Table S1. Full and partial GvaELV insertions within *G. variegatus* genome

Contig	Insertion fragment
G_variegatus-3.0.2-377.4	5’LTR-*gag*-*pol*-*ORF1*-*tat*-*env*-*ORF2*-3’LTR
G_variegatus-3.0.2-2016.3	5’LTR-*gag*-*pol*-*ORF1*-*env*-3’LTR
G_variegatus- 3.0.2-5697.1	5’LTR-*gag*-*pol*-*ORF1*-*env*-*ORF2*-3’LTR
G_variegatus-3.0.2-10092.17	5’LTR-*gag*-*pol*
G_variegatus-3.0.2-2153.2	5’LTR-*gag*-*pol*
G_variegatus-3.0.2-10866.4	5’LTR-*gag*-*pol*
G_variegatus-3.0.2-8117.15	*gag*-*pol*
G_variegatus-3.0.2-152925.1	*gag*-*pol*
G_variegatus-3.0.2-8117.16	*gag*
G_variegatus-3.0.2-10866.7	*pol*-*ORF1*-*tat*-*env*
G_variegatus-3.0.2-8117.13	*pol*-*ORF1*
G_variegatus-3.0.2-49750.1	*pol*-*ORF1*
G_variegatus-3.0.2-10866.5	*pol*
G_variegatus-3.0.2-10866.6	*pol*
G_variegatus-3.0.2-146880.1	*pol*
G_variegatus-3.0.2-17753.5	*pol*
G_variegatus-3.0.2-8117.14	*pol*
G_variegatus-3.0.2-10092.18	*ORF1*-*tat*-*env*
G_variegatus-3.0.2-8117.12	*ORF1*-*tat*-*env*
G_variegatus-3.0.2-8117.11	*env*-*ORF2*
G_variegatus-3.0.2-49750.2	*env*-*ORF2*-3’LTR
G_variegatus-3.0.2-3870.9	*env*-*ORF2*-3’LTR
G_variegatus-3.0.2-10092.19	*env*-*ORF2*-3’LTR
G_variegatus-3.0.2-77437.1	*env*-*ORF2*
G_variegatus-3.0.2-104494.1	*env*-*ORF2*
G_variegatus-3.0.2-132409.1	*env*
G_variegatus-3.0.2-10866.8	*env*
G_variegatus-3.0.2-132314.1	*env*

Table S2. The contigs of *G. variegatus* genome that contains solo-LTRs

G_variegatus-3.0.2-8479.2
G_variegatus-3.0.2-2048.1
G_variegatus-3.0.2-1170.10
G_variegatus-3.0.2-3930.9
G_variegatus-3.0.2-12197.5
G_variegatus-3.0.2-2846.1
G_variegatus-3.0.2-8406.15
G_variegatus-3.0.2-4589.3
G_variegatus-3.0.2-1584.17
G_variegatus-3.0.2-493.25
G_variegatus-3.0.2-112676.1
G_variegatus-3.0.2-77.12
G_variegatus-3.0.2-4172.5
G_variegatus-3.0.2-145895.1
G_variegatus-3.0.2-7305.2
G_variegatus-3.0.2-2848.6
G_variegatus-3.0.2-110.51
G_variegatus-3.0.2-9564.8
G_variegatus-3.0.2-778.18
G_variegatus-3.0.2-2717.70
G_variegatus-3.0.2-2506.24
G_variegatus-3.0.2-2257.8
G_variegatus-3.0.2-8914.13
G_variegatus-3.0.2-9564.7
G_variegatus-3.0.2-668.1
G_variegatus-3.0.2-7320.22
G_variegatus-3.0.2-6707.5
G_variegatus-3.0.2-1260.29
G_variegatus-3.0.2-2283.10
G_variegatus-3.0.2-6765.16
G_variegatus-3.0.2-5771.10
G_variegatus-3.0.2-11917.7
G_variegatus-3.0.2-14988.4
G_variegatus-3.0.2-3285.7
G_variegatus-3.0.2-2506.28
G_variegatus-3.0.2-6246.24
G_variegatus-3.0.2-4124.7
G_variegatus-3.0.2-1032.24
G_variegatus-3.0.2-3741.2
G_variegatus-3.0.2-7903.10
G_variegatus-3.0.2-2382.24
G_variegatus-3.0.2-9815.3
G_variegatus-3.0.2-1527.22
G_variegatus-3.0.2-1410.62
G_variegatus-3.0.2-92.27
G_variegatus-3.0.2-1206.42
G_variegatus-3.0.2-5385.19
G_variegatus-3.0.2-2433.8
G_variegatus-3.0.2-1280.15
G_variegatus-3.0.2-3400.33
G_variegatus-3.0.2-1207.29
G_variegatus-3.0.2-663.4
G_variegatus-3.0.2-11862.5
G_variegatus-3.0.2-11110.6
G_variegatus-3.0.2-5691.3
G_variegatus-3.0.2-1019.36
G_variegatus-3.0.2-6976.10
G_variegatus-3.0.2-762.20
G_variegatus-3.0.2-2360.10
G_variegatus-3.0.2-31940.1
G_variegatus-3.0.2-2851.5
G_variegatus-3.0.2-1272.13
G_variegatus-3.0.2-9787.2
G_variegatus-3.0.2-2783.23
G_variegatus-3.0.2-5377.4
G_variegatus-3.0.2-985.12
G_variegatus-3.0.2-4076.14
G_variegatus-3.0.2-846.30
G_variegatus-3.0.2-728.9
G_variegatus-3.0.2-1130.20
G_variegatus-3.0.2-6208.4
G_variegatus-3.0.2-93.30
G_variegatus-3.0.2-9418.9
G_variegatus-3.0.2-7812.1
G_variegatus-3.0.2-6820.1
G_variegatus-3.0.2-1721.22
G_variegatus-3.0.2-2986.3
G_variegatus-3.0.2-3330.11
G_variegatus-3.0.2-2410.7
G_variegatus-3.0.2-22810.4
G_variegatus-3.0.2-9850.2
G_variegatus-3.0.2-1543.12
G_variegatus-3.0.2-269.14
G_variegatus-3.0.2-55185.1
G_variegatus-3.0.2-4052.5
G_variegatus-3.0.2-3530.18
G_variegatus-3.0.2-1646.44
G_variegatus-3.0.2-960.14
G_variegatus-3.0.2-749.10
G_variegatus-3.0.2-644.8
G_variegatus-3.0.2-295.43
G_variegatus-3.0.2-169378.1
G_variegatus-3.0.2-4237.2
G_variegatus-3.0.2-4540.15
G_variegatus-3.0.2-2585.7
G_variegatus-3.0.2-8664.13
G_variegatus-3.0.2-1035.34
G_variegatus-3.0.2-4587.7
G_variegatus-3.0.2-2434.52
G_variegatus-3.0.2-12462.12
G_variegatus-3.0.2-9761.9
G_variegatus-3.0.2-1899.3
G_variegatus-3.0.2-1448.2
G_variegatus-3.0.2-12001.2
G_variegatus-3.0.2-733.20
G_variegatus-3.0.2-15830.7
G_variegatus-3.0.2-99674.1
G_variegatus-3.0.2-8284.2
G_variegatus-3.0.2-11676.11
G_variegatus-3.0.2-1570.8
G_variegatus-3.0.2-509.7
G_variegatus-3.0.2-186.46
G_variegatus-3.0.2-45573.1
G_variegatus-3.0.2-5978.7
G_variegatus-3.0.2-808.25
G_variegatus-3.0.2-3938.16
G_variegatus-3.0.2-8517.13
G_variegatus-3.0.2-10076.10
G_variegatus-3.0.2-154409.1
G_variegatus-3.0.2-8557.11
G_variegatus-3.0.2-851.14
G_variegatus-3.0.2-2084.29
G_variegatus-3.0.2-828.32
G_variegatus-3.0.2-44618.1
G_variegatus-3.0.2-5300.22
G_variegatus-3.0.2-1125.12
G_variegatus-3.0.2-204.33
G_variegatus-3.0.2-1223.11
G_variegatus-3.0.2-1068.43
G_variegatus-3.0.2-128317.1
G_variegatus-3.0.2-1189.17
G_variegatus-3.0.2-9755.6
G_variegatus-3.0.2-1019.35
G_variegatus-3.0.2-103.8
G_variegatus-3.0.2-985.11
G_variegatus-3.0.2-9489.7
G_variegatus-3.0.2-1838.3
G_variegatus-3.0.2-176.29
G_variegatus-3.0.2-34492.2
G_variegatus-3.0.2-2914.23
G_variegatus-3.0.2-5986.3
G_variegatus-3.0.2-3980.18
G_variegatus-3.0.2-96066.1
G_variegatus-3.0.2-41646.1
G_variegatus-3.0.2-34583.1
G_variegatus-3.0.2-3770.2
G_variegatus-3.0.2-90522.1
G_variegatus-3.0.2-2585.6
G_variegatus-3.0.2-3980.17
G_variegatus-3.0.2-11464.20
G_variegatus-3.0.2-5606.17
G_variegatus-3.0.2-93.31
G_variegatus-3.0.2-12462.13
G_variegatus-3.0.2-41149.2
G_variegatus-3.0.2-757.27

Table S3. The retrovirus sequences used in this study

Virus	Full name	Accession No.
OLV	Ovine lentivirus	NC_001511
MVV	Maedi-visna virus	NC_001452
CAEV	Caprine arthritis-encephalitis virus	NC_001463
BIV	Bovine immunodeficiency virus	NC_001413
HIV-1	Human immunodeficiency virus 1	NC_001802
HIV-2	Human immunodeficiency virus 2	NC_001722
EIAV-USA	Equine infectious anemia virus	NC_001450
EIAV-Ireland	Equine infectious anemia virus	JX480631
EIAV-Liaoning	Equine infectious anemia virus	AF327877
FIVcat	Feline immunodeficiency virus	NC_001482
FIVsubC	Feline immunodeficiency virus	AF474246
FIVoma	Feline immunodeficiency virus	U56928
FIVpuma	Feline immunodeficiency virus	EF455614
FIV-PPR	Feline immunodeficiency virus	M36968
FIV-lion	Feline immunodeficiency virus	EU117992
JDV	Jembrana disease virus	U21603
SIVagm	Simian immunodeficiency virus	NC_001549
SIVmnd2	Simian immunodeficiency virus	NC_004455
SIVhoest	Simian immunodeficiency virus	AF188116
SIVsyk	Simian immunodeficiency virus	L06042
SIVcol	Simian immunodeficiency virus	AF301156
SIVtal	Simian immunodeficiency virus	AY655744
SIVgsn	Simian immunodeficiency virus	AF468659
SIVrcm	Simian immunodeficiency virus	HM803689
SIVsmm	Simian immunodeficiency virus	M80194
SIVcpz	Simian immunodeficiency virus	AF115393
SRLV	Small ruminant lentivirus	HM210570
MELV	*Mustelidae* endogenous lentivirus	Ref. 1
RELIK-*Oryctolagus*	Rabbit endogenous lentivirus K	Ref. 2
RELIK-*Lepus*	Rabbit endogenous lentivirus K	Ref. 3
PSIVgml	Prosimian immunodeficiency virus	Ref. 4
GvaELV	*Galeopterus variegatus* endogenous lentivirus	This study
PSIVfdl	Prosimian immunodeficiency virus	Ref. 5
ALV	Avian leukemia virus	NC_015116
RSV	Rous sarcoma virus	NC_001407
LDV	Lymphoproliferative disease virus	KC802224
JSRV	Jaagsiekte sheep retrovirus	NC_001494
SRV-1	Simian retrovirus 1	M11841
Python-molurus	*Python molurus* endogenous retrovirus	AF500296

HERV-K	Human endogenous retrovirus K	AF164614
MMTV	Mouse mammary tumor virus	NC_001503

References

1. Han GZ, Worobey M. (2012). Endogenous lentiviral elements in the weasel family (Mustelidae). Mol Biol Evol. 29, 2905-2908.
2. Katzourakis A, Tristem M, Pybus OG, Gifford RJ. (2007). Discovery and analysis of the first endogenous lentivirus. Proc Natl Acad Sci U S A. 104, 6261-6265.
3. Keckesova Z, Ylinen LM, Towers GJ, Gifford RJ, Katzourakis A. (2009). Identification of a RELIK orthologue in the European hare (*Lepus europaeus*) reveals a minimum age of 12 million years for the lagomorph lentiviruses. Virology. 384, 7-11.
4. Gifford RJ, Katzourakis A, Tristem M, Pybus OG, Winters M, Shafer RW. (2008). A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. Proc Natl Acad Sci U S A. 105, 20362-20367.
5. Gilbert C, Maxfield DG, Goodman SM, Feschotte C. (2009). Parallel germline infiltration of a lentivirus in two Malagasy lemurs. PLoS Genet. 5, e1000425.

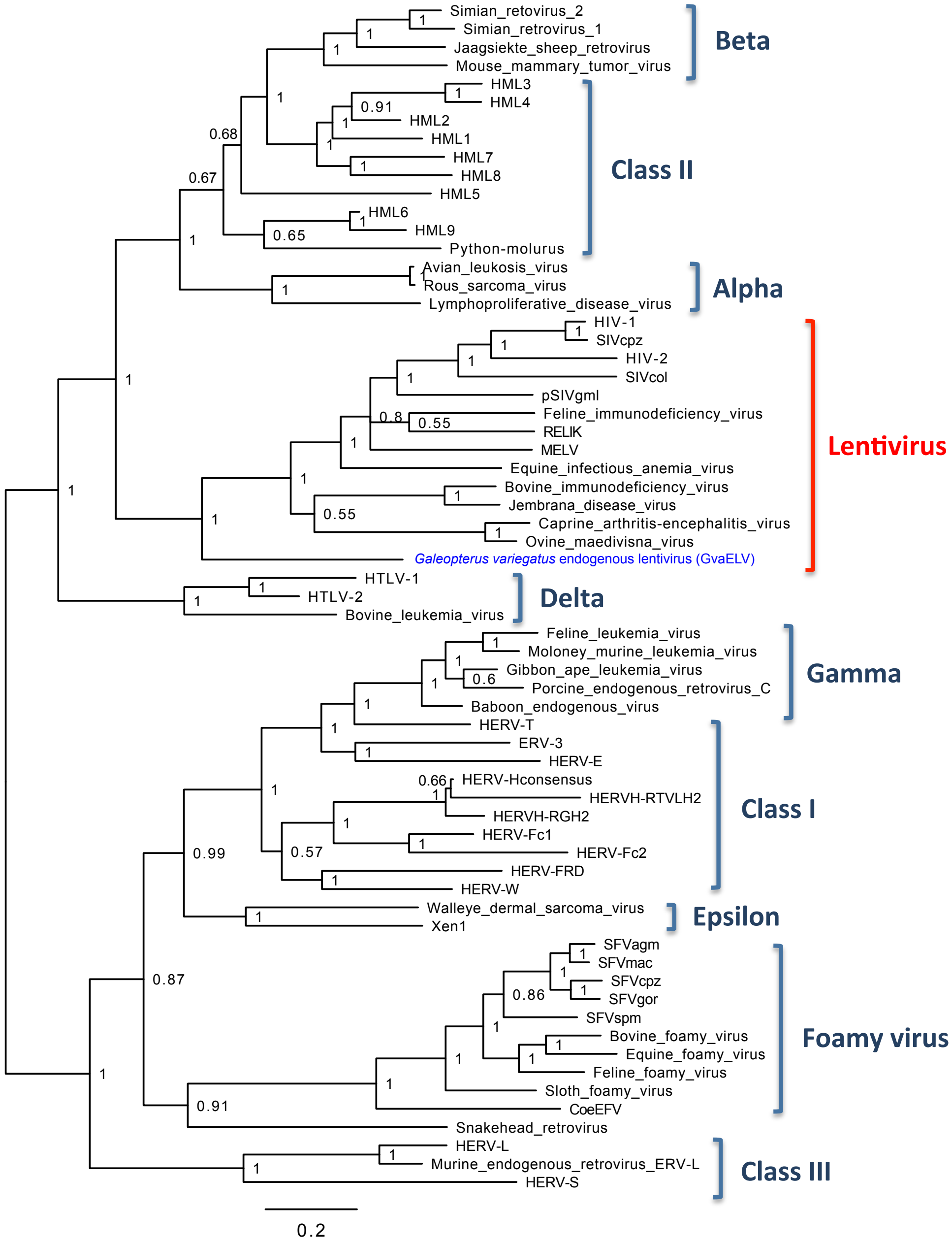

Fig S1. Retrovirus phylogeny. The phylogeny is 50% majority-rule consensus tree inferred using a Bayesian approach available in MrBayes. The node labels are posterior probabilities. The lentivirus reported in this study, GvaELV, is highlighted in blue.

Fig S2. The reconstructed GvaELV genome sequence

```
5’LTR->
TGGAGGAAAATGAAGGAGAACACCTGTTAAAGATGGAGGCACTCTTGCTCTCCAGGATTTAGACTGTCAGCAGAAATGCAAAAA
GACTGGGTCAGCAAAAATTCGCAAAGCTTAACCGCAAACCACAGACTTCCTATTGAGTTACCTAGGAGATGGGGAGGTATGGGA
GGGGTTTACAGCTGGCTAAAACATATATAAGACATGCTTTTACCCAATAAAGTTGCTTAGTGCGTGCGCTAAGCCCCAGAACTT
TGTGTCTGGTTTCTGGTTGACCACCTCCCCAGGTGGCAGCCCCTAGCTCAAGGCGTGCTCTGAGCAACTCACCTTGCTGCACTC
TCTTGGCAGGGCTGTGTGTGGCCCTGCTCCCTGAGGGACTCCTGGCAGAGGCTACATGTAGTCCTGCCTCCTGAGACCCCCAGG
               <- 5’LTR
ACCAAGAGAAAGGTGCAGACTTGGCGCCCAATGTGGGGCTTGAAGGCTGACCCAAGGACCGGACGGAGGTCTGCAGGACAGTGG
                                               gag ->
AGAAGGTGAGTCTGTGCGGCTAGGTCCGGGTAGGAGTAGCTTATAAGCAGGATGGGGACTGTCACGTCAGCAAGGCTTGTTACT
                                                   M  G  T  V  T  S  A  R  L  V  T
TCAGTCATACAAGGCATTAAGCATTGCCAATCAGTAGAGAAACTGAAGGGAACAGGGAAGACAACACACTGGGCACCTGAGGAT
S  V  I  Q  G  I  K  H  C  Q  S  V  E  K  L  K  G  T  G  K  T  T  H  W  A  P  E  D
ATTTATAGGGCATTTCGGCAGGCACATGTTTATTGCCCAACGTTTCCTGAAACTGGGACACTAGATGTCAGGGGACTTAGGGAA
I  Y  R  A  F  R  Q  A  H  V  Y  C  P  T  F  P  E  T  G  T  L  D  V  R  G  L  R  E
GGGATAAGTGCCCTTCAGAAAGCAAGAAAAGAGCACCCAGAGTTGTCTGTAGCTGTGCAAGCCCTGCAGGTATTTCTGTGTAGG
G  I  S  A  L  Q  K  A  R  K  E  H  P  E  L  S  V  A  V  Q  A  L  Q  V  F  L  C  R
GGAATGTTGGGAGTGGCGGTAAGGACCACTGAGGAAGTAGAAAAGGTATATTTCGAGTTTCGAGATGCTGGCCAGATCTCAGAC
G  M  L  G  V  A  V  R  T  T  E  E  V  E  K  V  Y  F  E  F  R  D  A  G  Q  I  S  D
TCAAGAAATATACAAGGAGAAGGTAAGCAGGTGTACCCAGTAGTCCAAGGTCCAGGAGGAACTATGTGTTCTCCTCTGCACCCT
S  R  N  I  Q  G  E  G  K  Q  V  Y  P  V  V  Q  G  P  G  G  T  M  C  S  P  L  H  P
TGAGTTATGGCAGATTTCTTTGACATCATTAGTCATCATGGATTACTGAGCTCAGAAACTATCACTCGAATGCAAGTTTTATGT
*  V  M  A  D  F  F  D  I  I  S  H  H  G  L  L  S  S  E  T  I  T  R  M  Q  V  L  C
ATGCAGGATTTATGCCAGGTAGAAGTAAAAGGCTTGATTGCTCAAGCACCTGGGAATGCTGCACAAAAATCAATTCTAAAAGAA
M  Q  D  L  C  Q  V  E  V  K  G  L  I  A  Q  A  P  G  N  A  A  Q  K  S  I  L  K  E
TTGTTAGCAAAATATAAAGATGAATGGGATCAACAAAACCCCATACAGCAAGATGTTTTTCCAAGACAATTGACAGGGGACAAA
L  L  A  K  Y  K  D  E  W  D  Q  Q  N  P  I  Q  Q  D  V  F  P  R  Q  L  T  G  D  K
ATAGTAGGAGAATCCCCAAACAATAATTTACCAGATGCACAAGCAGGACCCCAATGGGTAATAGCTAGATCAACCTTCCAGGAG
I  V  G  E  S  P  N  N  N  L  P  D  A  Q  A  G  P  Q  W  V  I  A  R  S  T  F  Q  E
TGGGTTGTCAATGCACTTAGGGAAGGAGTTAAATTAGTTCAGAAACCACCTGCAGTAACAAAAATTGTACAAGGTCCTAGAGAG
W  V  V  N  A  L  R  E  G  V  K  L  V  Q  K  P  P  A  V  T  K  I  V  Q  G  P  R  E
CCATATAAGGATTTTTTGGATAGGCTATTTGAAGCATTGGAAAAGGATTTTTTACCTTTAGAAGTTAGTAGATACTTACAGGAT
P  Y  K  D  F  L  D  R  L  F  E  A  L  E  K  D  F  L  P  L  E  V  S  R  Y  L  Q  D
AAATTGGCTTTTGAAAATGCAAATGAAGATTGTAAAAAGATGATGGCTCACTTGCCTCCGGATGCAGCAATTGCTGACAGGATA
K  L  A  F  E  N  A  N  E  D  C  K  K  M  M  A  H  L  P  P  D  A  A  I  A  D  R  I
CATTGCTGCAGATCCGTGGGTACAGTGACCCACAAACAAACCATGATGGCAGAGGCTTTTGCAGCAGCTCTCTTTAAAGGAAAA
H  C  C  R  S  V  G  T  V  T  H  K  Q  T  M  M  A  E  A  F  A  A  A  L  F  K  G  K
GGGAACCAGCAAGGGCTTAGATGTTTTAATTGCAATAAACTAGGCCATATGAAAAAAGATTGTAAGCAGCAAAGAAAGGCTACA
G  N  Q  Q  G  L  R  C  F  N  C  N  K  L  G  H  M  K  K  D  C  K  Q  Q  R  K  A  T
                                                   pol ->
GCTAATAAAGATGTTACATGTTTCAAATGTGGAAAAAAGGGACACATTGCCAGGTTATGCAGAACAAAAAACGGAAAGCAGGC
A  N  K  D  V  T  C  F  K  C  G  K  K  G  H  I  A  R  L  C  R  T  K  N  G  K  A  G
                                                        M  Q  N  K  K  R  E  S  R  L
TCAGGACAGGGAGCCCAGGCATTGGTATCACAGCCTGTTCTGCCATCAGCACCAGTGGAGGAAGAGACAACCAAGTCACTTTAC
S  G  Q  G  A  Q  A  L  V  S  Q  P  V  L  P  S  A  P  V  E  E  E  T  T  K  S  L  Y
  R  T  G  S  P  G  I  G  I  T  A  C  S  A  I  S  T  S  G  G  R  D  N  Q  V  T  L  P
                               <- gag
CCCAAATTACCAGTTTGCAACAAAGACCAATAACATTGATACAGTGTGCAGGGACAGAGGTTAAGGTTTTATTAGATACAGGAG
P  K  L  P  V  C  N  K  D  Q  *
  Q  I  T  S  L  Q  Q  R  P  I  T  L  I  Q  C  A  G  T  E  V  K  V  L  L  D  T  G  A
CAGATATTTCAATTTGGGCAAAACCACCCCCAATAGGAAGAAATATAGGCCAGCAATTTGTAATGGGAATAGGGGGAAGTCAAA
  D  I  S  I  W  A  K  P  P  P  I  G  R  N  I  G  Q  Q  F  V  M  G  I  G  G  S  Q  T
CAGGAATTAAATATGATCAAGTTAAAATGACATGGAATAATAAAACAATTTTAGGTGAAGTAGTAGTGTGTCCCACGCCTATAA
  G  I  K  Y  D  Q  V  K  M  T  W  N  N  K  T  I  L  G  E  V  V  V  C  P  T  P  I  N
ATTTATTAGGGAGAGATAATTTAAGCAAATTTGGAGTACACTTGGTGTGTACAACGTTAAATCAACAGTTAACACCAGTGAAAT
  L  L  G  R  D  N  L  S  K  F  G  V  H  L  V  C  T  T  L  N  Q  Q  L  T  P  V  K  F
TTAAATTGTTACCAGGATTTAGAGGACCACAGGTTAAGCAATGGCCATTATCTGAAGAAAAAATTAAAGCATTACAAGAAATTT
  K  L  L  P  G  F  R  G  P  Q  V  K  Q  W  P  L  S  E  E  K  I  K  A  L  Q  E  I  C
GTGCTGAATTAGAACAGGAAAAAAAGATACAAAGAGTAGGTCCAGAAAATCCTTATAATACACCTGTATTTTTAATAAAAAAGA
  A  E  L  E  Q  E  K  K  I  Q  R  V  G  P  E  N  P  Y  N  T  P  V  F  L  I  K  K  K
AATCTGGTAAGTGGAGGATGCTTATGGATTTTAGGGAATTAAACAAAGGAATAGATGTAGGAACAGAAACTCAGTTAGGTATTC
  S  G  K  W  R  M  L  M  D  F  R  E  L  N  K  G  I  D  V  G  T  E  T  Q  L  G  I  P
CTCATCCAGCAGGGTTTTCACAGGTAAAGCACTTTACTGTTATAGATATAAAAGATGCTTATTTCACCATCCCATTAGATGAAG
```

```
  H  P  A  G  F  S  Q  V  K  H  F  T  V  I  D  I  K  D  A  Y  F  T  I  P  L  D  E  E
AAATAAGTGAGTTAACAGCATTTACAGTACCTCAAAAGAATAATGCAGGGCCAGGTATAAGATTTAAATGGAAAGTTTTACCAC
  I  S  E  L  T  A  F  T  V  P  Q  K  N  N  A  G  P  G  I  R  F  K  W  K  V  L  P  Q
AAGGGTTGATAGGGTCTCCAAAAGGGTATCAATGTACTTTAGAAAAAATAATAACACCATGGAAAGCAAAACACCCTGGCATAA
  G  L  I  G  S  P  K  G  Y  Q  C  T  L  E  K  I  I  T  P  W  K  A  K  H  P  G  I  T
CATGTTATGTATATATGGATGATATTTTAATAGGAACAAATTATACAGCTCTCAAACATAAACAAGTAGTCAAAGAGTTACAAG
  C  Y  V  Y  M  D  D  I  L  I  G  T  N  Y  T  A  L  K  H  K  Q  V  V  K  E  L  Q  E
AGGAATTTAAAAAGTGGGGATTTGAAACCCCACAAGATAAAATACAGAATAGTGATGGGGACAAAGCAGTAAAATATTTAGGAT
  E  F  K  K  W  G  F  E  T  P  Q  D  K  I  Q  N  S  D  G  D  K  A  V  K  Y  L  G  Y
ATATTTTAGAACCCAAAACGTGGAAGCCACATAGCTATCAGCTATTGGATAGTAAACCAAAAACATTAAATGATGTTCAACGAG
  I  L  E  P  K  T  W  K  P  H  S  Y  Q  L  L  D  S  K  P  K  T  L  N  D  V  Q  R  V
TATTAGGAAGAATTAATTGGATAAAACAAATATATCCTAAGATACAAACAGAAAACATTCAGAAGTTATTAACAGGAGATCAAA
  L  G  R  I  N  W  I  K  Q  I  Y  P  K  I  Q  T  E  N  I  Q  K  L  L  T  G  D  Q  N
ACATATTAAGTAAAAGATCATGGACTAAAAAAGCTCAATATGAGGTAGATGTAATACAACAAATGTTAACACAGCAGGAAGGAA
  I  L  S  K  R  S  W  T  K  K  A  Q  Y  E  V  D  V  I  Q  Q  M  L  T  Q  Q  E  G  M
TGCCTTATTATCAAGAAAAGGAACTAGTTAGTCTTAGTATTGTATGGAATCAAGTAGCTCAAGGAGCTGTTCATCAAAAGCATA
  P  Y  Y  Q  E  K  E  L  V  S  L  S  I  V  W  N  Q  V  A  Q  G  A  V  H  Q  K  H  K
AGGTGCTCTGGTGGGGATACGGAACTAGCAAGGACAGAAAAATAAAAACCAAAGCACAAAAATGTGTGGCTTTAGCAAATCGTA
  V  L  W  W  G  Y  G  T  S  K  D  R  K  I  K  T  K  A  Q  K  C  V  A  L  A  N  R  M
TGGTTTCAGAAATAATCATACGATTGGGAAAAGTTCCAGATGAAGTTTGGTTGCCCATAGAAAAAGCTACTTTTGAGTATTTAT
  V  S  E  I  I  I  R  L  G  K  V  P  D  E  V  W  L  P  I  E  K  A  T  F  E  Y  L  L
TACAAGAAGGATTAGAGGGAACAGAAATTTGGTTATTAAATTGCAAATTTGGTATGGCACCTAACAAAATTTTTGATAAAAGTG
  Q  E  G  L  E  G  T  E  I  W  L  L  N  C  K  F  G  M  A  P  N  K  I  F  D  K  S  E
AATGTTTAAGTTATCCTTGTATATGGAAAAGTGCTGTTCCTGTCACAGGCCCTACAATTTATATAGATGGATCTAGAAAAAAGG
  C  L  S  Y  P  C  I  W  K  S  A  V  P  V  T  G  P  T  I  Y  I  D  G  S  R  K  K  G
GACAGCCAGCTCGTAGTGCTATATGGGAGAATGCAAATAATAATAAAGTAACAGAATTTGAGGGAACAGCTCAGCAAGCAGAAG
  Q  P  A  R  S  A  I  W  E  N  A  N  N  N  K  V  T  E  F  E  G  T  A  Q  Q  A  E  V
TTAGGGCATTTTTATTAGCATTACAATTAGAATATAATCAAATGAACATAGTTACAGACAGTAAATATTTATGGACATGTATTA
  R  A  F  L  L  A  L  Q  L  E  Y  N  Q  M  N  I  V  T  D  S  K  Y  L  W  T  C  I  K
AGTTAATGTCTGACTATGCATGGAAAGAGGAAAATCCAATTTGGCAAGAAATCTGGAAATTATTACAAACAAAGAAAATATATT
  L  M  S  D  Y  A  W  K  E  E  N  P  I  W  Q  E  I  W  K  L  L  Q  T  K  K  I  Y  L
TACAATGGGTGCCAGGACATCAAGGGGTGCCAGGGAATGAAGAAGCTGACAGTTTAGCCCAGGGAAATGTAGTTCTTACTGCAG
  Q  W  V  P  G  H  Q  G  V  P  G  N  E  E  A  D  S  L  A  Q  G  N  V  V  L  T  A  E
AAGAGGAGGACTTACTGCCACAGCAGTCAGAACTTTGTTTCCCTGGGTGGCAAATAAAAGCAGCCAGAGATTGTTATGTGGCAG
  E  E  D  L  L  P  Q  Q  S  E  L  C  F  P  G  W  Q  I  K  A  A  R  D  C  Y  V  A  A
CCCAGGAAAAAAGACAAATATTTACAGGAATACACCTCCAGTTGGATGAAATGCAAATAGCACAGATGCAATTAAATTATGAAT
  Q  E  K  R  Q  I  F  T  G  I  H  L  Q  L  D  E  M  Q  I  A  Q  M  Q  L  N  Y  E  L
TAGCTATAAAAGGATTGTTATTATTACAGTCAACTGTTAAAGGAAGTTATTCAACAGAATTAATATTACAAGTTGTAAATTTTA
  A  I  K  G  L  L  L  L  Q  S  T  V  K  G  S  Y  S  T  E  L  I  L  Q  V  V  N  F  T
CTAAAAAGGGAATAAATATTCAAAAGGGTGATCCAGTTGTTATTTTATATATTATAGCTGCTGAAGAGATGGGGGAAGCAAAGT
  K  K  G  I  N  I  Q  K  G  D  P  V  V  I  L  Y  I  I  A  A  E  E  M  G  E  A  K  L
TACAAAAGTCAGCTCCAGATGACATGGTCTGTTTCACCTCAGAAACTTTTCAGGAAGCAGAACAGTGGCATGATGTAACACATG
  Q  K  S  A  P  D  D  M  V  C  F  T  S  E  T  F  Q  E  A  E  Q  W  H  D  V  T  H  A
CAAACCCCCGACAATTAGCAAGAAAGTTTCATCTTCCTTTAGGATTAGCAAAGCATATTACACAGTTATGCAAAGACTGCCAAC
  N  P  R  Q  L  A  R  K  F  H  L  P  L  G  L  A  K  H  I  T  Q  L  C  K  D  C  Q  Q
AGCAGGATACGGCTACAGGAGAGGGATATAATCAGTTAAAAAACCAAGAAGGTGTTTGGCAGATGGATGTTACACACATGTTTG
  Q  D  T  A  T  G  E  G  Y  N  Q  L  K  N  Q  E  G  V  W  Q  M  D  V  T  H  M  F  E
AGTTAGGACTTTATAAATATGTGTATGTAGCAGTTGACATACAAACAGGAATGATATGGGTTAGTCCACAAAAGGGAGAAACAG
  L  G  L  Y  K  Y  V  Y  V  A  V  D  I  Q  T  G  M  I  W  V  S  P  Q  K  G  E  T  A
CAAAACATACACAAACAGCTCTGCTTCAGATCATTCATATGGCAGGGCTCCCAAAGGAAATCCAAAGTGATAATGGACCAGGTT
  K  H  T  Q  T  A  L  L  Q  I  I  H  M  A  G  L  P  K  E  I  Q  S  D  N  G  P  G  F
TTGTAGCAGATAGAGTACAAGTGATGTGTAAACAATTAGGGATTAAATGGCATCATGGCATTCCATACCATCCTCAGAGTCAGG
  V  A  D  R  V  Q  V  M  C  K  Q  L  G  I  K  W  H  H  G  I  P  Y  H  P  Q  S  Q  G
GAAAAGTAGAAGGTACACATAAATTAATTAAACAACACTTCAAAAAGGTAAAAGAGTTATTTGAAGACCCCATTAATGCTCTGT
  K  V  E  G  T  H  K  L  I  K  Q  H  F  K  K  V  K  E  L  F  E  D  P  I  N  A  L  L
TGTGGACAGTATTTTGCCTCAATTTTGAAAAAAGGGGTGGAGATGGGTATACAAGAGCAGAGGCATGGGTGGAGCAACAGAATT
  W  T  V  F  C  L  N  F  E  K  R  G  G  D  G  Y  T  R  A  E  A  W  V  E  Q  Q  N  L
TAAAAAATCAGTTAGAAACATGTAAGCATATGCAAACTTTAACTAATTCCTGTAAATTTTTGTATAAGGATCCCAAGTCACCTC
  K  N  Q  L  E  T  C  K  H  M  Q  T  L  T  N  S  C  K  F  L  Y  K  D  P  K  S  P  R
GAACAGGGTGGAGAGGCCCAGGCACTCTGCAGTGGGAAGGAGCAGGCGTCAGAGTATTAGAACTTACAACTGGAGAGGTAAAAG
  T  G  W  R  G  P  G  T  L  Q  W  E  G  A  G  V  R  V  L  E  L  T  T  G  E  V  K  V
                                           <- pol

                 hypothetical ORF1 ->
TAGTTCCTAAACATCTAACAAAAAGAATTTTTCAGATGGAGCCTAGAGTTCAGATGAATCCTGAAGTTACTGTTCAGTGGATG
  V  P  K  H  L  T  K  R  I  F  S  D  G  A  *
                                  M  E  P  R  V  Q  M  N  P  E  V  T  V  Q  W  M
GATGCCTTTTGTTTTGGACAAGAAGAAGCCATTATTATAGCTTACATCATTAATGGCCAAGAGTTACACCCTGGTGAACTTATG
```

```
D  A  F  C  F  G  Q  E  E  A  I  I  I  A  Y  I  I  N  G  Q  E  L  H  P  G  E  L  M
GTAGTGCATAAGCAAACTATTCCTAGATGGTTAACACAGGTTACTCGTTCTATGTGGAATAAACATAGAAAAAGTACAGAACAG
V  V  H  K  Q  T  I  P  R  W  L  T  Q  V  T  R  S  M  W  N  K  H  R  K  S  T  E  Q
AGATTAATCCAAATACAAAAAAGATATTGGATCAATACCTTTTGTGGACCTAAAAGATATGTCTTAGTGAAAGGATGGAACAA
R  L  I  Q  I  Q  K  R  Y  W  I  N  T  F  C  G  P  K  R  Y  V  L  V  K  G  W  E  Q
AAGGTAAAATGGGTCAAACACATACAGTTGTTGCCAAAGTTTGGGAACATATGTATGTGGTGCCCTGGTGTTAAAGTTCCCCAT
K  V  K  W  V  K  H  I  Q  L  L  P  K  F  G  N  I  C  M  W  C  P  G  V  K  V  P  H
TGGAGTGAACCATATGGTGCTCCAAAGCCTCGTGGATCTTATTTAGTGCAGTTAAGACAATTAATAAATCATATGGTAGAAGAA
W  S  E  P  Y  G  A  P  K  P  R  G  S  Y  L  V  Q  L  R  Q  L  I  N  H  M  V  E  E
ACTCATCCTAGCCCCAGAGGAACTTGTGGATGGCCTTACAGATGCCCACCATGTGTTCTTCAAGCAGCTATTATATTACTAGAA
T  H  P  S  P  R  G  T  C  G  W  P  Y  R  C  P  P  C  V  L  Q  A  A  I  I  L  L  E
GAGGATTATTGGAAAGATGAGGAAGTTATTCCAAAGTTACAAATGATATGTTATAGTAGATGTTTAAAACAACTGTGGAAAGAA
E  D  Y  W  K  D  E  E  V  I  P  K  L  Q  M  I  C  Y  S  R  C  L  K  Q  L  W  K  E
ATGACTTATCAAATTTCAAAAAGAAACATTAGTACATGGCATGCTAGGATGCAAACACCAGTAAGTACAAAATGGCAATCACCA
M  T  Y  Q  I  S  K  R  N  I  S  T  W  H  A  R  M  Q  T  P  V  S  T  K  W  Q  S  P
AAACATGGCCCTCATGTCATTACAGGAATCCCTTGGCAAAGACTAGCAACAAAATTTTGTGTTTGGGAGTCTAACGGGGATTTA
K  H  G  P  H  V  I  T  G  I  P  W  Q  R  L  A  T  K  F  C  V  W  E  S  N  G  D  L

             <- hypothetical ORF1
                            tat ->
AGATCAGAAGCATAAAGAGTAACTTGTGGTAAATGTGTTGTGAATGTAGATGCCTTTGCTTGCACTGCTATTATCATTGCATTT
R  S  E  A  *

                                M  C  C  E  C  R  C  L  C  L  H  C  Y  Y  H  C  I  L
TATGCTTTCAGCAAAAATACTTAGGGATAACACATGAAGGAGGTCCCAGAGGAATAAAACCTTGTTGTCATAAATGTTATAATC
  C  F  Q  Q  K  Y  L  G  I  T  H  E  G  G  P  R  G  I  K  P  C  C  H  K  C  Y  N  P
CAAAACAGGGAGGTAAATACTGTGGGACAAAAAACAGTTTTGAAGTACATTCACAACCTTGTTTTCATAAATGTTGTAGATGAC
  K  Q  G  G  K  Y  C  G  T  K  N  S  F  E  V  H  S  Q  P  C  F  H  K  C  C  R  *
CTATTTCAAGTGGAAGGAAAAGAAAGTACAATACTGGTCCAAAACGAGGGAGCAAATCCTCTGGGACAAGGAACAATTCTATAT
TAGATTAAGAATACAGATCCCAAGAACACTTAGTGAAGGAAGGGAAAATAATCCTCAGCTGCTACCCTCGTTGTTGCAAAAGGA
                                    env ->
ATTGTGAGCAACAGAGTTAGATCCAGTCCCTGAAGAAACCATGGCCTCTTCCTCCTTACCATTAATAAATAATGAGGAAGATGA
                                        M  A  S  S  S  L  P  L  I  N  N  E  E  D  D
TTTTTATATGGTAATGACTCCTAGACTGTTAACAAAAGGAACTCAAACTGACCAAATAAAAATGAGATCCCAAAAGGCAGGACC
 F  Y  M  V  M  T  P  R  L  L  T  K  G  T  Q  T  D  Q  I  K  M  R  S  Q  K  A  G  P
ATCCAAGGAGGATAAAATAATGGAAGAATTAAGGGTGCTTCAGCTTCAACAGGAGAAAGAGGAGAGATGGAGACAACAAGAGCT
 S  K  E  D  K  I  M  E  E  L  R  V  L  Q  L  Q  Q  E  K  E  E  R  W  R  Q  Q  E  L
GTGTAACACCAGCTCATCTGGAGAGGACAAGGACGTGAGTGTGCCTTTGTTACCGGCAGGAAAAGGTAAAAATATTATATGGGT
 C  N  T  S  S  S  G  E  D  K  D  V  S  V  P  L  L  P  A  G  K  G  K  N  I  I  W  V
CTGTATTTTAGCACATTTGTGGGTCCCCACATGTCAAGGTTTTAGGCTAGGGTTGTTAAAATGGTCCAGTGCATCTACACTTGT
 C  I  L  A  H  L  W  V  P  T  C  Q  G  F  R  L  G  L  L  K  W  S  S  A  S  T  L  V
AAGTCCTCAAAAAATCCCTTGGTTTTGTGCAGCAGCCAATACCCAATTAGGATGTCTTCCCAAAAACCATTACATAACACAGCT
 S  P  Q  K  I  P  W  F  C  A  A  A  N  T  Q  L  G  C  L  P  K  N  H  Y  I  T  Q  L
GCATTGGGCAGGAAACAATACCCATATGTGGGGACTTAGAGGACACAGAAATGGTACAGAATATGGAAACAACAATTGGGGGCT
 H  W  A  G  N  N  T  H  M  W  G  L  R  G  H  R  N  G  T  E  Y  G  N  N  N  W  G  L
GCAAATAGGAAAACAATGGGTATGGATTTTTGAAAAATATTTAAATGCTCAATGTCCAAAAAATATGGCTTATTATAAGCCATT
 Q  I  G  K  Q  W  V  W  I  F  E  K  Y  L  N  A  Q  C  P  K  N  M  A  Y  Y  K  P  F
TAATTGTACAAATGCCACCTGGCCTTTAAATGTAACTGAGGATTATTGGGAAGCAAACTGTAGTGGACTGGAATGGATTTCAGA
 N  C  T  N  A  T  W  P  L  N  V  T  E  D  Y  W  E  A  N  C  S  G  L  E  W  I  S  D
TTCCACTGGCAGAATACTTCAAGGACATAAGCCTATGATCTTTAATACCACTTGGGCATTAAAGCAAGGACTTGTAAATCAAAT
 S  T  G  R  I  L  Q  G  H  K  P  M  I  F  N  T  T  W  A  L  K  Q  G  L  V  N  Q  I
CAGTTGTAAACCAAATCATACTGCTCAAGGGTCTCAGTGTAAGAAAAGCAATAAAGCTTGGTGTTGTAATTGGAAAGACTGCAA
 S  C  K  P  N  H  T  A  Q  G  S  Q  C  K  K  S  N  K  A  W  C  C  N  W  K  D  C  N
TCAATTTCAAGATATTTTAGAAGCATGCCCTGAAAAGCTCAGGTTGCTCACTTCTTTTACCAAAGGAGTACAGTTATGTGTCCA
 Q  F  Q  D  I  L  E  A  C  P  E  K  L  R  L  L  T  S  F  T  K  G  V  Q  L  C  V  H
TAGACCTTATTATTTTTGGAGTTGTAACAAACACAAAATTCAAAGAGAGTATGACACAAACAGTAATGTTAGTGGTGAATTTAA
 R  P  Y  Y  F  W  S  C  N  K  H  K  I  Q  R  E  Y  D  T  N  S  N  V  S  G  E  F  K
ATGGGGGAATGAAACAATTGTGGGCAGTAGTACCAACAGACAAAATGAAACTTTACAGCTAGTAAAATGGGAATGTGACGGAGA
 W  G  N  E  T  I  V  G  S  S  T  N  R  Q  N  E  T  L  Q  L  V  K  W  E  C  D  G  E
AGCTCAAGTCACTCAGTGCTTACCCGAGTTAGGTGAACTTACAACTGATAGTACATTTATAGGAATAGGATGGAAAAACTGGAT
 A  Q  V  T  Q  C  L  P  E  L  G  E  L  T  T  D  S  T  F  I  G  I  G  W  K  N  W  I
ATATCCTCAATTGGTCTGGCAGGCCAAATTTGTAAATTTGACAAATAATAATACATGGAGAATTCCAACAAGTGAAACAAATGA
 Y  P  Q  L  V  W  Q  A  K  F  V  N  L  T  N  N  N  T  W  R  I  P  T  S  E  T  N  D
CCAAAGCATCAAAACTGAACAGTCCTGGCTACATAGTATAAAATTTAACGAAACTCATCCAGGACAAGTAAATTGTACTATTAT
 Q  S  I  K  T  E  Q  S  W  L  H  S  I  K  F  N  E  T  H  P  G  Q  V  N  C  T  I  M
GCATTTATCTAGTTATGCTAATAAGAAGCAATACATGCCTTTTAATAGCTTAGCATTTTGCAATTTATCCTTGCCAGGGCAAGT
 H  L  S  S  Y  A  N  K  K  Q  Y  M  P  F  N  S  L  A  F  C  N  L  S  L  P  G  Q  V
```

```
CATGAAAAGGTGTTTATCAAACATTAGTAATTTGCCAGAGGAAGGGAAATTTGATTGTGCTATCCCACATCCTTGGGTAGCAGG
 M  K  R  C  L  S  N  I  S  N  L  P  E  E  G  K  F  D  C  A  I  P  H  P  W  V  A  G
GGATTTCCAAAGATGGACTTGGCAATACACAAAAAGGCAAGGGGATGATGAGCCATCAACACCACATTTGTCCTTTCCATGTGA
 D  F  Q  R  W  T  W  Q  Y  T  K  R  Q  G  D  D  E  P  S  T  P  H  L  S  F  P  C  E
GATGAAAGACAAACAATTATTTCAATTAGAGTGCAATTTTGTAAATAAAACTAAAAATGCAAATGATACTCAATTTTGTCAAGA
 M  K  D  K  Q  L  F  Q  L  E  C  N  F  V  N  K  T  K  N  A  N  D  T  Q  F  C  Q  D
TTGTTTAGAATGTAGATCCAGGGATACTAGGATCTTACTATGGACTCACAAATTAATGCATGTACGGTCATCAGGGGAATACAA
 C  L  E  C  R  S  R  D  T  R  I  L  L  W  T  H  K  L  M  H  V  R  S  S  G  E  Y  K
ATATACAACATGTAGTTTCCAATTTATAGGAATTCATTGTGACTCATTAAGAGTTAAAGGCCAGGAAGGCACACAATGGAAAAA
 Y  T  T  C  S  F  Q  F  I  G  I  H  C  D  S  L  R  V  K  G  Q  E  G  T  Q  W  K  K
ATTTGACATAGAAAGTAGTCAGTGTGACATGTTGTTCAATGTATCAGCACCAGAGGGAGCAGCTGGCACACCAAACAGTACATG
 F  D  I  E  S  S  Q  C  D  M  L  F  N  V  S  A  P  E  G  A  A  G  T  P  N  S  T  C
TGCTCGTTCACGGTCAGCCATTTGTTTTGCAACTACAGGAGGCCCAGAGGAATGGATTCAACAGGCTTTTGACAGACATATTAA
 A  R  S  R  S  A  I  C  F  A  T  T  G  G  P  E  E  W  I  Q  Q  A  F  D  R  H  I  K
GGTGGGCTGACAGGAAAATATAATTCCCATTCCTATAGATGATAACAAGTTAGAAATTGTTAAGCCAAAGAGATCCAAGCGTGG
 V  G  *  Q  E  N  I  I  P  I  P  I  D  D  N  K  L  E  I  V  K  P  K  R  S  K  R  G
AGCTGTTTTATTGGCTGCTTTCGGAATTGTTACTGCCATCACTGCTGCGGCTGGTATGGCTACAGGTGGGGTGGCCCTCCATCA
 A  V  L  L  A  A  F  G  I  V  T  A  I  T  A  A  A  G  M  A  T  G  G  V  A  L  H  Q
ATTACAGGAACTCAAGGAACTAGCAGACAAAAATGTCCAGGTGATGGTGGATCTGATAAAGACCCAGGAAAATCATTGGGTTTT
 L  Q  E  L  K  E  L  A  D  K  N  V  Q  V  M  V  D  L  I  K  T  Q  E  N  H  W  V  F
TACGAGGGATCTTACATTTGGACTTTTGGAGCTAGAAGCTCGAGTACGCTGGATTGAGGAGGTGCTGGCCACTGACAAAGCCTT
 T  R  D  L  T  F  G  L  L  E  L  E  A  R  V  R  W  I  E  E  V  L  A  T  D  K  A  F
TCAAACATTAAACTGTGAAATATCAGCCGGCAGTCTTTGCATTTTCCGAGAGTGGTACAATACAACACAGCTCAAAGAATGGAT
 Q  T  L  N  C  E  I  S  A  G  S  L  C  I  F  R  E  W  Y  N  T  T  Q  L  K  E  W  I
CAATAGCACCAGCTGGTACAACATGTCAATTAATGAGTGGGCAGCAAGATTACAGTTTTTTCAAAAGGGCATAAAAGAAATAGA
 N  S  T  S  W  Y  N  M  S  I  N  E  W  A  A  R  L  Q  F  F  Q  K  G  I  K  E  I  D
CAAATTAGTTCTTAAAAATAGAGTAGGAAAAAAGAATAGAGTAGGTTTAACACACTTAAGACATAGTTTACAGAAGGTACAAGA
 K  L  V  L  K  N  R  V  G  K  K  N  R  V  G  L  T  H  L  R  H  S  L  Q  K  V  Q  E
GCAGTCAACATTAGAGGCATGGATTAGTACAGTGCCTGATTGGTTGTCTTCCATATGGAGGGCATTAAAAAAGAGTTGGTGGTG
 Q  S  T  L  E  A  W  I  S  T  V  P  D  W  L  S  S  I  W  R  A  L  K  K  S  W  W  W
GCTCATCATCGGCATAATTTGTTTGGTCATCATTATTCCCTTGTTACCTAGTTGTATTAAAGCAATTATAAATTATATCAGAGG
 L  I  I  G  I  I  C  L  V  I  I  I  P  L  L  P  S  C  I  K  A  I  I  N  Y  I  R  G
                  Hypothetical ORF2 ->
                                      <- env
ATATGTAATATTAAATATGTTTGATATACAGGAGCAATGACCACGAGAGAGATGATGGACAACAAGATAAAAAGACACAAGAAA
 Y  V  I  L  N  M  F  D  I  Q  E  Q  *
                                    M  T  T  R  E  M  M  D  N  K  I  K  R  H  K  K
AGAAAACTCAAACAAAAGGAAAAACAAAGAGCAGCAACACTTGCCCAAACACCATCCCCTTCGCTTTTAGGCTTAAAAGCCGCT
R  K  L  K  Q  K  E  K  Q  R  A  A  T  L  A  Q  T  P  S  P  S  L  L  G  L  K  A  A
                                                 <- Hypothetical ORF2
GCTGCAGCACTTGGATATCATCTTATTCCGGCAATAGAAAAAGAGATTTGAACCAAAATCATGCTATTGTTGCTGTTACTGTTG
A  A  A  L  G  Y  H  L  I  P  A  I  E  K  E  I  *
AGGCTGCATATTAATCAGATGGCTTTTTGCTGTGTGTATTTAGACTTTGCCAACATTTGCAAGCCTGCTTTTAAAAGAAAAAGG
      -> 3’LTR
GTGGACTGGAGGAAAATGAAGGAGAACACCTGTTAAAGATGGAGGCACTCTTGCTCTCCAGGATTTAGACTGTGTCAGCAGAAA
TGCAAAAAGACTGGGTCAGCAAGAATTTGCAAGGCTTAACCACAAACCACAGACTTCCTATTGAGTTACCTAGGAGATGGGGAG
GTATGGGAGGGGTTTACAGCTGGCTAAAACATATATAATACATGCCTGTAGGCAATAAAGTTGCTTAGTGCATGTACTAAGTCC
CAGAACTTTGTGTCTGGTTTCTGGTTGCCCACCTCCCTAGGCGGCAGCCCCTAGCTCAAGGCATGCCCTGAGCAACTCACCTCG
CTGCACTCTCTTGGCAGGGCTGCGTGTGGCCCTGCTCCCTGGGGGACTCCTGGCAGAGGCTATGTGTAGTCCTGCCTCCTGAGA
                       <- 3’LTR
CCCCCAGGACCAAGAGAAACGTGCA
```

```
CAEV_vif       MqnssRhQqkkRN-------------------KkpGPElPlalWihIAeSIN-gDssW-Y
SRLV_vif       MqnsfRpQqrrRN-------------------KepGPnlPlplWKhIAeSIg-rDssW-Y
OLV_vif        MLssyRnQkkyKqnKi----------------RevGPQlPlwaWKEIAfSIN-qEpYW-Y
MVV_vif        MLssyRhQkkyKknKa----------------ReiGPQlPlwaWKEtAfSIN-qEpYW-Y
HIV-2_vif      MeeDrnwivvp---------------------twrvP-grmekWhaL---Vk-ylkYrtk
HIV-1_vif      --MEnRwQvmi---------------------vwqvdrmrirtWKsL---Vk-hhmYvsg
BIV_vif        --MErtlQsvvg--------------------RrrGssnrgrgknsL---IS-tpsYaLH
JDV_vif        --MErtiQspmg--------------------RrrGssgrrkrnanI---IS-ppaYaIY
FIVsubC_vif    -MsEedwQvskRlfavlQggvhsamlyiselpemekeQykkefkKrL---lD-kEtgfIr
FIVpuma_vif    -MqpyRvQsskRNmRyfK--------------enncsEdqikqfKnk---lglqElrWIr

CAEV_vif       iTmRLQqmmWgKrgNKLqyknEDREYEnWEiTswgWkmhLRRvkqwIqdnrrgsP-W--Q
SRLV_vif       iTvRLQqmmWNKrgNKLqyrnEDKEYEnWEiTswEWkmhLRRvkhrIqengrgNP-W--Q
OLV_vif        sTiRLQGlmWNKrgHKLiFvKEEngYEyWEtTnkQWrmeLRRdlrLIaqinfrNa-W--Q
MVV_vif        sTiRLQGlmWNKrgHKLmFvKEnQgYEyWEtSgkQWkmeIRRdldLIaqinfrNa-W--Q
HIV-2_vif      dleeVryvphHKvgw-----------awWt--csrvifpLqgkshLeiqaY-----W--n
HIV-1_vif      kargwfyrhhyEspH-----------pris---sEvhipL-gdarLVittY-----W--g
BIV_vif        pppRfryprWefvrq-----------teys-----mtAcvRKgklVLtYqYai---W--K
JDV_vif        papQyryprWefvmN-----------Dlys-----qtArLqKeeiIItYrYav---W--a
FIVsubC_vif    rlrKaEGikWSfhtr-----------DyymgyvkElvAgsstpdsLrlYiYisNPLWhgK
FIVpuma_vif    klryVEGilWSfhtr-----------EwysdmvrElvA---gtgpLklYcYvshPIWk-R

CAEV_vif       YKvg-gtWKsigv----WFlqaGdYRkVD----------rhFW------waWRiliCSCr
SRLV_vif       YKva-gtWKsigv----WFlqaGdYKkVD----------qhFW------waWRilmCSCr
OLV_vif        YKsq-ekWniigi----WYdspGeYRDkE----------kqFW------fhWRiamCSCk
MVV_vif        YKsq-geWKtigv----WYespGdYKgkE----------NqFW------fhWRialCSCn
HIV-2_vif      ltpe-kgWlsshaVrltWYtekf-wtDVtP---------Dcad------ilihstyfSC-
HIV-1_vif      lhtgerdWhlgqgVsieWrkkry-stqVDP---------elad------qlihlYyfdC-
BIV_vif        -----rvWtietg-----FtdpslF--MtPagthTteEIghld------lfWlrY-CSCp
JDV_vif        -----reWKiqtg----fldlgyl---MtPagthTtgELNeld------lfWvrYtl-Cq
FIVsubC_vif    YRpglknfnkewpfvnmWiktgfmwdDIEkqkIcvGgEISpgWgPGmVgiaiKaFs--Cg
FIVpuma_vif    YRpt-feWnpcwpygnlWltekymwE-IqqddIwTGkvtSqFp-PGyIalivKaYtCkCd

CAEV_vif       KEkfDIREFmrGRhrw---DlCKSCaQgeVVKHTrtkSLeRLvLLqmveq--------Hv
SRLV_vif       KErfNIkEFLrGghrw---DlCKSCaQgeIIKHTrPrSLeRLvLLhiveH--------Hv
OLV_vif        KErwDIRDFmvGKhrw---DlCKSCiQgeIVRHTePrSLQRLALLhivrN--------Hv
MVV_vif        KtrwDIREFmiGKhrw---DlCKSCiQgeIVKnTnPrSLQRLALLhLakD--------Hv
HIV-2_vif      ftagEVRrairGEkll-------SCcnypqahkaqvpSLQyLALvvvqqNdrpqrkgtar
HIV-1_vif      fsdsaIRkaLlGhivs-------prcEyq-agHnkvgSLQyLALaaLitp--------kk
BIV_vif        hEmppwlDFLrGtlnl--RisCRralQasVltsTPrhSLQRLAaLqLctN--------ac
JDV_vif        hrspkwRElLlGEmth---tsCRrtaQaaVVsHTkPhtLQRLAgLtLvcN--------qn
FIVsubC_vif    ErkiEatpvmiiReeidpKkwCgdCwnlmclRnsPPgtLQRLAMLa-cgr--------ka
FIVpuma_vif    KrdltyREiilGKwyl---kwCadCwaliVVRnTPsltLQRLAaLaLgr---------kl

CAEV_vif       FQvlPlWrARrsSttdfPwcrdttgYThaWsvqecwlmeyLlEde----
SRLV_vif       FQvlPlWrARrsStidfPwcrdtsgYThaWsaqecwlmeyLlEde----
OLV_vif        FQimPlWrARrvTvqrfPwsgteglYdtlvytgllghginI--------
MVV_vif        FQvmPlWrARrvTvqkfPwcrspmgYTipWslqecwemesIfE------
HIV-2_vif      kQWrrdhwrglrvAredhrslkqggsepsaprahfpgvakVlEila---
HIV-1_vif      i--------------kpPlpsvtkltedrWnkpqktkghrgshtmngh-
BIV_vif        lcWy-------------Plg-rindtTplWlnfssgkeptIqqlsGHp-
JDV_vif        lcWy-------------Pvg-tvtrnSplWmhfttgkeptIqqlsGHp-
FIVsubC_vif    kcWrgccnqRfvSpyrtPadleviqYkpgWnllwlgel-----------
FIVpuma_vif    YsWyckppyRffeArvtPldhrilisSar--qedlykldkgsdnanrsm
```

Fig S3. Alignment of representative lentivirus Vif proteins.

```
OLV_rev             qAyRqvReiRYTRv-------------------tvvieapVDleE----kqReerdGssg
EIAV-Ireland_rev    --maegRdsRYqeemipkeeskgkeekgrnDwWkiapqtpLDndDwccIlrQslPeektP
EIAV-liaoning_rev   -----------------------------------dpqrpLDndEwcrIlrQslPeekiP
FIVpuma_rev         -------------------------------------------------kYsiyvsGiqP
JDV_rev             MmeegrKeepeeRgekstmrDllqravdkghltarppvhhVklgrsntVsiaecarGyrP
HIV-2_rev           MSeRadEe---------------------------glqgkLr---llrLlHQtnPypqgP
HIV-1_rev           MAgRsgd----------------------------sdeelIrtvrlikLlYQsnPppnpe
MVV_rev             MAsKesKpsRtTRr---------gmepplrEtWnqvlqelVkrqqqeeeeqQglvsGlqa
SMLV_rev            -mdagaKhiRFTgeetwcevtmgeegk------kkqegcnkEqqDiqnIkYpkiPtGhsh
CAEV_rev            MvlmakvysyYSsriirmhsDk-----------scdmcmstscaDiqnskYpdiPtGhsh

OLV_rev             sE--------------nlehekRtspRsf-iqiW---RatvQawktspwgkgWKk-ilym
EIAV-Ireland_rev    sQTCiARRel------GpgpvqstpskRe---RW--LRGqiQqaEslqeqleWRi-rgvq
EIAV-liaoning_rev   sQTCiARRhl------GpgpvscvpgRRd---sW--LRGqvQhaEalqeqleWRi-rgvq
FIVpuma_rev         cd------gq------snrnkRRnKqRRrrrrRW--rtiMnsleDR--------------
JDV_rev             cR--------------GrrpaRRppiRRh-psfWgtLRGLvseaQRrqedrmsdlenrMa
HIV-2_rev           ----------------GtasqRRnRrRRr-rrQW--LR-LvalanKlcavpdpptdspLd
HIV-1_rev           ----------------GtrqaRRnRrRRw-reR---qRqIhsisERilgtylgRsaepvp
MVV_rev             sKadqiytgnsgdrstGgiggktkKkRgw--yKW--LRkLrareknipsqfy----pdMe
SMLV_rev            l---------------GnksrRRrRksgf--wRW--LRGIrRqrntpkdgnk----egLv
CAEV_rev            h---------------GnksrRRrRksgf--wRW--LRGIrQqrnKrksdst----esLe

OLV_rev             tlLplLtlQiwmEetgwngd----------------------------------------
EIAV-Ireland_rev    qsaeaLr-Evn--qgiwkElqwtrrlrgDyss----------------------------
EIAV-liaoning_rev   qtakeLE-kvn--keiwrElqytrrqhgDyGs----------------------------
FIVpuma_rev         -----fE-kLf-------------------------------------------------
JDV_rev             eleerfE-DLaLvdsggknpaapaqsvsppssnpfayslshfskskrvdcgekgnrwgrp
HIV-2_rev           raIqhLQ-rLTi-qelpdpptdlpesnsNqGl----------------------------
HIV-1_rev           lqLppLE-rLTLDcnedcgtsgtqgvgspqil----------------------------
MVV_rev             snMvgmE-nLTLEtqlednalynpathigdma----------------------------
SMLV_rev            ecVgaLa-ELTLDgvmeeEsaeaahhstDdGn----------------------------
CAEV_rev            pcLgaLa-ELTLEgamekgpaeaarpsaDdGn----------------------------

OLV_rev             -------------------------------------
EIAV-Ireland_rev    --fysskrEErrWgEesKprilKPGdSKrRRKhl---
EIAV-liaoning_rev   -fgdyrrrEEerWgESs-prvlKPGdSKrRRKhl---
FIVpuma_rev         ----------------------------genRpq---
JDV_rev             gafpgagisEldWiESgdggderPkggRypRggntps
HIV-2_rev           -------------aET---------------------
HIV-1_rev           -------vEsptvlESgtKE-----------------
MVV_rev             ----mdgrEwmeWrESaQKEkrKgGlSgqRtnaypgk
SMLV_rev            ------ldkwtaWrtppQK------------------
CAEV_rev            ------ldkwmaWrtpqK-------------------
```

Fig S4. Alignment of representative lentivirus Rev proteins.