SUPPLEMENTAL METHODS

Assessment of HIV-1 sequence diversity and divergence.

Nucleotide sequence diversity and divergence were estimated using the program SENDBS(1) using the HKY 85 plus gamma model. Parameters examined for each twin included: the average pairwise diversity across the 1995 and 2000 time points (dt), the average pairwise diversity in 1995 (dx), the average pairwise diversity in 2000 (dy), the average nucleotide divergence between the 1995 and 2000 time points (dxy), and the net nucleotide divergence between samples (da). The net nucleotide divergence was defined as the average diversity corrected for within-sample variation (1). Estimates of standard errors were calculated using 1000 bootstrap replications.

Statistical comparisons

Quantitative variables were compared using the Wilcoxon rank sum test. Categorical variables were compared using Fisher's exact test. Correlation was measured using Pearson's correlation coefficient.

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure 1. T cell receptor Vβ spectratype profiles of the twins. T cell receptor Vβ spectratype profiles of the twins from 12/00 (A) and 6/01 (B) are shown.

Supplemental Figure 2. HIV-1 sequences of the twins aligned against clade B consensus, NL4-3, and JR-CSF sequences. Sequencing of proviral DNA from 2000 was performed as described in the main text. Consensus amino acids for each twin were derived when at least three sequences were available for comparison. The consensus sequences for each twin are aligned against the Los Alamos HIV Database HIV-1 clade B consensus, NL4-3, and JR-CSF sequences (these latter two strains are utilized in our laboratories). Regions corresponding to recognized Clade B consensus peptides or regions of overlap of consecutive peptides are boxed.

- indicates same amino acid as clade B consensus
x indicates lack of consensus >50%

SUPPLEMENTAL TABLE

Env

Twin	1995 diversity, dx (%)	se	2000 diversity, dy (%)	se	Total Divergence dxy, %	se	Net divergence, da %	se
1-05	4.4	0.59	0.53	0.79	23.12	2.57	20.65	2.49
1-06	6.42	0.77	0.64	0.54	8.96	1.1	5.43	0.96

Between twins

	dxy %	se	da %	se
1994	12.48	1.38	7.1	1.13
2000	19.47	2.38	18.89	2.4

Nef

Twin	1995 diversity, dx (%)	se	2000 diversity, dy (%)	se	Total Divergence dxy, %	se	Net divergence, da %	se
1-05	1.7	0.35	1.4	0.32	15.6	1.7	14.0	1.7
1-06	0.6	0.18	1.9	0.47	5.5	0.9	4.2	0.8

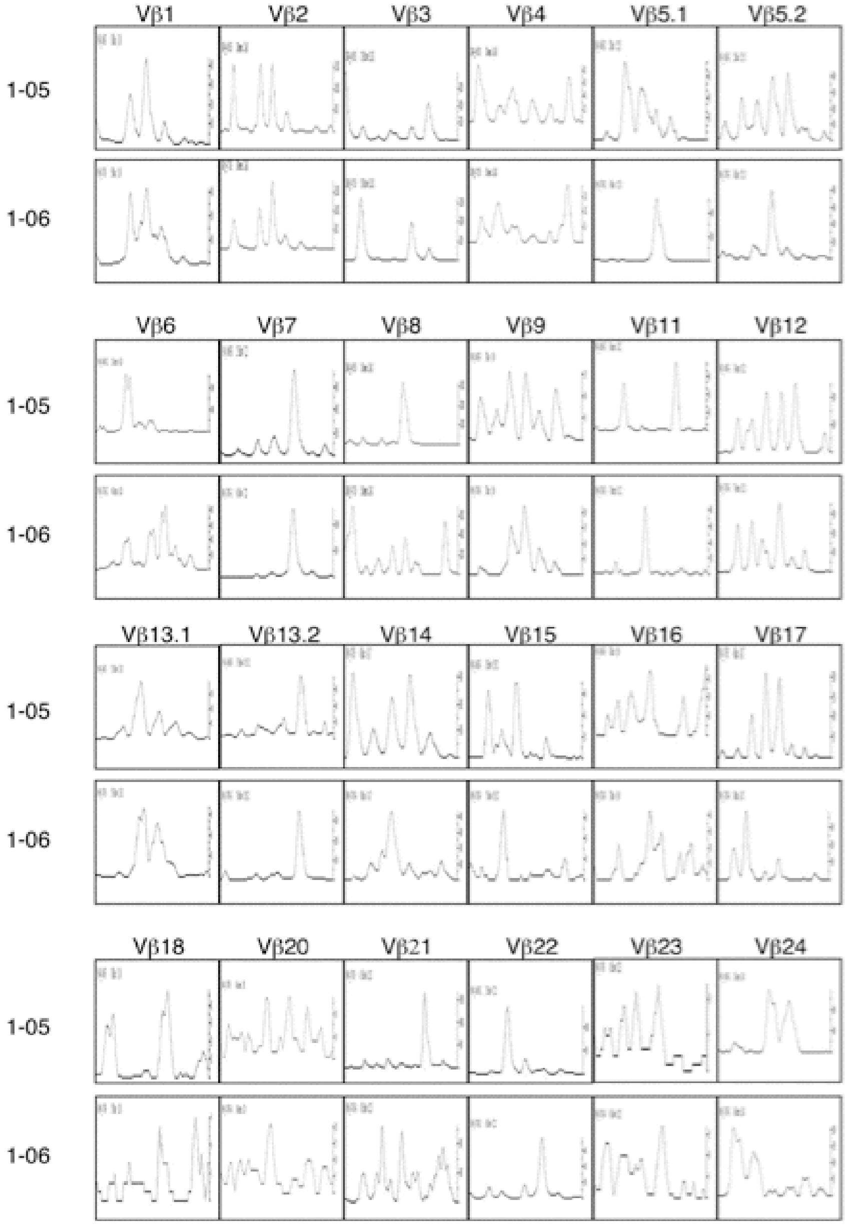
Between twins

	dxy %	se	da %	se
1994	9.6	1.3	8.5	1.2
2000	15.9	1.7	14.3	1.7

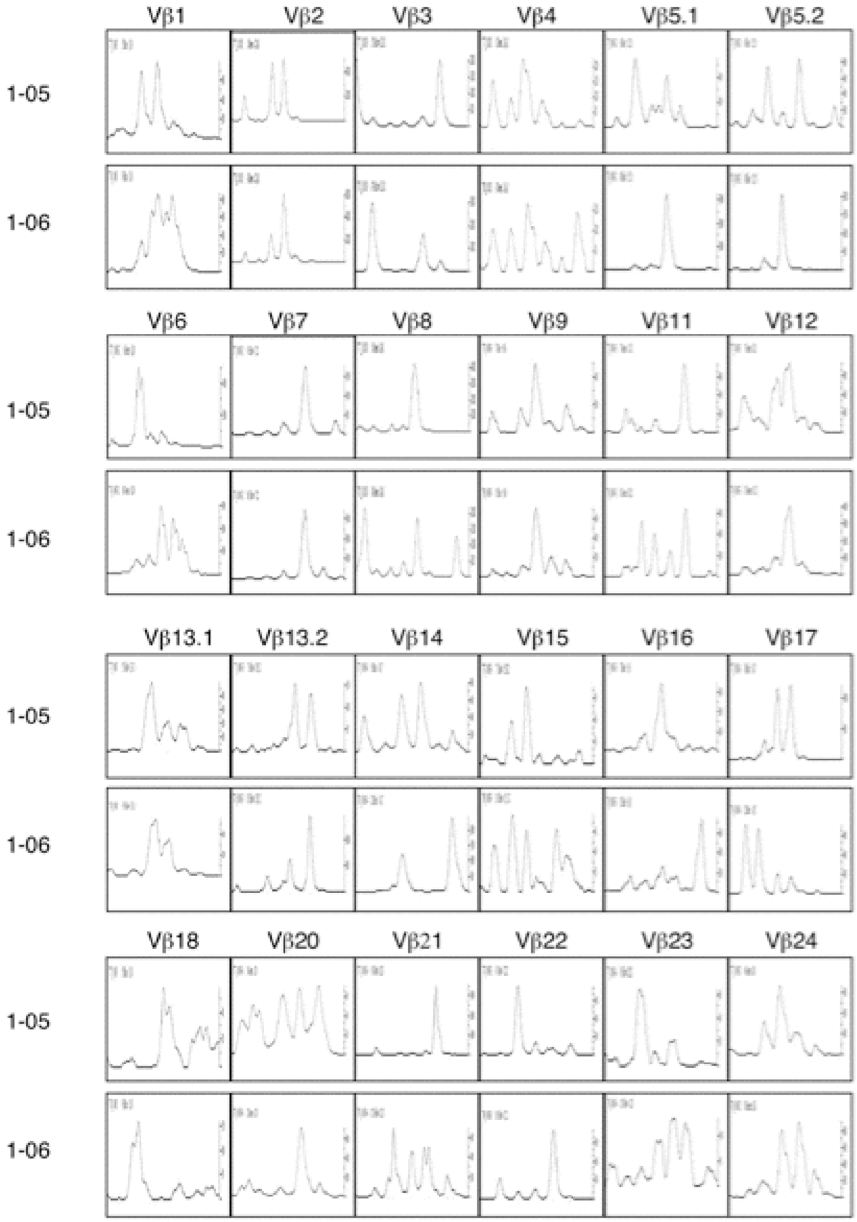
Supplemental Table 1. Diversity and divergence of HIV-1 sequences. Nucleotide sequence diversity and divergence were estimated using the program SENDBS using the HKY 85 plus gamma model. Parameters examined for each twin include: the average pairwise diversity across the 1995 and 2000 time points (dt), the average pairwise diversity in 1995 (dx), the average pairwise diversity in 2000 (dy), the average nucleotide divergence between the 1995 and 2000 time points (dxy), and the net nucleotide divergence between samples (da). The net nucleotide divergence is defined as the average diversity corrected for within-sample variation. Estimates of standard errors are calculated using 1000 bootstrap replications.

SUPPLEMENTAL REFERENCES

1. **Nei, M., and L. Jin.** 1989. Variances of the average numbers of nucleotide substitutions within and between populations. Mol Biol Evol **6:**290-300.



SUPPLEMENTAL FIGURE 1A



SUPPLEMENTAL FIGURE 1B

SUPPLEMENTAL FIGURE 2

Gag Full Length

```
ConsensusB_Gag    MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI
NL4-3_Gag         ---------- ----K----- -------Q-- ---------- ---------- ----------
JR-CSF_Gag        ---------- ---------- ---------R ---------- ---------- --S-------
Twin5_Gag         ------I--- ---------- ---------- ---------- --------S- ----Q--K--
Twin6_Gag         ---------- -----L---- S--------R ---------- ---------- --------K-

ConsensusB_Gag    LGQLQPSLQT GSEELRSLYN TVATLYCVHQ RIEVKDTKEA LEKIEEEQNK SKKKAQQAAA
NL4-3_Gag         ---------- ---------- -I-V------ --D------- -D-------- ----------
JR-CSF_Gag        --------K- -----T---- ---------- ---I------ --------T- -M--------
Twin5_Gag         ME------K- --------F- A--V------ --D------- -D-------E ----V-----
Twin6_Gag         -V-------- ---------- A--V------ ---I------ -D-------- ----------

ConsensusB_Gag    DTGNSSQV** ***SQNYPIV QNLQGQMVHQ AISPRTLNAW VKVVEEKAFS PEVIPMFSAL
NL4-3_Gag         ----N---** ***------- ---------- ---------- ---------- ----------
JR-CSF_Gag        --------** ***------- ---------- ---------- ---I------ ----------
Twin5_Gag         --------** SQV------- --A------- ---------- ---------- ----------
Twin6_Gag         ---SC--ASC GQT------- --I------- ---------- A--------- ----------

ConsensusB_Gag    SEGATPQDLN TMLNTVGGHQ AAMQMLKETI NEEAAEWDRL HPVHAGPIAP GQMREPRGSD
NL4-3_Gag         ---------- ---------- ---------- ---------- ---------- ----------
JR-CSF_Gag        ---------- ---------- ---------- ---------- ---------- ----------
Twin5_Gag         ---------- ---------- ---H------ ---------- ---------- ----------
Twin6_Gag         ---------- ---------- ---------- ---------- --PQ---V-- --L-D-----

ConsensusB_Gag    IAGTTSTLQE QIGWMTNNPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI RQGPKEPFRD
NL4-3_Gag         ---------- ------H--- ---------- ---------- ---------- ----------
JR-CSF_Gag        ---------- ---------- ---------- ---------- ----V----- ----------
Twin5_Gag         ---------- ------H--- ---------- ---------- ---------- K---------
Twin6_Gag         ---------- --Q---A--- ---------- ---------- ----A----- ----------

ConsensusB_Gag    YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK ALGPAATLEE MMTACQGVGG
NL4-3_Gag         ---------- ---------- ---------- ---------- ----G----- ----------
JR-CSF_Gag        ---------- ----T----- ---------- ---------- ---------- ----------
Twin5_Gag         ---------- ---------- ---------- ---------- ---------- ----------
Twin6_Gag         ---------- ---------- ---------- ---------- ---------- ----------

ConsensusB_Gag    PGHKARVLAE AMSQV**TNS ATIMMQRGNF RNQRKTVKCF NCGKEGHIAK NCRAPRKKGC
NL4-3_Gag         ---------- -----**--P ----I-K--- ---------- ---------- ----------
JR-CSF_Gag        ---------- -----**--P ---------- -----N---- ---------R ----------
Twin5_Gag         ---------- -----**--- -N-------- K-----P--- ---------R ----------
Twin6_Gag         ---------- -----NLA-- ---------V -----I---- ---------R ----------

ConsensusB_Gag    WKCGKEGHQM KDCTERQANF LGKIWPSHKG RPGNFLQSRP EPTAPPEESF RFGEETTTPS
NL4-3_Gag         ---------- ---------- ---------- ---------- ---------- ----------
JR-CSF_Gag        ---------- -E-------- -------Y-- ---------- ---------- ------A---
Twin5_Gag         ---------- ---------- ---L------ -------T-- --S------L -----KA---
Twin6_Gag         ---------- ---------- ------AY-- -----P---L ---------- K-R-------

ConsensusB_Gag    QKQ******E PIDKELYPLA SLRSLFGNDP SSQ
NL4-3_Gag         ---******- ---------- -------S-- ---
JR-CSF_Gag        ---EQKQ**- ---------T ---------- ---
Twin5_Gag         -N-******- K--------T ---------- ---
Twin6_Gag         ---EPPQKQ- Q--------T ---------- ---
```

Pol 57-908

```
ConsensusB_Pol    PQITLWQRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD
NL4-3_Pol         ---------- ---------- ---------- ---------- ---------- ------G---
JR-CSF_Pol        ---------- ---------- ---------- ----D-D--- ---------- ----------
Twin5_Pol         ---------I ---------- ---------- ---------- -----I---- --LL------
Twin6_Pol         ---------F --V----E-- ---------- ---------- -----I---- ----R-----

ConsensusB_Pol    QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNFP ISPIETVPVK LKPGMDGPKV
NL4-3_Pol         ---------- ---------- ---------- ---------- ---------- ----------
JR-CSF_Pol        --P-D----- -V-------- ---------- ---------- ---------- ----------
Twin5_Pol         --P------- VE-S--I--- ---------M --L------- ---------- ----------
Twin6_Pol         --QV------ ---A--I--- ---V------ ---------- ---------- ----------

ConsensusB_Pol    KQWPLTEEKI KALVEICTEM EKEGKISKIG PENPYNTPVF AIKKKDSTKW RKLVDFRELN
NL4-3_Pol         ---------- ---------- ---------- ---------- ---------- ----------
JR-CSF_Pol        ---------- ---------- ---------- ---------- ---------- ----------
Twin5_Pol         ---------- ---I---K-L -ED------- ---------- -------D-- ----------
Twin6_Pol         ---------- ---I------ ---------- ---------- -----N-DR- ----------

ConsensusB_Pol    KRTQDFWEVQ LGIPHPAGLK KKKSVTVLDV GDAYFSVPLD KDFRKYTAFT IPSINNETPG
NL4-3_Pol         ---------- ---------- Q--------- ---------- ---------- ----------
JR-CSF_Pol        R--------- ---------- ---------- ---------- ---------- ----------
Twin5_Pol         ---------- ---------- ---------- ------I--- E--------- ----------
Twin6_Pol         -K-------- ------G--- ---------- ---------- ---------- ---L------

ConsensusB_Pol    IRYQYNVLPQ GWKGSPAIFQ SSMTKILEPF RKQNPDIVIY QYMDDLYVGS DLEIGQHRTK
NL4-3_Pol         ---------- ---------- C--------- ---------- ---------- ----------
JR-CSF_Pol        ---------- ---------- ---------- -------I-- ---------- ----------
Twin5_Pol         ---------- ------S--- T--------- ---------- ---------- ----E-----
Twin6_Pol         ---------- ---------- X--------- -------I-- --------X- ----------

ConsensusB_Pol    IEELRQHLLR WGFTTPDKKH QKEPPFLWMG YELHPDKWTV QPIVLPEKDS WTVNDIQKLV
NL4-3_Pol         ---------- ---------- ---------- ---------- ---------- ----------
JR-CSF_Pol        ---------K ---------- ---------- ---------- ---------- ----------
Twin5_Pol         -K---EY-WK ---Y---R-- -Q----R--- ---------- ---------- ----------
Twin6_Pol         ---------K ---F---Q-- ------H--- ---------- ---------- ----------

ConsensusB_Pol    GKLNWASQIY AGIKVKQLCK LLRGTKALTE VIPLTEEAEL ELAENREILK EPVHGVYYDP
NL4-3_Pol         ---------- -----R---- ---------- -V-------- ---------- ----------
JR-CSF_Pol        ---------- ---------- ---------- -----K---- ---------- ----------
Twin5_Pol         ---------- P--------- -IK------- IV---K---- ---------- ----------
Twin6_Pol         ---------- T--------- ---------- IV---A---- ---------- ----------

ConsensusB_Pol    SKDLIAEIQK QGQGQWTYQI YQEPFKNLKT GKYARMRGAH TNDVKQLTEA VQKIATESIV
NL4-3_Pol         ---------- ---------- ---------- ------K--- ---------- ----------
JR-CSF_Pol        -----V---- ---------- F--------- -----T---- ---------- -----N----
Twin5_Pol         --E------- ------S--- ---------- ----K-KA-- ---------- ----------
Twin6_Pol         ---------- ---------- ---------- ----K-KS-- ----R----- ----T----X

ConsensusB_Pol    IWGKTPKFKL PIQKETWEAW WTEYWQATWI PEWEFVNTPP LVKLWYQLEK EPIVGAETFY
NL4-3_Pol         ---------- ---------- ---------- ---------- ---------- ---I------
JR-CSF_Pol        ----I----- --------T- ---------- ---------- ---------- ----------
Twin5_Pol         --------R- ---------- --D------- ---------- ---------- ----------
Twin6_Pol         ---------- --------T- ---------- ---------- ---------- ----------

ConsensusB_Pol    VDGAANRETK LGKAGYVTDR GRQKVVSLTD TTNQKTELQA IHLALQDSGL EVNIVTDSQY
NL4-3_Pol         ---------- ---------- ------P--- ---------- ---------- ----------
JR-CSF_Pol        ---------- --------S- ---------- ---------- ---------- ----------
Twin5_Pol         ---------- --------NK -------I-- ---------- -Y-------- ----------
Twin6_Pol         ---------- --------N- ---------- ---------- -Y-------- --------??

ConsensusB_Pol    ALGIIQAQPD KSESELVSQI IEQLIKKEKV YLAWVPAHKG IGGNEQVDKL VSAGIRKVLF
NL4-3_Pol         ---------- ---------- ---------- ---------- --------G- ----------
```

```
JR-CSF_Pol        ---------- ---------- ---------- ---------- ---------- ----------
Twin5_Pol         ---------- ---------- --E------- ---------- ---------- --S-------
Twin6_Pol         ?????????? ?????????? ?????????? ?????????? ?????????? ?????????-

ConsensusB_Pol    LDGIDKAQEE HEKYHSNWRA MASDFNLPPV VAKEIVASCD KCQLKGEAMH GQVDCSPGIW
NL4-3_Pol         ---------- ---------- ---------- ---------- ---------- ----------
JR-CSF_Pol        ---------D ---------- ---------I ---------- ---------- ----------
Twin5_Pol         ---------- -----N---- ---------- ---------- ---------- ----------
Twin6_Pol         ---------- -----N---- ------I--- -------C-- -------P-- ----------

ConsensusB_Pol    QLDCTHLEGK IILVAVHVAS GYIEAEVIPA ETGQETAYFL LKLAGRWPVK TIHTDNGSNF
NL4-3_Pol         ---------- V--------- ---------- ---------- ---------- -V--------
JR-CSF_Pol        ---------- ---------- ---------- ---------- ---------T ----------
Twin5_Pol         ---------- ---------- ---------- ---------I ---------- ----------
Twin6_Pol         ---------- ---------- --------S- ---------- ---------- ----------

ConsensusB_Pol    TSTTVKAACW WAGIKQEFGI PYNPQSQGVV ESMNKELKKI IGQVRDQAEH LKTAVQMAVF
NL4-3_Pol         ---------- ---------- ---------I ---------- ---------- ----------
JR-CSF_Pol        ---------- ---------- ---------- ---------- ---------- ----------
Twin5_Pol         ---------- ---------- ---------- ----N----- ---------- ----------
Twin6_Pol         ---------- ---------- ---------- ---------- ---------- ----------

ConsensusB_Pol    IHNFKRKGGI GG
NL4-3_Pol         ---------- --
JR-CSF_Pol        ---------- --
Twin5_Pol         ---------- --
Twin6_Pol         ---------- --
```

Env gp120 267-463

```
ConsensusB_Env    EEEVVIRSEN FTDNAKTIIV QLNESVEINC TRPNNNTRKS IHIG**PGRA FYTTGEIIGD
NL4-3_Env         --D-----A- ---------- ---T------ ---------- -R-QRG---- -V-I-K-*-N
JR-CSF_Env        --K-----D- ---------- ------K--- ---S------ ---------- ----------
Twin5_Env         --QI----A- -SN---I--- --K-L-K-H- M--------- ---A------ -----Q---N
Twin6_Env         -K--I---D- -S----I--- ---------- --------NR LS-------S ----RQ-V--

ConsensusB_Env    IRQAHCNISR AKWNNTLKQI VIKLREQFG* NKTIVFNQSS GGDPEIVMHS FNCGGEFFYC
NL4-3_Env         M--------- ----A----- AS-------N ----I-K--- -------T-- ----------
JR-CSF_Env        ---------- -Q-------- -E------N- ------TH-- ---------- ----------
Twin5_Env         ---------V T---K--I-L TD--K-K-*- -------H-- ---------- ----------
Twin6_Env         -------L-S T------G-- -K------K- ----K-EK-- ---------- ----------

ConsensusB_Env    NTTQLFNSTW NGTWNDTEGN IT******** LPCRIKQIIN MWQEVGKAMY APPIRGQIRC
NL4-3_Env         -S-------- FNSTWS---S NNTEGSDTIT -------F-- ---------- ----S-----
JR-CSF_Env        -S-------- -D-EKSSGTE GNDTII---- ---------- ---------- ----K-----
Twin5_Env         --SK--SG-- --SSWTGAHW SNSTNDTII- -------F-- ------R--- ----K---K-
Twin6_Env         -L-K------ -VNGTIGTN- TNNNDTII-T ---------- ---K------ ----E---K-

ConsensusB_Env    SSNITGLLLT RDGGN***** ***NET
NL4-3_Env         ---------- ---------- -NN-GS
JR-CSF_Env        ---------- ----K----- -----S
Twin5_Env         I--------- ----KVDNKT THE---
Twin6_Env         ---------- -----DKNGT GNGT——
```

Vpr 4-96

```
ConsensusB_Vpr    APEDQGPQRE PYNEWTLELL EELKNEAVRH FPRIWLHGLG QHIYETYGDT WAGVEAIIRI
NL4-3_Vpr         ---------- ---------- ----S----- -------N-- ---------- ----------
JR-CSF_Vpr        ---------- ---------- ---------- -------S-- -Y-------- ----------
Twin5_Vpr         ---------- ---Q------ ---------- ---------- ---------- -E-------T
Twin6_Vpr         ---------- -H-----N-- ---------- --GP--Y--- -Y-------- ------M-CM

ConsensusB_Vpr    LQQLLFIHFR IGCQHSRIGI TRQRRARNGA SRS
NL4-3_Vpr         ---------- ---R-----V ---------- ---
JR-CSF_Vpr        ---------- ---R------ ---------- ---
Twin5_Vpr         ------V--- ---R------ -P---T---- ---
Twin6_Vpr         ---------- ---R------ -P---T---- ---
```

Tat 1-72

```
ConsensusB_Tat    MEPVDPRLEP WKHPGSQPKT ACTNCYCKKC CFHCQVCFTT KGLGISYGRK KRRQRRRAPQ
NL4-3_Tat         ---------- ---------- ---------- --------M- -A-------- --------H-
JR-CSF_Tat        ------S--- ---------- ---------- -L-------- ---------- -------P--
Twin5_Tat         ---------- --------G- --N------- ---------S ---------- ----------
Twin6_Tat         ---------- --------R- ---------- --------I- ---------- -------P-P

ConsensusB_Tat    DSQTHQVSLS KQ
NL4-3_Tat         N-******** **
JR-CSF_Tat        --******** **
Twin5_Tat         ----N-A--- --
Twin6_Tat         GG------P- --
```

Rev 1-25

```
ConsensusB_Rev    MAGRSGDSDE ELLKTVRLIK FLYQS
NL4-3_Rev         ---------- --IR-D**** *****
JR-CSF_Rev        ---------- D----D**** *****
Twin5_Rev         ---------- ---------- L----
Twin6_Rev         -------R-- D--QA----- --H—
```

Nef Full Length

```
ConsensusB_Nef     MGGKWSKRSV VGWPTVRERM RRAE****** ****PAADGV GAVSRDLEKH GAITSSNTAA
60
NL4-3_Nef          -------S-- I---A----- ----****** ****------ ---------- ----------
JR-CSF_Nef         -------H-- P--S------ ----PATDRV RQTE---V-- ---------- ----------
Twin5_Nef          --N----M-* *--TR----- ----PAA*** ***E------ --------R- --------X-
Twin6_Nef          --S-C--C-- P--AA----- --TD****** ****----R- R-----P-Q- --S------T

ConsensusB_Nef     NNADCAWLEA QEEEEVGFPV RPQVPLRPMT YKGALDLSHF LKEKGGLEGL IYSQKRQDIL
120
NL4-3_Nef          ---A------ ---------- T--------- --A-V----- ---------- -H--R-----
JR-CSF_Nef         T--------- Y-D------- ---------- --A-I----- ---------- ----------
Twin5_Nef          T--------- --D------- T--------- --A-V-I--- ---N------ V---R-----
Twin6_Nef          S--A--MP-E --D------- ---------- F---V----- -R-------- -------E--

ConsensusB_Nef     DLWVYHTQGY FPDWQNYTPG PGIRYPLTFG WCFKLVPVEP EKVEEANEGE NNSLLHPMSL
180
NL4-3_Nef          ---I------ ---------- --V------- --Y------- D------K-- -T-----V--
JR-CSF_Nef         ---I------ --------A- --V-F----- --------D- ---------- --C------Q
Twin5_Nef          ---I------ --------K- --V------- --------D- DX----TX-- ---------Q
Twin6_Nef          ---------- ---------- --V----C-- --------D- --R----A-- S-----S--Q

ConsensusB_Nef     HGMDDPEREV LVWKFDSRLA FHHMARELHP EYYKDC
NL4-3_Nef          ---------- -E-R------ ---V------ --F-N-
JR-CSF_Nef         -------K-- -------K-- L--V------ ------
Twin5_Nef          ---E------ -------L-- -R-------- ----N-
Twin6_Nef          --TE---E-- -E-R------ ---------- -FF---
```