT-AT-GG--TA-CT--AGGC--TAGG-----T--A--A--A--TT-G--	7189
SIVSYK173TAT--T-CC--TT-CA-----T--AAGT--T--AT--CGCA--A--GAG--ACAGA-----C--A--CT--G--T--A--C-----T--AC-G--ACT-AT-GT-G--	6885
ENVY_Y_Y_L_L_H_C_N_T_S_V_I_S_A_A_C_E_K_Q_T_F_Q_P_F_P_I_Q_Y_C_A_P_P_G_Y_S_L_L	ENV



B_FR.HXB2R	AATGTAATAATAAGACGTTCAATGGAACAGGACCA...TGTACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTAGCAGAAGA	7030
ENV	K_C_N_N_K_T_F_N_G_T_G_P...C_T_N_V_S_T_V_Q_C_T_H_G_I_R_P_V_V_S_T_Q_L_L_L_N_G_S_L_A_E_E_E	ENV
A_UG.U455	-G---GG--CCTGA-----A--G---...-C-GG-----C-A---G-----AG--	6475
C_ET.ETH2220	-G---GAG--A---C-----...CCAT-----C-----AA--G-----AT-----T--A-----GT--	6433
D_ZR.84ZR085	-G---CG--A-G-----C-----C-A-----T-----G-G-----G-----G-----	6572
FL_BE.VI850	-G---G--A-GA-----G---G---...-C-AG-----AA--G-----T--T-----C-----G	6352
G_SE.SE6165	-G---GGG---GA-----A-----T-----A-----G-----A-AC-G-----T-----G-A-	6454
H_CF.90CF056	-G---C---A-A-----TT...-----T-----A-----G-----A-----A-C-----C-	6361
J_SE.SE92809	-G---G--AT-----T-----C-A---T-----C-----A---G-----A---C-----A-----G--	6343
CRF01_AE_TH.CM240	-G---G--AT-----G---G---...-A-----T-----C-----A---G-----T-----G--	6610
CRF02_AG_NG.IBNG	-G---GG--AGG-----G---...-C-A-----C-----C-A---G-----A-----G--	6554
CRF03_AB_RU.KAL153	-G---CG--A-A-----C-----G---...-----T-----A-----T-----G--	7047
CRF04_cpx_CY.94CY03	-G---G-G-A-AT---C---TT-G---...-C-----T-G---G---C-T-----A---G-----T-----ACG---	6418
N_CM.YBF30	-G---G-AGGA-AC-T-G---AT---AGC...-----G-T-T-----A-A---GA---C---GT-AA-CC---A-CT--AATAC---T-G	6623
O_MVP5180	-G---G-C-CAGAC-T-----C-T...-CCAC--A-TTCAGTG--TACT-----C-C-A---AC---AGT-----AA-AC-G---G-CA--CT-TAG---A	7065
O_CM.ANT70	-G---C-GC-CAGAA-T-----CA---...-C-G---CA-A-CGGT--TACT-----C-----AC---AGT---G-AA-A---G-CA--CT-TA---G-A	7065
HIV2_A.ST	G---C-G---CC-AT-ATTCA--CTTT-AG--CAAT--T-T-G--AGTAG-T-CTAC-----AGGATG--GGAAA-GCA-ACC-C---TGGT-TGGC-T---C-AGG-----A-TAG	6966
HIV2_A.BEN	G---G---CC-AC-ATTCA--CTTT-A---TAAG-C-T-A-AGTAG-T-CTTC--C---AGGATG--GGAAA-GCA-ACT-T---TGGT-TGGC-T---C-AG-----A-TAG	7539
HIV2_A.ALI	G---CG-G---TC-AT-ATTCA--CTTT-C---CAAT-CT-T-A-AGTAG-T-CTAC--C---AGGATG--GGA-A-GCA-ACT-T---TGGT-TGGC-T---C-AGG-----A-TAG	7520
HIV2_B.D205	G---G---CC-AC-ATTCA--CTTCATG--CAAC-C-GT-G-AGTAG-GTCCTCC--C---AGAATG--GGAAA--CAGTCC--T---ATGGT-TGGC-C---T-CAAGG-----GA-CAG	7506
HIV2_B.EHO	G---C-G---CT-AA-ATTCA--CTTCATG--CAAC---GT-G-AGTAGT-TCT-TG...TACAGAATG--GGAAA--CAGACC--T---ATGGT-TGGC-C---T-CAAGG-----A-CAG	7498
STMM83293	G---C-G---CA-AT-AT-CA--CTTT-C---TAAT-CT-T-G-AGTAGTCTCTCG---TAGAATG--GGA-A-CA-ACT---TGGT-TGGC-T---A-C-AG-----A-TAG	7096
SMM251	G---G---G-C-CA-AT-ATTCA--CTTTATG--TAAA--T-T-G--GGTGGTCTCTTC--C---AGGATG--GGA-A-CAGACT-T---TGGT-TGGC-T---A-C-AG-----A-TAG	7439
SMM9	GG---G---CC-AT-ATTCA--CTTT-CT--TAAC---T-G-AGTAGTTCTTC--C---AGAATG--GGAAA--CA-ACC-T---TGGT-TGGT-C---T-C-AG-----A-TAG	6934
SIVCPZANT	G---G-AG-G---GAT-T-CA---GT--GATG...-A-----TCAGT---T-C-T---A-C---A-G-GG---ATGGT-A-A---A-C-TACCA-AC-A-CAC	6713
SIVCPZGAB	---C-G---AGAC--TTCA---A---TAA...-T---A-T---G-T---AA---G-TA-C---GT-A-TA-T---A-T---T---G-A-	7091
SIVCPZUS	---G-A-AGAT---A---AG--GAG...-A---T-T---G-G---T---C-A-AA---G-GA-T-A-G-CA-AA---T---ACTA--A-	7074
SIVSTAN1	---GG---CT-AT-AT-CA--TTT-TGTT...-G---TTCAGT---TAGT---TGGATTG--G-ATA-TAC--GAGCT-AGC-T-TGGCA---AG-----A-CAG	7105
SIVGRI677	---G---GAT-AT-GA--CTTT-CT---AAG-C-AG---TTCAGT---G-T---TAGATT--C-ATA-TACTA--A-T-AGGA-AGGA---T---AG-T---A-TAG	6846
SIVVER9063	---CC-G-C-TGAT-AT---TTATAAG-A-AAT--GT---TTCAGT---C---TGGACTT--G-ATA-AC--T-T-AGG--A---G---A-CTATT-T--GA-TCG	7145
SIVVER155	---G---TG-GAT-ATGCA--TTTAAAG--GAA--T-T---TTCAGT---G-T-C---AC-TT--G-ATA-AC---A-C---GGT--T-A-G---A-CTATT---A-TCG	7125
SIVVERTYO	---G---G-AGAT-AT---GTATAA--A-AAT-CT---ATCAGT---G-T---GGCTT--G-ATA-AC---GA---AGGT--T-C-G---A-CTATCAT--GA-TCG	6617
SIVVERAGM3	---C-CG-T-TGAT-ATGCA--GTTTAAAG--AAC--T-T---TTCAGTG--G-T---TA-CTTG--A-ATA-AC---GA-T---GG---T---G---G-CTACT--GA-TCG	6643
SIVLHOEST	G---C---G-AC-ATTAAA-TT-A-T-A-AA...-GA---A-A-CAG-TAC--C---TGGATAT--GCTTAGTAGT---AGT-G-TTCT-TGGC-T---A-C-AATCATACAGG--	7213
SIVSUN	G---G---G-A-ATTAAATGC-ACT-AGAA...-A---A-A-CTG-TACC-CT---TA--TAT--G-CAAGTAC--TAGT-GCTTTT-TGGC-T---A-C-AG-CATA-G---	7275
SIVMNDGB1	G-A---G-G---TTAAAT---AAT-A-TTG...-GC--A-ATCAG--G-G---T--GCAC-TAGTAG-CAC---AGT-GCTTTT-TGGC-T---A-C-A-GCATA-G---G	6541
SIVSAB1C	G---GCAG--CAGAT-ATTCA---CATAAG...GCA---GG---A-CAGT-TC-GCT-C---TAGGTT--C-ATA-TACT---AGC-AGG-A-AGGAA-T---T---TAT-T--C-A-TAG	7316
SIVSYK173	---C---G---CA-AT-TG-A---GAC-ATGTC...-----A-CTG---TC---C---G-A-T-C-ATA-T-G-CTAGC--ATGGT-C-AAC-G---G-CCTACAA--C-A---	7012
ENV	K_C_N_D_T_N_F_E_G_D_D_V...C_T_N_V_T_A_V_S_C_T_Q_E_F_N_T_L_A_S_T_W_F_Q_L_N_G_T_Y_K_A_K_D	ENV

```

B_FR.HXB2R      GGTAGTAATTAGA.....TCTGTCAATTTACGGACAATGCTAAAACCATAATAGTACAGCTGAACACATCTGTAGAAAATTAATTGACAAAGCCCAACAACAATACAAGAAAA      7139
ENV             _V_V_I_R_....._S_V_N_F_T_D_N_A_K_T_I_I_V_Q_L_N_T_S_V_E_I_N_C_T_R_P_N_N_N_T_R_K_      ENV

A_UG.U455       AA--AGG-----.....AA-----AA-----A-----TGT--ATC-----A---C-----T-C-----TT---T...-----      6581
C_ET.ETH2220   -A-A-----TG.....-T--AA-----C-G--AA-----C---TA-----T--TGA-----C-----G-----G---T-----G---      6542
D_ZR.84ZR085   -A-C-----A-----.....AA-----C-----AA-----T-----T-----T-T-TC-----G---T-----A-GGA-----C---      6681
FL_BE.VI850     TA-----C-----.....CAA--A-T-AA-T-----A-----C-T-TGA-----C-G-----G-----G-----      6461
G_SE.SE6165     AA--AA-G-----.....AA-----A-----A-C---GT-----T-T-A-A-----C-----C-----T-----TG---      6563
H_CF.90CF056   -A-CA-----.....A-AAA--A-T-A-----A-C---A-----T-G--C-A--A-C-----CA-----C-G--T-----T-----G---C-      6470
J_SE.SE92809   CA--A-----.....AA--CA-T-A-----C---A-----A-T-T-A-A-----GTG--TAC-----T--T-----T-----G---      6452
CRF01_AE_TH.CM240 -A--A-----C-----.....AAG--C--AA-----C-----G-C-T-T-A-----C-----C-----TC-----C---      6719
CRF02_AG_NG.IBNG A---G--C-----.....AA--A--AA-----C-----T--GCT-ATC-----A-----C-----T-----      6663
CRF03_AB_RU.KAL153 -----.....A-----A-----A-----AGC-----G-----      7156
CRF04_cpx_CY.94CY03 -----.....AAA--A--A-----A-C---AT-----TGCA-AGG-----A-----C---TGG-----      6527
N_CM.YBF30      AA-T--T-----.....AT--T-G-CAC-GT-ATC-GT-G-G-ATG---TGAGA-A-GCC---A-----G--AGGA--T-----G--GG-      6720
O_MVP5180      -A--AG-----TG.....GGAAAA--A-T--A-ATCA--A-G-AT-C-----ACC--A-----TC--A--A-C--G-CC--C-T-----GAAGGA-TTGCAGAGGT-C-      7174
O_CM.ANT70      AA--AG-----G-TG.....G-AAAAG--A-TTT--AGG--GA-----AT--C-----GACC--A--T-TA-CC--A-C--G-CC--GA-----AC-A-TA...GAC-T-C-      7171
HIV2_A_ST      AAC-TAT--CTAT.....TGGC--GG-GA--T--AGG-CT-T--T-GCT--A-CAA-T-TTATAA-C-CAC--GCG---AG---AGGA--T--G--GTTT-      7072
HIV2_A_BEN     AAC-TAT--CTAT.....TGGC--GGT-GA--T--AGA-CT-T--C-GCT--A-CAAAC-TTATAA-C-TACT--GC---AG--G--AGGA--T--G--GTTGT-      7645
HIV2_A_ALI     -AC-TAT--ATAT.....TGGC--GAA-AA-----AGG-CC-T--GCT--A-TACAT--TATAA-T-GTC---AC-C---AG--G--AGGA--T--G--GGTTGT-      7626
HIV2_B_D205    -AC-TAT--CTAT.....TGGC--GGT-AA-----AGG-CT-T--GCT--A-TTCAT--TTATAA-C-GAC---GC-C---A--G--AGGA--A-TGGTTGTG      7612
HIV2_B_EHO     AAC-TAT--CTAT.....TGGC--GG--G--T--AGG-CT-TT--T-GCT-GA-CAA-T-TTACAA-T--AC--G-GC---G---AGGG---A--GTCTT-      7604
SMM83293       AAC-TAT--CTAT.....TGGC--GGT-G--T--AGG-CT-TA--T-GTT--A-TAA-T-TTATAA-C--AC---G--A--G---AGGA--T--G--GTTT-      7202
SMM251         AAC-TAC-----TAT.....TGGC--RG--GAAG-----AGA-CC-TA--T-GCT--A-TAA-T-TTATAA-C--AC---G-GA---G---AGGA--T--G--GTCTT-      7545
SMM9          TTC---G-A-TG.....AA--GTCGCAA--AT--ATC--TGCTTGA-T-G-T-G-A-AGAATT-GA-AACT--AC---CA---T---AGGA--T--GG--GT--G-      7040
SIVCPZANT      CA--ACTG-A-----.....GTG-AG--AAGT-AA-A--A--G-TGT-TGG-----AGTAGA-G-A--TAGTC-A-----CAT---AGGA-----GGG-      8222
SIVCPZGAB     T--TACTG-A-----.....AGT-AAA--TT--AG-C-TT--TC-----AT-CTCAGAGGGA--CA-T--G-C---TT---AGGA-----GTGGG-      7200
SIVCPZUS      AAC--A--AT-G.....C AAAAGCATGGAG-G-GCA-T--T--GTG-TA---A-C--A-TAAAC-TTATAAGT-GA-G-AGTG---G---AGGA--T--G--GTCTT-      7174
SIVSTAN1      AAC--AG--AT-G.....CAGAAA-GAGGAAATGATA-TG--A-AGTT-TA---AGT-GA-TAA-TTTTACAAC-T-GAC-G-G-GA--CCG---TGTT--T--A--GTGTTG      7217
SIVVRI677     AAC-CA--AT-G.....CAGAAGCAT-GAG-G-ATA-TG--T--GT--TAG-TC-TT-TA-CAA-T-TTACAA-T-GAC-G-C-C--C-A--G--AGGA--T--G--GTCTT-      6958
SIVVER9063    AACCCAG--CT-G.....C AAAAACATGGAG-G-GCA-TG-CT-AGTGTTA--CT-GC-CA-TAA-C-TTATAACC-GAC-G--CA--C-A--G--AGGG--T--G--GTCTT-      7257
SIVVER155     AACCCAG--AT-G.....CAGAAACAT-GGG-A-ATA-T--CA-AGT-TTG--CT-GT-CA-CAA-C--TATAA-C--TC-G-C-CC---GG---AGGA--A--G--GTCTT-      7237
SIVVERTYO     AACCCAG--AT-G.....CAGAAACAT-GAG-A-GCA-TG-CT-AGTGTAG-GT--T-TA-TAAAC-TTACAA-C--AC-G---C---C-A---AGGA--A--G--GTCTT-      6729
SIVVERAGM3    T-AGC-T---CC-CTAACT.....C CAATAAGATG-AAG--C-A-AT-GTGCAAGTTTGTGTAT-A-----GCAGGA--ATGGGG-T-ATT--GA--C-T---AAAGGA--T-GAT--GA-GT-      6755
SIVLHOEST     T-AGT---ACC-ATCAAT.....AACAAGGTA-GAC-AGGGG-T--GGGA-AAT-TGTATGG-A-----GCAGCT--GTGGGGT--T---C-G-C-T---GAAAGGG--T-GGT--CA-GT-      7337
SIVSUN        A-A-T-G--ACCCATAGATGATAAAATATAGGGGCC-A-AGG-A--TCATC-A-GGAAGTTTGT-TAT-AG--G-CAGGA--ATATGGCT--A-G--AG-A--CAC---AAAGGA--T-GGT--GT-GT-      7396
SIVSAB1C      AAC--A--CT-G.....CAGAAAAATGGCAATT-CA-TG--T--GTT-TT--T-GGC-CA-TAGAT-TTTTAA-T--ACT--C-GA---G---AGGA--T--G--GT-TTG      6671
SIVSYK173     TAA--T-GATTT.....ATCAAACAA--GGAT-AAA-TG-GT-AGT--TT--TT---T-CAGAAGCACTTAGAT-GC-G---TA--GA---AGGA--TG-GT-T-TC---      7428
ENV            _K_V_R_F_....._I_K_Q_K_D_K_N_E_S_V_I_I_L_V_P_E_A_L_R_L_Q_I_I_C_E_R_P_G_N_E_S_I_K_      ENV

```

	loop tip	V3 loop end <-	
B_FR.HXB2R	AGAATCCGTATCCAGAGA...GGACCAGGGAGAGCATTGTTACAATA.....GGAAAAATA...GGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCAAAATGGAATAACACTT		7248
ENV	R I R I Q R G P G R A F V T I	G K I G N M R Q A H C N I S R A K W N N T	ENV
A_UG.U455	-AT-A-A---GGT-T--TATA--T---ACA-----CTA-GT--C-.....	-T-----ATA--GG---A-----TG-C-----AGGG-C-----GA--A	6696
C_ET.ETH2220	--T-AA-G--A.....-ACA-A-----CTA-G--C-.....	--G-C---ATA--G---A-----GA--A-----A--C	6648
D_ZR.84ZR085	--G-CA-C--A.....-A-----CA--GC-CTA--C-.....	A-GT-T-C-ACAA-G-TA--AG-----T-----T-----G--T-----C-	6787
FI_BE.VI850	G-T--A-A-T-A.....-ACA-A-----CTA-G--C-.....	--GC---ATA--G-C-A--A-G-----G-A--C-----C	6567
G_SE.SE6165	--G-AA-A--G...G-ATT-----ACA-A-G-----CTA-G--C-.....	--TGCC--ATA--G-C-A-----TG--C-A-AG-----AG-GG-C-	6675
H_CF.90CF056	--T-A-A-T-A.....-G---AC-----CTA-G--C-.....	--TG-C-CATA--G---A-----T-----A--G-C-----G---	6576
J_SE.SE92809	G-T--A-AC-G.....-ACA--TGC-CTACG--C-.....	--G---ATA-----A--G-A-----GA-AG-G-T---G---	6558
CRF01_AE_TH.CM240	--T-AAC--A.....-AC--T-----CTA-G--C-.....	--G-T-ATA-----A--A---T---G-G--A-G-A-----AGT--	6825
CRF02_AG_NG.IBNG	G-TG-A-A--A.....-CA-----CTA-G--C-.....	--TG-C-ATA--GG---A-----TG-C--A-A--G-----A---	6769
CRF03_AB_RU.KAL153	G-T--T-A--A.....-TA-G--C-.....	--G-C-ACA--G--A-----T-A-----A-----A-	7262
CRF04_cpx_CY.94CY03	--TG-A-A--A.....-G---ACT-A--GGTA-G--C-.....	--TG--ATA--G--A-----G--AATG-T---G---C-	6633
N_CM.YBF30	CAGG-G-AG--A.....-T-CT-TGA-----TA-AC--G-.....	--A-----GTA--G-C-T-----T-C-----TG-CTC--A--A-CT---G-ACCA-TG-	6826
O_MVP5180	GAT--ATA--CA.....-T---TGAG--GGCGC-GT--GACACTTAAAAGAAGTAAC--T-C-TCACC--GATCA--GGT--TT-----CATA--A--AGA-TGT---G-A--TG-CC		7292
O_CM.ANT70	GAG--GA--A--A.....-T---TG--C-GGTAC-GC--G.....GGAATAGGG--C-GC-GGAAAC-GCTCA--GGC--TT--C-GTA--A-GCCA-TG-T--GGA--A-TA-		7283
HIV2_A_ST	CC--AACAC--AT-TC.....--TT-TG-----CACTCC.....	.CAGCC--CAATA--GACCC--G---TGG--CTGGT-C-AAG-C-AGTGGAA-G-AGC--TGA	7175
HIV2_A_BEN	CC--AACAC--TAT-TC.....--TT-AT-----CACTCC.....	.CAGCC--CAACAC--GGCCT--G--G--TGG--CCGGT--G--AG-CAG-TGGA--G-AGC--TGC	7748
HIV2_A_ALI	CC--AACAC--TAT-TC.....--TT-AT-----CACTCC.....	.CAGCC--CAATAA--GACCC-----TGG--CTGGT-C-AAG-C-A-TGGA---AGC--TGC	7729
HIV2_B_D205	CC--AA-A-C-GT-TC.....--ACT-CTT--C...CATTC.....	.CAGCCT--CAATAAG-GACCC-----TTGG--CTGGT--AGG--AACTGGACAG-AGC--TAA	7715
HIV2_B_EHO	CC--AA-A-C-GT-TC.....-T-TTCTC--C...CATTC.....	.CAGCCT--CAATAA--GACCT-A-----TTGG--CTGGT--AAG--AACTGGACAG-AGC--TAC	7707
STMM83293	CC-G--ACC--TAT-TC.....-CTTG-TC--C...CATTCG.....	.CA-CCC--TAAT--AG-GACCA-A-----TTGG--CTGGT--G-AG--A-TGGA-AGGAGCA-TAA	7305
SMM251	CC-G--ACC--TAT-TC.....-ATTG-TT--C...CACTC.....	.CA-CC--CAAT--AT-GGCCA-AG--G--TGG--TGGT--G-AG--AA-TGGA--G--GCA-TAA	7648
SMM9	CC-G--ACC--TAT-TC.....-TTG-TC--C...CATTCG.....	.CA-CCC--AAT--AG-GACCA-A--G--CTGG--CTGGT--G-AG--A-TGGAAG--GGC--TCC	7143
SIVCPZANT	-ATC-A-AA--A.....-A-TGA-T--CTA--ACG--G-.....	-A--TG-ACA--G--C-CT--GA--GTTT--CAG-C-A--AGA-GCT---G-AC-AG-AC	6928
SIVCPZGAB	GAGG-G-AG--A.....-G---A-TGA-C--TA--AT--G-.....	-A--TG-GTA--G--CC--TCT--CT-C---G--A-A-G-GA-C-C-----CGA--GG	7306
SIVCPZUS	-ATG-A-AAC-A.....-A-TGA-T--TA--AC--G-.....	.CC--G--GTA--G--G-A--G-----CTCA-A-CTG-C---G-G--ACGA	7280
SIVSTAN1	CC-G-AACC--AT-GC.....-TCTC-TG--C...CACTC.....	.CA-C--TATAATAC--AT-A-----TGG--CC-TT-CCAGG-T-ATTGGAAGGAGC-TGGA	7320
SIVGRI677	CC-G-AACA--AT-GC.....-TT-T-----CACTCT.....	.CAG--TATAATACC--GGT-A-A-----GTGG--CC--T-CCAAG--ATTGGAAGGGCATGGA	7061
SIVVER9063	CC-G-AACC--AAT-GC.....-TTT--TT--C...CATTCG.....	.CAG--TACAAATAC--GC-T---G--CTGG--C-TT-CCAAG-CAGGTGGA-AG-AGC-TGGA	7360
SIVVER155	CC-G-AACG--AAT-GC.....-ATT--TC--C...CACTC.....	.CAG--GTATAATAC--GAC-A--G--G--CTGG--CC--T-CCAGG-CAATTGGAAGGAGCCTTGA	7340
SIVVERTYO	CCGG-AACG--AAT-GCG.....-ACTG-TT--C...CACTCT.....	.CA---TACAACATG--GC-T---G--TTGG--C--T-CCAGG-CAATTGGA-AGG-GC-TGGC	6832
SIVVERAGM3	CC-G-AACA--AT-GC.....-CT--TG-----CATCT.....	.CAG--GGTACAATAC--GGC-----TTGG--C--T-CCAGG-CAACTGGA-AGGAGC-TGGA	6858
SIVLHOEST	TCC-CAAT--G-TCT-C.....-ATATTT---TA-TATGGC.....	.CT-G--CATGCA-C-GAT-A---TT-----G---AT--GAAG-CCA-TGGG-A-GA-TGTT-A	7443
SIVSUN	--T-CTATA-GTAGT-C.....-ACTGCTG---TA-TATGGG.....	.CT-G--GCC-GGTTCT--GT-A---TG-----G---AT--GTAG--CA-TGGG-A-GAGC-TT-G	7502
SIVMNDGB1	--T-CT-CATCAGCT-C.....-ATT-TT---TA-CATGGG.....	.TT-G--CCTGAAAAG--T-A-AGA--GCATG--C-C-T-C-AAG--CGTTGGG--TTAGCACT--	6777
SIVSAB1C	CCTG-AACC--AT-GCT.....-CTTG-TC--CCA--GC.....	.CAG--TATAACACC--GGT--A-----TTGG--CTGGT--G-AG-TAATTGGA-AGGAGC-TGGA	7531
SIVSYK173	-AT--T-AAT-GGCTGC.....-CTATTTC-CCCAGT--G-.....	.CA-GGG-A-TTGAA--CCGG--GG-C---A-AC--GC-T-CT-C---T--C-G-A---TGG--AG	7230
ENV	N I Q L A A G Y F L P V I	Q G K L K T G R D A K R A F C R V T G N W T	ENV

B_FR.HXB2R	TAAAACAGATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAATC.....	7304
ENV	L_K_Q_I_A_S_K_L_R_E_Q_F_G_N_N_K_T_I_I.....	ENV
A_UG.U455	--C----G----GAAC-----AGA--A-G--AAT--C--A-C.....	ATAAT 6748
C_ET.ETH2220	--C--A--G--AAGGAA-----CA-A-G--C--CCCT-----	AAAACAATAGA 6700
D_ZR.84ZR085	--CG---G-----A-----G--A-C-TT-GAACC-A.....	ACAAAAATAAT 6839
FL_BE.VI850	--GG-GT-TG--AAGGCAG-----AGTCG--T--CCCT-----	ACAGCAATAAAA 6622
G_SE.SE6165	--C-GA-TG-C--GCAG--C--G-GA--ATC--AAT--G-GC--GCGAG--AC.....	ATAAC 6733
H_CF.90CF056	--C-C--G--T--CAC-----G--AT--C--GAAC-----GA.....	ACAATAAG 6628
J_SE.SE92809	--CGTAGAG-----CA--C-----C--AAT--A-CA.....	ATAAAA 6607
CRF01_AE_TH.CM240	-----G--A--GAA-----A--G--C-----G--C-----	ATAAT 6874
CRF02_AG_NG.IBNG	--C-T--G--T--CAC-----AGACGT--C--AAG--C--CC--C.....	ATAAT 6821
CRF03_AB_RU.KAL153	-----T--T-----A-----G-----A-C--T-----	GT 7314
CRF04_cpx_CY.94CY03	-----GT---AG-GAAG---G-A-AG--TC--CCCT---A-C--T--A-	6686
N_CM.YBF30	GG--TAGA-C-AGAGAGG--A--AGA--ATCC-G--G--A--C--C.....	ATAACCTTCAG 6884
O_MVP5180	--C-----A-C-----TA-GG-ATTT-A-T-TTG-AAACC-A-CAG-G-AT.....	GTTACCATAAT 7353
O_CM.ANT70	-----A-C-----GAA-GG-ATTT-----T-G-AAAC---CAGGT-GT.....	ATTAACATGAC 7344
HIV2_A_ST	AGG-GGT--AGCTA-C-CTTGC--A-C-T-CCAGGTAT--AGGA-CC-ACGAC-CA.....	GAAAAAATTCGTTT 7245
HIV2_A_BEN	AGG-GGT--AGCAA-C-CTTG--CA-C-T-CCAGATAC--AGGA-TC-ATGAT-CA.....	GGGAAAATTAACCTT 7818
HIV2_A_ALI	AGG-GGT--AG-AA-C-CTTG--A-C-T-CCAGGTAT--AGGA-CC-ATGAC-CA.....	AACCAATTAACCTT 7799
HIV2_B_D205	AGG-GGT--A-AGG-C--TCA--A-C-T-CCAGGTAT--AGGAGGTG--A--ATATCACA.....	AGCGTAAAGTTAGT 7791
HIV2_B_EHO	AGG-GGT--A--AG-C--TTAA--ATC-T-CCAGATATTCAGGA-C-----AT---	TCACAGATAAGGTT 7777
STMM83293	A-G-GGTT-AG-AA-C-TTGG-C-A-C-T-CCAGGTAT-CAGGA-CC-ATGAC-CG.....	GCAAAAATAAGGAT 7375
SMM251	A-G-GGT--A-CAG-C--TTG-C-A-C-T-CCAGGTAT-C-GGA-CT-AC-AT-CT.....	GATAAAATCAATTT 7718
SMM9	GGG--GT--AG-MAMC-TTGG-C-A-C-T-CCAGGTAT--GGGA-CT-ATRAK-YT.....	RAACAAATTAAGCT 7213
SIVCPZANT	GT--CA-A-C--AGCA-GTTC-TGCG--G--T-GGAA---AGTAG-C-AC-A.....	ACCAATGCCAAAAC 6995
SIVCPZGAB	--GG--G-AG--AAA-AGGCT--GC-ACCTC--CAAATCGA-CAGC-G-C.....	AAACATAAC 7364
SIVCPZUS	G----T-C-CTTTAGAG-T-A--A-A-GG--GCAAACCTG.....	ACAAAGGTAGA 7332
SIVSTAN1	GGG--GT--G-AAA-CA-T-G-GGAGCTT-C-AAA-AG--AT--CG-GGG-C--AT.....	AAACACAAGGCAGAT 7390
SIVGRI677	A-G--GTC-G--AAGAAGT-AAG-A--TGA--AA-CTT-CAG-AGT.....GC--A.....	GAAAAATATACATCT 7128
SIVVER9063	AGG-GGTT-A--AAGAA-T-A-C-A-TT--C-AAA-AT-GGT--CGGGG--CC-AT.....	GATACAGAGAAAAT 7430
SIVVER155	AGG--GTACA--AGGAA-T-G--A-TT--C-AAA-A-CGGT-CC--GGC-CC-AT.....	GATACAAAACAAT 7410
SIVVERTYO	GGG--GTA-A-CAA-AA-T-G--GAGTT--C-AAA-AC-GGT----GG--CC-AT.....	AATACGACACAT 6902
SIVVERAGM3	A-G--GTA-A-AA-GAA-T-G--A-TT--C-AAA-AT-GAT-CC--GG--CC-AT.....	GATACGAAAGAT 6928
SIVLHOEST	AT--T-TAGGCAAA-TGCT-AAGGAGTTGA-TGCA-A-GCC-TG--TTAT--C-GAAGGGACA.....	GGGACATGTGATTCTAAGAAAACAACCTTGTGGACGGAAATTAAGAGCTTACCAATAGC 7564
SIVSUN	CC-CCTTAGGGAAACTATTGAGGCAG-TGG--CC-TT-GC---C-TG-GT-ATGG-TGTACATTTGATAATATAACAAAGACATGCAGATTTTACCAATGGTACAGATTTTAAAAAGATGATTTAAGTTTTAA 7632	
SIVMNDGB1	GG-GT-TAGCTAAAGAACT-AAT-A-TT-A-TGACTCC-TC--AGTG-ACCAG-C-----	6833
SIVSAB1C	AGG--GTC-A--AA-CA-T-G-G--CT--C-CCAAA--GT--GTGGC-C--AT.....	GACCAAAAACAAGAT 7601
SIVSYK173	A-TTTTTT-AGCAGGTAC-TGA-CAG-CCACGAAAACCTGG--A--TGT--CC-A.....	ACCAC 7291
ENV	E_F_F_K_Q_V_H_E_Q_A_T_K_T_W_K_N_V_T_N.....	T T ENV

B_FR.HXB2R	.TTTAAGCAATCCTCAGGAGGGGACCCAGAAATTTGTAACGCACAGTTTTAAATTTGGAGGGGAATTTTTCTACTGTAATTC AACACAACCTGTTTAATAGTACTTGGTTTAAATAGTACTTGGAGTACTGAA	7433
ENV	F K Q S S G G D P E I V T H S F N C G G E F F Y C N S T Q L F N S T W F N S T W S T E	ENV
A_UG.U455	C---GCTAGC-----TAT-----AC--A--T--C-----A-----T--C--CA--T--GGC-----C-T----AA-GGC--C-TG-CA-A-GACATG	6878
C_ET.ETH2220	A-----C-----T-----AC--A--T--C-----A-----T--C--A--T--A--C-----C--AAAAC-AG-ACTGTT-AAT--C-G-AC-	6830
D_ZR.84ZR085	T---A-C-----T-----AC--A-----C-----A-----T--GG-----G-A---AA--TATCAGGACAT-----A-GG	6969
FL_BE.VI850	A-----C-----T-----AC--T--T-----A-----A-----CG--A--T--GG-----GAC--AG-C-CC---A-GGCACT.....	6743
G_SE.SE6165	C-----CTC---TG-----T-----AC--A--T-----C-T---A--A--G-----T-----A--T--GGC-----TCA-GCCT--G-GATCA-A-A-T--GAGA-T	6863
H_CF.90CF056	C-----C-AA-----ATG--G--AG--A--T-----A-----A-----T--C--A--T--GGG-----G---GAA-TGCA---AATTAC--ATC-	6758
J_SE.SE92809	C---CATC-C-----TAT-----G--A--T-----A-----A-----A--T-----A--T--A-----C-----G---GA---G-A--GCATTGAAG--AC-	6737
CRF01_AE_TH.CM240	C---C-A-C-C-----A-T-T-----AC--T--TCA-----A-----A-----T--C--A-----A-----A-----CC-AGGA-A-GAAACC-TGG-G-GG	7004
CRF02_AG_NG.IBNG	C---GCTA-CC---T-----TGT-----AC--A--T-----A-----A-----T--C--A--T--A-----C-----GA-----C-A-A-C-C-G-CA-C	6951
CRF03_AB_RU.KAL153	C-----T-----T-----T-----A-----A-----A-----A-----A-----C-----A-----A-----C-----A-----A-GG-ACTGAAGAGTT-	7438
CRF04_cpx_CY.94CY03	..---GCT-C-C--GT-----T-----AC--A--T--C-----AA--A-----T--C--A-----C-----ACACA-GC-A-A-GG-ACA-AC-T-AC-	6815
N_CM.YBF30	GGC-CGAG-GAGGAAT-A--A---T---G-GAC-CACTTA-TG-C-----A---A--G-----T---CA-TT-CA--T-A-----GAGGAA-TAC---CGAG--AG-TGAGC--...	7011
O_MVP5180	A-C-CGAG-A-TAGT--T-A-TG-----G-AAGCCATTTACA-----C---CAT--A-----C--T--T---CA--T-TGGGA-----CTA---TTA-C--CT---AAA-TCCGGATGC	7483
O_CM.ANT70	A-C--T--CAG-AGC--T-A-T-T---GG-AACCCATTTACAC-----C---CAT--A-----C--T--T---CA--G-TA-GA-----TA---C-TT-CATG--ACGGAACC-CCTG-AGT	7474
HIV2_A_ST	TA-AGC--C-GGAGA-C-CTCA-----G-G-C-TACATGT-GAC--C-CA--A-----C-----C--ATG--TTGGT-CC-C--T-GGTAGAAAAC-GA-CG-A-CA--CACAG...	7372
HIV2_A_BEN	TACG--A-CGGGAG---CTCA---G--G-C-TTTATGT-GAC--C-CA--A-----C-----C--CATG--TTGGT-CC-C--T-GGTAGAAAGAC-G-ACCAAACAC-GCGC...	7945
HIV2_A_ALI	TACA--A-C-GGAAG---CTCA-TG---G-G---TATATGT-GAC--C-CA--A-----C--C-T--C--CATG--TTGGT-CC-C--T-GGTGAAAAC--A-CGGG-CA-GAACAG...	7926
HIV2_B_D205	A-CAG-A--TGAAA---TTCA-T-----CAAC-TATATGT-GACC-----A---A---C--A-----ATG--CTGGT-C--A--CT-GGTAGA-AAC--A-CA-A-ACA-C-CGG...	7918
HIV2_B_EHO	AGCAG---CG-GAG-A-CTCA-T-----G-AAG-TATATGT-GAC--C---A-G--A--G--C-----ATG--TTTTT-C--A--CT-GGTAGAAA--GA-C-GGGCTC-AG-GA...	7904
STMM83293	AG-GGCT-CT...GGG---A-T---GG-CACCTTCATGT-GACA---CC-T-A---C--T--T---CAATG-ATTGGT-C-A--T-G-TAGAAAAC-GA---A-CAGAA.....	7496
SMM251	AACGGCT-CT...GG---A-T-G---G-ACCTTCATGT-GACA---CA---A--G--CC-----AATG-ATTGGT-TC-A--T-GGTAGA-GA--GGGA-GTAACT-CC.....	7839
SMM9	AACAGCT-C...GG---A-T---G-ACCTTCATGT-GACA---C---A---C--A--T--C--AATG-ATTGGT-TC---T-GGTAGAAGY--TA-AA-A-G-TTC.....	7334
SIVCPZANT	AA-ATG-AC--T-CA--AT--A-T-T---G-AAA-GT---TT-G---CCA--A-----C--T--T---G--AT---CTTG---C---GCC--A-ACACGGGA-ACCTCATC-CA-AC-G-	7125
SIVCPZGAB	AC-C--CAG-G---T---A-T---GG-CAC-CAT--T-TG-----A-----A-----T-----A--T-GC--A-A---C-GAC-ACATTACA---G-C-T.....	7482
SIVCPZUS	G--A-TT-C-AATG-----A-----G-G---ATATG-TGC-----A-----T--T---A---TT-CC-A---C-TG--C-ACAAC---CCGACAAC-CC---...	7459
SIVSTAN1	T-GGCTAAGTAGACA-T-G-A-T---GCA-CT-ACATTT-GC-C---CCA--A--G-----T---C-C-TGATTGGT-TG---T-GCTGAATAA-G-GTC--A-A-TG-A-GAA-T	7520
SIVGRI677	GAGA-G-.....AT-T-G-A-T---TCA-CG-ATTTT-G-C---CA--T-----T---GATGGACTGGT-TA-C---TA-CTAAACAA-CGA-CAGAAGATGCAGAA...	7249
SIVVER9063	A-A-CT---GAGACAGTTT--A-T---GCA-C--ATTTGT-G---CCAG--A--G--C-----C--GATGGACTGGT-TC-A---TA-TTGAATAA-C-ATCAGTAGATGCA-A-C-T	7560
SIVVER155	C---TT---AGACA-TTT--A---GCA-C--AT-TAT-G-C--C---CA---A-----C-----AATGGACTGGT-T-A---TA-CTGAATAA-TTA-CAGTGGATGC-GA-C-T	7540
SIVVERTYO	A-ACCT---AGACA-T-G-A---GCA-C--ACTTGT-G---CA---A-----C--T--T---GATGGATTGGT-T-A---TACTTAAATAA---A-CATGGGATGCATAC...	7029
SIVVERAGM3	T-A-CT---GAGACT-TTT--A-T---GCA-C--ATTTAT-G---CAG-----C-----T---AATGGATTGGT-TC-A---TACCTGAATAA-CG--CAGTAGATCCGACC-T	7058
SIVLHOEST	TAA--T-ACTAGGCAT---CA--TT--C--CA-AG-T-TTA-TGCA--C-----A---A--G--TT---GT---TAGGA-A--CC-AGAATGGAACAA---A-A-T-AGAT-AATGG...	7691
SIVSUN	ACCA-T-G-GAAGCAT---CA-TG--C--CA-AG-T-TTG-TGA-G-C-----AA--GA-G--TT---T---AG-A-T--C--AGTATGGAATGAC-CA-CCT-CAAT-AATGG...	7759
SIVMNDGB1	...TGTA--AAT-TTACTA-CACTGG---GGAGAAC-AA--A-ACACGG-CAAGCA-AA--G---GC-A-A--C-TAAAG--T-TTAA-A-AG---A-TA-ACTACATCAG-AGA-A-AGCAG-A---	6959
SIVSAB1C	T---CTA---AGACA-T-G-A-T---TC--A--TTCTT-TT-----CA--T-----C-----AATGGATTGGT-CC-C--CTA-TTGAATAA---GTCAGTAGACCCA.....	7725
SIVSYK173	A-GG-G-TCCCAGC---G-A-T-T---GG-ACG---A--TT-G--CC-A---G--A--G--C-----GT-T-TA-G--C--GC--AC-T-AC-AA-GGA-ACG-AA-T-AG-A-A-T	7421
ENV	W R S Q P G G D L E V R T H W F Q C G G E F F Y C N V S K L F A N I T N G N A S K N N	ENV

B_FR.HXB2R	GGGTCAAATAACACTGAAGGAAGTGACACA.....	ATCACCCCTCCCATGCAGAATAAAAACAAATTATAAACA	7500
B_FR.HXB2R	_G_S_N_N_T_E_G_S_D_T.....I_T_L_P_C_R_I_K_Q_I_I_N_	ENV
A_UG.U455	---C---GG-----	---A-T---A---T-----G-----T-	6930
C_ET.ETH2220	AAT-T---C.....	---A---A-----	6876
D_ZR.84ZR085	TTAAATG---CA-TC.....	---A---A-----	7021
FL_BE.VI850	---T-----TC-G-----G----	6780
G_SE.SE6165	---A-T.....	---A-----T-AG-----G-G-GA-	6906
H_CF.90CF056	AATGAC-CA--GGGAA-C-A--AC.....	---T-A-G-----T-A-----G-G-GA-	6819
J_SE.SE92809	AATGAC-CA-GTNAC-C-ACT.....	---AA-A-----T-A-----G-G-GA-	6795
CRF01_AE_TH.CM240	T-TAATG-C-CT.....	---A-T-----AG-----G-----T-	7053
CRF02_AG_NG.IBNG	CACA--GGGTCA-A---CACT.....	---A-T-----A-----G-----G----	7009
CRF03_AB_RU.KAL153	AATAAC-C-G-AGGA--TAT.....	G---A-----C-----T-G--GG-	7496
CRF04_cpx_CY.94CY03	A-TA--G--TCT--AA-TTC--CC.....	---A---A-----C-----T-G--GG-	6876
N_CM.YBF30	---T-T-G--T--T-----G--G--TT	7048
O_MVP5180	CA-GAG-TC--AGGGAGCAATGAGAC--ATAAAAATGGT.....	---TA-A-T-----AGT--G--GC-AG--GAT	7556
O_CM.ANT70	-TTAGT--GTT-G-C---T-ACA-TGGC.....	---T-A-T-----A-C-G--GG-GG--GGT	7538
HIV2_A_STCAC-ATTATG-G---CAT---G---A-T-T-	7412
HIV2_A_BEN-A-TAT.....-CAT---G-G-A-T-T-	7976
HIV2_A_ALICAC-ATTATGCA--G--CAT---G--A-T-T-	7966
HIV2_B_D205AGA-ATTATGCG---TCAT--C-G--G-A-C---	7958
HIV2_B_EHO-ATTATGCGT---CAC-C-G--G-AG-C---	7941
STMM83293ATGAGAGATTGGAACAAAAACAAAAGGAACAACAAAAGAGG-ATTATG-A-----CAC-C-G--G-A-T-T-	7575
SMM251CAGAGGCCAAAAGGAACGGCATAGAAGG-ATTA-G-G--G-TCAT-T-G--A-C---	7903
SMM9AGATGGACAAGTCAAAACCAGAAAGACGATACCAGAAA-ATTATG-G---TCAT--TRG--G-A-C---	7410
SIVCPZANT	-CC.....C---TAGCA-AT-----T-G-G--AG-T--TC	7165
SIVCPZGAB-A-TA-G-----T-G--G-AG-T-GTT	7519
SIVCPZUS---A--TAAG-T-----G-----G---TC	7496
SIVSTAN1GTAGATGTAGAAGGTAATAATTGCACCACTGGAAAGGATAAACCGTGCTACAAAAGG-CATATG---C--CAT--T-GGTC---G-C--TG	7614
SIVGRI677GGTACTAATAGGACCTGTGACAAAAGGGAAGCCAGGACCAGGACCATGTGTTTCAGAGA-CITATG-TG-C---CAT--CG--G-AG---TG	7343
SIVVER9063AATCAATGTAGTAATGAAACAAAAAAGGGAATGCCCCAGGACCATGTGTACAAAGG-CITATG-TG-T---CAC--C-G-TCTG-C---TG	7654
SIVVER155AATCATTGTAAAAACAACCGAGGAAAGGTCGAAGTCCAGTCCCTGTGTACAGAGA-CITATG-TG-C---CAT--CCG-TCTG-C---TG	7634
SIVVERTYOCATAATTTTGTAGCAGCAAAAAGAAAGGACACCGCACCAGGACCATGTGTACAAAGG-CGTATG-TG-T-A-CAT--C-GGTCG-A---TG	7123
SIVVERAGM3AATCCGTGTAATGGTACGAAGGAAAGGTAAGGCACCAGGACCCCTGTGCACAAAGA-CATATG-TG-T---CAT--CG-TCTG-C--T--TG	7152
SIVLHOESTTATCCTGGGC-AAC---CAC---AGC-C--TG-TG	7731
SIVSUNTATCCTGGGC-AAT---CAC---AGTG-G--G-TG	7799
SIVMNDGB1	AT-ATG-TG-TG--ATGTCA-G---A-TGTTCTCTGTAATGTAACAAGAATCATGAGGGCATGGAATGATCCTAATGAGAAGAAGTGGTATCCTTATGC-T---TCA---T-GG---AG--G-TG	7089	
SIVSAB1CGACCATAACAACCTGTGCAAAAATAATACAAAGCCATGTGGCAGAGA-C-TATG-T-C---CAC---GG--G-AG---TG	7810
SIVSYK173TATGCCAGTAACC-TCGT--GT---TGCC---TCGC---A-T--TG	7470
ENVY_A_S_N_L_R_L_S_C_A_I_R_Q_I_I_N_	ENV

B_FR.HXB2R	TGTGGCAGAAAGTAGGAAAAGCAATGTATGCCCTCCCATCAGTGGG.....CAAATTAGATGTTTCATCAAATATTACAGGGCTGCTATTAAACAAGAGATGGTGGTAAT.....AG	7606
ENV	M_W_Q_K_V_G_K_A_M_Y_A_P_P_I_S_G.....Q_I_R_C_S_S_N_I_T_G_L_L_L_T_R_D_G_G_N.....S	ENV
A_UG.U455	-----G-----C-----CAA-----GT--A--G--GA-----C-----A--A-----G---C-AAC.....	7037
C_ET.ETH2220	-----GG-----CG-----TGAA-----AT--A-TG--AG-----C-----A--A--C-G--C-T-----A-CC--A.....GAGCC	6985
D_ZR.84ZR085	-----G-----G-----TGAA--G-----AT-----AT-----A-----A-----GC-----GC-----A	7127
FL_BE.VI850	-----GG-----CG-----A--T-----TGCA-----A-C--CC--AAC-----TT-A--G-----	6884
G_SE.SE6165	-----GG--G--C-----GCA-----A-C--GA--AAC--GC-----A--AA-----G---AAT...AACAAATACA-A	7024
H_CF.90CF056	-----G-----CG-----CAA-----A-C--TG--GT-----A--AA--G--TT--C-AG--C.....GGGAATGGC--	6925
J_SE.SE92809	-----A-GGAC--C-----C-----TGCA-----A-C--CG--CA-----AT-A--G--C-T-----A--A--CAGG...GGGAATGGC--	6913
CRF01_AE_TH.CM240	-----GG--C--C-----T-----AG--AT--GT-----AA-A--G-----GT-AAT.....	7160
CRF02_AG_NG.IBNG	-----GG--C--C-----CAA-----AT--A--GAT-----C-----A--A-----A--G-----	7113
CRF03_AB_RU.KAL153	-----G-----G-----GCA-----GCA-----A-----A-----C.....CA	7602
CRF04_cpx_CY.94CY03	-----G-----G--GC-----C--T-C-----TGCA--G-----AGC--AC--CAGC--G-----AA-AA--G-----C.....	6980
N_CM.YBF30	-----ACA-GG-----G--T-----A--A--A--TC--G-----GTTC--AC--A-C-C-----T--A--G-TC--GA-TATAG-----GCC.....	7152
O_MVP5180	CA--AT--G--G--AGTCGAG--C--A-----CCC--C-----A-CT-A-C--CAT--C--C--A--T--AA--A-TC--CAGTT--CAACCATGG.....-A	7662
O_CM.ANT70	CA--ATA-GG-G-CAGTCG-G-C-C--A-----AA--T.....A-TC-A-C--ATG-----C--A--T--AA--A-CC--CA--TG--AACACATGG.....AAC--	7647
HIV2_A_ST	CC--C--G--G--AATG-A--TTG--T-GGGAA-----GT--A-CC--CAAC--T-CAG-G--CA-CA-AA-TGCT-AC-TT--C--A--AG-G.....	7516
HIV2_A_BEN	CC--T--T--G--AATG-A--TTG--T-GGGAA--G-----G-GT-GGCC--GA--CAG-A--CA-CA-AA-TGCT-AC-TT--CATA-A--AAAT.....	8083
HIV2_A_ALI	-C--C--C--G--AATG-A--TTG--T-GGGAA-----G-GT-G-CC--CAAC--CAG-A--CA-CT--A-TGCT-AC-TT--CAGC-A-GGC.....	8070
HIV2_B_D205	C--C--C--G--T--G--AAT--A--TTG--A-GAGAA--T-----G-GC-CTC--AAT--C-C-G-G--TA-C--CA--GCC-AC-TCA-CTCA-AC.....	8059
HIV2_B_EHO	C-----C--GA--T--G--AATG--TTG--A-GAGAG--T-----G--C-CTCC--AAT--C-C-G--CA-C--CA--GCC-AC-TT--CT-GATAG--AAG.....	8048
STMM83293	CA--T--T--G--AATG--T--TTG--A-GACAA-----G-CC--CT--AAT--C-C-G-G--CA-TA-AA--GC--AC-TT--CT-GAC--CAAT.....	7682
SMM251	CT--T--T--C--AATG--T--TTG--A-GAGAG-----G-CC-C-CG--AAC--C-CAG-G--CA-T--CA--GC--AC-T--T-GAC-G--GGA.....	8010
SMM9	C-----C-----C--AATG--R-TTG--T-GGGAA-----G-CC-G-C--AAC--C-C-G-A--TA-C--CA--GC-GA--T--T-GATC--GGC.....	7517
SIVCPZANT	AT--GGC-T--TTC-----GC--T--CTTAG-C--T-GG--A--G-----A-TG--TCC--A-T--CAT-GC--A--T--AA-TA-G--GGA-G-TC--AATATA-----	7269
SIVCPZGAB	CA--AT--GG-----G--G--A--A--C--A-----A-C--C-CT--CAAT--T--C-----T--A--T--C-----TTC--CAC-CC-GTG.....ACT-A	7628
SIVCPZUS	A--AT--G-----G--C-T-----A--A--AA--T.....GTGC--A--T--AAC-----A--G--AA--A-TC-TGAC-T-AGCATAA-CGCA.....GTCAAT-A	7608
SIVSTAN1	AC--T-C-CGC-CA-T--GAA--CC-----A--A--A-GGGAA-----CT-AGA--CA-----CAG-G--TCTA--A-GG--T--CTG--TA-AAC-GCAAA.....	7721
SIVGRI677	AT--T-C-CT--CTCT--AAGG-A-----T--A--A-GGGAA--T-----TT-GGAG--AAC--TCAG-C--G-CA--ATACG-GG--T--TA-AAC--CAAG.....	7450
SIVVER9063	AT--T-T-CCA--TC--GAA--G-----T--A--A-GGGAA-----TC-GGA--CA-----C-CGG-C--T--TA--AC-G--GA-CT-A--TA-AAC--AAC.....	7761
SIVVER155	AT--T-T-CTA--TC--GAA--CA-----T--A--A-GAGAA-----TT-GCAG--CA-G--C-CAG--T--A--AC-G--GAGCT-A-CTA-AA--CCAG.....	7741
SIVVERTYO	AT--CCT-T-CCC--TC--GAA--CT-----T--G--A-GAGAA-----TT-GCA--CAGG--C-CAG-C--T--A--AC-G--TGAGCTTA--TA-AA--G-AAA.....	7230
SIVVERAGM3	AT--T-C-C-C-TC--GGAA--CC-----A--G--A-GAGAA--G-----CT-GCA--CA-----C-CGG-A--G--TA--TC--G-GGAGCTA--TACAA--G-AAG.....	7259
SIVLHOEST	AT--GCT-C-A-----GAAG--T--CTA--A--T-CTCA--TTTCAACAATAG--C-----A-TCAT-GAG-A--AAA--T-C-TGAG-TG--AAAAT-GG-ACCACATGAGGACTTAG--	7861
SIVSUN	AT--GC-TC--G--C--GAAG--T--CTTG--C--A-CATCA--GTTAATAACAG--A-----G--AAC-GAG-G--AAGCATGG--C--CTG--AA-G-TGG-AGACTGGAGGTAATG--	7929
SIVMNDGB1	AC--AT--C-----G--AAG--A--TTA--A--T-CATCA--TTTAATAAT--C--A--G--A--CAT-GGG-A--AAA--TAC--TGA--TGC--AAAGATAG--AGTAATGAACAAAA..	7217
SIVSAB1C	AT--T-C-CCT-GTCT--AAG-CA--C-----GAGAA--C-----CT-GGAG--AAC--G-CAGCC--C--TATG--GA-TTGA-CTA-AAC-G-AAA.....	7917
SIVSYK173	AT--AGAT-T--A--TTG--A--CCTG--C--T-CGGCA--G.....T--A--A-T--C--G-C--T-CAG-AT-GACGGAC-T--ATACTACCCAGGA.....	7577
ENV	D_W_R_Y_V_R_K_L_I_Y_L_P_P_T_A_G.....H_I_K_C_T_S_N_V_T_A_V_L_T_D_I_E_Y_Y_P_G.....	ENV

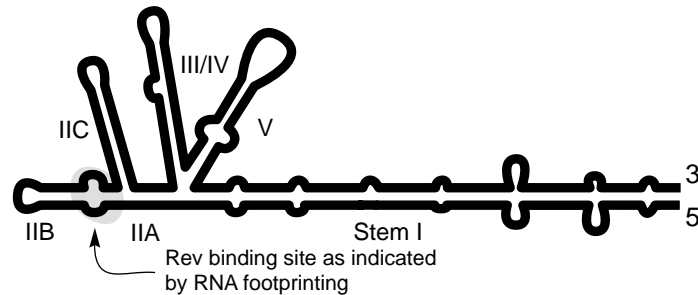
Accession	Sequence	Position
B_FR.HXB2R	CAACAATGAGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTTATATAAATAAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCA.....	7727
ENV	N N E S E I F R P G G G D M R D N W R S E L Y K Y K V V K I E P L G V A P T K A	ENV
A_UG.U455	.-T-CAA-AAAT---C-----A-----G-----C-----G-----	7157
C_ET.ETH2220	AC-T-GCACAAAA-A-----A-----G-G-A-G-C-----T--C-----	7106
D_ZR.84ZR085	T--T-C-C--AAT-C-C--T-----A-C-----C-----	7248
FL_BE.VI850	.G-G-G-A-CATT---C-----AG---A---AA---C---G---G---G---	7004
G_SE.SE6165	T-CA-G---AGT-----G---G---AC---CA-T-C-----G-----	7145
H_CF.90CF056	GTCTGCA--AAATT-T-C-----G-----C-G-A-----A-----	7046
J_SE.SE92809	TG-G--GCA-T---C-----AC---A---A-----G-G---G---C---A-----	7034
CRF01_AE_TH.CM240	.-T-CG--TAA---C-----A-C-A-A-----C-----C---A-----G-----	7280
CRF02_AG_NG.IBNG	.-T-G-ACAAAT--A-C-----G-----G-----C---T-----CGT-----	7233
CRF03_AB_RU.KAL153	G-G---TTA-----T-T---A-----T-----G-----	7723
CRF04_cpx_CY.94CY03A-TA-T-----C-----C-----G-----A-----AT-----	7097
N_CM.YBF30	.G---CCA--GAAACA--AG-ATAT--CTC-----A-C---GTTA-TCTC-----CAA-G-G---G-C-----T-GC-A---CA-----AGGT--A--T-----	7272
O_MVP5180	TTC-CA-GTGAAA-T-CAC-T---A-T---G-----AA-T-TA-----C-A---G-C-C-C-----C-G-AA--T-TA-T-----T--A--AATGTCAGACCA	7792
O_CM.ANT70	--G---CA-CAAT-TA-CA-T---AAT---G---C---AA-T-TA-----G-TC-C-C-----GGG-AA--T-TA-T-G-----ACGTATTGCAAGGCCA	7777
HIV2_A_STAA-C--CAAAT-TTA-CTTTA-T-C--GG-GCA-ACTA-ACC-TTA-----GGGGG-T---T-GA-G-G-AAC--A-T-CT-T---T-ACCA-T-----	7627
HIV2_A_BENCG-A-TC-T-C-AA--TTA-CTTTA-T-C--AG-GCA-ACTG-ACC-TTA--C-GGGAG-C-C---T-A-G---AAC--A-T-CT-C---T-AG-TCAG-----	8197
HIV2_A_ALIAA-C--CAAAT-TTA-CTTTA-T-C--GG-GCA-ACTA-ACC-TTA-----GGGGG-T---T---G-G-AAC--A-T-CT-C---T-ATCA-A-----	8181
HIV2_B_D205A-CAGTACC-C-AAT-TTT--TGA-T-C--AG-TCA-ACTG-AT--TTG--C-GGGGG-C---T---G---AAC--A-T-CT-C---T-AG-T-T-----	8173
HIV2_B_EHOAA-CTT-CTAAT-TTA--TGA-T-C--AG-TCA-ACTG-AT-A-TTG--C-GGGGG-C-C---T---G-G-AAC--A-T-CT-T---T-A-GTAT-----	8159
STMM83293AAT---CTAA--TCA--C-A-T-C--GG-GCA-ACTG-ATC--TTG--G-GGGAG-T-C---T---G---AACT--A-T-CT-G-C-A--G--T-TG-----	7793
SMM251AA-C-A-CTAGT-TCA-CATGA-T-C--GG-GCA-ACTG-ATC--TTG--G-GGGAG-T---T---G-G-CACT--GA-T-CT-G-C---AG-T-TG-----	8121
SMM9AATA--C-AAT-TCA-CATGA-T-C--GG-GCA-ACTG-ATC--TTG--G-GGGAG-T-C---T---G-G-ACT--A-T-CT-C-C---A-GT-T-----	7628
SIVCPZANTGAAACTG-TAAAGTGT-ACCT-CT-C-AGAG-AGCA--C-A-----GCG-G-G-CC-GG-CC-G-G---G-G---RT-C-GTC---C-A-A-CA-GN-----	7380
SIVCPZGABG-GTAA-CT-CA-T---AC-----A-C---AA-T-T-----GC-C---G---TCG--A-G-T---TC---G---A-A-A---AGAAGGCAT-----	7758
SIVCPZUS	-G-T-G-AG-AATATA-CAG-G-TG--AC-----C-CTTTA---A-A---C---G---G-G-C-GC-A---TA-----G---AGGT---C-----	7729
SIVSTAN1GAAAG--CTAATGTGA-ATTGAC--CTA-T-GA-A---TA--GCTTA--GGGA-G---GC-CA-G---A-C-TT-C-G-A--AG--AT-----	7832
SIVGRI677TCTGG-CCA--AAATGTGA-CCT-A-TCCTC-GG-AC-CAG-TA---GCGTAC--C-GGGAG-C---T---G-G-AAC--A-T-CT-T-T-T-AG-T-T-----	7564
SIVVER9063-ATAGA-C-AATGTGA-ATT-A-TCCCC-G-AGAA-GA-TC---GC-GCA--GGGAGC---T-G-G---CAC--A-C---T-T-----G--T-T-----	7872
SIVVER155AA-AG--CAAATGT-A-ATTGA-TCCCC-G-AGAAAC--TC---GCGGCA--GGGC-G--C-T-G---G-G---AC--A-T--T-T-----AG-A-TC-----	7852
SIVVERTYOAA-AGA-C-AA-GTGA-ACT-A-TCCCC-G-AGAACT-TC---GCGGC---GGGC-G--C-T---GG---CAC--A-T-CT-T-----AG-A-T-----	7341
SIVVERAGM3AA-AG--CTAATGT-A-ATT-A-TCCCC-G-AGAAAC--TC---GC-GCA--GGGC-GG-C-T---G---AC--A-T-CT-C-----AG-A-T-----	7370
SIVLHOEST	AGG--CCTAAGT...--AAAGTTCTTGCCTCC-TC-TGGGAA-CTA-TC-A-TTGTGGCA---GGG-C-----T-GA--T-GA-T-A-T-GT-T-T-T-AG-T-AG-----	7979
SIVSUN	ATCA--CATC--A-TTG-AGC-TTC-AGCC-CCTACCA-C-CCTTAA-TC-A-TTGTG--CAC-GG-GC-A-T-C---C-----G---AGG--A-T-GT-T-T-A--AG-T-AG-----	8050
SIVMNDGB1ATGC-A--TAAATTCCTGCCTCCCA-T--A-CCTCCA-TC-A-TTGTTC-T-TGG-GC-C-T---T---C-----AATG--A-T-CA-----T-AG-T-TG-----	7328
SIVSAB1CAA-AG--C-AATGTGA-GTT-TC-CCTC-G-TGAAAG--T---GCT-A---GGGAG-T-C---T---GG---CA-G--A-T--T-T-T-A--A--T-----	8028
SIVSYK173-TAC-T-GAA-TTTA--CC-AC--C-A--G--GA--TGT-----GCA--TC-G-TC-T---T-GA-TC---A---A-T--CT-T-G---TG-TCA-----	7688
ENVS T L N F T P T A N V E D V W R A D L F N Y K L I Q I K P I G F A P T D Q	ENV

Rev responsive element secondary structure in RNA, see Charpentier, J Mol Evol 266:950-956 (1997)
see HIV-1 complete genome annotation in Compendium for stem-loop positions

B_FR.HXB2R	...AAGAGAAGAGTGGTGCAG...AGAGAAAAAGAGCA...GTGGGA...ATAGGAGCTTGTTCCTTGGTTCCTGGGAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGG	7845
ENV	...K R R V V Q...R E K R A...V G...I G A L F L G F L G A A G S T M G A A S M T L T V Q	ENV
A_UG.U455	...-T---C-----A-C-----A-----T-----A-----G-G---A-----	7275
C_ET.ETH2220	...-A-G-----G-----GCA.....C-----A-----A-----A-----	7221
D_ZR.84ZR085	...-G-----G-A.....A-A.....C-----A-----G-----G-----	7366
FL_BE.VI850	...-A-CA-----C-----C-----TC-T-G-ACAGCA-GGAGCA-----G-G-G-A-----	7122
G_SE.SE6165	...-G-----G-----C-G-----G-C-----A-----G-G-A-----	7263
H_CF.90CF056	...-G-----G-----G-----CT-----A-----G-G-A-----	7164
J_SE.SE92809	...-A-----G-----G-----G-----A-----T-----G-A-----	7152
CRF01_AE_TH.CM240	...-G-----G-----G-----A-A-T-----A-----G-G-A-----	7398
CRF02_AG_NG.IBNG	...-A-----G-----G-----T-----C-G-----G-C-----A-----GCGG-A-----	7351
CRF03_AB_RU.KAL153	...-G-----G-----G-----G-----G-----G-G-A-----	7841
CRF04_cpx_CY.94CY03	...-G-C-----G-----T-----G-CA-----TC-T-----G-----G-----	7215
N_CM.YBF30	...-A-C-CACA---AGT-----GCCT-T---C---T-GC---T---TC-T-----G-----G-A-----	7393
O_MVP5180	ATA-TA-AC-TTCACACC-CTCAC-G-----A-----T-G---ATGC-A---T-G---G-GC-AA-T---T-----T-----GG---CAG-----G-A-----	7916
O_CM.AN770	GTC-TA-C-CTAGAACT--T-----A-----T-G---ATGC-A---T-G---G-GC-AA-T---T-----T-----GG---CA-----G-----A-----	7898
HIV2_A_ST	...-A---TACTCCTCT...GCTCC--TG-GG-ATAA-AGA-GT-T-TTC...-TGCTAGG---T-A-T-TC-CACGA-----GCTG-A-----G-G-CT---T---CT---T	7748
HIV2_A_BEN	...-GA---GTACTCCTCA...ACTCC--TG-GG-ACAA-AGA-GT-TGTTT...-TGCTAGG---T-G-T-TC-C-CGA-----TTCTG-A-----GCGG--CC-----T-A-CC---T	8318
HIV2_A_ALI	...-G---TACTCCTCT...ACTCC-AGG-GG-ATAA-AGA-GT-TGTTT...-TGCTAGG---T-A-T-TC-C-CGA-----TTCTG-A-----A-G-AG-TT-A-----T-T-CT---T	8302
HIV2_B_D205	...-GA---TACTCCTCA...GTA-A-CCG-GG-ATAA-AGA-GT-T...-G-TGCTAGGA---T-G-A-C-T-C-ATG-----TTCTG-A-----A-G-TT-----T-A-CT---T	8294
HIV2_B_EHO	...-A---TATTCTCTCA...GTG-C-CCG-GG-ATAA-AGA-GT-T...-C-T-TGCTAGG---TT-G-A-C-T-CGA-G-----TTCTG-A-----G-G-CT---T---CT---T	8280
SMM83293	...---TACACAAC...AGT-CCTC-G-CTAAGAGA-G-TCFTT...-TGCTAGG---T-G-T-TC-C-CGA-----TTCTG-A-----G-GC-----CT---T	7914
SMM251	...---GTACACTACTGGTGGC-CCTC-G-ATAA-AGA-G-TCFTT...-TGCTAGG---T-G-T-TC-C-C-A-G-----TTCTG-A-----G-G-GT-----C-CT---T	8245
SMM9	...-A---GTACACCACAACCTGGTGCCTC-G-ATAA-AGA-G-TCFTT...-TGCTAGG---T-G-T-TC-C-CGA-----TTCTG-A-----G-G-GT-----T---CT---T	7752
SIVCPZANT	...-A---GCC---AAA-AA-ACAACACTCC-G-CA-AA-AGA-GCATTGGA---GCTG-C---T-G-TC-TC-CA-T---T-C-T-A-----G---AG-----C---A	7507
SIVCPZGAB	ACAGT-GC---CA-AAAG-C...-C-G-----GCCT-C---T...-C-G---AC-----C---T-----GG-G-A-----	7882
SIVCPZUS	...-A---GCATACA---A-A-----GCCT-C---C---T-AC-----G---TC-T-----A---G-AGT-----	7850
SIVSTAN1	...-GA---TAT---T-G---CAACTC-----G-G-----CCCTTC---TGTTGGG---A-C-T---G---T---GCTG-A---T---A-AG-G-CAG---T-----C---T	7950
SIVGRI677	...-GA---TATACT-GC-CCACA-----G-----CC-TTC...-TGCTAGG---T-A-C-----T---T---CTG-A-----GG---CA-----A-C---T	7682
SIVVER9063	...-GAC-TTATAC-GA...GGACA---C-G-TC-AA-CGA---CCCGTTT...-TGCTAGG---A-C---A---T---T---CTG-A-----A---GG-G-CAG-C-----C---T	7993
SIVVER155	...-G-C---TACAC-GA...GGCCA---G-GGCA-AA-CGA---CCCGTTC...-TGCTAGG---A-C-----T---T---G-CTG-A-----A---GG-G-CAG-C-----C---T	7973
SIVVERTYO	...-G-C-TTATAC-GA...GGACAT-G-GGCA-AAGAGA---CCCGTTT...-TGCTAGG---T-A-C-----G---T---T---CTG-A-----A---GG-G-GC-GC-----C---T	7462
SIVVERAGM3	...-GA---GTATAC-GA...GGTCAT-C-G-C-AAGCGA---CCCGTTC...-TGCTAGG---A-C---A---T---T---G-CTG-A-----A---GG-G-CAG-C-----C---T	7491
SIVLHOEST	...-C-C---GTAC-CACCC...-G-G-GGCA-A-CAGA-CA-CGCTC-G-CTTTAGGAGC-T-G-AC-TC-CA-T-T---C-CTG-A---TT-TA-G-C-TA-A-T-C---T	8100
SIVSUN	...-C-CC-GTAC-CACCC...-AGG-GGA-AAGAGG-CA-C-CCTG-T-CCCTAGGAGCTT-G-CAC-TC-CA-T-T-G-T-CTG-A---TT-TA-G-C-TA-A-T-C---T	8168
SIVMNDGB1	...-A---CACACTT-A-CTGAACATC-T---GAGAAGAGA-GA-C---G---AT-CT-GGTA---TC-GC-CTCGCTG---TC-G-G---T---TG-GG-G-A---T---C---AT	7452
SIVSAB1C	...-GA---TACTACT-GC...CCC---G-CA-AA-AGG---CC-TTC...-TGCTAGG---A-T-----GCTG-A---T---GG-G-CAG-C-----C---T	8146
SIVSYK173	...CGTC-GTAT-A-T---CT...ATACC-GGGA-AA-AGG-CA-C-CCAT-G-C-CT-GG---G-A-AC-CTCG-T-G-C-CTG-A---G-GC---TG-G-CAG---A-AC-C---T	7812
ENV	...R R Y E L P...N T R E K R A A P L A L G F L G L L S A A G T A M G G A A T A L T L Q	ENV

RRE Structure

after B. Charpentier,
J. Mol. Evol., 266:950-62 (1997)



Rev responsive element secondary structure in RNA, see Charpentier, J Mol Evol 266:950-956 (1997)
see HIV-1 complete genome annotation in Compendium for stem-loop positions

B_FR.HXB2R	CCAGACAATTATTGCTGGTATAGTGCAGCAGCAGAACAAATTGCTGAGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGA	7975
ENV	A_R_Q_L_L_S_G_I_V_Q_Q_Q_N_N_L_L_R_A_I_E_A_Q_Q_H_L_L_Q_L_T_V_W_G_I_K_Q_L_Q_A_R_I_L_A_V_E	ENV
A_UG.U455	-----A-----G---C-----A---T-----A-----T-----T-A-----G-----	7405
C_ET.ETH2220	-----G-----A-----A-G-----AA---A-----A-----G-----G-----T-----A-----G-----A-A--	7351
D_ZR.84ZR085	-----G-----A-----A-----A-----A-----A-----G-----G-----T-----A-----G-----A--	7496
FI_BE.VI850	-----A-----A-----A-----A-----A-----G-----G-----T-----A-----G-----	7252
G_SE.SE6165	T-----C-----A-----AGG C-----A-----G-----G-----T-----A-----C-----A--	7393
H_CF.90CF056	-----G-----A-----A-----A-----AC-----G-----A-----G-----G-----T-----A-----G-----C-----	7294
J_SE.SE92809	T-----G-----C-----A-----A-----A-----A-----A-----T-----A-----G-----T-----A-----G-----	7282
CRF01_AE_TH.CM240	-----A-----A-----A-----A-----A-----G-----A-----A-----G-----T-----A-----G-----	7528
CRF02_AG_NG.IBNG	-----C-----C-----A-----G-----AA---A-----T-----A-----A-----G-----T-----A-----G-----C-----	7481
CRF03_AB_RU.KAL153	-----A-----T-----C-----A-----A-----A-----A-----A-----G-----T-----A-----G-----	7971
CRF04_cpx_CY.94CY03	-----C-----C-----G-----A-----T-----A-----A-----AG-----G-----T-----A-----G-----C-----	7345
N_CM.YBF30	--C-GAC---A---G---A---T-T-C-T---A-A-A---A---T---A---T---T-A---T---A---A---G---T---A-A---	7523
O_MVP5180	--CACAGTG--C--AAG---A---G---CC---A-G-AC---C-G-A-CT-C--AGGT-AT-T-A---T-T-GA-A---GA--TC-CC-G-AA-CT-A---	8046
O_CM.ANT70	--CACACT--GC--AAG---A---G---CC---A-A-A-AC---C-G---AT-C--AGG-AT-TR-A---T---GA-A---GA--TC-CC-G-A-CT-A---	8028
HIV2_A_ST	--TC-GACT---G-C-G---A---C-AC-GC--T--GAC-TGG-CA--AGA---AG-AA---G--G-C---A-CA-AA-T---G--ACT--A-C---	7878
HIV2_A_BEN	--C-GACT---C-G-C-G---A---C-AC-GC--T--GAC-TAG-CA--AGA---AG-AA---G--G-C---A-CG-AA-C---G--ACT--A-C---	8448
HIV2_A_ALI	--TC-GACT---G-C-G---A---C-AC-GC--T--GAC-TGG-CA--AGA---G-AA---G--G-C---A-CG-AA-T---G--ACT--A-C---	8432
HIV2_B_D205	--C-GACT---C-G-C-G---A---C-AC-GCCTG--GAC-TGG-CA--AAGA---G-A---GG-G-C---G-CG-AA-C---G--ACT--CA-C---	8424
HIV2_B_EHO	--C-GACT---C-G-C-G---A---C-AC-GC-CG--GAC-TGG-CA--AAGA---AG-A---GG-G-C---G-CG-AA-C---G--ACT--CA-C---	8410
STMM83293	--C-GACT---A---G---A---C-AC-GC--T--GAC-TGG-CA--AGA---AG-AT---G--G-C---A-C---A-C---A-T---G--ACT--A-C---	8044
SMM251	--C-GACT---G---G---A---C-AC-GC--T--GAC-TGG-CA--AGA---AG-AT---G--G-C---A-CA--A-C---A-T--GG-ACT--CA-C---	8375
SMM9	--C-GACT--G--G--G---A---C-AC-GC--T--GAT-TGG-CA--AGA---AG-AT---G--G-C---A-CT--A-C---A-T--G--ACT--A-C---	7882
SIVCPZANT	---GA-T--GY-CCA---T-A-A---GC---C---CAA-C-A-A-A-G-A---C-A-G---T-G---AG-A-A-A---G-T-A-C---	7637
SIVCPZGAB	--C---G--GC---G---A---T---C---T---AA-A-A---T---C---A---AT---A---AG-A-A-A-A-A-C---T-G--T---A---	8012
SIVCPZUS	--C---A---G---A---T---C---A-A-A-A---C---T---A---AT---T---T---A---T---A-C---G-A-T---	7980
SIVSTAN1	--CAG---C-TG-A-G---T---G---C---GC---AG---CA---G---GA-C-AA-G---CA-T---TG-G-AA-C---A-T--CC-CG--ACT--CC-T---	8080
SIVGRI677	--TC-G--T--GC-TG---G---T---G---C---GC---G-G-ACA---A---GT---A-G-G-CA-T---TG-G-AA-C---A-T--CC-CG--ACA---C---	7812
SIVVER9063	--TCAG--T--C-TG---G---A---C---GC---G-G---T---AA---A-G-G-CA-T---TG-G-AA-C---A-T--CC-CG--ACA---C-T---	8123
SIVVER155	--TCAG--T--C-TG---G---T---G---C---GC---G-G-GA---T---GA---A-G-G-CA-T---TG-G-AA-C---A-T--CC-CG--ACA---C-T---	8103
SIVVERTYO	--C-G--T--GC-TG---G---C---GC---G-G---T---GA---A-G-G-CA-T---TG-T-AA-C---A-T--CC-CG--ACA---CC-T---	7592
SIVVERAGM3	--TCAG--T--C-TG---G---C---GC---G-G---T---GA---A-G-G-CA-T---TG-G-AA-C---A-T--CC-CG--ACA---C-T---	7621
SIVLHOEST	--CA-GC-G-C---CAA-G---T---GC-G---GT-CT-G---AAA-G-AG-A---AAG---CA-A---AG-G--A-T---A--C-CC-GACA--CC-T---	8230
SIVSUN	--CA-GC-G-C---CAA-G---T---GC-G---GT-CT-G---AA---G-AG-G-AC-AAG---CA-A---AG-G--A-T---A--C-CC-GACA--CC-A---	8298
SIVMNDGB1	--TCAGTCT--GG-A---G---G-A-A-A-AC-G---T---A-CTC-A---CAA-GTCTG-A-C-AA---CA-A---AG-A--A-TT-A---A-TC-CC-GACCAG-T---	7582
SIVSAB1C	--TCAG--G--C-CG---G---C---A---C---TGC---G-A---CA---G---AA-C-AA-G---TA-T---TG-A-AA-C---A-T--CC-CG--ACT--C-T---	8276
SIVSYK173	--TCAGACT--G---G---G---C-A-G---GA---AG---A-G-A-C-C--AGGCT-A-T---TG---AA-C-TA-C--CC-GC--ACA---T-A---	7942
ENV	S_Q_T_L_L_A_G_I_V_Q_Q_Q_Q_K_L_L_E_A_V_E_A_Q_Q_H_L_L_G_L_T_V_W_G_V_K_N_L_N_A_R_L_T_A_L_E	ENV

Rev responsive element secondary structure in RNA, see Charpentier, J Mol Evol 266:950-956 (1997)
see HIV-1 complete genome annotation in Compendium for stem-loop positions

Rev responsive element end

B_FR.HXB2R	AAGATACCTAAAGGATCAACAGCTCCTGGGGATTGGGGTTGCTCTGGAAAACCTATTGACCACTGCTGTGCCTTGGAAATGCTAGTTGGAGTAATAAATCT.....	8078
ENV	<u> R Y L K D Q Q L L G I W G C S G K L I C T T A V P W N A S W S N K S </u>	ENV
A_UG.U455	-----C-----A-A-----C-----G-----A-----C-----CT-----	7508
C_ET.ETH2220	-----G-----A-----C-----C-----A-----C-----CT-----G-----	7454
D_ZR.84ZR085	-----G-----A A-----C-----A-----A C-----CT-----G-----	7599
FI_BE.VI850	-----A-----C-----C-----C-----AA-----C-----CT-----	7355
G_SE.SE6165	-----A-----C-----C-----C-----AA-----C T-----CT-----	7496
H_CF.90CF056	-----G-----C-----C-----C-----AA-----CT-----A-----	7397
J_SE.SE92809	-----A-----A-----C-----C-----AA-----C-----	7385
CRF01_AE_TH.CM240	-----A T-----A AC-----C-----A-----C-----CT C C-----G-----	7631
CRF02_AG_NG.IBNG	-----GA-----A A-----C-----C-----A A C-----CT C-----A-----	7584
CRF03_AB_RU.KAL153	G-----A-----C-----C-----A-----C-----	8074
CRF04_cpx_CY.94CY03	-----C-----A A-----C-----C-----AA-----A-----CT-----	7448
N_CM.YBF30	-----T G-----G AA-----AA TC A-----C-----A-----ACA A-----TAT CA-----AG C-----C C-----TA CTCT.....	7629
O_MVP5180	--CCCTTA--C--A--G--A--G--AAACC--A--C--TAAA-----A--C--TTA--AT--A--AAAA-----CA--ATCA--TCAGGA.....AGATATAAT	8152
O_CM.ANT70	--CC TA--C--A--G--A--AA CC A-----C--TAAA-----G--AG C-----TA--AT--A--AAAA-----AGA CA--TAGGA.....	8125
HIV2_A_ST	G A T-----C-----GGC AA-----AAATTC A-----TG GTTT G AAG C-----CA A-----A A-----GTAAA GACACC.....	7969
HIV2_A_BEN	G AG-----C-----GGCA AG-----AAATTC A-----TG GTTT G AGG C-----CA A-----A G-----GTAAA GAC CT.....	8539
HIV2_A_ALI	G A T-----C-----GGC GG-----AAATTC A-----TG ATTT G AAG C-----CA A-----A A-----GTAAA AC CC.....	8523
HIV2_B_D205	G A-----C-----GGCA AA-----AAATTC A-----TG TTC G AGG C-----TCA GA-----A A-----CCAAA GAAACC.....	8515
HIV2_B_EHO	G A T C A C-----GGCA AA-----AAATTC A-----TG TTC G AGG C-----TCA GA-----A A-----GTAAA GAA CC.....	8501
STMM83293	G A-----C-----GGCA AG-----AAATTC A-----TG GTTT G AGG C-----CA CA-----A A-----CCAAA GA CT.....	8135
SMM251	G AG T-----C-----GGC AG-----AAATGC A-----TG GTTT G AAG C-----CA A-----A A-----CCAAA GCAA T.....	8466
SMM9	G AG-----C-----GGC CRN-----AAATTC A-----TG TTT GG AGG C-----CA A K A A-----CCAAA GACACA.....	7973
SIVCPZANT	G AG-----GA-----A AT A CC C-----TG AC GG G CC TCA-----A G-----AA TCC-----GTA CTTCA GCAAA CATGTGCAAAGA CAGCAGTGAT	7767
SIVCPZGAB	---G---GC---G A T---CC G-----C-----GGCTG C-----TTAT CA G-----CAACTC-----CC GGG GCAA TCC.....	8118
SIVCPZUS	---T---G A---A---TC G-----C-----A-----ACA-----TTAT CA-----A CC-----C CCTCTCC.....	8086
SIVSTAN1	--A-----G-----GACA GG-----AAATT G-----A TG ATTC-----AGG G TCA-----A G-----G-----C TTC ACAAT.....	8171
SIVGRI677	G AG-----G-----GGCA GG-----AAATTC A-----TG GT G-----AAG A TCA-----CA A-----A-----GTA A AAC.....	7903
SIVVER9063	G AG-----G-----GGCA GAT-----AACG C-----G G AT G-----AGG C-----TCAT CA A A A-----C ATGGCAAAAC.....	8214
SIVVER155	G AG-----GGC-----GGCA GGT-----AAACGC-----G G GT G-----AAG A TCAT AA A A C-----CGTGG A AAT.....	8194
SIVVERTYO	G AG-----G-----GGCA GA-----AAACTCC-----G G AT G-----AAG A TCAT CA A-----GAG-----CCCTGG CAAAT.....	7683
SIVVERAGM3	G AG-----G-----GGC GGT-----AATGC-----G G AT G-----AAG C-----TCAT AA C A G-----C GTGG A AAT.....	7712
SIVLHOEST	GGAG TG G AC-----GC-----T CCTC-----G TCAAT G-----AGG-----TCAT AA-----GAG-----CATA A-----	8318
SIVSUN	GGAG TG GC A C-----AGC TT A CTTC-----G TCAAT G-----AAG A-----CAT AA-----C-----TA A-----	8386
SIVMNDGB1	G AT TA C-----C-----GCTT G-----TCTCAA-----G T AT GGC AGG G-----TCAT T-----AGAG-----C AA CAA C.....	7673
SIVSAB1C	G A-----G-----GGCA GGT-----AAC-----A TG TTC GG AGG G-----TCA GA G C T-----GTA A AAC.....	8367
SIVSYK173	G CG T C GA C-----GGCCA T-----TCA A-----A-----G ATTC G AGG C-----CAT A A-----A G-----G GAAAGCA TG A C C-----CAAC.....TTCTGT	8054
ENV	<u> T Y L R D Q A I L S N W G C A F K Q I C H T A V T W E K A C G N N S N </u>	ENV

B_FR.HXB2R ENV	CTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTTAATGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGG L E Q I W N H T T W M E W D R E I N N Y T S L I H S L I E E S O N Q O E K N E Q E L L	8208 ENV
A_UG.U455	-A--GG-C--A-----A--T-----C--C-A----G-A-----GT-G-----G-A-----T-TCAAC-----G-----G-----G-----T---C-----	7638
C_ET.ETH2220	-AA--GG-----G-A--T-----C-----T-----GT-----T--GA-A--T--AAT--GC-----TC-----G--C--A-----A---T---C-A--	7584
D_ZR.84ZR085	G-A--GT-T-----GGGA--T-----C-----A-----G-----GG-----T--AA--G-----T-----T-----A-----A-----	7729
FL_BE.VI850	-A--G-----A--T-----A-AG-G--G-----T--A-A--T--AAA--G-----G--A-----G-----A-----A-----	7485
G_SE.SE6165	TATA-TG-A-----G-A--T--T---A-A-----A-G--C-----T-A-CA--T--AG-C--C-----A--G-----G-----C-----	7626
H_CF.90CF056	-A-AGTG-A--C--G-CA--T--T-----A-----T-A-C-----G-----GAGGA--T--AGG--GC-----TC-----C-----G-----G--C---A--	7527
J_SE.SE92809	TAT--GG-C-----G-GA--T-----AC-A--A--G-----G-AA--T--AGTC-----G-A-----G-----T-----A--C-----	7515
CRF01_AE_TH.CM240	T-T--G-----CA--T--A--A-----G-----G-----A-CA--T--TGAGA--C--AC-----G-----A--G--C--G-----A--G--T--G--A--	7761
CRF02_AG_NG.IBNG	T-TA-TG-C--A--G-A--T-----AC-A--G-A-----G-----GA-A--T--TAATC-----G--G--G-----C--G-----	7714
CRF03_AB_RU.KAL153	--A--TG-----A--A--T-----A-----A-----T-----G-T-----T--AAT--A-----G-----A-----A-----	8204
CRF04_cpx_CY.94CY03	TATA-TG-T--A--G-CA-T-T-----T--C-A--T--A-----CAAA--T--TGGG--C-----A--G-----G-----G-----C-----	7578
N_CM.YBF30	TAT--TACA--C-----A--TTTA-----CAAC-A-----TGAGA--G-A-GA--C--TT--G-TG-C--TTTTGGAC-T--A--C--GG-A--G-A--A--GA-C-CA-----GA--TC-C-C---	7759
O_MVP5180	GAT--CAGT-----G-CA--CTT--A--CA-C-A--CA-C-C-A-----GTA-GCTC-A-T--T--TGATGA--AC--C-G-A--G--A--G-----T-A--C--G---	8282
O_CM.ANT70	AAC--AGC-----G-CAC-TTA--A--CA--A-----TC-GC-G--A--G--CATA-GCTC-ACC--T--TGAGGA--AC--A-GG-A--GTA-----G--C-A-----GA--A-G--GC---	8255
HIV2_A_ST	T-AACG-CTGA-----CA--T--A--CA--A-----GCA-CG--CCG--CCTAGAGGCAAAT--CAGTGAAGTTT--A--C-GG-A--T-----CATGT-T--C--CAAA	8099
HIV2_A_BEN	T-ATCG-CTGAC-----AA-T-T--A--CA-----G-A-C--G-CCG-T-CCTAGAGGCAAAT--CAGTCAAAGTTT--A--G-C-----TT--A-----ATGT-T--C--CAAA	8669
HIV2_A_ALI	T-AA--CTGA--G-CA--T--G--CAA-----ACA-C--G-CCGTT-CCTAGAGGCAAAT--CAGTGAACAGT--A--CGGG-A--TT-----ACGT-T--C--CAAA	8653
HIV2_B_D205	--CAC--CA-A-----A--T--A-----CAAC--A--AGC--G-CC-TTTC-TGGATGCAAAT--ACGG-TC-GT--A--GG-T--G-TA--A--G-----CATGT-T--C--CA-A	8645
HIV2_B_EHO	--TA-G-CAGAC-----A--T--A-----CAAC-A-----G--GC--G-CCG-TTC-TGGATGCAAAT--ACAAAA--C-A-----GG-A--G-TA--A-----CATGT-T--G--CA-A	8631
STMM83293	T--T--C-GA-----CA-T-T--A--CA--A-----G--A-GG--G--TT-CTGAGGCAAAT--ACACAAAC-GT--A-----G-T--GTT--A--G-----ATGT-T--G--GCAAA	8265
SMM251	--AAC--CAGAC-----CA-TGAT--T--CAA-----GC--A-GG--G--TTC-TGGAGGAAAAT--ACAG--C-CC-A-----GG-A-----TT--A-----G-----CATGT-T--C--CAAA	8596
SMM9	T--AC--CT-RC-----CA-T-T--YT--CAA-----A--A-C-GG--G--TTCCTAGAGGCAAAT--ACTSAA-Y-T-G-----G-A-----TT-----M-----ATGT-T--G--GCAAA	8103
SIVCPZANT	A-AC--TGT-----G-AA-T-T--A-----TT-G-AC-G-----CA--G-ACAG--T--TAATA-CT-AC--AT-G-A--TG-G--A-----G-GA--A--A--G-----AT-	7897
SIVCPZGAB	ACA--TG-C-----GGGA-TCTA--A--CA-C-A-----T-A-TT-G-GTCT--C-----G-GAA-----TTTTGTGC-GT--A--G--G-A--TCA--A-----G-A-----AGG--CC-----	8248
SIVCPZUS	TAT--TGCT-----GGCA-TCTA--T--CAA--A-----A--G-A-GA--C--TT--G-TACT--TTTTAGTC-T--A--C--G-A--G-A-----GA-T-CA-----A--TC-C-C---	8216
SIVSTAN1	. . .AC--CAGAC--G-CA-T-T-----CA--A-----G--CC--A-CTGCC-TGGA-G-AAAT--TAGTA-TACTC--TCA--G-AT-TG-G--AG-GC-G--A-----ATGG-TACT-ATCAAA	8298
SIVGRI677	. . .ACT-CT-AG--G-CA-T-T--T-----G--C-----TGCC-TGGA-G-AAC--ACTCAAC--T-G-----G-A-----T--G--TC-----CTGG-TCTG-ACCA-A	8030
SIVVER9063	A--ACG-CC-A-----C-AA--T-----T-----A--C-G--GGAG-G-TGGA-G--AAC--ACAGAACAGT-A-T-A--G-AAG-G-A--AG--G--A-----TT-G-T-CC-ATCAAA	8344
SIVVER155	. . .AC--CAGAG-----A-T-T-----T-----A--C-G--AG-AGGA-TGGAGG--AAC--ACAAAACA-T-G--C-GG-AAGG-A--AG-G-----TTGG-T-CT-ATCAAA	8321
SIVVERTYO	-G-ACT-C-GA-----C-AA-T-T--T-----A--C-----AGCTG--TGGA--AAC--TACGGGACA-T-G--TGA-GG-TAG-G-A--G-G-----CT-G-T-CC-ATCA-A	7813
SIVVERAGM3	AG-ACC-CTGA-----A-T-T--T-----C--A-----A--C-G--ATCGT--TGGA-G-TAAC--ACAA-ACA-T-A--G--G-CAG-GCA--G-G--G-----TTGG-T-C--ACCAAA	7842
SIVLHOEST	A-CAC--CC-A-----CCA-GGAT-----GA--A-----ATC-A-GG-AGCA-T--TGAC-AGAAC--TACT--TC-GT-AC--G-AT-C-C-ACAG-GCT-G-----C-GA-TA--TAA-A	8448
SIVSUN	A-CACG-CC-A-----C-A-GGAC-----A-----T--GC--G-A--A-TG--TGATGATAAT--ACAG-TC-TC-CC-G--GG-AT-TGTAACAG-GTT-G-----C--A-CA-G--CAAA	8516
SIVMNDGB1	A-CACT-CA-A-----CATCAGAA--T-----A--A-----G-C-AG--C-G-TT--CTGCA-CAAAAC--TACAGAAA-GT-AA--C-GG-AT-TG-T-GAG-GC--GA--CAC-T-T-----CA-A	7803
SIVSAB1C	. . .ACT-CCGAC--G-AA--T--G--CAA--A-----G--GC--AG-A--G--TGA-GCTAAC--CAGTAGGA-TT-A--GC--G-A--TG-A--G--C-----TTGG-TAGT-ATCAAA	8494
SIVSYK173	-CCA--CACAG-----AA-T-T--A--CACAGA-----ACA--GG-AG-----TG--TGA-CAT--TG-TGGA--GC-CAG--G-A--G-A--A--G--G-GA-----TG--T--TC--ACAA	8184
ENV	P_K_P_Q_W_K_N_M_T_W_H_R_W_E_Q_E_V_D_N_L_T_D_H_I_D_G_L_L_R_E_A_Q_E_Q_Q_E_R_N_V_H_D_L_T	ENV

B_FR.HXB2R	AATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAAACATAACAAATTGGCTGTGGTATATAAAAATTATTCATAATGATAGTAGGAGGCTTGGTAGGTTTAAAGAATAGTTTTTCTGTACTTTCTATAGT	8338
ENV	E_L_D_K_W_A_S_L_W_N_W_F_N_I_T_N_W_L_W_Y_I_K_L_F_I_M_I_V_G_G_L_V_G_L_R_I_V_F_A_V_L_S_I_V	ENV
A_UG.U455	C---G-C-G-----A-C---A-----CT-----T-G-C---TG---A-----AA---A-----A---G-----A-	7768
C_ET.ETH2220	C---G-C-----A-A-C-----A-----G---TG-A-----A-----G-----G-----	7714
D_ZR.84ZR085	---G-C-G-----G-----C-A-----A-----A-----A-----A-----	7859
FL_BE.VI850	C---C-----C-----G---T---C-----A-----G-----A---C-----	7615
G_SE.SE6165	CC---G-CC-G-----G-----GG---GA---A-----A---T-----T-AA-----G-----	7756
H_CF.90CF056	C---G-C-----C-----C-----G---T---C-----A-----T-AA-----A-----G-----	7657
J_SE.SE92809	C---G-C-G---A---A-----T---C-----C---C---A-----A-----T-AA-----A-----G---G-----	7645
CRF01_AE_TH.CM240	---G-----C-----G-----A---T-----T-AA-----A-----G-----	7891
CRF02_AG_NG.IBNG	C---G-C-G-----C-A-----G-----A-----G-A---T-----T-AA-----A-----G---A---A-	7844
CRF03_AB_RU.KAL153	C-----G---T---A-----A-----A-----	8334
CRF04_cpx_CY.94CY03	CC---G-C-G-----CC-----G-----A---A-----A---T-----A---C-----G-----	7708
N_CM.YBF30	---G---C---AC---C---GC---GGT---T---A-----A---GCT-----C---A---T---CA---CA-AAG-ATA---A-AA-----A-	7889
O_MVP5180	-GC-----G-----CTC-C-T-----G-----T---A---T-----A---GCT---C---G---CAC-AA---A---G-TA---A-GATAA---AA-C---	8412
O_CM.ANT70	-G-----G-----CTC-A-T-----C---G-----T---A---T-----A---GCA---C-----CAC-A---GG-G---G-TA-CA-GATA---AA---	8385
HIV2_A.ST	---A---GC---ATGT---TG-C-C---G---TT---CTCC---A-CAAA---TC-G-ATGGAG-TTAT---T---AA-AA---T---AA-ATA---G-ACAA---GT---	8229
HIV2_A.BEN	---A---GC---AT-T-C-TG-C-C---G---T---CTCC---G-CAA---TC---ATGGAG-GCAT---TG-AA-AA---C---CAA-CTA---G-GCAAT-GT---	8799
HIV2_A.ALI	---A---GC---ATGT---TACC-C---C---G---T---CGCC---G-CAA---TC---ATGGAG-TTAT---A---T---AA-A---C-C-T---AA-ATA---G-GCAA---GT---	8783
HIV2_B.D205	---A---A---GC---ATGT---TG-T-----G---C-C-CTCC---A-AAAA---CC-T---GGAC---TAT---C---G---A---TG-----AG-ATACA-TG-GCAG---GC---	8775
HIV2_B.EHO	---A---C-----AT-T---CA-T-----G---T-C-CTCC---A-GCA---C---C-GG---GGAT---TAT---AT---AA-A---TA-----CAA-ATACA-TA-ACAG---GC---	8761
STMM83293	-C---A---GC---ATGTA---TG-C-----G---C-T-CTC---G-AA-A---C---C---ATGGAG---TATT---TAT---A---A---ATG---G---CAA-CTA-A---A-GCAAT-GT---	8395
SMM251	-G---GA---GC---ATGTG---TG-C-----G---C-TG-TTC---A-AAA---C---ATGGA---TTATG---T-T---AG-AA---CTG---GA-CTA-A---G-ACAA---GC---	8726
SMM9	---A---GSK---ATNT---YG-C-----K---G---C-T-TTY---A-AAAA---C---ATGGTG---C-----TCT---AG-AA---A---GA-ATA---G-GCAG---GT---	8233
SIVCPZANT	-C---C-----AGCTCA-A-----G-----C-A-----A-----A---T---T-----CTA-T---AC-----TT-GC---TAT---G---AG-TGCT---	8027
SIVCPZGAB	-G---C-----C---C-----G-----A-----A-----A---T---C-T---GC-----AA-CA---G-----TA-A-GA-A---TT-C---AG---	8378
SIVCPZUS	---G---C-----T---C-A---C-----G---T---T---C---A-----A---TT---AG---C---A---A---AA-C---T---AGG---TGA---T---AC---	8346
SIVSTAN1	-G---G-G-T---A-TTC-G-----CATC---G---G-CT-TCA---T-C---GG---T---GGGGAT-TTAT---AT---AC-TA---TTG---T---G---G-C---GGCT-A-TTGGGGATGCA---	8428
SIVGRI677	---G---G-T---T---G-G-C---TCA---CTCAC-GT---C---T-AG-C---G-T---A---GGAT-TT-AG-GA---TGATTA-TC---A---T-T-CA-GG-TAT---TGGGGATGTA---	8160
SIVVER9063	G---GACATCT---T---AC-C---TC---CG-TT-CT---A---AAACAT-C-T---A---GGAT-CT-AG---T---TA-CA---AC-T---GT---C-G-A-A-AA---TA-AG-TGCA---	8474
SIVVER155	-G---GTCAG-C---T-G---T---TC---CG-TT-TT---A---AACAT-T---GA---GG-T-TT---GC---AT---G-TA---G---T-GC---ACA-AT---TA-A---TGCA---	8451
SIVVERTYO	-G---AC---GT---T---GA---C---TC---CG-TT-CT---A---TAACAT-T---A---GGAT-TT-AG---A---T---AA-AA---G---T---C---ACA-A---TA---GGATGTA---	7943
SIVVERAGM3	---AG---GT---T---GA---C---TC---CG-TT-CT---G---AACAT-C---A---GGAT-TT-GAT---CT---TA-TA---A---T---GC---A-A-A---TA---TGCA---	7972
SIVLHOEST	-G---C-AG---TT...-AC-T---G---GG-T-CT---C-C---T-TACT---G-T---G-ATGCAG-GT-A---TA---CTT-TAA-CA---G---G---T---GAGCTT-A---A-ACAA-AT---	8575
SIVSUN	-GC---C-AG---TT...-AC-C---G---GG---C-GT-TCA---T-TCTA---C-----ATGCAG-CC---TA---CATTAA-AA---CAGCT---TC---AG-TT-A-CA-ACAAACAGA---	8643
SIVMNDGB1	-G---GAG-CCTTA---TC---G-GCA-G---G---T-T-TTGG---G-TCAA-CT---GGGGAG-TT-CT---GTT---AA-TA---A-----TT-G---A---CT---TGGAA---C-A---	7933
SIVSAB1C	---G---TGTCT---T---GAC---T---TCG---G---C-C---A---T-TG-A-GG---G---GA---GC---C---G-TA---C---A-CA---TAGCT---G-TC-GC-A-TAA---A-AGGA---TC---	8624
SIVSYK173	-G---C-GG---ACTCG---GC-----G---T-GT-G-A---T-C-TT---C---GA---GGAT-TTATG-A---CT---A---GCT---GC-----AG-TT-AGTG-AGGA---CA---	8314
ENV	K_L_Q_E_W_D_S_L_W_S_W_F_D_L_S_K_W_F_F_Y_L_K_I_G_F_Y_V_I_G_A_L_V_L_L_R_L_V_S_F_S_V_G_I_I	ENV

Tat and Rev intron end \| exon start Tat end in HXB2 <-|

<p>B_FR.HXB2R ENV TAT REV</p> <p>A_UG.U455 C_ET.ETH2220 D_ZR.84ZR085 FL_BE.VI850 G_SE.SE6165 H_CF.90CF056 J_SE.SE92809 CRF01_AE_TH.CM240 CRF02_AG_NG.IBNG CRF03_AB_RU.KAL153 CRF04_cpx_CY.94CY03 N_CM.YBF30 O_MVP5180 Q_CM.ANT70 HIV2_A_ST HIV2_A_BEN HIV2_A_ALI HIV2_B_D205 HIV2_B_EHO STMM83293 SMM251 SMM9 SIVCPZANT SIVCPZGAB SIVCPZUS SIVSTAN1 SIVGRI677 SIVVER9063 SIVVER155 SIVVERTYO SIVVERAGM3 SIVLHOEST SIVSUN SIVMNDGB1 SIVSAB1C SIVSYK173 ENV TAT REV</p>	<p>GAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACC...CCGAGGGGACCCGACAGGCCGAAGGAATAGAAAGAAGGTGGAGAGAGACAGAGACAGATCCATTCCGA</p> <p style="text-align: center;">_N_R_V_R_Q_G_Y_S_P_L_S_F_Q_T_H_L_P_T_..._P_R_G_P_D_R_P_E_G_I_E_E_G_G_E_R_D_R_D_R_S_I_R_</p> <p style="text-align: center;">_P_T_S_Q_#_#_R_G_D_P_T_G_P_K_E_#_K_K_K_V_E_R_E_T_E_T_D_P_F_D_</p> <p style="text-align: center;">_D_P_P_P_N_#_#_E_G_T_R_Q_A_R_R_N_R_R_R_R_W_R_E_R_Q_R_Q_I_H_S_</p> <p>C-----C-----TC-G-----TTGC----T...-AGA---T-T-G-----A-G-A---C-----CA-G-A-----G----C</p> <p>---C-----C-----TTA---CA---T-G-----C-----CA-G-----A---C</p> <p>A-----G-C-----T-----G-----A-----G-----C-----CA-G-----GT-C-----</p> <p>A-----A-----C-----T-G-A-A-----TTA---G-----C-----G-----CA-G-A-----G-G---C</p> <p>A-----C-----T-G-A-----TTAC-ACCA...-A---A-----C-----G-----C-CA-G-----AG-G---</p> <p>A-----C-----T-G-T-----TTG---G-A...-AC-----C-----G-----C-CA-----G-GA---</p> <p>A-----C-----T-G-A-----TTA---A...-A-C-A-G-----G-----C-----G-----CA-G-----AC-----G---</p> <p>A-----C-----T-G-T-C-----CTTC-ATCAT...-A-A-A-----A-----C-----G-----C-CA-G-----AG-G---</p> <p>A-----C-----T-G-C-----TTAC-ACCA...-A-----A-----A-----C-----G-----CA-A-----G-G---</p> <p>-----G-----A-----G-----G-----A-----A-----A-----A-----C-----A-----C-----</p> <p>A-----C-----T-G-T-G-----TTA---AAC-AAC---T-----G-----C-----C---CA-----AG-----C</p> <p>AGCA-----T-CC-T-----G-----TTA---A...G-A-----A-----A-----G-----GC-T-----G-----G---</p> <p>---G-ACA-----CA---CC-C-----G-----T-----CTG---CA-CGG-A-GAA-C-GAAACGCCAGGAAG-AC-GG-----GT--A-A-GAGAC...-GCC--AG-GG-CAGCC</p> <p>---A-ACA-----CA---CC-C-----A-----T-----C-AA-ATCA...-AAGA-A-G-A-GA-C--A-G-A--C--G--G-----AGA--G--GCC--G-GG--A-CC</p> <p>A-G---C---AA---C---AGG-TG-T-TC-C-TCCC---C-GCTTACTT-CAA-A-TCCATAT-C--A-GA-CGG-A-CAGCC--CCAG-AA-A-ACAGA--AGAC-TTG-AA-GCGTT</p> <p>A-G---T---AA---C---AGG-TG-T-TC-C-TCCC---C-GGTTACTT-CAA-A-TCCATAT-C--A-GA-CGG---CAGCC--CCA-C-AA--ACAGA--AGAC-T-G-GG-GACA-T</p> <p>A-G---C---A---C---AGG-TG-T-TC-CCTCCC-T-C-GGTTAC-T-CAA-A-TCCATAT-C--A-GA-C-G-A-CAGCC-ACCAG--A-A-ACAGA--AGAC-TTG-GA-ACGTT</p> <p>AGCA---C---A---C---AGG---G---TC-CCTCCC-T-C-TCTTAT-TCAA-A-TCCCTAT-CG-A-GA-CGG---CAGCC--CCA-C-AA-A-ACAGA--AG--GTG-AA-GACG-G</p> <p>AGCA--GC---A---C---AGG--G---TC-CCTCCC-T-C-TCTTAT-TCAA-A-TCCCTAT-CG-A-GA-CGG---CAGCC--CCA-C-AA-A-ACAGA--AG--GTG-AA-ACGAG</p> <p>AGC--GC---A-A-T---CGG-TG-G-TC-C-TCCC-T-C-TCTTGTCGTCAG-A-TCCCTAT-C--A-GG-C-G-A-CAGCC-ACCA--AA--ACAGA--AG--GTG-GA-GAG-T</p> <p>AGC--AGT-A---G---AGG--G-G-TC-C-TCCC-A-C-TCTTATTT-CAGTA-CTCATA-C-ACA-GA-CCG-C-C-GCC-ACCAG--AA-CA-AGA--GAGAC-GTG-GAAGCG-T</p> <p>AGC--GT-A-A---T---AGG--G-G-TC-C-TCCC-T-C-GCTTRTGTTYAG-A-TCCCTAT-Y-R-C-GG-C-G-A-C-GCC-ACCA--AA--AGA-GAGAC-GTG-GR-GAG-T</p> <p>A-GA-AG-----CAT-TC-G-A-----T-CTAC--A-A-CAG-A-GATCC-GAGC-GCCAGAA--AT-G-----GT--AA-A-AGACAGG-TCAGGT-GAGGGCTTG</p> <p>C-GG-----C-C-----TC-C-----G-----TTA---GT...-A---A-AA-GA--T--G-A--C-C-----G-----C---CA-----AG-GA---</p> <p>AGCA-A-----T---TT-CC-C-----A-----TTT-----A...A-TC--A--A-----A-----C-----GC-C-----A-A-C---AC-T-----A-CGA--</p> <p>AGC--G-----T-C-TT-T-G-TCC-----T-A-TA---GCCTG...G---AGG---CAGCC--CA-C-C---A---AGACA-AGATTTC--CAG--GCA--</p> <p>C-GA-ATA-----AAT-TC-C...CCC---T---TA---AC-GTTCA...GC--A---GGCCAGACAA--G---GGCC---CAG-----A---CAG--CAG--A--TG--AA--</p> <p>CGG--G-----C-----T-C-----TCCA---T---TA---ACC-GTGG...A---GGCCAG-CAA--C---GAGCC-----AC-----C-AGA-TCT</p> <p>AGC--G-----T-C-T-T---TCC---T---TA---ATC-GTGG...A---AGCCAGACAA--C---GAGCC-----AGA-C--G--A-AG--A--T-CG-AT</p> <p>AGTG--G-----GTT-TC-----TCCA---T---TA---ACCAAGTGGG--A---GGCCAGACAA--CC-ACGAGCC--G-----C-ACAG--GAT--A--TAGAGTCT</p> <p>AGC--G-----T-C-T-TC-T-TCCA---T---TA---ACC-GTGG...A---AGCCAGACAA--C-A-GGGCC--G-----C-AGCG--AGA---C---GAG-CT</p> <p>AGTA-A-A-GT-TAG--G---AGGGTGC-C---CCCTCTGTTT-TA-TGA-CAGGACTAC-A-T-GGAGA-AGAAGAGA-CCA-GA-C-GCC---CA-A-A---AGAGA-AG--CAGATA-AGAGACT</p> <p>CT---GA-GT-T-----AGGGTG--G---CCC-TCTG-TT-TG-TGA-CAGGATTG-CTCCAGGAAAC-T-C---A---CCC-C---CA-----AA-A---AGAGACAGAGA-AGA-AGA--C...</p> <p>A-G---GT-----C---CG---TG-C-TT-CA---GA-TG--AG-AG-A-CTATAAC-CAA--GGCCAGACAA--G--A-GA---AGCA-CA-CTT---ACT--GAGAGC---AC---GAGAAC</p> <p>T-GG-AGT---AA-A-G---G---TC-C-C-CC-TCC---TCTTCT-ATTAA...-A...ATC-ATAT-CAC-TG--CAAGGG-C-GCC---CAAA-A---AGAGC-AG-CAG-G-TGGG-AAAC</p> <p>A--G-ATT-GTT-GGA--C---GTT--TA-T-T-CAGA-CC-TAC--AGGGC...C---AAGGA-CCAG-CAAACCC-C-GAC-----CA-T-GAGACAGAGA--GACTCAA-G-CTCT</p> <p>ENV TAT REV</p> <p style="text-align: center;">_K_N_L_L_G_G_Y_V_P_I_L_Q_N_P_T_Q_G_..._R_K_D_P_G_K_P_A_D_E_E_E_G_S_G_D_R_E_G_L_N_V_S_</p> <p style="text-align: center;">_L_P_R_#_#_G_R_T_Q_A_N_P_Q_T_K_K_K_A_V_E_T_E_K_D_S_T_S_L_</p> <p style="text-align: center;">_E_P_Y_P_G_#_#_E_G_P_R_Q_T_R_R_R_R_R_R_Q_W_R_Q_R_R_T_Q_R_L_</p>	<p>8465 ENV TAT REV</p> <p>7895 7841 7986 7742 7883 7784 7772 8018 7971 8461 7838 8016 8539 8512 8359 8929 8913 8905 8891 8525 8856 8363 8157 8505 8473 8540 8284 8601 8578 8073 8099 8705 8770 8063 8748 8435 ENV TAT REV</p>
---	--	--

	Tat end in HIV-1 <-	
B_FR.HXB2R	TTAGTGAACGGATCCTTGGCACTTATCTGGGACGATCTGCGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACCTTCTGGGA.....	8582
ENV	_L_V_N_G_S_L_A_L_I_W_D_D_L_R_S_L_C_L_F_S_Y_H_R_L_R_D_L_L_L_I_V_T_R_I_V_E_L_L_G.....	ENV
TAT	-\$	TAT
REV	I_S_E_R_I_L_G_T_Y_L_G_R_S_A_E_P_V_P_L_Q_L_P_P_L_E_R_L_T_L_D_C_N_E_D_C_G_T_S_G_#.....	REV
A_UG.U455	-----G---T---A-GA-GC-----C-----A-----A-----CGC-----G---AGC-----	8012
C_ET.ETH2220	-----T---A---A-T-T-----C-----T-----A-----CA-----	7958
D_ZR.84ZR085	--G-----T-CA-----A A-----A A-----A A-----C G-----	8103
FI_BE.VI850	-----C---T---A-T-GC-----A A-----G-A-----CA-A-A-----C G-----CAGGGGCCTG.....	7859
G_SE.SE6165	-----G---T---AC-G-----C-----CA-----CCA-----G-----CA-----	8000
H_CF.90CF056	-----T---AC-G-G-----C-C-----CA-----G-T-----A-----GT-----C-----G-----	7901
J_SE.SE92809	-----T-C-A-T-GC-----C-----A-----T-----A-----CG-----C-----G-AC-----	7889
CRF01_AE_TH.CM240	-----G---T---A-GC-----A-----G-----AC-----C-G-----CG-----	8135
CRF02_AG_NG.IBNG	-----G---T---A-GC-----A-----A-----A-----A-----C-G-----C-----	8088
CRF03_AB_RU.KAL153	-----A---T---A-----C---A-----T-----T-----A-----A-----C-G-----	8578
CRF04_cpx_CY.94CY03	-----T---C-----C-----A-----A-----G-A-----A-----A-----G-----C-----G-----	7955
N_CM.YBF30	-----G---T---CA-T-G-----G-C-C-----TGA-----CT-----C-----C-----AC-G-G-----C-C-----C-----	8133
O_MVP5180	--GCCACCA--T---CA--AGT-G-ACACG--CA--CAA-AAT-T-G-GG-CT-----T-----CA---A-A-CAGGGA-CCG--C-GA-C-CTAC-----	8656
Q_CM.AN70	-CGCC-C-A--G-T---C---GT-G-ACACG--C--CA-A-CAA-AATAT-G-GG-CT-----T-----CA---GCA-CAGGGA-CCA--A-G-GA-CAGCTA---A-G.....	8629
HIV2_A_ST	GG.....GA-AATTGGTGGCC---CCGATAAGATAT-TA-AT-T---GA--C--C-G-TGATTGCGCTCTTGAACAGAC-ATACAAC-TCT-C-GG-ACTT---ATCCA-GAGCTTCCAGACCC	8480
HIV2_A_BEN	GGT.....-A-GACTTGTGGCCT---CCAATAAACTATGTG-A-T---GA--CA-CTA-TGACTCGCCTTTGA-CGGGC-ATACAAC-TCT-C-GA-ACTT---ATCCAAGAAGCTCCCG....	9046
HIV2_A_ALI	GGG.....GA-AGATTGTGGCC---CCGATCGCATATTTA-AT-T---GA--CA-CTG-TAGCTCGCCTTTGA-CGGGC--TACAGC-TCT-C-GG-ACTT---ATCCA-GATCTCCCGATCC	9034
HIV2_B_D205	GAC.....-A-AGAT-CTGGCC---C-GATAGAATAC-TT-AT-TT-AC-GC--C-G-TGA-GAACCTCTTAA-TTGGC-ATACAACGGCT-C-GAACCTT---GT--AAGACCTTCCAA....	9022
HIV2_B_EHO	GGC.....-A-AGAT-TTGGCC---C-GATCGAATAC-T-AC-T---CGA-TC--C-A-TGA-GGACCTCTTGA-TTGGC-ATACAACGGCT-C-GAACCTT-T-ATCAAAGACCTTCCAA....	9008
STMM83293	GG.....AT-AACT-CTGGCCT--C-AATAGAATAT-CT-AT-T---GA--C--C-G-TAGTACGCCTTTGACTTGGC--TACAAC-ACTTC-GG-CTGT--ATCCA-GATCTACCAGACCC	8646
SMM251	GGC.....AA-AGCT-CTGGCCT--C-GATAGAATAT-TT-AT-T---GA--C--C-A-TGATACGCCTTTGACTTGGC-AT-CAGC-ACT-C-GAACCTTG--ATC-A--GCATACCAGATCC	8977
SMM9	GGC.....AA-AGAT-TTGGCCT-R-C-GATAGAATAT-TT-AT-T---GA--C--C-G-TGATACGCCTTTGACTTGGC-AT-CAGC-GTT-C-GG-ATTGG--AT--A-GANCTGCCAAATCC	8484
SIVCPZANT	CAGCACGGGTTT-C-GCACTCT-GT..	8182
SIVCPZGAB	-----G-A--G-G-----C-----C-C--A-A-----G-GA-T-GG-----GA-----C-AG---GC--GC-A--GTG--ACAAC-CA-GACGT-----	8622
SIVCPZUS	-----C-G-----T---A-----G-----A-----T---A-----CCTGA-----CT-----A-----GAG-----A-A--C-G-G---CA--AC--A-C-----	8590
SIVSTAN1	GACAA-TCGCCGC---CA-TGAAAGAA-CTTTATTG-CCAAC--AGGAG-GA--CAAGCGG-GG-GA-AGCTT-GCGGCAG-AT---CCAATTG-T-CT-GACAATCAGCAGTT-G.....T	8658
SIVGRI677	--GCA-G-A-AG---CAA---C-TCGA--AT-A-CAACT--T-G--CAA-T--AAG---G-AG-TTGAGA-T-AGGACTTGG-GTTACAAC-TCT-CC-GACCCTC--CA-ATTC.....	8401
SIVVER9063	--GGCA---A-----TGGAA-GAGACAAA-AGA-GCAGTT---AAAC--T-GAC--AT-GG-T-TGCAG-TTCAG-ACCTGGC-ATACAAC--CTTC-GACCCTC--AA-TCAT.....	8718
SIVVER155	--AGCA---A-A---TTGG--G--GAGA-AA-AGGACCAGTT--T--AG-CAAT-GAC--AT-GGTT-TTGACACTCAG-A-CTGG--ACACAAC-GCT-CC-GACCCTC---CAAG.....	8695
SIVVERTYO	--GGTA---A-AC---AA-AGCAGGTGTAT-C-GTTGACCCGCT-G---ACGAGGC---A.....ACTTGGC-ATACAAC-GTT-CC-GACCCTC--CA-TCAG.....	8172
SIVVERAGM3	--GGCA---A-A---TGGCA--GCAGAG--A-GAGCAACT-T--AA-C-AT-GAC--AT-GGTG-TCGA-C---CAG-A-CTGGC-ATACAAC-GTT-CC-GACCCTC--AG-TCAT.....	8216
SIVLHOEST	A-C.....TAC---AACTTA--G-A-T-C-AGAAAAGATCT-C---GCCA-TGT-GAACGTG---GGAA-GA-CC-T--CAAGACTCGC--TT-G-GACTCTCCTGAAGTGG.....	8813
SIVSUNTACAT-AACTT-GAGCAG-CCA-GA-AGAATCTTTG-CACCA-CT-GG-CAGTAG--T-GGACGAGCCT--GAGAGACTC-C-GTTAGT--C-C-TCTCAAGTG-CTGAAAGCTTGG....	8887
SIVMNDGB1	--GAA-G-A-A-----AAACAGATC--T-ATA--GGAC-T--C-AGC-TTGC-AGGGAG-GA-CAATT-G-CTTTGGC-GAAAATC-GAG-G-AGCA---AAT-TGGGT-CCTCGCAGTCCAA....	8189
SIVSAB1C	AGC-AC-G-A-G--A-ACA-TTGGCAGA-A-ATTCT--A-ACA-----T-A-AGATTGAT-ACTT-GC--GA-T--GAC-AGTTGGT-CT--CA--CT-CAGCAAC--CCAT.....	8865
SIVSYK173	ACCT-C-G-A-----GAA-CTT-GAGGCAA-CTTTGGAAGCAG--C-GCAG--T-G--GGAC-GTCT-C-GCAGCTT-CAGATCTCTGATTAGACAGCTTA	8532
ENV	_T_F_S_R.....E_S_L_R_Q_S_L_E_A_G_Q_Q_L_W_R_T_V_C_S_S_F_R_S_L_I_R_Q_L	ENV
TAT	_P_S_A_#.....#_N_L_#	TAT
REV	Y_L_Q_Q_#.....#_I_F_E_A_I_F_G_S_R_T_A_A_L_E_D_S_L_Q_Q_L_Q_I_S_D_#	REV

```

B_FR.HXB2R .....CGCAGGGGGTGGGAAGCCCTCAAATATTGGTGGAACTCTCTACAGTATTGGAGTCAGGAACATAAAGATAGTGTCTGTAGCTTGCTCAATGCCAC 8677
ENV .....R_R_G_W_E_A_L_K_Y_W_W_N_L_L_Q_Y_W_S_Q_E_L_K_N_S_A_V_S_L_L_N_A_T 8677
REV .....#_O_G_V_G_S_P_Q_I_L_V_E_S_P_T_V_L_E_S_G_T_K_E_$ 8677

A_UG.U455 .....CGCAGCAGCCTCAAGGGACTGA-ACT-----G-G-----CT-----T-G-T-----G--G-----A-T-----A--CT-----TG-C--TGT 8128
C_ET.ETH2220 .....CGCAGCAGTCTCAAGGGACTA-AG-----A---T-----CT-G-A-GC--TG-G-----G--T--G--A-G-----A--ATC-----T--A--- 8074
D_ZR.84ZR085 .....CT-----G-A-----GA--G-----A-----G-TG---A-- 8198
F1_BE.VI850 .....A-G-----T-----CT-G-----ACG-G-----A-C--C-T-T--A--- 7954
G_SE.SE6165 .....CGCAGCAGCCTCAAGGGACTGA-ACT-----GT--T-----T-----T--GTT-----G-CAG-----A--T-----TG-A-AGT 8116
H_CF.90CF056 .....A-A-----A-----CTC-----T-----A--C-----G-A-----A--GAT-----T--CA--- 7996
J_SE.SE92809 .....T-----GAT-----CT-GT--C--TG--TG-----G-G-----A--T-----T--A--- 7984
CRF01_AE_TH.CM240 .....CACAGCAGTCTCAAGGGACTGA-AC-----G-----CT-G-----T-GTTA-----G-C-----A-T-----A--TCT-----TG---T-- 8251
CRF02_AG_NG.IBNG .....CACAACTGTCTCAAGGGACTGA-ACT-----G-----CT-----TA--TCA--C--GT-----A--AT-----T--A-A-T 8204
CRF03_AB_RU.KAL153 .....CT-----T-----G-----A--ATC--A-AGG-A--T 8673
CRF04_cpx_CY.94CY03 .....ATA-----G-----CT-----CT---G-T-----G-A--G-----A--AT--T-T--A--- 8050
N_CM.YBF30 .....CAGAGTCTCAGCAGGGGACTG-AACTACT-AAT---CT-AGA-C-C-C-T---GGAA-A--TGCA---G-AA-A--GT---G-G-----A-C-----T--A-A--- 8249
O_MVP5180 .....CTGGGACTGTGGATCTGGGA-AA-A-ACAATT---TTGT-G-CT---TG-AGC-G-AA-G-A---CTA--A--T-G-A-----ACA-A-C-----TG--A-T-T 8772
O_CM.ANT70 .....CTTGGACTGTGGATCTAGGG-AG-A-ATAATTA-T-TTTG-G-AT---TGCAGC-G-AAC--A-C--CTA--A--T-GC-----ACA-----AG-CA-ACT 8745
HIV2_A_ST .....TCCAACCTAATCTCCAGAGTCTTCGGAGAGCATTGACAGCA-TCA-A--CTGG--G-G--T-AACACAGCCTA--G-A--G--G-CG--TGGA-CC-AG-AGCGTTCGAGC---CGCG-GG--T-- 8610
HIV2_A_BEN .....ACCCGCGGACTGATCTCCAGAGTCTAACAGCAATCA---CTGG--G-G-CT-AA-GC-GCC-AA--G-A--G--T-CG--TGGA-CC-AG-AGC-TTCCAAGCA---CGCG-GGA-T-- 9168
HIV2_A_ALI .....TCCAACCGATCTTCCGGAGTCTCAGAGAGCGTGACA-CAATCA---CTGG--G-G-CT-AAAGCAGCCTA--G-----G--T-CG--TGGA-CC-AG-AGCGTTCGCGGC---TGCA-GGATTG- 9164
HIV2_B_D205 .....ATCCTCCACCAAATCTCA-CCAATCTCC--C-A---GGCTCCC-GTTGCCTA---C-----G--TCAGCTGGT-CC-AG-AGCGTCCGAGCGGCAGCA-GG--T-- 9132
HIV2_B_EHO .....ACCCTC-AAACA-TACTCC--C-G---GGCT-CCACCTGCATA--C-GA--G---TCAGCTGGT-CC-AG-AGCAATCCAAGCAGCAGC--GG--TG- 9109
SMM83293 .....TCCACCAACCTTCCAAGGATCTCCAGGATACTG-AG--AATTA-A---T-G--G-CTAG-AGCAGCCTATT-----G--T-CATCTGGA-CC-AG-AGCA--GCAAGCAGCATGG-GA--TG- 8776
SMM251 .....TCCAACCAACTCTCCAGAGGCTCTCTGCGACCTA--A---TTC-A---T---GGAC-GAATC--CCTA-----A--G--T-GAGCT-TT-CC-TG-GGCG-TCCAAGC-GGCTGG-GAT-TG- 9107
SMM9 .....TCCAGCCAGTGTCTCCAGAGGCTCTCAAGGACGCTG-AG--A-CCC-T---T-A--G-GT-GA-ATA-CCTA-----G--T-G-GTT-C--CC-AG-AGCA--GCARR--G-TGG--ATTGTG- 8614
SIVCPZGAB .....CATCTAATACTTACAGCCTCA-ACT-CTAC---GAGA--GTGCTGTG-T-GGAA-TA-T--A---G-AA--G-----A-T-----A-A---C---TG---A-- 8738
SIVCPZUS .....CAGAACATCAACAAGGGACTG-AACT-TT-AAT---CT-AGAGC-CGC--C---GGGG-AA-CGCT--C---GCAAG--G-----AGT-----ACA-----AT-AG--A-A-- 8706
SIVSTAN1 .....TATTGAGACTTTACCAGATCTCCGAGGAGCCTCAGC-CTCTTCTTC--CTG--T-G-C-GGA--CC-ATA-A-T-----G--T-G--C--T-C-AG-GG-A--A-CA--G-CTT-TG-G--TTT 8788
SIVGRI677 .....A---GGACAGCAGT-GGATA--G-----G--CTC--C-----CC-AG-GGCA--AACAG-GC-TGCTC-A--TCT 8478
SIVVER9063 .....T-G-GCGCT---C-GTA-A-C-A---G--CT-GG-----C-A-CCGCA--ACAAGAAGCA-GTCTC--TGT 8795
SIVVER155 .....T-G-AGCGCC--C-GTA-T---A---G--CT-GG---G--C--AGCCGCA--ACAAGAAGCAG-T--CCATCT 8772
SIVVERTYO .....T-GGA-AGCT-TTC-ATA--G--A--G--CTCGCA---C-A-CCG-C--ACAAGAAA-A---C-AA-TCT 8249
SIVVERAGM3 .....T-GGAGCGCT-TCC-GTA-A---A---G--CT-GG-----C--AGCCGCA--ACAAGAAGCAG-TGTC--TTT 8293
SIVLHOEST .....CTCAAGGAAGGGGAATACTCTCTATCACTA-T---C-GAG---GT-GGCT--C--T-GT-GATCCT--TCTTC--AA-TGGGC-ACGCCT-TGGCAG-C-AGCAG--GATGG-T 8935
SIVSUN .....GGAATAGTCTAGCA-AA-ACATC-ACC-CCTG--A-GT-TC-T---C-C-A--TAC-AC--CCTTC--TC-TGGGC-ACGCCT-TGGCAG-C-C-CAGAG-TGGCT 8997
SIVMNDGB1 .....G-AGC-ATC-GAAG-T-GGGT-CGCTACTCGTCAGCT-TGC-GCA-T-AG-TGG--GGCGGTT-CA--GG 8260
SIVSAB1C .....CGTGCCTC-AGGACATCCAAC-GAGAAC--G-C-GCTAACAGC--AT--TG-A---G--T-G--A--G--C--AGCAGCA--A-CGTAC-CACTGTG--TGT 8969
SIVSYK173 .....CTATCACCTGGGATTCATCAGCTATGGTTCAACAGGACTCAAATC-C---AGCTTCGCTCG-AA--G-AA-A-CGG-T-GGGT-GCGGTATTTGGC-AGCA-TATAC-CCGTACC..... 8651
ENV .....T_I_T_W_G_F_I_S_Y_G_F_N_E_L_K_I_A_A_A_S_L_G_R_E_I_R_D_W_V_A_A_I_W_Q_A_I_Y_A_A_T..... 8651
NEF .....M_G_S_T_S_S_K_S_Q_Q_L_R_S_E_G_K_Y_A_I_G_W_R_L_F_G_K_Q_Y_T_P_L_#..... 8651

```

	gp160 and gp41 Env end <-	
B_FR.HXB2R	AGCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAGGAGCTTGTAGAGCTATTCCGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGGATTTGCTATAAG.....	8796
ENV	<u>A I A V A E G T D R V I E V V Q G A C R A I R H I P R R I R Q G L E R I L L \$</u>	ENV
A_UG.U455	---AG-----G-T---T-----A--G---AC-AT-G-----T-A-----C---A---A---GC-----C.....	8247
C_ET.ETH2220	--A---T---G--A-A-----AT-----T-A---A-AT--G-----T-CT-A-----GCAGC---A---A.....	8193
D_ZR.84ZR085	--A-----A-----A-----CA---AGGA-----A---G-G-TT-----C-C-----T--A---GC-----A.....	8317
FI_BE.VI850	--A---T-----A-----AA-----TT-G--A---G-----G--T-A-----GCA-----GC---G--A.....	8073
G_SE.SE6165	--A-----A-TT-----C---A-----C-T-A-----C--G---A---AGC-----A.....	8235
H_CF.90CF056	--A-----A-----G-A-----T-A---G---A---G-----T-----T--A-GC-----TA.....	8116
J_SE.SE92809	--A-----A-A-----AA-C---A---C---A---T-----TT-----C-----T-A-AGC-----A.....	8103
CRF01_AE_TH.CM240	--A-----C--G-G-T-----C-----C-----G-----C---T-----C-----C-----C.....	8370
CRF02_AG_NG.IBNG	--A---T-----A-CT-----A-C-----A--G---A---TGG-----A-----C-----GC-----C.....	8323
CRF03_AB_RU.KAL153	--A-----G-T-----A---G---A---TT-----G--TA-----G---C-----GCA---A-GC---A---C.....	8792
CRF04_cpx_CY.94CY03	--A-----A-----A-----C---GA-----T-A-----C-----C-T---AGC---T---A.....	8169
N_CM.YBF30	--T--T-T-----A-A-A-----A-----T--C---A---ATAG-A--G-GA-ATTA-----C---A---C-A---AGCAC--A---C.....	8368
O_MVP5180	T--AG-GT---T--CA-TT---T--CG-CA-C--CTT--GTC---A---ATAG-ACA--GAT-C-TT---C-A-----T---A-TGCA---A--C--AG-G---C.....	8891
O_CM.ANT70	T--AG-G-----CA-TT---T--CG-CA-A--C-C--GGA---A---ATAG-A-C--GA-----TA---C-A-G---T-----A---A-G--AT-G---C.....	8864
G-GAGAGA-TC-TA-AA-C-CCTGGAGAG-CT-CTGG-GGAC-C-GGG-CA-AT-G-G-G-GA-A-TTGCAG-C-A---G-C-G---GGCA---TCGCC-C-G-G---G.....	8729	
HIV2_A_BEN	GAGAGAGA-TC-T-G-GC-C-TGG-GAT--T-ATGG---C--CG-G-C-CATCG-G--G-GA-A-T-GCAG-T-A---G-C-G---AGCA---CTGCC-C-G-G---G.....	9287
HIV2_A_ALI	GAGAGAGA-TC-TA-AA-CACCTGGAGAGACT-GTGG-GG-C--G--GT-G-TCG-G--GAGG--A-T-GCAG-C-A-G-G-C-G---GGCA---TTGCC-C-G-G-A.....	9283
HIV2_B_D205	G-GAGA-A-TC-T--GAGT-CCGG--G-CCT-ATGG--G-CTC-CAGGA---GGCG---A---CAT-GCG-C-C-G-C---A-AC-T---CTC-CCC-CT-G-G---G.....	9251
HIV2_B_EHO	G-GAGAGA-TC-T--GAGC-C-G-GAGG-CCTCGTGG-G---CC-CAGGA---GGCAG--AG--CATTGCA---C-G-G-C---A-GGCT---CTGCC-CT-G-G-T.....	9228
STMM83293	G-GAGAGA-TC-T--GAGC-C-GGGAGAGACT-ATGG--ACTC-TGG-A-G-T-G-G--AGG--CG-GGC---C-A-GC-C---A-GC-T---CTC-CAC-CT-G-G---G.....	8895
SMM251	GA-AGA-A-TC-T--G-GC-C-TGGAGAGACT-ATGG--GACTC-TAGGA--G-G-A---TGG--C-T-GCA-C---G-G-T-G-A-GC-T--GCTC-CGC-CT-G-G---G.....	9226
SMM9	RCGAGAGA-TC-T--AAGC-C-TGGAGAGACT-ATGG--GACTC-GGG-A-G-T-G-A--G-GA-A-T-GCA-C--MR--C-C-C-G-A-GC-T--GCTC-C-C-CT-G-G---G.....	8733
SIVCPZGAB	---T--T-----A-A-T-----AA-A-----CTT-T---TTA-ACT---AT---A-AA---C---GC-C---A-GC---AGC---A-T---CACTGCATT.....	8866
SIVCPZUS	---A-----A-A-T-----AA-----T--AC-AG-A-GCTC-T-CT-G---ATA-----A-----T--G-A--C-A-G--A-GC--AT-----G.....	8823
SIVSTAN1	G--G-GC--T--CGCAAAGC-CC-GTCG--C-C-ATGGA-T-CTTGCAG-TCC---A-C-G-A--C-T-G-ACAT-A-----GC---A-AAC---C--TGG--CAAC--G---G.....	8907
SIVGRI677	G--G-GG--T--CGAGG--A-CCTGG-GC--AC-GGGT-CTA-T--C-G-TCC---A-C-G--AG-CAT-A--GT---G-GC-G-A---C-T---AAG-CC--GGG--A.....	8597
SIVVER9063	G--ACGCTTT-C-CAAA-C-CAGGTC-CAAAA-ATGGCTT-CTTGCAG-TCCA---A--GCAC--CAT-AG--GT-A-----G-C---A-GC-T--GGAA--CC-TAAT--G---G.....	8914
SIVVER155	G--G-GCTTT-C-CGCA-C-C-G--C-CCA-A-ATGGCTT-CTTGCAG-TCC---A-C-G-A--CAT-A--TCT-A-----G-GC---A-GC-T--GGAAG-CC-TAAT--G---G.....	8891
SIVVERTYO	G--AGGC-TT-C-CAAA-C-CATGTC-CCA-A-ATGGCTT-CTTGCAG-TCC---A--GAAC--CGT-A--GT-A-----G-GC---A--C-T--GGAA--CC-TAAT--G---G.....	8368
SIVVERAGM3	G--ACGCCTT-C-CAAA-C-C-GGCT-CCA-A-ATGGCTT-CTTGCAG-TCC---A--G-A--CAT-A--TCT-A-----G-GC---A--C-T--G-A--CC-TAAT--G---G.....	8412
SIVLHOEST	G-TGGA-AAC-CTCAGA--ATTCAGAGCT--C-C-GG--GAA-TGCAG-A-GAACA-AG--CAGT-ATCGAG--CTGA---A-AC-TC--ATTGGGCA--A...AA-AG--GGA.....	9051
SIVSUN	G-GTTTCGCACC-TAT-C---CT--GAGC--A-C-GGA-C-CGTGCAG-ACCAGCA-GGAGAGAG-ATCATCTGCT-AG-AGGC-G---TAGAAC---TAGTCTCGGAGAAA--GGA.....	9116
SIVMNDGB1	-TGGC--CA--C-AA-AGAATCTTACAG--TGC-G---C-A-CATT-CT--GG-G-CGTGC--AT-G-----TC-----G-T---AA-TGCA---C-ACCC--TAAC--A.....	8379
SIVSAB1C	G-T-CA---T-CGA-AAGC-CC-GT--TCA--ACG-C-T-CTTGCAG-TCCAT-GT--G--AG-CAT-GCACAT-A-----GC-C--A-AGC-T--G--ATGG--TAAT--G---G.....	9088
SIVSYK173AG-C---T-GTCGAG--CG--GCAGC-T-G--GC-G-GC-TC---A-----G-TCTACC-TAAC-G.....	8724
ENVR R V V E A V A G L P R R L R Q G L E I Y L N \$.....	ENV
NEF# D E L S R P L Q P C R G G F D K A W R S T L T E.....	NEF

```

B_FR.HXB2R      ... ATGGGTGGCAAGTGGTCAAAAAGTAGTGTGATTGGATGGCCTACTGTAAGGGAAAGAATGAGACGAGCTGAGCCAGCA.....GCAGATAGGGTGG      8887
NEF             _ M_G_G_K_W_S_K_S_S_V_I_G_W_P_T_V_R_E_R_M_R_R_A_E_P_A.....A_D_R_V_
A_UG.U455      .....AG--CAGAG-G-A-----GAG-T--A-----GA-A--CCTG--.....GCA--A-AG-A--A-      8338
C_ET.ETH2220   ...-G--C-AT-----T-----CCAG-A-----G-A--A-----A-----CT-----GCA----GG-A--A-      8284
D_ZR.84ZR085   .....A-----A-AG-----G-A-----A-AA-A--T--AGGGAAAGAAGAAGACCTGAGCCA.....GCA----G-----      8432
FL_BE.VI850    .....G-----C--A-AG-----G--G-----A-A-CCCTA-----GCA----AG-----      8164
G_SE.SE6165    .....A-----A-----CA-AG-----CGAG-----A--AACA-CCCTA-----GCA----AG-A--A-      8326
H_CF.90CF056   .....A--A-----GA--GG--G--T--A-----G-----GA-C-----GTA--AG--A-      8207
J_SE.SE92809   .....GAA--A-----CAG--C-----A--A-AG-A-A--CCT-----GCA----G-A-----      8185
CRF01_AE_TH.CM240 .....AA-T-----CA-AG-G-----CAG-C-----A--A-AG-A-A--CCT-----GCAA---AG-A--A-      8461
CRF02_AG_NG.IBNG .....-C--CA-AG-G-----AG-T-T-A-----A-A-CCCTA-----GCA--ACAG-A--A-      8414
CRF03_AB_RU.KAL153 .....G--A-----CA-AG-G-----CAG-T-----A-----CCTG-TC-----GCG--AGAG-A--A-      8886
CRF04_cpx_CY.94CY03 .....A--A-----CA-AG-----GAGA-----G-----CGAG-T-AGCCAGAAAGAATGAGGCGAGCTCAAGCTGAGCCAGCA---CAG-A--A-      8296
N_CM.YBF30     .....AAAG-TT-----G-C--CC-AG-A-----AGAAA-C-A-----A--CAAAC--A--A-----CCAGCAGTAGAGCCA---TAG-A-CA-      8474
O_MVP5180     .....GAATGCA--AGC---C-AAT-TGCA---T-AGAA--A-T-----T-CTCT-T-AT-----CCTCAACAACCATGT--CC-G-A--A-      8997
O_CM.ANT70     .....AAA-GCA-T-AG---G--AAT-TGAG---G-AG-A--A-----A--A-AGAA-TTTC.....CCTGAGTCTGAACCATGC---CC-G-A--A-      8973
HIV2_A_ST     .....GGACGGCGGTATCAAC      8745
HIV2_A_BEN     .....GGACAGC-GTATCA-C      9303
HIV2_A_ALI     .....GGACAGCGGTATCA-C      9299
HIV2_B_D205   .....      9251
HIV2_B_EHO     .....      9228
STMM83293     .....CCTCAGATATAAT-      8911
SMM251        .....GGACAGA-ATACAATC      9242
SMM9          .....CCTCAGA-ATATTGT-      8749
SIVCPZGAB     AAG---AAC---A---T-----C-G-A-----GAG-C--AAG---A-GGA---CCAA---GCA----GG-A--A-      8960
SIVCPZUS      .....AA-----T-G---CA-AG-A-----AGAA-C--AA-C--T-----A-A-C-A-A-.....GCAGCA----AG-A--A-      8920
SIVSTAN1     .....      8907
SIVGRI677     .....      8597
SIVVER9063    .....      8914
SIVVER155     .....      8891
SIVVERTY0     .....      8368
SIVVERAGM3    .....      8412
SIVLHOEST     .....      9051
SIVSUN        .....      9116
SIVMNDGB1     .....      8379
SIVSAB1C      .....      9088
SIVSYK173     .....      8724
NEF           .....NEF

```

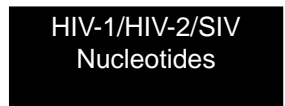
B_FR.HXB2R	GAGCAGCATCTCGAGACCTGGAAAAACATGGAGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTTGCTTGTGCCTGGCTAGAAGCACAAAGAGGAGGAGGAG.....GTGGGTTTTCCAGTCACACC	9011
NEF	G_A_A_S_R_D_L_E_K_H_G_A_I_T_S_S_N_T_A_A_T_N_A_A_C_A_W_L_E_A_Q_E_E_E_E.....V_G_F_P_V_T_P	NEF
A_UG.U455	-----T-----A---TT-A--T---T-----G-----C--T-----T--T---T---CAG-----G-----G-----A-GA--C.....A--C--C-----T-GG--	8462
C_ET.ETH2220	-----G-----T-A--C---T---G---C--T-----C---C--C--C--AT---C--A-----G-----G-----A--AGAG.....A--C-----G---	8411
D_ZR.84ZR085	-----T-----G-----G-----G-----A-----T-----C-----G-----G-----G--A---C---GAG.....C-----G---	8559
F1_BE.VI850	-----TG-----T-A--C--G--GG-----T---C-----TAG-A---T---C--AC-TG-----G-----A.....A--C-----G---	8288
G_SE.SE6165	-----T-----A---TT-A--T-GG-----C-----C-----C--AT---C--A-----G-----G-----CTCAGAA.....A--C-----G---	8453
H_CF.90CF056	-----TG-----TT---T-G--GC--G--G---T--AT-----T---T---CG--A--GCC-----G-----C--G---GAA.....A--C-----GG---	8334
J_SE.SE92809	-----TG---A---TT---CT--G---G-----C-----A-----AC-----G-----ACA.....A--A-----A---	8309
CRF01_AE_TH.CM240	-----T-----A---T--A--T-----A-----C-----T-----ATAGAT-----A-----T-----GAG-----C-----A--C-----TG---	8579
CRF02_AG_NG.IBNG	-----A---TT-A--T-G-----C-----C-----CAA--T--CC--A-----G-----G-----T---A--T.....A--C-----GG---	8538
CRF03_AB_RU.KAL153	---C---T---A---TT---T---GT-----G-----C--T-----AT-----A---C-----G--G--G-----A---A--C-----GG---	9010
CRF04_cpx_CY.94CY03	-----TG---A---T---C-----G-----T---A-----T---C---ACAAAA-----G-----A-----A--AGAG.....A--C-----GG---	8423
N_CM.YBF30	-----T---A---T--A--CT--T--GA--G--C-----C--TA--A-----TAG--AC--AT---AAAG--ATA--T-----G--A--A--GAA.....A--C-----ACGC---	8601
O_MVP5180	---T--TC--CA--G--GT--A--C--CTAGA--G--G--AT---TC--C--C--TC--TCAA--A-----A--CCT--A--TC---CAGC--CA--A--T---T.....A--C--C---A--G---	9121
O_CM.ANT70	---CAGATC--CA--G--AT--A--C--GCTAGA---GG--AC---TC--C---TC--TCAA--A-----A--CCT--A--TC---AGT--CC--A-----A-----AG---	9097
HIV2_A_ST	AG-G--AT--T--ATGA--TACCCC--TGGAGA--CCC--GCA..GAAG--GG--G--A--GCT--GTA---GCAACAAAA--ATGGATGAT--T--ATTCT--T--T--T--C...CTA--A--GG--C--T-----	8869
HIV2_A_BEN	AG-G--AG--ACATGA--AGCCC--TGGAGAAACC--GCA--GAA--AC--G--A--ATTTGTAT--GGCAGCAAAA--ATGGATGAT--T--ATTCT--T--T--T--C...CTAA--A--AG---T--T-----	9430
HIV2_A_ALI	AG-G--AT--T--ATGA--TACCCC--TGGAGAACC--GCA--GAA--AG--G--A--A--TTGTA--GCAACAGAA--ATGGATGAT--T--ATTT--T--T--T--TAGCCTA--A--GG--CT--T-----	9429
HIV2_B_D205G--CAA--ATGCCC--GGGG---A---GGGAGGGCAACAAGA---AGATGA--AGT--ATG--G--C--T--A.....GTG--GA--CCA--ATGTAAG--C---	9343
HIV2_B_EHOG--CAA--AG--C--TTGGGAG---A---GGGAGGAAACAAGA--T--AGATGAG--T--ATG--G--CA--T--A.....A--GG--C--GT--A--G---	9317
STMM83293	A--GTCA--T--ATGA--TACCCCTGGGA--AAACC--GGCAG--GAA--GC--AGCT--AATAT--GACAGCAAAACATGGATGAT--T--ATG--G--A--T--TA--T...CTA--A--AG--AG---ACAT---	9038
SMM251	AG--GGCAG--A--ATGA--TACTCC--TGGAGAAACC--GCTGA--GAA--AAG--A--TT---ATA--GAAAACAAAA--ATGGATGATAT--ATG--G--A--T--T--C...TTG--A--GG--AT---G--GG---	9369
SMM9	AG--GTCA--TCATGA--TACCCCTGGAGGAACC--GCA--GAA--AGCA--A--TT--G--TAT--GACACAAAAACATGGATGAT--TG--ATA--T--A--T.....TTA--A--G---G--GC---	8870
SIVCPZGAB	---A--TT--GAAG---A---G--C-----T-----T-----G--C--CC--AG--T---CAAA--CTA--T-----G--A--ATG--CA--T--A--A.....A--A--C---A--G---	9084
SIVCPZUS	---C--T--TC---A---T--A--C--G-----G--T-----T---C---A-----C--AG--A---CCAAA--CTG--A---T---T--AGATGACTA--TC--T--AAGTGAA--A--C--C--T--TCGC---	9050
SIVSTAN1A--G--G--AG--GC--CCCCTG--AGG--AG--AAATTGGCA---A--GACA--A---GAT--CTGGG--CCA---A--AGAA.....C-----A--G---	9005
SIVGRI677GT--AA--TG--C--CCAG--CGGCCGCCCTGC--A--AA--GAGACA---TTGAT--GTGGG--T--T--T--A--AGAA.....A--C--C--T--GCA---	8692
SIVVER9063GAGGAAT--GC--A--TG--C--CCAG--GG--AG--CGCCTGC--A--AA--GAGAC--AA---GAT---TGGTC--T--T--A--TGAA.....A-----T--A--A---	9018
SIVVER155GAGAAAT--GC--A--TG--C--CCAG--AGG--AG--CG--CTAC--G--AC--GGGACCAA---GAT---TGGTC--T--A--A--TGAA.....A-----A--G---	8995
SIVVERTY0AAAT--GC--A--TG--CGGAAG--GGA--AGGAGCCTTC--A--AA--GAGACA---GA---GTGGTCG--T--T---AGAA.....A-----A--G---	8472
SIVVERAGM3GAGAAAT--GT--AA--TG--C--CCAG--AGG--AG--AAATTAC--A--AA--GAGA--AAA---GAT---TGGTCT--T--A--A--TGAA.....A-----A--G---	8516
SIVLHOESTACTC-----TTC-----AGAA.....A-----A-----G--AG---	9092
SIVSUNGGCCC--AATGG--ATA--AC--A---GAA.....A--AA--C---G--G---	9163
SIVMNDGB1GATG--GAAACT--GACTTAACA--AGC--G--AT--C---A-----AGAA.....T--A-----T--GTGT--G	8447
SIVSAB1CC--GG--G--CTCCT--T--GA---GAAGAAGG--GT--AA--ATG--AAC--CAA--GAA--ACAGCCCTCGTGGTATG--T---T--A--AGAA.....A--C-----A--GG---	9195
SIVSYK173C--CATT---CAC--CGGG--CG--ACAG--GA--TGGG--ACATTCCGG--G--CAA--AATTCT--AC--G--A--ATATAG--GC--G--ATG---GA--CAC--GTCTC...A---C--T--GTGT--P	8843
NEFP_I_D_P_H_G_P_D_R_D_W_G_H_S_G_G_Q_K_F_S_P_G_D_I_V_Q_D_E_G_D_T_G_L.....V_G_F_P_V_T_P	NEF

	-> 3' LTR U3 region	
B_FR.HXB2R	TCAGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATTCCTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTAC	9141
NEF	Q V P L R P M T Y K A A V D L S H F L K E K G G L E G L I H S Q R R Q D I L D L W I Y	NEF
A_UG.U455	A-----T-----T-T---C--TT-----T--T-----A-G-A-----A-----G---T	8592
C_ET.ETH2220	-----T-----AT-C---C--T-----T-----A-G-A---G--G--T-----T---G---T	8541
D_ZR.84ZR085	-----G-----C-T-----C-T-----G-A-----G-T-----G-A-----T--G---	8689
FI_BE.VI850	-----G-----C-T-----C-T-----T-----T-----A-G-A---GG---C-----G-T---	8418
G_SE.SE6165	A--A-----G-----TT---GT---T-T---C--TT-----T-----T-----A-G-A---G---C-----C---G---T	8583
H_CF.90CF056	-----A-----A-----C-T---G---T-T---C--T-----T-----T-----A-GCA---G--C-----T-A---G---T	8464
J_SE.SE92809	-----A-----A-----A-----G---T-----TT-----T-----T-----A-G-A---G---G-----G-TC-T	8439
CRF01_AE_TH.CM240	G-----C-----T---G---T-T---C--TT-----T-----T-----A-G-A---G---G-----CT-A---G---T	8709
CRF02_AG_NG.IBNG	A-----G-----G-----T---G---C---C-----T-----T-----A-G-A---A-----G---T	8668
CRF03_AB_RU.KAL153	G-----G-----G-----A-T---G---T-C---C-----AG-----T-----T-----A-G-A---A-----G---T	9139
CRF04_cpx_CY.94CY03	A-----G-----G-----C-TT-A-G---T---C-----T-----T-----A-G-A---G---G-----G---T	8553
N_CM.YBF30	-----A-----G-----A-C-T-ACAG---T-T---TC-TT-----T-----G-TGG---AG-A---G-----T-A-C-C---G---T	8731
O_MVP5180	--A--G---C---G-----C-TT-A---CT-T-C-C---TT-----A-----T-----T-----T-AG-GC---A---G---C---A-T	9251
O_CM.ANT70	--A--G---C---G-----C-T-A-G---AT-T-C-C---TT-----A-----T-----T-----T-A-GC---A---G---T---G-G-T	9227
HIV2_A_ST	AAGA---A---GA---A-T-G-TTG---AAG---A-GTCA-T-GA-----GTA-T-AGTG-T-G---GTAGAG---A-CA-A-ACT-AG-A	8999
HIV2_A_BEN	AAGA---ACGG---GA---C-T-ATTG---AA---A-GTCA-T--A-----A-----C---A-GT-T-AGTAGG---G---TAGA---A-CA-A-ACC-AG-A	9560
HIV2_A_ALI	AAGA---AA---A---A---ATTG---A---A-GTCA-T-AA-----G-----A-GT-T-AGTG-G---TAGA---T-A-CA-A-ACT-AG-A	9559
HIV2_B_D205	AATA---AC-G---G-----A---ACT---A---CA-GTCT-T--A-----C-----A-GTA-T-TAGTG-G---CAGA---AT-A-CACA-ATT-TG-G	9473
HIV2_B_EHO	CGG---C---AC-G-G---A-T---ACT---A---CA-GTCT-T--A-----A-----A-TT-CT-TAGTG-G---G-TA-A-A-A-CACA-ACT-AG-A	9447
STMM83293	AAGA---C---AC-T-GGA---A-TT---TTG-AA---C-GTCA-T--A---AGC-----A-TTA-T-AGTG-G---TAGA---AT-G-CA---ACT-AG-A	9168
SMM251	AA-A-T---CC---G-----ATTG-AA---A-GTCT-T--A-----A-TTA-T-AGTGC---TAGA---T-A-CA---ACT-AG-A	9499
SMM9	AAGA---C---CG---ATC---A---ATTG-AA---CA-GTCT-T--A-----G-TTA-T-AGTATT---TA-A-AT-A---R-ACT-AG-A	9000
SIVCPZGAB	A---T---AACG-----T-A---T-T---TCA-----T-G-T-AGG-----G---C-C---G---T	9214
SIVCPZUS	--A---A---G-C---ACA---AT-C---T-AG-TT-----T-G-T-T-AG-----G---A---C---G-T---	9180
SIVSTAN1	A-GA-G-----A---A-T---TTG-A---T-CTC-----A-----G-CA-TA-TGG-A-C-AG---G-GC-A---GA---C-ATGC-CTA	9135
SIVGRI677	--GA-C-C---AG---C-T-ATT---A-G-CT-TC-----TC-----T-A-TA-T---TG-----G-A-G---GA-T---ATGC--TG	8822
SIVVER9063	AAGA---A-G-G-A---ATT---A-G-CT-TCG-----A-----T-A-TTA-T---G---G-GA-TAGA---CA-C---ATGCTCTT	9148
SIVVER155	AAGA---G-AC---A---A---ACTT---A---T-TCG-----A-----T-A-TTA-T---G-T-G---A-TA-G---A---ATGC-CT-	9125
SIVVERTYO	AAGA---C---G-A---T-CTT---A-G-T-TCG-----A-----T-A-TTA-T---G-T-G---A-TA-G---GA---ATGC-CT-	8602
SIVVERAGM3	AAGA---G-GC---A---C-T-ATT---G-G-CT-TCG-----T-----T-A-TTA-T---G-C-G---GA-TC-G---AA-C---ACGC-CT-	8646
SIVLHOEST	A---AG---GC-TT-TAGCCCT---C-T---A-TCT-A---C-CTCA-T--A-----T-----T-AGTTG---G---G---C---T-AAT---C-ACTGTG-G	9222
SIVSUN	A---AG---AC-TT-TAAGCCA---A-T---CAGCTAA---CTCA-T--A-----T-----CTGGT---AG-CC---G---A---T-G---T---ATGCAG-A	9293
SIVMNDGB1	C--A-TT-C---GTGCCAT---A---A-ATCTGA---CT-CTCT-T--A-----G---A---TA-T-TAG-AGG---G---A---A---C-ATGCAG-G	8577
SIVSAB1C	-TGCT---AC-C-G---A-T---TTG-AA---CT-G-----A-TTA-T-AGTG---G---A-GA-G---T-G---C-ATGCTCTT	9325
SIVSYK173	--AAC---AC-C---A-T-A-C---ATTG-AA---C-CTCT-T--A---A-C---A---C-G-CA-G-A-T-T-GTG---AG---G-T-A-T---C---ACC-AC-G	8973
NEF	Q T P L R T L T Y K L A I D L S H F I K N K G G L Q G M N Y C E K R D E I L H L Y L Q	NEF

Nef end premature in HXB2 <-			
B_FR.HXB2R	CACACACAAGGCTACTTCCTGATTAGCAGAACTACACACCAGGGCCAGGGGTCCAGATATCCACTGACCTTTGGATGGTGTCTACAAGCTAGTACCAGTTGAGCCAGATAAGATAGAAAGAGGCAATAAAG	9271	
NEF	<u>H T Q G Y F P D \$ Q N Y T P G P G V R Y P L T F G W C Y K L V P V E P D K I E E A N K</u>	NEF	
A_UG.U455	-----A-T-----G-----T-----AA-----C-----A-A-----T-----C-G-AG---G---A---T-C-GG--	8722	
C_ET.ETH2220	A-----T-----G---A-----A-----T-----C---AG-G-AG-G---AAT---G---	8671	
D_ZR.84ZR085	-----T-----G---A-----A-----T-A-----G-----T-G-----T-----GGT-G---A---C-G---	8819	
FI_BE.VI850	-----G-----G-----A-----T-----G-----T-----AG--G-G---A-----G---	8548	
G_SE.SE6165	A-T-----A-----A---G-----G-----ACT---TC-C-----T---A-----A-G-T---CAG--G---G---A-----	8713	
H_CF.90CF056	A-----T-----C-G---A-----A-----AG---T---C-----G-----T-----AA-T---C-GG--G---C-----G---	8594	
J_SE.SE92809	A-----T-----C-G---A-----A-----TC---T---A-----T-----AGCG-AG---G---A---T---G---	8569	
CRF01_AE_TH.CM240	A-T-----T-----G---A-----A-----TC-----TGT-----T-----C-A-AGAG-AG---G---A---C---	8839	
CRF02_AG_NG.IBNG	A-----A-T-----A---G-----ACT---TC-----G-----T---A-----A-G-T---CAG--G---G---A-----G---	8798	
CRF03_AB_RU.KAL153	-----G-----G-----A-T---TC---A---A---G-----C---C---C-G-AG---CAG-G---	9254	
CRF04_cpx_CY.94CY03	-----T-T-----G-G-T-----A---AG---TC---TG-----T-----A---T---C-GG--G-G-----C-G---	8683	
N_CM.YBF30	-----AT-C-----C-G---T-----AA-T---C---CG-A-----T---A-----T-GTCAG-T---AG-AG---T---G---	8861	
O_MVP5180	-----T-G-A-T-----G---TGT-----G---A---ACCT---TC-----A-----TG-TT---A---G-----GTCAG---AG--GC---GAGACTGGG---TA	9381	
O_CM.ANT70	A---T-G-A-T-----G-----A---AAC-G-TC-----A-----TG-TT---A-----GTCAGA---AG--GC---AGACTAGGA---TA	9357	
HIV2_A_ST	A-GGA-G-G---AATAA-TGGA---C-G-----T---T-AT---A---A---G---AA-TT---G---TA-GG-----A---TGTCCAC-AGAG-G---T---A---G-G-G-	9128	
HIV2_A_BEN	A-AGAGG---GATAA-A-A---G-----T---T-AT---A---A---G---C---A---TA---C-G---CTG-GG-----T---A---A---TCTCAC-AGAG-C---AAG--G-G-	9689	
HIV2_A_ALI	A-GGA-G---GATAA-T-A---G-----T---T-AT---AA-A-G-C---GA---TT---G---CTG-GG-----A---TGTCCAC-AGA--GG---A---C-G-G-	9688	
HIV2_B_D205	A-TGA-G---ATTG-GT---GC-G---A---T---T-AT---A---A---G---CAAATA---T---CTG-GG---G---A---GTGCCAGCAGCAGCCCGA-AGG-GG-G-	9603	
HIV2_B_EHO	A-TGA-G---ATTG-GT---GA-G---AT---A---A---CAA-TT---C---CTA-GG---G---A---AA-TATGATAGCAGA-CC---A-G-GG---	9576	
STMM83293	A-AGA-G---GATAG-T-A---G-----T---G---A---A---C---AA-CAG---CTA-GG-A---A---TATGTCA---TGAG-C-C---AGG--G-T-	9298	
SMM251	A-GGA-G---AT-A-A-A---G---G-T---CT---A---AA-T---C---AA---A---C---CTA-GG-AT---C---T---AA-TGT-TCAG-TGAG-C-C---AGG--G-G-	9628	
SMM9	A-RGA-G---AATAA-A-A---G---A---T---G---AA---C---TA---TT---G---CT--GG-AT---WC---G---AA-TGTCTCAG-TGA---CTC---AAG-CG-G-	9129	
SIVCPZGAB	-----T-----C-G-----A---A---AACA---TC---CTGT-----T---A---G---CC-GACAGAG--GC--G---C---A-----G---	9344	
SIVCPZUS	-----AAT-----C-G---T---C-----T---A---A---C---A---A---G---TT---T---C---TC-CACAGA---GG-AG---GC-----	9310	
SIVSTAN1	A-TGA-TGG---ATTA-AGA---G---AGCA---TT-C---T---CAC---GA-G---C---GATG---C---C---TT---TTG---T---G---G---A---TGTCTCCC---GA---CCC---.G--G-G-	9261	
SIVGRI677	A--GAGTGG---AATAA-AGA---G---AGCT---T---C---G---A---A---G---C---GAGAGT---G---C---TC---TT-----C---G---C---TGC---G---GAG-C---CGCAA-TG-G-G-	8951	
SIVVER9063	A-TGA-TGG---AAT-A-TGA---G---A---CGC---GGT---A---A---C---A---C---GCTG---C---TT---T-----G---A---ATGC---G---AGAG-C---A---TG-G-G-	9277	
SIVVER155	A-TGA-TGG---AAT-A-TGA---G---A---CGC---GGT---AA---A---T---A---A---C---GAG-TG---C---TC---T-----G---G---CC-TGC---G---GA---C---AA-ATG-G---	9254	
SIVVERTYO	A-TGA-TGG---GATAA-TGA---G---A---TGC---GGT-GAAG---A---AA---A---TC---TAAATG---G---TC---TT---T---G---G---C---CTT---C---G---GA---C---A---ATG-G---	8731	
SIVVERAGM3	A-TGAGTGG---AAT-A-TGA---G---A---TGCT---GGT---GA---A---AA---C---AGATG---C---C---TC---TT---AT---G---A---C---TGC---G---GA---C---A---TTG-G-G-	8775	
SIVLHOEST	A-TGAGTGG---TCT-A-TGGA---C---TCAT---GAT---A---AACA---C---T---A---TG-GGC---T---A---G---CTTGT---G---TATA---TC---TCAG-T-	9352	
SIVSUN	A-TGAGTGG---T-A-AA---GC-G---AG-T---CAA---T---A---G---AA-G---C---CTA-GG---C---CACCATT---G---AGACAGG---TC---T---C---TC	9423	
SIVMNDGB1	A-TGAGTGG---A-TTGAA---GA-G---AC-G---T---GA---T---AAC---TAA---A---TCCTG-TT---G---A---GAGCAG---C---TAGG---T---TATGCAGCT-	8706	
SIVSAB1C	A-TGAGTGG---ATAG-AGA---GA-G---T---T---GAT---TAC---G---C---AA-TG-----TT---A---G---T---G---C---TGTCAG---GAG---CTA---AAT---G---G-G-	9454	
SIVSYK173	A-TGAG---T---AAT-A-AG...-CAG-ATC---T---CT-----AAC-C-G-C---GT-A-T-----TA-GGG-A---T---AA---AT---GGGATGCTG---TATG-GG-G-	9099	
NEF	<u>N E H G I I # # R I N Y T S G P G Y R Y P L I F G W L W E L V P N E I E G C L E Y E E</u>	NEF	

B_FR.HXB2R GAGAG...AACACCAGCTTGTACACCTT.....GTGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTAGAGTGGAGGTTTGACAGCCGCTAGCATTTCATCACGTGGCCCGAGAGCTGCA 9392
NEF G_E...N_T_S_L_L_H_P.....V_S_L_H_G_M_D_D_P_E_R_E_V_L_E_W_R_F_D_S_R_L_A_F_H_H_V_A_R_E_L_H 3097

A_UG.U455 -----A--TC-AC-----A-AT--AA---AG-A---TGAA---A-----AT---A-----TAC--G---AA-A--AGA--TTAT----- 8843
C_ET.ETH2220 -----A-T---C-G-----C-----A---G---TGAA---C-----A--A-----A-----T-A-----CGCAGA--A-----C-----A-- 8792
D_ZR.84ZR085 -----AGAC...A-T---C-----A-CT--A---A---G---GA---AG-----T---A--A-----A-A-----G-G---AA-----AA---AAA-T- 8940
FL_BE.VI850 -----A-T-----C.....A---AA---A---G---TGAA---C-----AC-GAG---A-----A---G---C-GAGA---A-A---A---AGA-- 8669
G_SE.SE6165 -----A--TC-A---C.....A-CT--A---A---G---GAA---C-----C-G-T---A-----TA-----CGGAGA--A-A----- 8834
H_CF.90CF056 -----A---C-----C.....A-----A---G---TGAC-G-----C-GAT---AA-----T-A-----GACA---T-----TAAA-- 8715
J_SE.SE92809 -----A...A-T-C-----C.....A-AT--A---A-A-G---TGAA---A-----C-----A-----TC-----CGGAGA--A-A----- 8690
CRF01_AE_TH.CM240 -----A...A-T-C-----C.....A--A---A---A-A-G---GAA---A-----C-GAT---A-----TGC-----CGAA--A-----A-A-- 8960
CRF02_AG_NG.IBNG -----A...A-T-A---C.....A-CT-T-AA---A---G---GAC---T-----C-GATC---A-----TA-A-G---CAGA---ACA---TA----- 8919
CRF03_AB_RU.KAL153 -----A...C-A-----C.....A-AT--AA---C-A-----TGA---A-A-----AT---A-----T---G---TC-AACA---AGA-- 9375
CRF04_cpx_CY.94CY03 -----T-T---C-G-----C.....A-A---A---A---G---TGAA---A-----A---A-----T---G---ACA-G---A----- 8804
N_CM.YBF30 -----C...ATGC-C-C-----C.....A-AT-T-AA---AGCA---TGATC-T-A-----G-T---C-A-----TC-----AGAAGA-T--A-AA----- 8982
O_MVP5180 C-A-TGAAG-TG-T--TC-TC---T-A.....CTT-TAAT---AGCT--G---TG-AC-CG-G---GA-AC--A-A---CA---T---ATCAT---GC--AACA--TA-A---TGC-AAA-- 9505
O_CM.ANT70 C-TGTGAG-GGG-T-ATC-CC-G--T-A.....CAT-TGCC---AT-T-A---TA-AC-T-A---A-AC-GAT---A-----T-ATCT---GCAACACC--T--T---TATGATAACT-- 9481
HIV2_A_ST -----CTCA-T---AG-G--T---AGCACA...ACAAGCA--T-T-----C-TG---ACA---TT-----CC-ACG---T---AGCT--A---TTTATT-GAT- 9244
HIV2_A_BEN -----GC--A-T---AG-----AGCACA...ACAAGCA-ACAT---TGA-C-TG-G---GACA---T---CA-----TC-ATG-G-C-A-A-CT-AA---TTCACT--T- 9805
HIV2_A_ALI -----CTCA-T--C--C-----AGTACAA...ACAAGCA--CAT---A-C-TG-G---GACA---TT---A-----CCTAAG-G-TCA-G--T-AAA---TTTATT--A- 9804
HIV2_B_D205 A.....C-CATT--C-AA-G---GGCACAG...A-CTCCTCATG---ATCC-TG-G---GACTC-TATC---CA-----TTC--T--G---A-G-T-T---TTCA-TAG-TT 9721
HIV2_B_EHO -----C-CATT-TC--G-G--T---AGCACAG...ACCTCC-CATG-----CC-CGA---GACCC-T-TC---CA-----TC---T-----A-G-CT-T---TTCAGCAG-TT 9692
STMM83293 -G.....CACATTATC--G-G--T---AGCACAG...ACA--CA-TG-----TTG-G---AC-G-TT---AA---TCCA-TAT---CCA-AC-T-T-A---TTT-TTAGA-- 9416
SMM251 -----GGCATTAT--AA-G--G---AGCTCAA...ACTTCCAA-TG-----TTG-G---G--TC---C---A---TCCAAT---C-ACAC-T-T-A---ATAT-CTAGAT- 9744
SMM9 -----CACATT--A-G-T---AGCACA...ACTTC-CA-TG-----CTG-G---G--AC-G-CA---A---TCCARAGT---T-A-R-CT-TAA---ATTT-TTGA-- 9245
SIVCPZGAB -----TAAC--TG-CTG---CAT-C-ATA.....T-T-A-----A-TGA---C-A---G--C--G-TC---C-C-----A-G-G-C--AAGA--TA-T--A---A-AA-- 9465
SIVCPZUS -----C-A--TAC-C-G-----C.....A--T--A---A---A-TGAAG-TG-C---C---ATC---CA-----CTGAA--G--TCGGAGA--AGA--TAA----- 9431
SIVSTAN1 -----GACA-T-T---C-C-----CCCAGA-TG-AT--GAAAGC---TTG-A---GACA--G-T---AA---A-TCCA-T-----G---GT-----A- 9362
SIVGRI677 -----GACA-T-TC--A-G--T---AGCACAGA--G-GGAAG--CCTGAT-GAATAGATC-TG-----C--G-TC---A-----CCGAAGT-G--GG-GG-GT-----G 9058
SIVVER9063 -----GACATTA---GCG-T---GCACAAA-A-AAGA-G-CCCCGAT-GCAT-AGTC-TG-----GACC--G-TC---A-----CCTATGT---G-G-GT-T-----C-CAA- 9390
SIVVER155 -----GGCATT---G-----AGCACAAC--CATGAAG--CCTGAT-G-ATAAATC-TG-----A-A--G-CA---A-----TCCAATGT-G-TG---GT-----C-CCTC 9367
SIVVERTYO -----GACATT--C-AG-C--T---AGCGCAGA--G-AGAAG--CCAGAT-G-AT-AGCC-TG-----GA-C--G-T---A-----TCCTATGT-G--A-A-GT-----C-CCA- 8844
SIVVERAGM3 -----GACATT--C-G-G--T---AGCACA...GGAAG--CCTGAT-GAAT-AACC-TG-----C--G-TC---A-----TCC-ATGT---G-C--AT-T-----C-CTA- 8888
SIVLHOEST AG.....CACAAAT--C-T--G--TT-AAGCCAG...---GGG-TCC---A---CTG-G---GAG-C-TATC---CAC--CA-TCCAATGT-G--G-GG-CTTTA-A---TCA-AAA-- 9470
SIVSUN CT.....TGTCGAAGC-C-AC-G--TT-CAGCCAG...-AAGGG-TT-AT--A---TTG-G---GAGAC-TATC---CA---TCC-ACT-G--C-A-G--TTTAG---AT-C-AAA-- 9541
SIVMNDGB1 -----ATCATC-----T-CTCCAGT-AT-T-CT--G-AAGATCCA--AGGA---CCCTCA---GGTCTG---CCC-CAT-TTG-CTA-GA-T--G--G---TTAACATATCACC---AG 8828
SIVSAB1C -----CA-T--C--C-G--T---GCTCAA...G--GCCTATGA-----G-CTG--A---GACCC-G-T---A-----CCA-TA-----G-GG-CT-T--T--TT-GAG--- 9570
SIVSYK173 -----C-T--ACT-C-CC-C--T---AGCC.....TCTGG-C--G---T-CGAG--TGG-AGA-C-TCA-GTGG-AC-TCAA-CCCCA-CTGG 9182
NEFH_T_L_L_L_H_P_A.....S_G_Q_G_S_S_S_M_G_E_P_H_V_E_L_Q_P_P_P_G NEF



	TCF-1 alpha	
B_FR.HXB2R	TCCGGAGTACTTCAAGAACTGCTGACATCGAGCTT	9427
NEF	_____P_E_Y_F_K_N_C_\$_____	NEF
A_UG.U455	-----T--AT--AG----AAAGAC-GCT-.	8875
C_ET.ETH2220	-----T-A--AG-----CA-C.	8824
D_ZR.84ZR085	-----C.	8975
F1_BE.VI850	-----T--A-C-AG---AGACTGCT--CACAA.	8724
G_SE.SE6165	-----A--AG-----CA--AG-.	8869
H_CF.90CF056	-----AAAG-CTG--ACACAGAA--T-G.	8750
J_SE.SE92809	-----TT-A--AG-----AA--AG-.	8725
CRF01_AE_TH.CM240	---A-----AT--AG-----AG--AG-.	8995
CRF02_AG_NG.IBNG	-----A--AG-----CA--AG-.	8954
CRF03_AB_RU.KAL153	-----T-A--AG-----CA--CTGC.	9410
CRF04_cpx_CY.94CY03	-----TT-A--AG-----CA--AG-.	8839
N_CM.YBF30	-----TT-A-----AG-GA-.	9022
O_MVP5180	C-A--CT---CCC--G-AACTGACA-TGCGG.	9539
O_CM.ANT70	C-A--CT---C---GGA--A-A-A-.	9515
HIV2_A_ST	C-A--G-G--TGG-T--AAG-C-GGC-TGC.	9370
HIV2_A_BEN	C-A-AG-G-TGG-C-AAG-C-GGATTGC.	9931
HIV2_A_ALI	C-A--G-A--TGG-T--AAG-C-GGC-TGC.	9927
HIV2_B_D205	---A--AG-G---GG-T--CAG-C-GGG-T-C-.	9851
HIV2_B_EHO	C-A--G-G--TGG-T-TCAG-C-GGAATGC-.	9822
STMM83293	C-A-AG-G-TGGTTCTAAG-C-GGCTTGC.	9507
SMM251	C-A-AG-G-GG-AGCAA-TCAGGCTT-TC-.	9834
SMM9	C-A-AG-G-TGGT-GTCAG-C-GGCTTGT.	9336
SIVCPZGAB	-----A--AG---ACTTT-GACT-GCGCATGCGCACAGAAGTCTGACTCTGCGG.	9529
SIVCPZUS	---A-----CG-----A--G-TGTA-C-GC.	9481
SIVSTAN1	C---CAG---T---G--ATGCATGG--TT-T-AAGAGGAAGTAGCCTA.	9459
SIVGRI677	C---CATG--T--G--ATGCACG-A-AT--AAAGCCTAGTGTGACG.	9111
SIVVER9063	CAGAC---T--TG-AG--ATGCATGCAAT--TGAAGAGGAAGT.	9474
SIVVER155	AAG-----T-CAG--TA-ATTCAAC--T-GGTACAG.	9452
SIVVERTYO	---G-----T-CTG--ATGCATGGG-TG-TGAAGAGGAAGT.	8943
SIVVERAGM3	CAGAA-A-TC---CTG--ATGCATG---TT-GCAAGAGGAAGT.	8987
SIVLHOEST	G-T-C-A-GA-AC-A-TACTGCTTTTCGCTTT-G.	9587
SIVSUN	C-T---AG-G-----C-TGTTACTAG--TGCGAG-.	9631
SIVMNDGB1	-GTTTCAATAAGGCTA-GAGTA-TGA-CATCT.	8947
SIVSAB1C	C-A---C-GG-GCC-TCAGCACA-GGCTA-C-AC.	9639
SIVSYK173	CTATACCCAGGTTG-G-GATGGCC-GCTTGC.	9244
NEF	_____Y_T_P_G_W_E_M_A_R_L_#_____	NEF

	NF-kappa-B-II	NF-kappa-B-I	Sp1-III	Sp1-II	Sp1-I	TATA Box	
B_FR.HXB2R	.GCTACAAGGACTTTCCGCTGGGACTTCCAGGGAGGCGTGGCCTGGG.....				CGGACTGGGAGTGGCGAGCCCTCAGATCTGCATATAAGCAGCTGCT		9525
A_UG.U455ACACAGAA-T-CT---AGGGACTTTCCA-TG--GACTTTCCGAGG.....				GAGGTGTGGTTT-GGC--A-TTG--GAGTGG--AGCCCT-A-ATGC-GCAT-TAA--A		8979
C_ET.ETH2220GCC-G-A-T-TC-ACT-----				GCGTTCAGGAGGAGGGTCTGGGC--GACT-----C-A-----G-----		8925
D_ZR.84ZR085		8975
FI_BE.VI850	AT--A-----A-----	A--G--CG--CCAGA-----	TCA-----G-----C--	8823
G_SE.SE6165GC-GACAA-----CCTG--AC--TT--G-----	AGGCGCGCCCTGGGA--AGT-----	T-A-----AG-----C--		8969
H_CF.90CF056AAA-----	GCT-G-A-T-TC-GG-A-----	GCGGTGATCTGGG--GA-----	C-A-----G-----		8843
J_SE.SE92809	GA-GAGT-G-CT-----	G--GT-----	T-A-----G-----		8824
CRF01_AE_TH.CM240CT-CTA--AC-----	GA-GTGT-G-C-----	G--AGT-----	A-T-A-----G-----A-----C--		9092
CRF02_AG_NG.IBNGGAC-----A-----	G-----T-----TT-----	A--AGT-----	T-A-----C-G-----		9052
CRF03_AB_RU.KAL153GAC-T-A-G--G-T-ACA-----	GC--G-A-T-TC-AG-----	AGGTGTGGTTTGGG--AGT-----	T-A-----G-----		9520
CRF04_cpx_CY.94CY03GC-GACAAA-----	GCCCG--AC--TT--A-----	GAGGCGCGCCTGGGA--GT-----	T-A-----G-----A-----C--		8940
N_CM.YBF30	CTGACA-G--ACT--ATACT-----	GCCAG--AC--TT--A-----	AGGTGT--TTG-----	CTTG--CTCAGAG-----A-----C--		9124
O_MVP5180	ACT-CT-ACA-TGC--G.....	GACTTTCAGCGTGGGAGGATAAGGGGC--TTC-----	T-A-----G-----	T-A-----AG-----C--		9642
O_CM.ANT70	AC--GA-GATTG--GA-ACTGT--A-----	CA-A-AC--CTGACTCGCGGGACTTCCAGTGGGAGGACAGGGGGC--TTC-----	T-A-----	AG-----C--		9645
HIV2_A_BEN	AAGG-A-CTA-CTGAGA-TGCA-----	AAG--GC--TAACCA-----	GGGAGGACATGGGA--AG--C--T-G-GAA--CC-T-AT-CT-T--T--A-T-T-C-C--		9483	
HIV2_A_ALI	CTAAGA--CA-CTGAGG-TGCA-----	AAG--GC--TAACCA-----	GGGAGGACATGGGA--AG--T-G-GAA--CC-T-AT-CT-A--T--A-T-T-C-C--		10044	
HIV2_B_D205	ACTG-A--CA-CTGAGA-TGCA-----	AAG--GC--TAACCA-----	GGGAAGGACATGGGA--AG--T-G-GAA--CC-T-AT-CTC--T--A-T-T-C-C--		10039	
HIV2_B_EHO	AAGG-A-CTA-CAGACA-TGCA-----	AAAGA--GC--TAACAT-----	GGGAGGACATGGGA--GT--TTG-GAA--CC-AT-A-ATCT--T--A-T-T-C-C--		9963	
STMM83293	GAAGGA--CTAG--GATACTGCA-----	AAG--GC--TAACA-T-----	GGGAGGACATGGGA--AG-A--AG-AACGCCCAT-AAGAC--T--A-T-T-C-C--		9934	
SMM251	AAGG-A-CAA-CTGAGA-AGCA-----	CAAG--GC--TAA.CA-----	GGGAGGACTGGG.A--A--T-G-GAA--CC-TT-T-CT-T--T--A-TA--A-C--		9618	
SMM9	A-GG-A-CTC-CTGAGATAGCA-----	CAAG--GA--TTA-----	TGGGG.A--AG-C--TCG-GAA-ACC-A--TTCT-GA--T--A-TATCA----A		9934	
SIVCPZGAB	AAGA-A-CAA-CTAGGA-AGCA-----	CAA--G-AT-TT.AC-----	GGGAGGTCTGGGGA--AG--CTG-.AA--CC-A--T-TTCT--T--A-TA--A--A		9447	
SIVCPZUS	A--GACTT--CCGGG-A-----	AGGTGGTCCGGGG--GT-----T-----TTTG--CTCAGAG-----		9618	
SIVSTAN1	CTGCTG-CT-AGGAC-TTCTAA-----	A..G--ACGTT--AA-----	GG.GTGGTTCAGGG--A--A--CGTG-TTTTA-----A-GA-----C--		9591	
SIVGRI677	CATA-AG-A-TG--GAT-AGCA-----	T--T-C-AGGGC-----	T--T--CGTG--TTTA-----CTT--CTCAGAG-----A-----A--		9554	
SIVVER9063	TTGCGGTT-----CA-----	CA-T--GTG-AT-G-A-----	GCGGTACAGGGGC--TACT-----TTT--CTCAGAG-----A-----A--		9221	
SIVVER155	T-AC-TTT-AA--GCTGA-A-----	A--GACTT--CCAGG-C-----	GGCCATGGGC--TAC-----TTTTA--CTCAGAG-----A-----A--		9581	
SIVVERTYO	GATG--TTAAGAAC-G-TGAC-----	CAC--GACTTT-CAA-----	GCGGGACATGGGC--TAC-----TTTA--CTCAGAG-----A-----A--		9562	
SIVVERAGM3	GCTG--GG-ACTT-CCAGCGAA-----	A--C--GACATGGGC-----	T-C-----TTTA--CTCAGAG-----A-----A--		9038	
SIVLHOEST	GCTG--GG-ACTT-CCAGC.AA-----	C--GTCATGGGC-----	TAC-----TTTA--CTCAGAG-----A-----A--		9081	
SIVSUN	C--A-TGGCTTG--GA-GAGCA-----	AC--GCC--GGA-T-----	TT--A--TTGCTGCAT--AGCAGCTGCT--		9669	
SIVMNDGB1	CTAGCGC-T-CG-GC-AGAG-A-----	G-AC--GAG--GGA-C-----	TTA--A--TT-CTGCAT--AGCAGCTGCT--		9713	
SIVSAB1C	ACTGG-GCTT-CGCGCTA-GAA-----	AAC--GAG--GGA-C-----	TC--CAT-C-G--AT--AGCAGCTGCAT--		9029	
SIVSYK173	CAT--AG-ACTG--GA-T-AGA-----	T--A-ACTGGGC-----	TACTG--A-T-GCTT-CC--G-GCTGCT-AAT--AGCAGCTGCT--		9736	
	A--G--GTCAG-AC-TA-TAA-AACA-C--TA--A--A-TCCATGGT-A.....	CAAGCTCGGCCAGGGGAGGA-CCTG--C-G-G-AA-GCT-AC-CT-CTGCAT-A-AGCCG-TGCT--			9362	

B_FR.HXB2R	3' LTR U3 end \ / 3' LTR R repeat start TTTTGCCTGTACTGGGTCCTCTCTGGTT...AGACCAGAT..CTGAGCC.TGGGAGCTCTCTGGCTAACT.....	9588
	<pre> TAR element stem ^^^ bulge stem </pre>	
A_UG.U455	GC-GCTT-TCG-CT-TA--GGG-CTC-C TTGTT...AGACCAGAT..-GA-CC.TGGGAGCT--CTGGCTAGCG.....	9047
C_ET.ETH2220	---C--T---C-----A-G-.....-T-----	8988
F1_BE.VI850	---C-----A-----T-----G-----	8886
G_SE.SE6165	-C-C-----T-----T-----G-A-----	9032
H_CF.90CF056	---C--T-----T-----G-----	8906
J_SE.SE92809	---C--T-----T-----G-----	8887
CRF01_AE_TH.CM240	---C--T-----T-----G-----	9154
CRF02_AG_NG.IBNG	-C-C-----T-C-----G-----G-A-----	9115
CRF03_AB_RU.KAL153	---C-----T-----G-----	9582
CRF04_cpx_CY.94CY03	-C-C-----T-----G-----	9003
N_CM.YBF30	-C-C--T-----T-C-...G-----TA--T-----ATAT--G-----	9182
O_MVP5180	---CC--T---C---TAG-TAGA...G---G---C---C---CTC-----	9703
O_CM.ANT70	---C--T---C---GG-TAGA...G---G---C---C---CTCTA-----	9708
HIV2_A_ST	AC-C--A-TGTA-TCAGTCGCTCT-CGGAG--G-TG-CAGAT----C----G-TCTCTC-AGCAC.....TAGCAGGTAGAGCCTGGGTGTTCCCTGCTAG	9583
HIV2_A_BEN	-C---A-TGTA-TCAGTCGCTCT-CGGAG--G-TG-CAGATC---C--A--G-TCTCTC-AGCAC.....TAGCAGGTAGAGCCTGGGTGTTCCCTGCTGG	10144
HIV2_A_ALI	GC---A-TGTA-TCAGTCGCTCT-CGGAG--G-TG-CAGAT----C----G-TCTCTC-AGCAC.....TAGCAGGTAGAGCCTGGGTGTTCCCTGCTAG	10139
HIV2_B_D205	-C-C--A-TGTA-TC.....	9978
HIV2_B_EHO	--G--A-TGTA-TCAGTCGCTCT-CGGAG--G-TG-CAGATC---C---G-TCTCTC-AGCAC.....TAGCAGGTAGAGCCTGGGTGTTCCCTGCTAG	10034
STMM83293	-A-A--TCTGTA-TCAGTCACTCT-CGGAG--G-TG-CAGAT---C---G-TCTCTC-AGCAC.....TAGCAGGTAGAGCCTGGGTGTTCCCTGCTAG	9718
SMM251	---C--TCTGTA-TCAGTCGCTCT-CGGAG--G-TG-CAGAT---C---G-TCTCTC-AGCAC.....TAGCAGGTAGAGCCTGGGTGTTCCCTGCTAG	10034
SMM9	---C--TCTGTA-TCAGTCACTCT-CGGAG--G-TG-CAGAT---C---G-TCTCTC-AGCAC.....TAGCAGGTAGAGCCTGGGTGTTCCCTGCTGG	9547
SIVCPZGAB	---C-----T-CAC-...G-----T-----AC-AGTGA-----	9681
SIVCPZUS	C-C--T-----C-C-...AA-----C-----G-----	9652
SIVSTAN1	CGC--G--TGTAAGTCAG---TAC-AGG---C-----T--G---TG-G-C-----	9619
SIVGRI677	CGC--G--TGTAAGTCAG---TAC-AGG---C...A-----T-T--G---T--G-C-----	9286
SIVVER9063	CGC--G--TGTAAGTCAG---TAC-AGG---C...-----T-T--G---T--G-C-----	9646
SIVVER155	CGC--G--TGTAAGTCAG---TAC-AGG---C...-----T-T--G---T--G-C-----	9627
SIVVERTYO	CGC--G--TGTAAGTCAG---TAC-AGG---C...-----T-T--G---T--G-C-----	9103
SIVVERAGM3	CGC--G--TGTAAGTCAG---TAC-AGG---C...-----T-T--G---T--G-C-----	9146
SIVLHOEST	CCGCTTG-TCGGGA-TC--TG--CTGGGACC-G---TCGCTCCT-GGTGA-C-CACTGG-CAGCT-AG.....CAGTCAGCCCTGGGAGGCTACTACT	9762
SIVSUN	-CGCCTG-TCGGGA-TCTCTG--CAGAGGATAGG-GACGAC-CTGAGAGTCTCA-CTAGA-CTGGCT-.....GTCGCTCCTGAGTGAGCACTTGCT	9806
SIVMNDGB1	-CGCTTG-TCGGGA-TCTCTA--ACAGAGGCT-AGG-T-GTA-CTCTGAGCA--T-C-CT-A-AGC-AGGACCAGAG.....TCTCTGAGTGACTGGGTC	9123
SIVSAB1C	GCGCCTG-AA-ACA-TCTCTG--T-GAGGCT-C-AGAT-GAGCCT-GGTGTCTCTGG-AA-T--G-ACCAGCTTGAGCCTGGGTGTTGCTGGGGAAGGCAAGCCTAAGAGCACTCTGGTGAGGTCTTG	9866
SIVSYK173	G-AGCTA-CCG-G-A--TG-AACTAGCTCTCTC-AGCGCCTG-GAG-CCT-AGAGA--GGTAGAGCGG.....CCCCGGAGCCCTTGCCAGTAAAGGCTAGTGG	9463

TAR elements of HIV-2 and SIV have extra stem-loop, see Berkhout, Nucleic Acids Res 20(1):27-31 (1992)
 Poly-A signal

B_FR.HXB2RAGGGAACCCACTGCTTAAGCCTC...AATAAA.GCTTGCTTGAGTGCTTC...AAGTAGTGTGTGCCCGTCTGTTG.TGTGACTCTGG 9669
 |||||

A_UG.U455AG--A-C--ACTGCT-A-.....-GTGC-TG.....-TA-T----- 9129
 C_ET.ETH2220G-----G----- 9031
 FI_BE.VI850G-----G----- 8903
 G_SE.SE6165GG--A-C--ACTGCT-A-.....-AA-CC.. 9074
 H_CF.90CF056-A-----T... 8953
 J_SE.SE92809G--GCAT..GCAAGCC- 8943
 CRF01_AE_TH.CM240AG--A-C--ACTGCT-A-.....-A-TG----- 9203
 CRF02_AG_NG.IBNGGA----- 9197
 CRF03_AB_RU.KAL153-A-----NNNN.NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 9665
 CRF04_cpx_CY.94CY03 9050
 N_CM.YBF30 9182
 O_MVP5180TAGCT----G-----CG-----AGAA..C---GTGTGC-CAT-TGT-CAACCTG-TG---A- 9788
 O_CM.ANT70GCT-----G--CG---CG-----AGCTT..GGTCT-GAA--TTGATC-CA-GTTA-CAGTC-AAT 9791
 HIV2_A_STACTCTCACCAGTGTCTGGCCGGCACTGGGCAGA.CGGCTCCAC-CTTG-TTGCT-A--AGAC-TCTT---GCTGCCAG--AGAA--A.. 9672
 HIV2_A_BENACTCTCACCAGTGTCTGGCCGGCACTGGGCAGA.CGGCTCCAC-CTTG-TTGCT-A--A-CTCTT---GCTGCCAG--AGAA--AGGT..T-AGT-----T---A---C-CCTA--CG-CGCCT 10271
 HIV2_A_ALIACTCTCACCAGTGTCTGGCCGGCACTGGGCAGA.CGGCTCCAC-CTTG-TTGCT-A--A.C-TCTT---GCTGCCAA--AGAA--AGGT..T-AAG-----T---A---C-CCTA--CG-CGCCT 10265
 HIV2_B_EHOACTCTCACCAGCACTTGGCCAGTGTGGGTAGAGTGGCTCCAC-CTTG-TTGCT-AT-TATCTCTC---ACTGCCAA--AGAA--AAGA..C-AGT-----T---A---C-CCTA--CG-CGCCT 10160
 STMM83293ACTCTCACCAGCACTTGGCCAGTGTGGGTAGAGTGGCTCCAC-CTTG-TTGCT-AT-TATCTCTC---GCTGCCAA--T--AAG-AAG..C-AGT-G---T---A---C-CCTA--CG-CGCCT 9846
 SMM251ACTCTCACCAGCACTTGGCCAGTGTGGGCAGAGTGGCTCCAC-CTTG-TTGCT-A--A-CTCTT---GCTGCCAT--T--AAG-AAG..CCAGT-----T---A---C-CCTA--CG-CGCCT 10162
 SMM9ACTCTCACCAGCACTTGGCCAGTGTGGGCAGAGTGGCTCCAC-CTTG-TTGCT-A--A-CTCTT---G---GC-AG-T--AAG-A.. 9638
 SIVCPZGABA-----TA-TT.G-C-----TA--A--C.A-AC----- 9762
 SIVCPZUSGTA-----T.....TA-CT-CA.. 9732
 SIVSTAN1TAACCTGGTCTGGCCATCCAGGGGTAAG--CCTC-TTG--TAGA-G--TATTGCT...G----GCTTAAGAGTC..T-TCT-A-..CAAGT-C--CA-TGC-C-C-CT-CC 9731
 SIVGRI677TAACCCGGT.TGGCCACC.GGGGGTAAGGA.CTC-TTGGCTTCAT-TAGCTC...C-...TCGCTTA-TCG-T.AT-T-G-A-.TCAAGT-CTCA--C-C-C-GAGCC 9394
 SIVVER9063TAACCTGGT.TGGCCACC.AGGGGTAAGGA.CTC-TTGGCTTGG--AGCT...ACT----GC-T-AGAG-T..T-TCT-A-.TCAAGT-C-CTCATTGAC-C---ACT 9753
 SIVVER155TAACCTGGT.TGGCCACC.AGGGGTAAGGA.CTC-TTGGCTTGG--AGCT...ACT----GC-T-AGAG-T..T-TCT-A-.TCAAGT-C-CTCATTAA--C---ACT 9732
 SIVVERTYOTAACCTGGT.TGGCCACC.AGGGGTAAGGA.CTC-TTGGCTTGG--AGCT...ACT----GC-T-AG.. 9170
 SIVVERAGM3TAACCTGGT.TGGCCACC.AGGGGTAAGGA.CTC-TTGGCTTGG--AGCT...T---CG--GC-T-AGGCAG..-GACTTGTGAGAGA-A-TGCC--G--CAT-CAGCA 9253
 SIVLHOESTTAGCACCACCTGGGCCTGATCAGCCTAGGGAGCTAAGGGCTCCT-G-TTGCTAC-TTAAAG-CTTC---AAG---TCATT--TAGAAAGT-CAAGC-AGTGT-T-A-AGTAGTC-C-TT---CTT 9892
 SIVSUNAAGCACCTCACCCGAGCCTGATCAGCTTGGGGAGCTAGAGGCTCCTG-TT---AGT-CAAG-CTTC---GC-GC--AGT-TTACTG..C-AGCAA---T---G-TC-CTCT-TCTC--TC 9933
 SIVMNDGB1TGAGCACCTCACTCGGGCTGATCAGCTCGAGGTAGTGA-CTCCTTG-TTGCT-GCTATTGTCTTC---GTAACCTAGATTGA-AGCA 9215
 SIVSAB1CGTAGACCTCGCCTGGGCAGTGGCCATTGGCAGTAGCAGAGACTCCG-TTGCTTGCTT-AT-GCCTC---GT-----AGT-TACAAGTA---C-AGTGTGTG--CAT-TA-TCCTCAGGA-AA 9996
 SIVSYK173GTCCGGGTTTGGCCGCCGGTCTTCGGCAGCTCTCCGTTGCTT-C-TTGTCT-A-GCTGAT-AGCCCA-----GTG-G-TAAACGCAAG-AACGTCT-CTC-CATCTGTTG-GCTGG-TCACCT-C-GAA 9593

	3' LTR end	
B_FR.HXB2R	TAACTAGAGATCCCTCAGACCCCTTTTGTAGTGTGGAAAAATCTCTAGCA	9719
A_UG.U455	-----AC-A--A-T---.A-----	9178
CRF02_AG_NG.IBNG	-----	9201
CRF03_AB_RU.KAL153	NN	9715
O_MVP5180	AG-TC	9793
O_CM.ANT70	ATGACCCTA-C	9802
HIV2_A_BEN	GGT-ATTC-G-GTTCATCTGAG-AAC-AGACCCTG-TCTGT-AGGACC-TTCTCGCTTTGGGAATCCAAGGCAGGAAAAATCCCTAGCA	10359
HIV2_A_ALI	GGT-ATTC-G-GTTCACCTGAG-AAC-AGACCCTG-TCTGT-AGGACC-TTCTCGCTTTGGGAAACGGAGGCAGGAAAAATCCCTAGCA	10353
HIV2_B_EHO	GGT-ATTC-G-G-TC--CT.....AGAAACCCTG-TCTGT-AGGACC-CTTCTGCTTTGGGAAACCAAGGCAGGAAAAATCCCTAGCA	10242
STMM83293	GGT-A-CTCGGTA----AGA--C-GGTC-GTTA-GACCCTT---GC	9892
SMM251	GGT-A-CTCGGTA---G-TAATAAGA--A-CC--GTCTGTAGGACCCCTTCTGCTTTGAGAAACCGAAGCAGGAAAAATCCCTAGCA	10249
SIVCPZGAB	-----TTAAA---TAGTCAAG- .-----	9811
SIVCPZUS	-----A-TC--A-AAGTAG---.-----	9781
SIVSTAN1	-C-GGGA-CCCT-G.....TTAC-GG	9752
SIVGRI677	-CTAG--GTGAA--.TCTCTTAC-G	9418
SIVVER9063	C--GC--G-GAA--GTTCTTAC-GG	9779
SIVVER155	CTTGA-CG-GAGAAGTTCCTTAC-GG	9758
SIVVERAGM3	C--G-GAG-GAAGAC-CTGATGGAA-CAA-CA--GA---GT	9294
SIVLHOEST	CGC-CT-GTTCAGTGATCT-T-AC-AGAGAGA-TG-AGCCT-G-GATC-GGGACAGGGCTTCCAG	9957
SIVSUN	---ACCCT-T--GACTGCT-T-AGC---GG--CAATT-CTGG-CGCCTAGTGATCCGGGGCTGACGGCTAGCA	10006
SIVSAB1C	CCTGGTT-CTAAG	10009
SIVSYK173	GGTG	9597