Supplementary Data Contents

(please use links below or bookmarks to navigate)

S1 - Mutational Models

S2 - Retrovirus Structures

S3 - Motifs

S4 - RetroTector output HERV-Fc1

1. LTRID
2. RetroVID
3. ORFID
   - 3.1. ORFID-Gag
     - 3.1.1. Gag Putein
   - 3.2. ORFID-Pro
     - 3.2.1. Pro Putein
   - 3.3. ORFID-Pol
     - 3.3.1. Pol Putein
   - 3.4. ORFID-Env
     - 3.4.1. Env Putein
4. XonID
5. Chainview

S5 - Genes in Reference Retroviruses

S6 - RetroTector vs. RepeatMasker

S7 - RetroTector vs. HERVd

S8 - Detailed comparison of ReTe with RepeatMasker

S9 - ReTe vs RM, galGal3 canFam2 and hg18

S10 - ReTe vs RM galGal3

S11 - ReTe vs RM canFam2

S12 - ReTe vs RM hg18

S13 - ReTe vs RM galGal3 smaller set

S14 - ReTe vs RM canFam2 smaller set

S15 - ReTe vs RM hg18 smaller set

S16 - Description of some proviral chains missed in older or recent versions of RM

S17 - Two ERVs missed by RM.pdf

MUTATIONAL MODELS USED TO RECONSTRUCT DEGRADED RETROVIRAL SEQUENCES

In order to test the detection limit of RetroTector with regard to degraded retroviral sequences, such as old integrations, a test set was created. This test set consists of a retroviral sequence that has been mutated to varying degrees according to four different models.

Data set
The test set sequences are all based on the complete genome of HIV-1, isolate MN (HIVMNCG, Genbank accession number M17449), with degradation ranging from 1 to 60 or 90 percent mutations, with one percent intervals. To avoid biases and increase variability, all test set sequences are based on the original, unmutated HIVMNCG, as opposed to letting the highly degraded sequences build upon the less degraded ones. For each level of degradation, and for each model, 20 sequences were generated.

RetroTector is designed for the analysis of genomes, so each test set sequence was surrounded by an insertional repeat and 16000 random bases, to be in a proper context.

Random model
The first model uses equal probabilities for all positions to be substituted, and equal probabilities for all nucleotides (i.e. 0.25). This is also known as the Jukes-Cantor model.
Mutation ranging from 1 to 60 % with 1 % intervals, and 20 sequences generated for each level, yields a set of 1200 sequences.

Kimura model
This model uses substitution frequencies according to Kimura's two parameter model, where transitions are twice as probable as transversions. All positions in the sequence have the same probability of mutation. As above, 1200 sequences were generated.

Indel model
Most endogenous retroviruses seem to lack function. Without function, there is no selection to keep the ERV intact. A reasonable hypothesis is that the mutational decay of non-functional retroviral integrations is similar to that of pseudogenes. In this model, substitutions follow the same pattern as in the Kimura model. Insertions and deletions are added according to the frequencies found by analyzing data sets of human and murid pseudogenes (Gu and Li, 1995; Ophir and Graur, 1997).

Ophir and Graur (1997) state that deletions occur on average once every 40 substitutions and insertions once every 100 substitutions. The mean size of deletions is 4.67 and the standard deviation is 0.90. Therefore, the upper limit for number of bases deleted is set to 4.67 + 2*0.90, or approximately 7, since this should cover 95% of all cases, assuming a normal distribution. Gu and Li give a formula for the size distribution (which is not normal, but mean + 2*SD should still cover most cases) of deletions:

F(probability of bases deleted)=0.52*(number of bases deleted)^-1.86

Insertions have a mean size of 8.03 and a standard deviation of 2.46, so the upper limit is set to 13 bases. The formula according to Gu and Li is F=0.54*(No. of bases inserted)^-1.95

When the model calls for a deletion, a random number between 0 and 1 is generated. According to the formula above, the number of bases to be deleted is determined by the following boundaries:

x <= 0.52	1 base deleted
0.52 < x <= 0.66	2 bases deleted
0.66 < x <= 0.73	3 bases deleted
0.73 < x <= 0.77	4 bases deleted and so on.

Mutation degree lies between 1 and 60 %, with 1 % distance between sets. Twenty sets gives a total of 1200 sequences.

Exogenous model

An active retrovirus has exogenous as well as endogenous phases, and is subjected to selectional pressure during each of these. Mutations that adversely affect the ability of the retrovirus to infect and replicate, are eliminated from the population.

This model uses the 217 complete genomes from HIV-1 and SIV that are aligned on codon level in the HIV Sequence Database at Los Alamos Laboratories, US. The statistics for substitutions and gaps for each position in this alignment is calculated, and these statistics are then the basis for the model.

When the model calls for a substitution, the frequency of nucleotides in the corresponding position in the alignment determines the probability of which nucleotide to substitute for. Deletions and insertions are also dependent on the large alignment from Los Alamos: bases can only be deleted if there is a gap in some of the other sequences, and the probability is ruled by the frequency of gap in the alignment. Only whole codons are deleted, using these borders (they follow the rules from the indel model):

x <= 0.7	One codon removed
0.7 < x <= 0.9	Two codons removed
0.9< x <= 1	Three codons removed

Insertions are dependent on the other sequences in the alignment, that decide the probabilities for different bases to be inserted. Only whole codons are inserted.

The model was run with mutation ranging from 1 to 90 percent mutation, with 20 sequences for each level, i.e. 1800 reconstructed sequences.

RESULTS

Random model

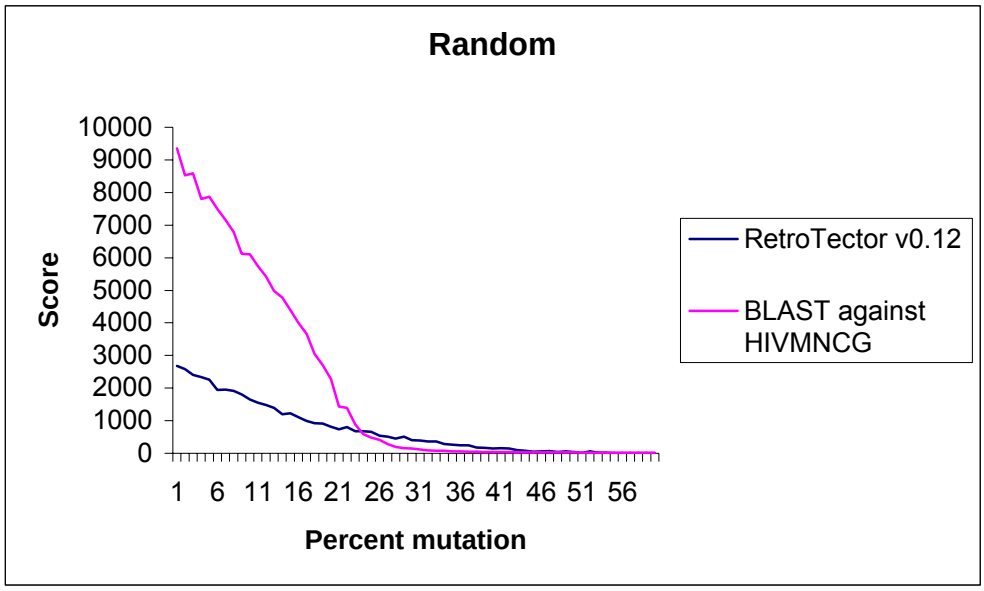


Figure S1.1: Average scores for 20 sets of random mutation.

When looking at the trend in this diagram, it is clear that BLAST has a steeper slant. This shows that BLAST is more negatively affected by the degradation of the sequence (figure S1.1).

After normalising by dividing each score with the score for the original, unmutated HIVMNCG, the graphs show that for low degrees of mutation, BLAST and RetroTector perform equally well, which was expected. However, when the mutations increase (around 15 % and up), RetroTector is much better at recognising the sequence as a retrovirus (figure S1.2). As discussed elsewhere in this work, a reasonable limit for recognition as a "retrovirus" is a RetroTector chain score of 300.

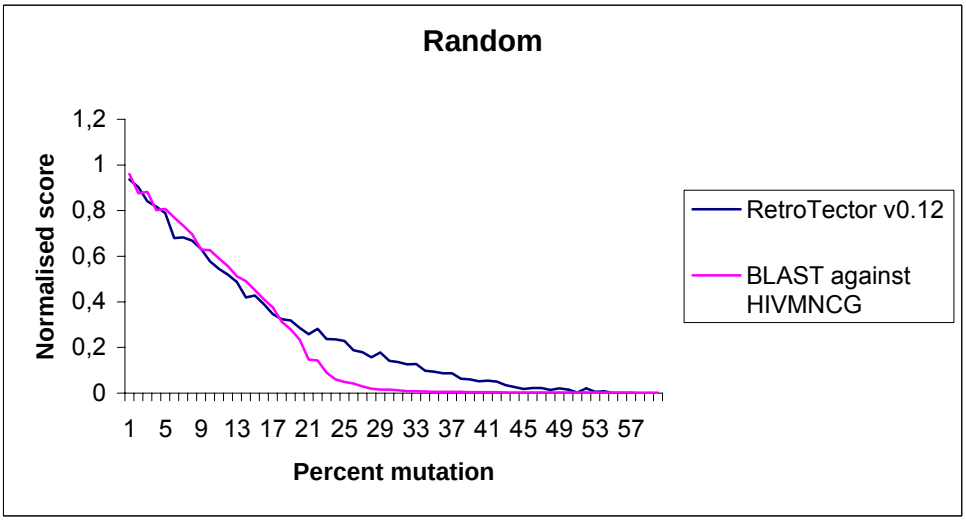


Figure S1.2: Average normalised scores for 20 sets of random mutation.

Now the objection can be raised that RetroTector scores the sequence according to rules for a general retrovirus, whereas BLAST compares to the original sequence HIVMNCG, and therefore cannot take advantage of conservative mutations like RetroTector. To address this, one sequence from each degree of mutation was matched against the entire non-redundant nucleotide database from GenBank (figures S1.3 and S1.4). Due to time restrictions, all 20 sets could not be analysed in this way. Since the mutation models are ruled by random variables, stochastic variations in the sequences produced must be taken into account.

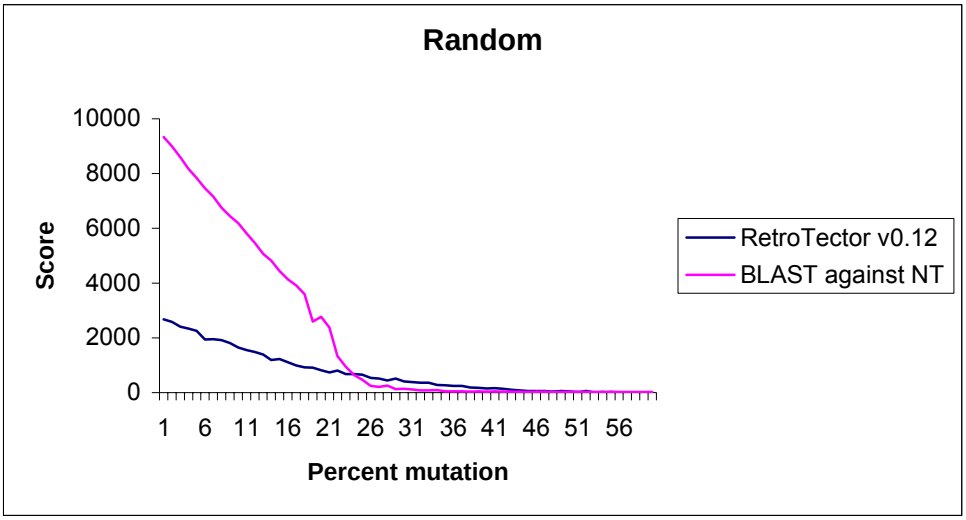


Figure S1.3: Score for one sequence matched with BLAST against the nt database, compared to the average score for 20 sequences analysed with RetroTector.

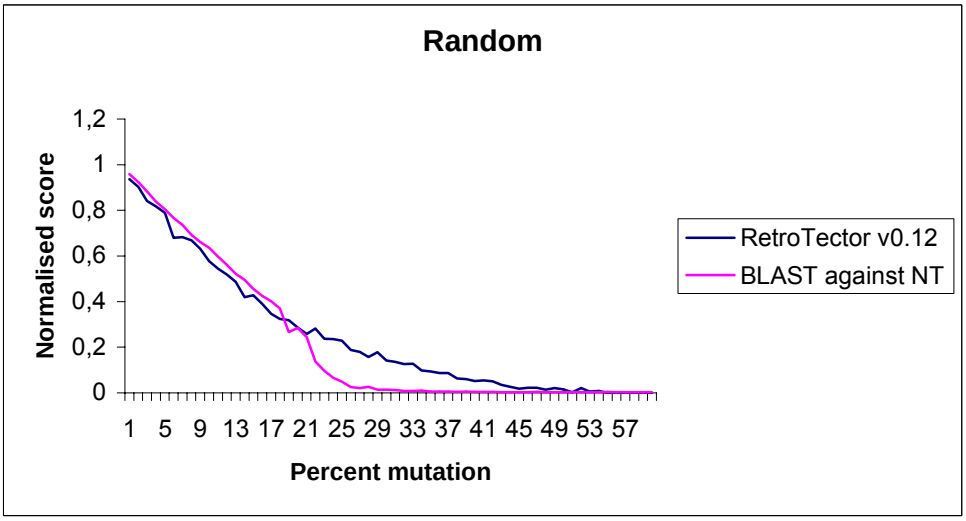


Figure S1.4: Normalised score for one sequence matched with BLAST against the nt database, compared to the average normalised score for 20 sequences analysed with RetroTector.

The trend is the same even when BLAST can choose between all sequences in the nonredundant nucleotide sequence (NT) database from Genbank. RetroTector still is better at

recognising retroviral properties when the degree of mutation increases above around 20 %. When the degree of mutation is low, BLAST obtains the highest score when matching against the original HIVMNCG. When mutation increases, other accessions take over, mainly sequences coding for one or a few specific HIV-1 proteins like env. When the sequence is highly degraded (above 40 % degradation), BLAST reports the highest scores against completely unrelated entries in the database, like mouse BAC clones and human clones. In this model, the highest scores came from matching against HIVMNCG.

Kimura model

When the mutation model takes a step closer to reality, the trend seen with the randomly mutated sequences is even clearer. BLAST scores drop fast and vanish at around 30%, while RetroTector detects the sequence as retroviral up till 45% mutation. (figures S1.5 to S1.8).

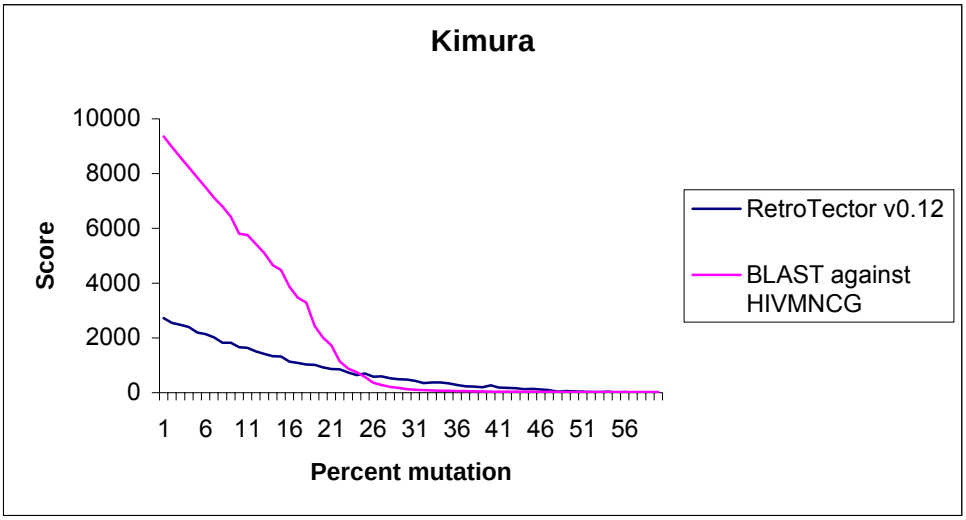


Figure S1.5: Average scores for 20 sets of mutation according to the Kimura model.

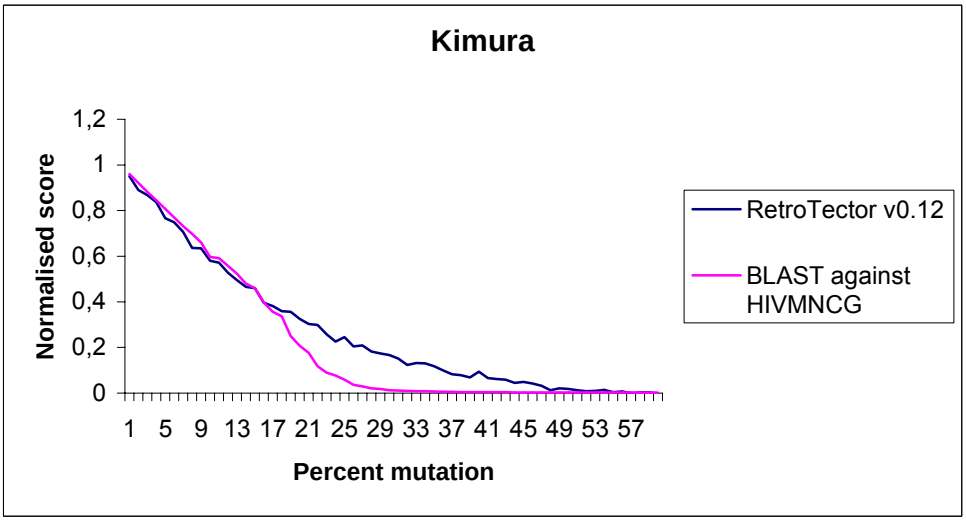


Figure S1.6: Average normalised scores for 20 sets of mutation according to the Kimura model.

When analysing one sequence from each percent mutation with BLAST and the nt database, 37 highest-scoring hits is against HIVMNCG. A few hits are to other whole-genome HIV-1 sequences, and quite many against parts of HIV-1 genomes like env or gag, as in the random model. As the sequence is almost totally degraded, a few hits to completely unrelated sequences appear.

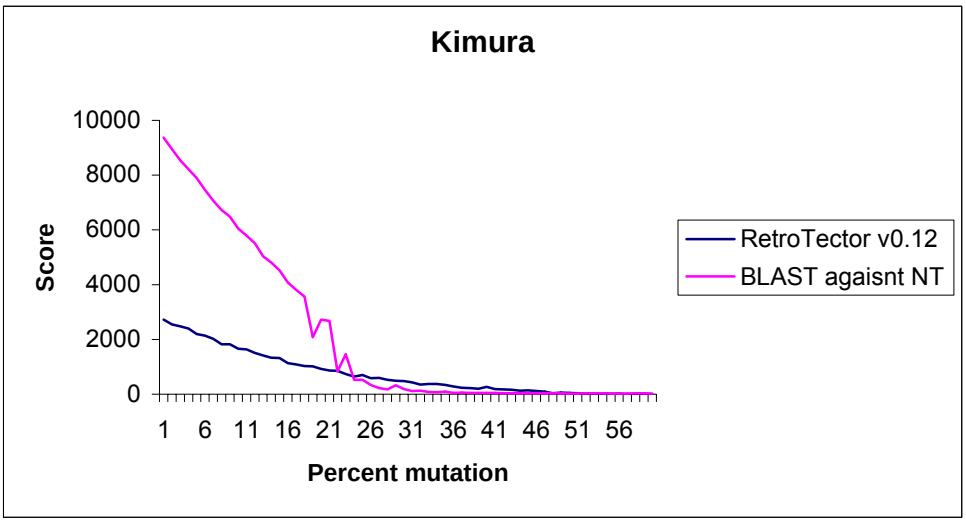


Figure S1.7: Score for one sequence matched with BLAST against the nt database, compared to the average score for 20 sequences analysed with RetroTector.

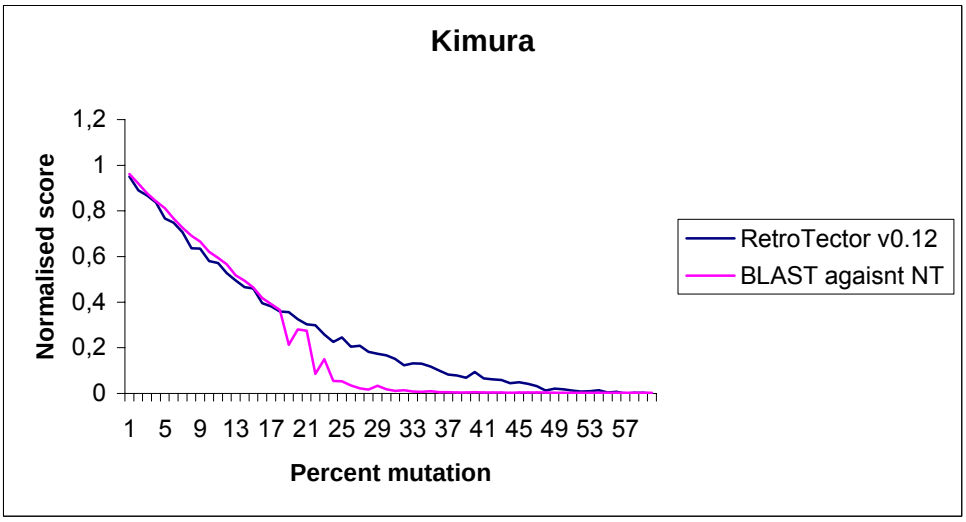


Figure S1.8: Normalised score for one sequence at each level of mutation analysed with BLAST and the nt database, compared to the average normalised score for 20 sequences analysed with RetroTector.

Indel model

In this model, not only substitutions alter the sequences, but deletions and insertions allow the sequences to degrade even faster. This is reflected in the scores that drop much faster than in

the previous two models. Here, RetroTector shows a clear dominance even for sequences with a low degree of mutation (figures S1.9 to S1.12).

The indel model is the model that comes closest to mimicking the real situation for endogenous retroviruses. A mutation level of 20 % would reflect an element that integrated 100 million years ago, assuming a mutation rate of 0.2 % divergence per million years (Li, 1997). BLAST does not recognize an element this degraded, but RetroTector is still able to detect it (fig S1.9).

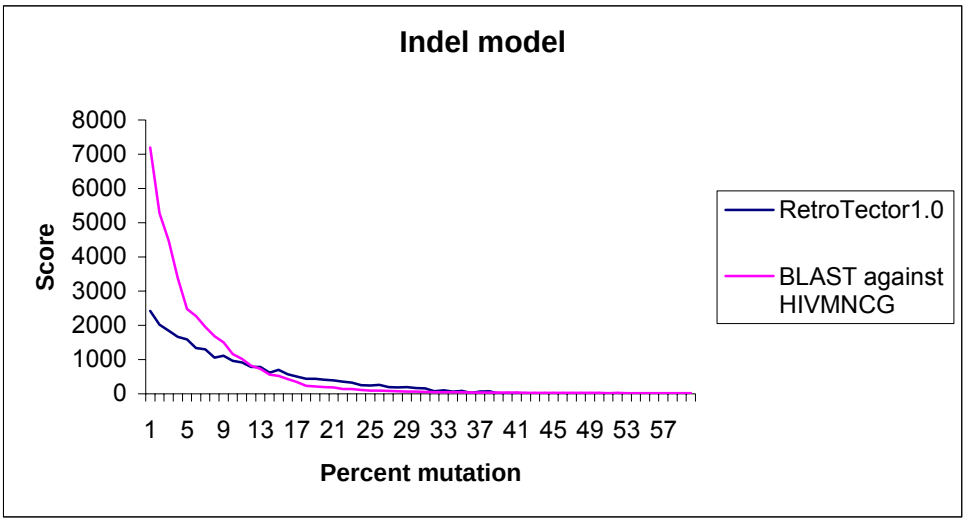


Figure S1.9. Average score for 20 sets of sequences mutated according to the indel model.

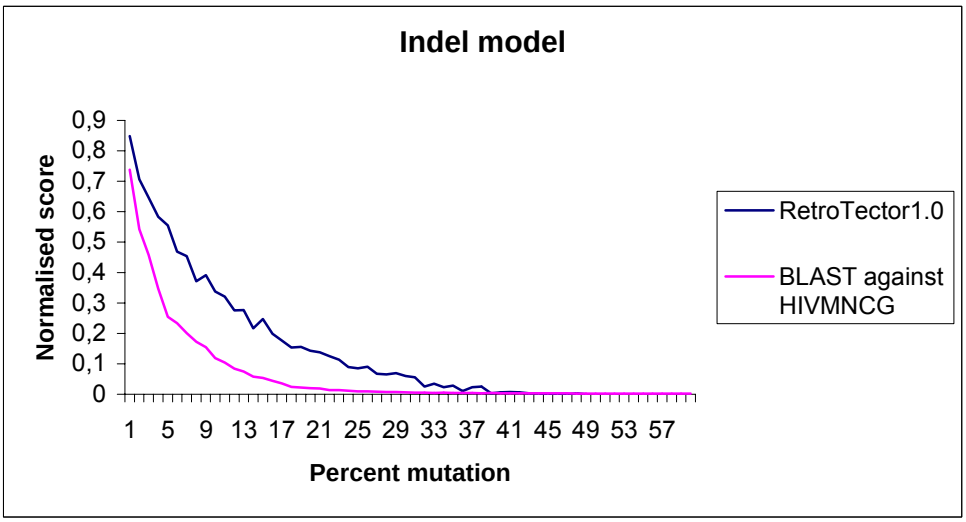


Figure S1.10: Average normalised score for the indel model sequences.

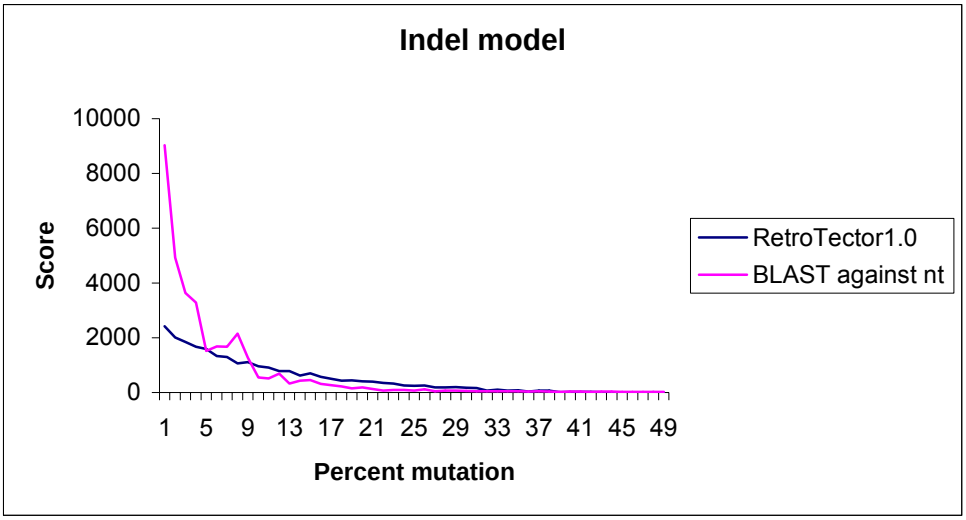


Figure S1.11: Score for one sequence analysed with BLAST and nt compared to the average score for 20 sequences analysed with RetroTector.

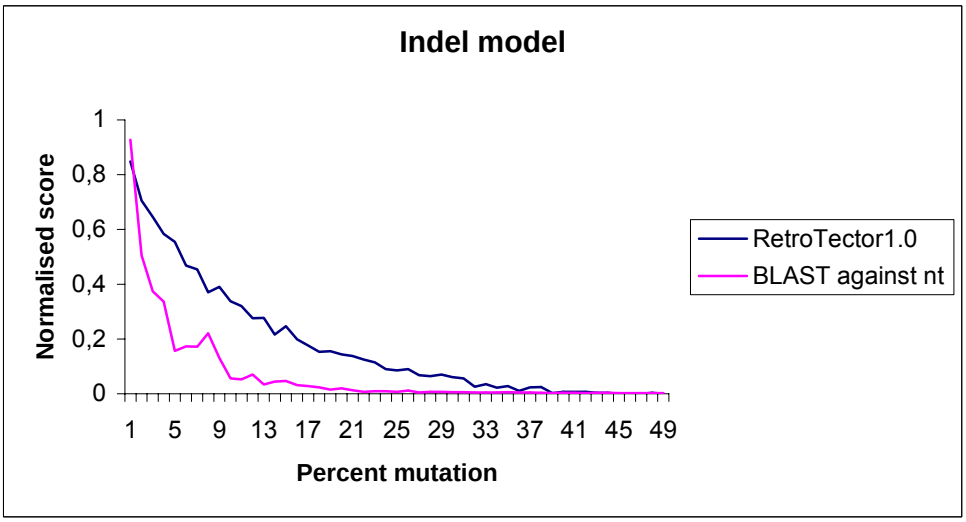


Figure S1.12: Normalised score for one sequence analysed with BLAST and nt, together with the average normalised score for 20 sequences assessed with RetroTector.

Puteins

RetroTector attempts to reconstruct the retroviral proteins. Study of the Pol puteins shows that as the mutation of the sequence increases, it becomes more difficult for RetroTector to achieve a perfect replica of the original protein, since this information is lost during the mutation process. Yet, at a mutation level as high as 20 %, over half of the residues in the putein are identical to the original protein.

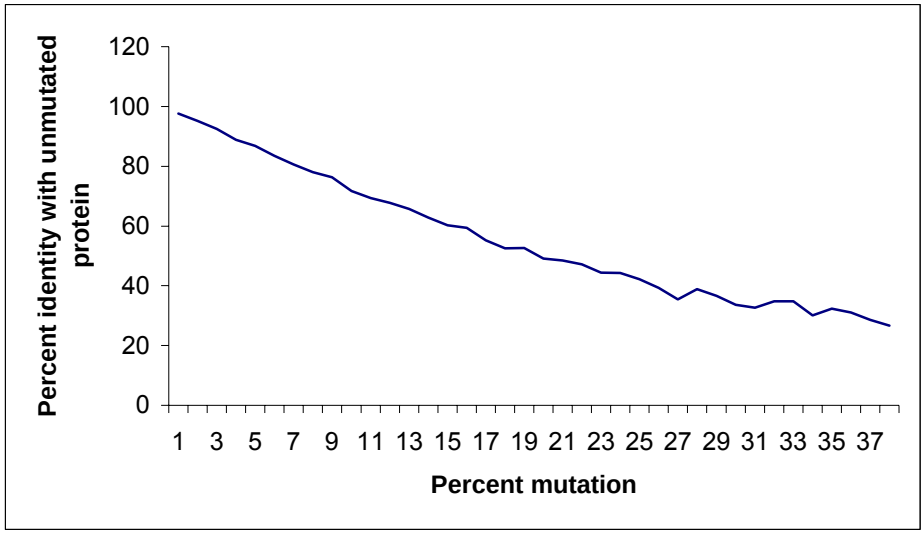


Figure S1.13: Sequence identity for Pol puteins in the indel model.

Exogenous model

This model uses real data to mimic mutation of an exogenous retrovirus under selectional pressure. Our expectation was that BLAST would find it easier to keep up with RetroTector when most mutations are discarded due to the statistic from Los Alamos. However, this is the model that truly demonstrated the power of RetroTector when compared to BLAST – RetroTector scores stays on a constant high level, with a normalised score close to 1, whereas BLAST scores drops as in the previous studies (figures S1.14 to S1.17).

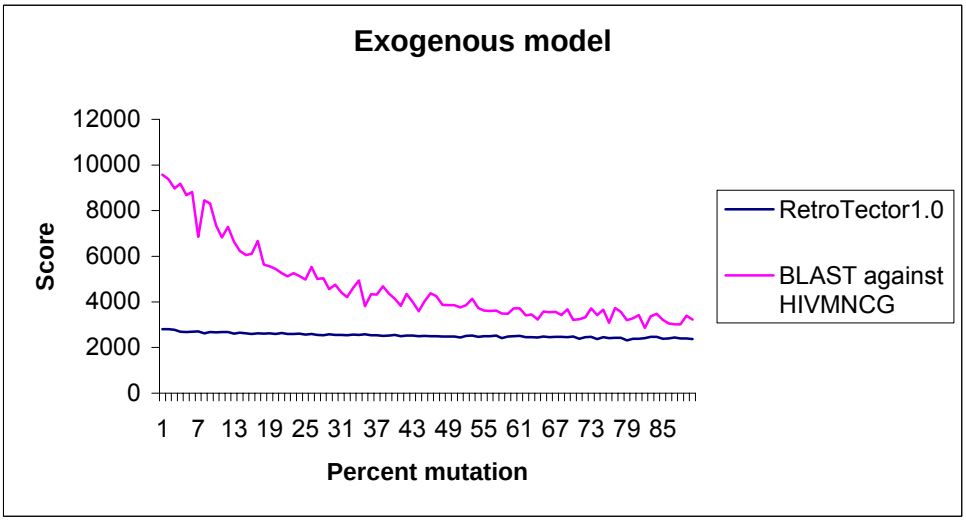


Figure S1.14: Average scores for sequences mutated according to real data, reflecting the properties of exogenous HIV-1. Note that the RetroTector score remains virtually constant.

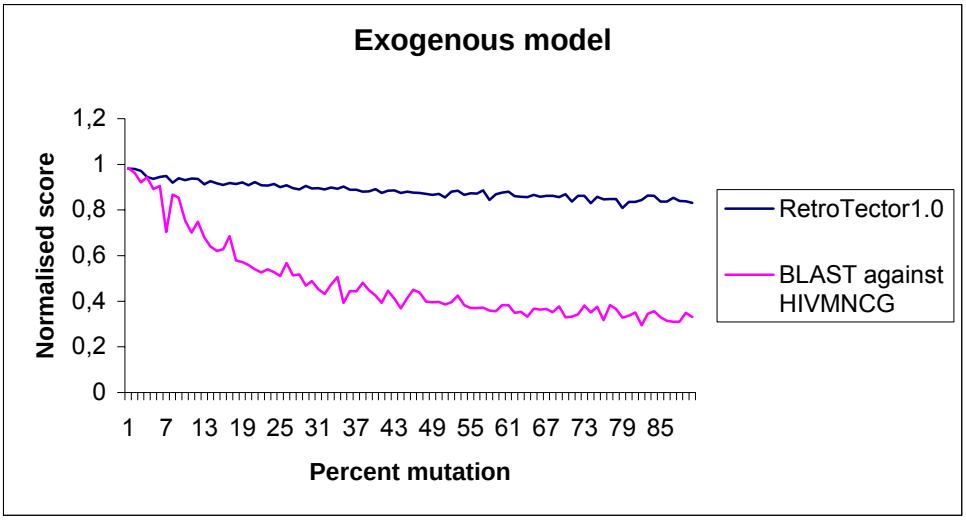


Figure S1.15: Normalised score for the exogenous model sequences.

When BLAST is free to utilise the entire nt database, the scores still drop, although not as much as in the previous models. HIVMNCG is the highest scoring match in 35 out of 90 cases, and other HIV-1 sequences yield the rest.

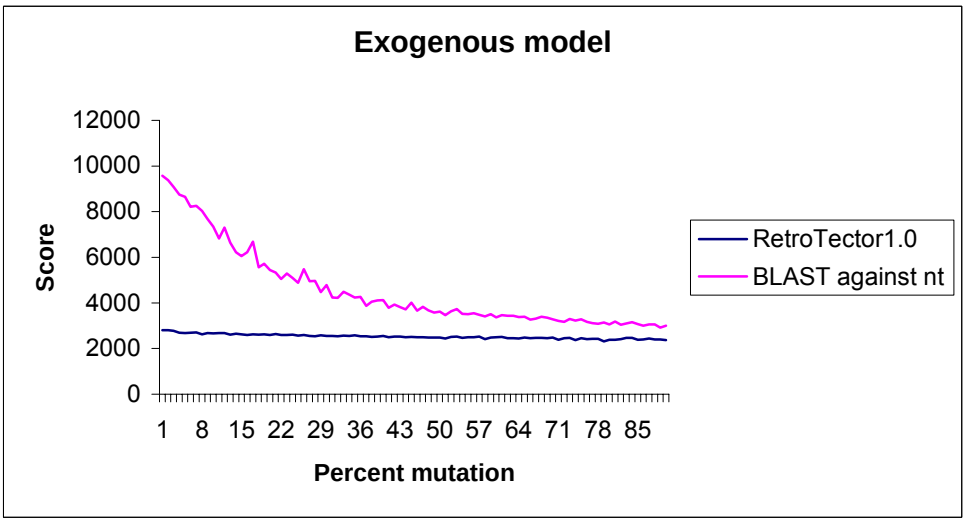


Figure S1.16: Average score for sequences matched against nt with BLAST, compared to the average result from 20 sequences scored with RetroTector.

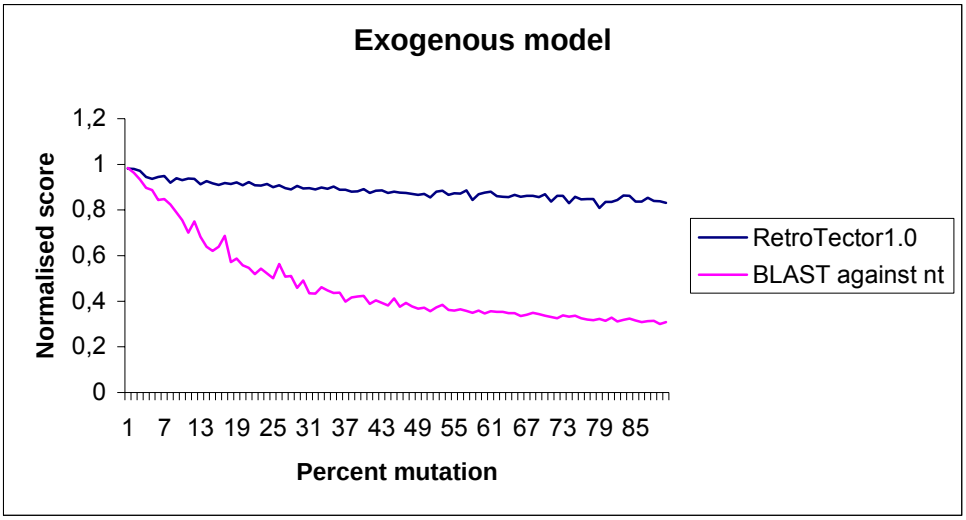


Figure S1.17: Normalised values for the scores in the exogenous model.

Puteins

RetroTector is able to reconstruct the Pol protein almost perfectly throughout the exogenous model simulations. This is encouraging. It is also expected, since the mutation model does not allow mutations that would be detrimental to the virus function, and the Pol protein has many conserved portions.

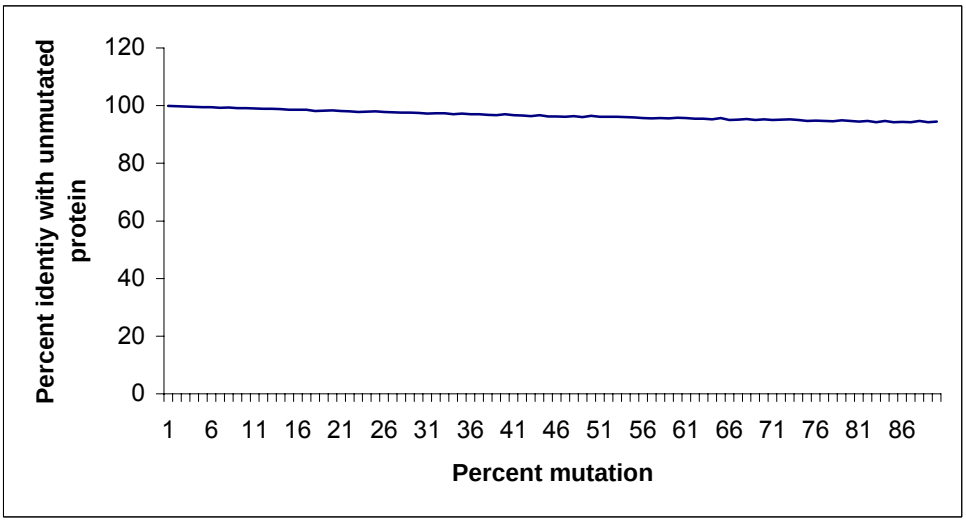


Figure S1.18: Sequence similarity for Pol puteins in the exogenous model.

Sample alignment

CLUSTALW alignment with three sequences:
The original HIVMNCG Pol protein
A putein reconstructed from a sequence with 10 % mutation according to the indel model
A putein reconstructed from a sequence with 75 % mutation according to the exogenous model

```
HIVMNCG           PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALIEICTEMEKEGKISKIGPENPYNTPV
Exogenous75%      PISPIDTVPVKLKPGMDGPKVKQWPLTKEKIKALTEICTEMEKEGKISKIGPENPYNTPV
Indel10%          PISPIQTVPTKLKPGMDRPKVKQN---IDKIKTLIEICTGMEKEGKISKIGPEN-IQYSI
                  *****:***.******* *****     :***:* **** **********  : .:

HIVMNCG           FAIKKKDS----TKWRKLVD-FRELNKKTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAY
Exogenous75%      FAIKKKDS----TKWRKLVD-FRELNKKTQDFWEVQLGIPHPSGLKKKKSVTVLDVGDAY
Indel10%          FAIKKKDSQRZVTEWRKLVVIFRELNKKTQDFWEVQZGIPHPA---KKKSVTMLYVGYAY
                  ********    *:*****  ************** *****:   ******:* ** **

HIVMNCG           FSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQN
Exogenous75%      FSVPLDKNFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKKN
Indel10%          FSVPL----RKRTAFTMPSINNETPEIRYQZNVLSQGCKGSPAIFQSSD-KILEPFTKQN
                  *****    ** ****:******** **** ***.** **********  ****** *:*

HIVMNCG           PDIVIYQYMDDLY--VGSDLEIGQHRAKIEELRRHLLRWGFTTPDKKHQKE----PPFLW
Exogenous75%      PDIVIYQYMDDLY--VGSDLEIGQHGTKIEELREHLLKWGFTTPDKKHQKE----PPFLW
Indel10%          PDIVIYQYRMVCMZDLSTDLEIGQHZAKIEELRRHLLTWGFTTPDKKHQDAHICNPPFLW
                  ********       :.:******* :******.*** ***********.     *****

HIVMNCG           MGYELHPDKWTVQPIVL----PEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRG
Exogenous75%      MGYELHPDKWTVQPIML----PEKDSWTVNDIQKLVGKLNWASQIYP-IKVKQLCKLLRG
Indel10%          MGYVLHPDKRTVQSIVLNMAKPEKESWTVNDIQKLVRKLNWAYQFYAGIKVKQLCKLLRE
                  *** ***** ***.*:*    ***:*********** ***** *:*. **********

HIVMNCG           TKALTEV-IPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEVQKQGQGQWTYQIYQE
Exogenous75%      AKTLTEV-IPLTKEA--ELAENREILKEPVHGVYYDPSKDLIAEVQKQGQGQWTYQIYQE
Indel10%          TKALTEVRVPPTEEVELELAENREIQ--YVHGVYYDPSKDLIAEVHEQGLGQWTLN-FQE
                  :*:**** :* *:*.  ********    ****************::** **** : :**

HIVMNCG           PFKNLKTGKYARMRGAHTNDVKQLTEAVQKIATESIVIWGKTPKFRLPIQKETWETWWTE
Exogenous75%      PFKNLKTGKYAKRRGAHTNDVKQLAEAVQKIAKESIVIWGKTPKFRLPIQKETWETWWTD
Indel10%          RFIN-RNWQMCKNGGCPHZRCKISNRGMQZIATESIVIWGKTPKLDYPTKRNIGNMVDRD
                   * * :. : .:  *.     *   ..:* **.**********:  * :::  :    :

HIVMNCG           YTXATWIPEWEVVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKKGKAGYVTNRGRQ
Exogenous75%      YSQATWIPEWEVVNTPPLVKLWYQLEKDPIVGAETFYVDGAANRETKZGKAGYVTDRGRK
Indel10%          TZ-ATYISEWEVVNTPSLVKLWDQLEKEPIVGAQTFYVDAAANTETERGKAGYVTNRGRQ
                     **:*.********.***** ****:*****:*****.*** **: *******:***:

HIVMNCG           KVVSLTDTTNQKTELQAIHLALQDSGLEVNIVTDSQYALGIIQAQPDKSESELVSQIIEQ
Exogenous75%      KVVSLPETTNQKTELQAIHLALQDSGLEVNIVTDSQYALGIIQAQPDKSESELVSQIIEQ
Indel10%          KIVSLTDTTNQKTELQAIHQPLKDSVLEVNIVADSZZALGIIZAQPDKGESELVSQLIEQ
                  *:***.:************ .*:** ******:**  ***** *****.*******:***

HIVMNCG           LIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQEDHEKYHSNWRAMAS
Exogenous75%      LIKKEKVYLSWVPAHKGIGRNEQVDKLVSTGIRKVLFLDGIDKAQEEHEKYHSNWRAMAS
Indel10%          LIKTEKVYVAWVPAHKEIGGNEQVDKLAECWKQESTIFRLIAKSQEDHEKYRSND-SMTS
                  ***.****::****** ** *******..   ::  ::  * *:**:****:**  :*:*

HIVMNCG           DFNLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYI
Exogenous75%      DFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYM
Indel10%          DFNLPPIVAKEIVASCDKCQLKGEGIHGQVHCSAGIZQLNCTHLEGZVIPVAVHVASVYI
                  ******:*****************.:****.**.** **:****** ** ******* *:
```

```
HIVMNCG           EAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGPNFTSTTVKAACWWTGIKQEFGIPYN
Exogenous75%      EAEVIPAETGQETAYFLLKLAARWPVKIIHTDNGTNFTSSTVKAACWWTGIQQEFGIPYN
Indel10%          EVEISPAEAGQETAYFLLKLAGRWSVKTIHTDN-ANLTSTTLRPPFWWTGIKZEFGIPYN
                  *.*: ***:************.**.** ***** .*:**:*::.. *****: *******

HIVMNCG           PQSQGVIESMNKELKKIIGQVRDQAEHLKRAVQMAVFIHNFKRKGGIGGYSAGERIVGII
Exogenous75%      PQSQGVIESMNKELKKIIGQVRDQAEHLKRAVQMAVFIHNFKRKGGIGGYSAGERIIDII
Indel10%          PQSQGVIESMNKELKKIIVQVKDQAEH--RTVQLAVFINNYKKKGGIRGYSVGERIVGTI
                  ****************** **:*****  *:**:****:*:*:**** ***.****:. *

HIVMNCG           ATDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDNNDIKVVPRRK
Exogenous75%      ATDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDNNDIKVVPRRK
Indel10%          ATDIQAKELQKQIT--QNFRDYYRNSRDPVWKVPAKLLW-GEGAAVTQDNDDRKVVPRRK
                  *****:********  **** ***:****:** ****** ****.* ***:* *******

HIVMNCG           AKVIRDYGKQTAGDDCVASRQDED
Exogenous75%      AKIIRDYGKQTAGDDCVAGRQDED
Indel10%          AKAIRDYGKQTADDDCVASRQDZD
                  ** *********.*****.*** *
```

REFERENCES

Gu, X. and Li, W.H. (1995) The size distribution of insertions and deletions in human and rodent pseudogenes suggests the logarithmic gap penalty for sequence alignment. *J Mol Evol*, 40, 464-473.

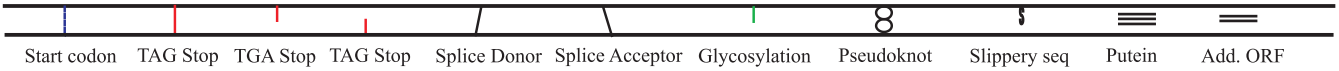
Li, W.H. (1997) *Molecular Evolution*. Sinnauer Associates, Inc., Publishers, Sunderland, MA, USA.

Ophir, R. and Graur, D. (1997) Patterns and rates of indel evolution in processed pseudogenes from humans and murids. *Gene*, 205, 191-202.

Selected RetroTector© outputs

Selected endogenous retroviruses (ERVs) and exogenous retroviruses (XRVs) extracted from the human genome (hg16), and from annotations in GenBank.
(Graphic display of LTRs depend on the availability of full LTRs [U5-R-U3].)
(Graphic display quality of ERVs depend on their ages [accumulated mutations])

Legend



Reading frames are indicated as 1, 2 and 3

Motifs

5LT **5'LTR**

PBS **Primer binding site**

Gag **Group specific antigen**

MA1, MA2	Matrix
CA0, CA1, CA2	Capsid
NC1, NC2	Nucleocapsid (zinc fingers)

Pro **Protease**

DU0, DU1, DU2	dUTPase (deoxyuridine triphosphatase)
PR1, PR2, PR3	Protease

Pol **Polymerase**

RT1, RT2, RT3, RT4, RT5, RT6	Reverse transcriptase
RT7	RNaseH (Ribonuclease H)
DL1, DL2	dUTPase (deoxyuridine triphosphatase)
In1, IN2, IN3, IN4, IN5, IN6, IN7	Integrase

Env **Envelope**

SU2, SU3	Surface unit
TM2, TM3, TM4, TM5	Transmembrane protein

PPT **Polypurine tract**

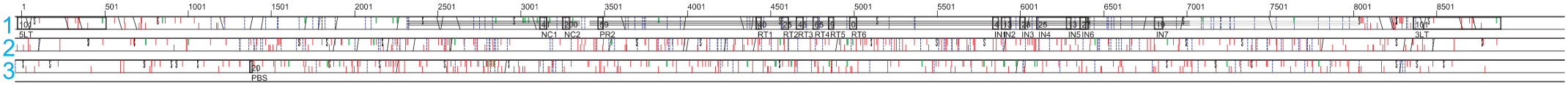
3LT **3'LTR**

Retrovirus Abbreviations

FLV	Feline Leukemia Virus
HERV	Human Endogenous Retrovirus
HFV	Human Foamy Virus
HIV	Human Immunodeficiency Virus
HTLV	Human T-Cell Leukemia Virus
MMTV	Mouse Mammary Tumor Virus
MoMLV	Moloney Murine Leukemia Virus
MPMV	Mason-Pfizer Monkey Virus
RSV	Rous Sarcoma Virus
WDSV	Walleye Dermal Sarcoma Virus

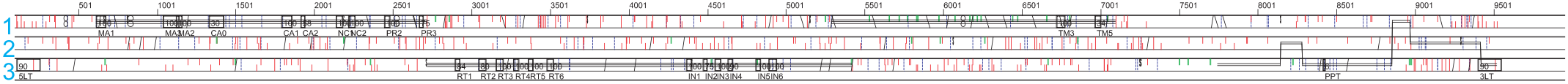
Errantivirus

Cer1/Gypsy-U15406



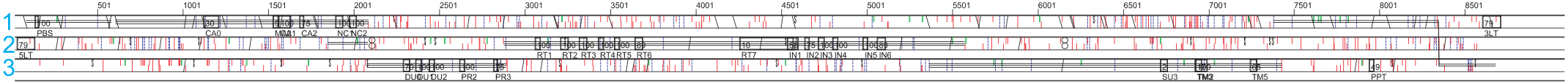
Alpha

RSV-J02342

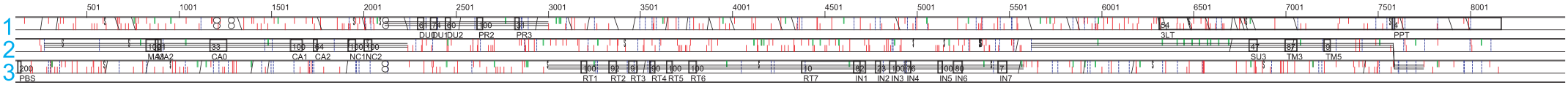


Beta

MMTV-NC_001503

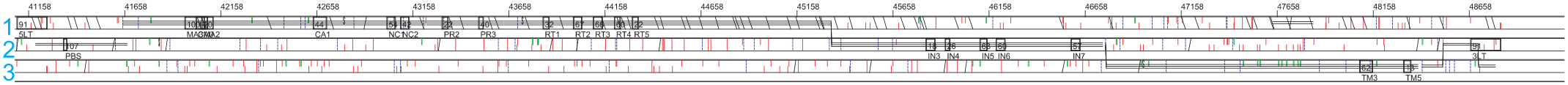


MPMV-NC_001550



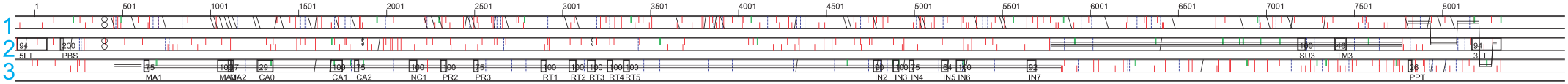
Gamma-like HERV

HERV-Fc1-AL354686

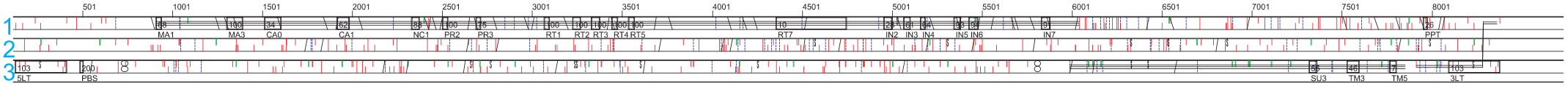


Gamma

MoMLV-J02255

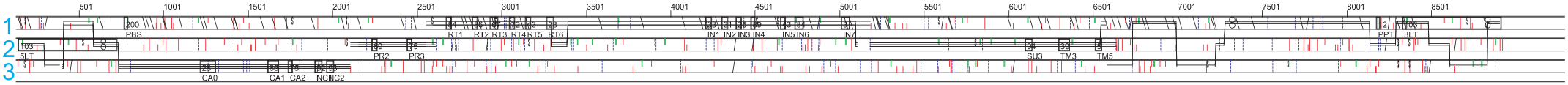


FLV-NC_001940



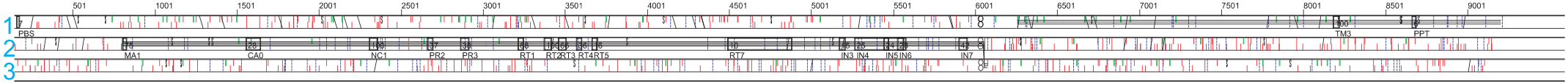
Delta

HTLV2-M10060



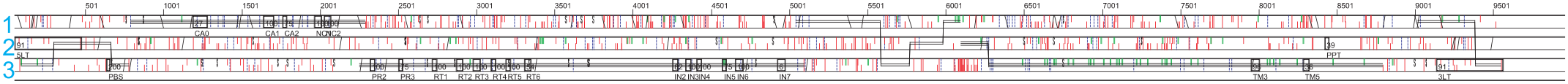
Epsilon

WDSV-NC_001867



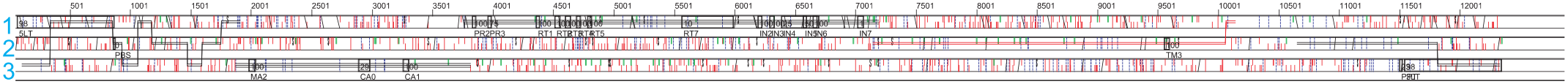
Lenti

HIV1-K03455



Spumalike

HFV-NC_001736



```
{ For RetroTector 1.0
{ Copyright ((c)) 2000-2005, Jonas Blomberg & Göran Sperber. All Rights Reserved.
{  Last changed 2005 12 06
{          !    ! !          !    !      !    !    !    !                                    !
Motifs::
          1ANN  C ABCDELSGO       CA     CA0  050  0.75 CAStartNNData.txt                  CA Start NN
        101P    D A             CA     CA1  100       adImQGPsESFvdFAnRlikav              RSV
        102P    D B             CA     CA1  100       ntVRQGSKEPYpDFVaRLqdVa              HERV-K10
        103P    D B             CA     CA1  100       tkIvQGPqEPFsdFVaRmteaa              IAPcri
        104P    D B             CA     CA1  100       agLKQGNEESYeTFIsRLeeAv              MMTV
        105P    D B             CA     CA1  100       tgVKQGPDEPFaDFVhRLitTa              MPMV
        106P    D C             CA     CA1  100       rtITQGkDESPAaFMERLlEGF              BAEV
        107P    D C             CA     CA1  100       seVIQGkEESPAkFHERLcEAY              HERV E (4-1)
        108P    D C             CA     CA1  100       kgITQGPNESPsAFLeRLkeAy              MuLV
        109P    D C             CA     CA1  100       fdIQQEKDEGPiRFLdRLkeQm              RTVL-I =HERV-I
        110P    D C             CA     CA1  100       teVVQGPGEXPGaFLECLqEAY              S71
        111P    D D             CA     CA1  100       asIlQGLeEPYhaFVeRlnial              HTLV-I
        112P    D L             CA     CA1  100       ldIrQGPkEPFrdYVdRfyktl              HIV-1
        113P    D G             CA     CA1  100       llNaEVraDALhaFIsGLkkal              Gypsy
        114P    D G             CA     CA1  100       akktlydevcLnAFIsRIrepl              Zam
        115P    D G             CA     CA1  100       vffapptsvrsselHIGLseip              RIRE3
        116P    D S             CA     CA1  100       vYeILGLNarGQsIr                     HSRV
        201P    D A             CA     CA2  075       aksQpdiQ                            RSV
        202P    D B             CA     CA2  075       eQATqECR                            IAPcr
        203P    D B             CA     CA2  075       EnANslCq                            MMTV
        204P    D B             CA     CA2  075       EdANpaCq                            MPMV
        205P    D C             CA     CA2  075       qaaldigK                            BaEV
        206P    D C             CA     CA2  075       QsAPdiGr                            MuLV
        207P    D C             CA     CA2  075       aQASdirK                            S71
        208P    D D             CA     CA2  075       SnANkeCq                            HTLV-I
        209P    D L             CA     CA2  075       QnANpdCk                            HIV-1
        210P    D G             CA     CA2  075       REAEASIE                            Gypsy
        301P    D L             DL     DL1  100       GVIDEgyrGEI                         FIV
        401P    D L             DL     DL2  100       rqKIAqlii                           FIV
        501P    D B             DU     DU0  100       LPegtvglILGR                        HERVK10
        601P    D B             DU     DU1  100       SvVDSDykGEI                         HERVK10
        602P    D B             DU     DU1  100       GiIDSDytGEI                         HML1
        603P    D B             DU     DU1  100       GvIDSDfQGEI                         MMTV
        701P    D B             DU     DU2  100       rDRIAqllll                          HERVK10
        702P    D B             DU     DU2  100       gERIAvvavt                          HML1
        703P    D B             DU     DU2  100       gERIAqllll                          MMTV
        801P    D A             IN     IN1  100       lHvRsHSeVPG                         RSV
        802P    D B             IN     IN1  100       gHiRAHStLPG                         JSRV
        803P    D B             IN     IN1  100       tHiRAHTpLPG                         HML3
        804P    D B             IN     IN1  100       sHicshiqLPG                         HML6
        901P    D A             IN     IN2  075       dlHtalHig                           RSV
        902P    D B             IN     IN2  075       elHaltHvn                           HERVK10
        903P    D B             IN     IN2  075       efHkrfHvt                           IAPha
        904P    D B             IN     IN2  075       esHalhHqn                           MMTV
        905P    D C             IN     IN2  100       ldflHqltHlsfskt                     MLV
```

906P	D	D	IN	IN2	100	paelHsftHcg	HTLV-1
907P	D	L	IN	IN2	100	AqedHekyHsn	HIV-1
908P	D	S	IN	IN2	100	vlqaHnlaHtg	HSRV
909P	C	G	IN	IN2	100	lqeahraPyaiHpggtkm	Ananas comosus
1001P	D	A	IN	IN3	100	lvksCshCqkt	LPDV
1002P	D	A	IN	IN3	100	vvqtCphCnsa	RSV
1003P	D	B	IN	IN3	100	ivqhCtqCqvl	HERVK10
1004P	D	B	IN	IN3	100	ivtqCqnCcef	IAPha
1005P	D	B	IN	IN3	100	ivksCstCpqf	JSRV
1006P	D	B	IN	IN3	100	ivklCpnCpdw	MMTV
1007P	D	B	IN	IN3	100	ivkqCpiCvty	MPMV
1008P	D	C	IN	IN3	100	vtsaCkvCqqvnaga	BAEV
1009P	D	C	IN	IN3	100	itetCkaCaqv	MLV
1010P	D	D	IN	IN3	100	patlCetCqkl	BLV
1011P	D	L	IN	IN3	100	ivasCdkCqlk	HIV-1
1012P	D	S	IN	IN3	100	qlgrCqqClit	HSRV
1013P	C	G	IN	IN3	100	ltplCteqkvqadlgt	Gypsy
1100SPA	C	ABCDELSGO	IN	IN4	025	!(4)D(51-58)D(35)E	DDE (Integrase)
1201P	D	A	IN	IN4	100	lWQtDfTlep	RSV
1202P	D	B	IN	IN4	100	lWQmDvTHvp	HERVK10
1203P	D	B	IN	IN4	100	llQkDlTHvp	HML7
1204P	D	B	IN	IN4	100	lWQmDvTHvs	MMTV
1205P	D	C	IN	IN4	075	hWeidfte	MuLV, motif C
1206P	D	D	IN	IN4	100	iWQaDiTHyk	BLV
1207P	D	D	IN	IN4	100	IWQgditH	HTLV, motif C
1208P	D	L	IN	IN4	100	iWqlDcTHle	HIV-1
1209P	C	G	IN	IN4	100	kitmDFVtGlPrSQa	Ananas comosus
1301P	D	A	IN	IN5	100	kTDNgscft	RSV
1302P	D	B	IN	IN5	100	kTDNgpgyc	HERVK10
1303P	D	B	IN	IN5	100	kTDNtpght	HML3
1304P	D	B	IN	IN5	100	kTDNglaya	HML5
1305P	D	B	IN	IN5	100	kiDNgpayt	HML6
1306P	D	B	IN	IN5	100	kTDNapayv	MMTV
1307P	D	B	IN	IN5	100	ktDNgpgY	MPMV, motif D
1308P	D	C	IN	IN5	100	gSDNGPafvSQv	BAEV
1309P	D	C	IN	IN5	050	gTDNgpafv	MLV
1310P	D	C	IN	IN5	050	gtDNgpaF	MuLV, motif D
1311P	D	D	IN	IN5	050	nTDQganyt	BLV
1312P	D	D	IN	IN5	050	ntDNgpaY	HTLV-I,motif D
1313P	D	L	IN	IN5	075	hTDNgsnft	HIV-1
1314P	C	G	IN	IN5	100	vSDRDtrFvSh	Ananas comosus
1315P	D	S	IN	IN5	050	hSDQgaaft	HSRV
1401P	D	A	IN	IN6	100	PgNSQGqamVERanrl	RSV
1402P	D	B	IN	IN6	100	PYNSQGqaiVERtnrt	HERVK10
1403P	D	B	IN	IN6	100	LYNPQGqaiVDhthst	HML5
1404P	D	B	IN	IN6	100	PYNPrGqgiiEwahqt	HML6
1405P	D	B	IN	IN6	100	PYNPQGqaiVERahrt	IAP
1406P	D	B	IN	IN6	100	PYHPQGqaiVERthqn	MMTV
1407P	D	C	IN	IN6	100	AYqPQSsgKVERmnrt	HERV-E
1408P	D	C	IN	IN6	100	AYrPQSsgqVERmnrt	MLV

1409P	D D	IN	IN6	100		PYNPtSsgldERtngl	BLV
1410P	D L	IN	IN6	100		PYNPQSqgvVESmnke	HIV-1
1411P	D S	IN	IN6	100		PYHPQSgskVERknsd	HSRV
1412P	C G	IN	IN6	100		afHPQSdgqsERtiqtl	Ananas comosus
1501P	D C	IN	IN7	100		eprWkGPyiVLlttpt	BAEV
1601P	D A	MA	MA1	100		MeaVikVissacktycgkTS	RSV
1602P	D C	MA	MA1	075		MGQTvtTpls	MuLV, start of MA
1603P	D C	MA	MA1	100		MGNTprKten	RTVL-Ia, start of MA
1604P	D C	MA	MA1	100		MtlgLQThqpk	HERV-H, RTVLH2, 6 aa before pr
1605P	D E	MA	MA1	075		MGNssstp	WDSV
1700ANN	C ABCDELSGO	MA	MA1	025	0.87	GagStartNNData.txt	Gag start NN Changed from 0.90 050531
1801P	D A	MA	MA2	100		APPPPyvGsGlyP	RSV
1802P	D D	MA	MA2	100		iPPPYvepta	HTLV, end of MA
1803P	D S	MA	MA2	100		LRLTeGwWGqieRFqmV	HSRV
1900SPA	C ABCDESLGO	MA	MA3	100		K(1-5)K(10-25)!PPPY	Exp Wills acid motif
2001P	D A	NC	NC1	100		CytCgspGHyqaQC	RSV
2002P	D B	NC	NC1	100		CfnCgrmGHlkkDC	IAP
2003P	D B	NC	NC1	100		CfsCgktGHirkDC	MMTV
2004P	D B	NC	NC1	100		CfkCgkkGHfakNC	MPMV
2005P	D C	NC	NC1	100		CayrkeiGYwknkC	HERV-E.4-1
2006P	D C	NC	NC1	100		CayCkerGHwikDC	BAEV
2007P	D C	NC	NC1	100		CyqCgllGHfkkDC	ERV9
2008P	D C	NC	NC1	100		CayCkekGHwakDC	MLV
2009P	D C	NC	NC1	100		CtyCkqiGHwkkEC	S71
2010P	D D	NC	NC1	100		CyrClkeGHwarDC	BLV
2011P	D D	NC	NC1	100		CfrCgkaGHwsrDC	HTLV-I
2012P	D E	NC	NC1	100		CFFCKQpGHWKadCPNK	WDSV
2013P	D L	NC	NC1	100		CynCgkpGHlssQC	EIAV
2014P	D L	NC	NC1	100		CfnCgkeGHtarNC	HIV-1
2015P	D G	NC	NC1	200		CrvCgqeGHravRC	Osvaldo, D buzzatti
2016P	D G	NC	NC1	200		CyrCgepGHragAC	MarY1, T matsutake
2017P	D G	NC	NC1	200		CfpCgklVHaiaDC	peabody, P sativum
2101P	D G	NC	NC2	200		CwqCgriGVrtvAC	Osvaldo, D buzzatti
2102P	D G	NC	NC2	200		CfyCkkeGHrlnEC	TY3, S cerevisiae
2103P	D G	NC	NC2	200		CfnCgeeGHigsQC	peabody, P sativum
2104P	D G	NC	NC2	200		CheCqgyGHikaEC	Endovir1, A thaliana
2105P	D G	NC	NC2	200		CfrCnemGHiawNC	Cer1, C elegans
2106P	D A	NC	NC2	100		CqlcngmghnakQC	RSV
2107P	D B	NC	NC2	100		CyrCgkgyHrasEC	IAP
2108P	D B	NC	NC2	100		CprCkkgyHwksEC	MMTV
2109P	D B	NC	NC2	100		CprCkrgkHwanEC	MPMV
2110P	D D	NC	NC2	100		CpiCkdpsHwkrDC	BLV
2111P	D D	NC	NC2	100		CplCqdptHwkrDC	HTLV-I
2112P	D L	NC	NC2	100		CfkCkqpGHfskQC	EIAV
2113P	D L	NC	NC2	100		CwkCgkeGHqmkDC	HIV-1
2114P	D S	NC	NC2	100		rprtyQPQRYG	HHSRV
2115P	D G	NC	NC2	100		CyrCgepGHragAC	MarY1, T matsutake
2201N	C B	PBS	PBS	200		TGGcgcccaacgtggggc	tRNALys12-HSRV,MPMV,SRV1,VILV,
2202N	C BC	PBS	PBS	200		TGGcccatatggggattg	tRNAIle-RTVLI, RRHERV-I,HML5,ver020308
2203N	C BC	PBS	PBS	200		TGAtggcccgcatgagga	tRNAIle-HERV-I,HML5

2204N	C BL	PBS	PBS	200	TGGcgcccgaacagggac	tRNALys3-MMTV,HIV,EIAV,FIV,NMW
2205N	C C	PBS	PBS	200	TGGtgagccagccaggag	tRNAArg-ERV3, ver020308, Snakehead RV
2206N	C C	PBS	PBS	200	TGGttccctggccaggaa	tRNAGlu-HERVE
2207N	C C	PBS	PBS	200	TGGtgctgtgactcagat	tRNAHis-RTVLH, ERVfrd
2208N	C C	PBS	PBS	200	TGGtgtcagaagtgggat	tRNALeu-HERVL
2209N	C C	PBS	PBS	200	TGGtgccgaaacccggaa	tRNAPhe-HERV.F
2210N	C C	PBS	PBS	200	TGGtgctgagacccggga	tRNAPhe-IAPm
2211N	C G	PBS	PBS	200	TGGcgcagccggaaaact	tRNASer-ZAM,D melanogaster
2212N	C C	PBS	PBS	200	TGGtgtagtcgtcaggat	tRNASer-HERV.S
2213N	C C	PBS	PBS	200	TGGaggccccgctgggat	tRNAThr-HERV-T
2214N	C C	PBS	PBS	200	TGGcaaccacgaacggac	tRNATrp-HERV-W
2215N	C C	PBS	PBS	200	TTGgcgaccacgaaggga	tRNATyr
2216N	C C	PBS	PBS	200	TGGcgaccacgaaggga	tRNATyr, vs020818
2217N	C CD	PBS	PBS	200	TGGgggctcgtccgggat	tRNAPro-MLV,HTLV,HuersP
2218N	C C	PBS	PBS	200	TGGtgcattggccgggaa	tRNAPro-BaEV
2219N	C C	PBS	PBS	200	TGGgggcttgccctagat	tRNAPro-HERV-PT47D
2220N	C C	PBS	PBS	200	TGGgggaccacctgggat	tRNAThrPro-HERV.ADP
2221N	C O	PBS	PBS	200	TGGtatcagagcgggcactcta	tRNAxxx-SIRE1/Copia/TY1
2222N	C G	PBS	PBS	200	TGGcgccgttgccgggga	tRNAxxx-Cyclops2, P sativum
2223N	C G	PBS	PBS	200	TGGcgctagaaggagggg	tRNAAsn-Tat1-3, A Thaliana
2224N	C G	PBS	PBS	200	TGGtatcagaacaggtcg	tRNAMet-Peabody, P sativum
2225N	C G	PBS	PBS	200	TGGtatcagattttcagg	tRNAxxx-Cereba, H vulgare
2226N	C G	PBS	PBS	200	TGGtatcagagccccctt	tRNAMet-RIRE3, O sativa
2227N	C G	PBS	PBS	200	TGGtatcagagcaggcatc	tRNAMet-Endovir1, A thaliana
2301P	D C	Prot	PR1	100	QGCQGSGAPPEPRLTLVL	BAEV
2401P	D A	Prot	PR2	100	llDSGAdiTi	RSV, motif A
2402P	D B	Prot	PR2	100	llDTGAdkTc	MMTV, motif A
2403P	D B	Prot	PR2	100	liDTGAdvTi	MPMV, motif A
2404P	D C	Prot	PR2	100	TFLVDTGAQHSVLTKAN	BAEV
2405P	D C	Prot	PR2	100	lvDTGAqhSv	MuLV, motif A
2406P	D L	Prot	PR2	100	llDTGAddTv	HIV-1, motif A
2407P	D S	Prot	PR2	100	hwDSGAtiTc	HSRV, motif A
2501P	D A	Prot	PR3	075	ilGRdclq	RSV, motif B
2502P	D B	Prot	PR3	075	lwGRdimk	MMTV, motif B
2503P	D C	Prot	PR3	100	qilGRdvlsrlqasisi	BAEV
2504P	D C	Prot	PR3	075	llGRdllt	MuLV, motif B
2505P	D D	Prot	PR3	075	iiGRdalq	HTLV, motif B
2506P	D L	Prot	PR3	075	iiGRnllt	HIV-1, motif B
2507P	D S	Prot	PR3	075	vkGRkvea	HSRV, motif B
3200SPA	C C	RT	RT7	010	!DG(25-50)E(5-30)DS(35-70)N(3)D	RNaseH
2601P	D B	RT	RT1	100	wNSlVsvIqK	HML4
2602P	D B	RT	RT1	100	wNSPVFvIqK	HERVK10
2603P	D B	RT	RT1	100	wNSPVFvIKK	HML1
2604P	D B	RT	RT1	100	wNSPVFvIKK	HML6
2605P	D B	RT	RT1	100	wNtPVFvIKK	MMTV
2606P	C G	RT	RT1	100	yNSPtWVVdK	Gypsy
2607P	D B	RT	RT1	100	wNtPIFvIKK	MPMV
2608P	D C	RT	RT1	100	wNtPllpVKK	MLV
2609P	D D	RT	RT1	100	gNnPVFpVRK	BLV
2610P	D D	RT	RT1	075	gNnPVfpv	HTLV-I, motif ax

2611P	D L	RT	RT1	100	yNtPVFaIKK	HIV-1
2612P	D S	RT	RT1	100	mNtPVYpVpK	HSRV
2701P	D A	RT	RT2	100	iamDIsDcFFsiPL	LPDV
2702P	D B	RT	RT2	100	iiiDLkDcFFtipl	HERVK10
2703P	D B	RT	RT2	100	iviDLkDcFFtipl	HML1
2704P	D B	RT	RT2	100	vviDLkDcFFttpl	HML6
2705P	D B	RT	RT2	100	iiiDLqDcFFnikL	MMTV
2706P	D C	RT	RT2	100	svlHLkDaFFtiPL	HERV-H
2707P	D C	RT	RT2	100	svlDfkNFFciPL	ERVfrd
2708P	D C	RT	RT2	100	tviDLkvdFgmpPg	ERVftd
2709P	D C	RT	RT2	100	tvlDLkDaFFclrL	MLV
2710P	D D	RT	RT2	100	iclDLkDaFFqiPL	BLV
2711P	D D	RT	RT2	100	qtiDlrDaFF	HTLV-I,motif A
2712P	C G	RT	RT2	100	ttLDlKsgYHqiYL	Gypsy
2713P	D E	RT	RT2	100	tviDlsNaFFsvPi	WDSV
2714P	D L	RT	RT2	100	tvlDVgDaYFsvPL	HIV-1
2715P	D S	RT	RT2	100	ttlDlaNgFWahPI	HSRV
2716P	D S	RT	RT2	100	aaiDLaNgllpiPa	HERV-L
2801P	D A	RT	RT3	100	rFqWkVLPQGmtcSP	RSV
2802P	D B	RT	RT3	100	rFqWkVLPQGmlnSP	HERVK10
2803P	D B	RT	RT3	100	kYhWkVLPQGmlnSP	HML1
2804P	D B	RT	RT3	100	hyqWrVLPQGmlnSl	HML6
2805P	D B	RT	RT3	100	rFqWkVLPQGmknSP	MMTV
2806P	D C	RT	RT3	100	qltWtrLPQGfknSP	MLV
2807P	D C	RT	RT3	100	tismdsLaQGftdSP	ERVftd
2808P	D L	RT	RT3	100	rYqYnVLPQGwkgSP	HIV-1
2809P	D S	RT	RT3	100	qycWtrLPQGflnSP	HSRV
2810P	D S	RT	RT3	100	qciFTvLlQGyinSP	HERV-L
2811P	D G	RT	RT3	100	kyeFcrLPfGlrnas	Gypsy
2901P	D A	RT	RT4	100	lhYmDDLlla	RSV
2902P	D B	RT	RT4	100	ihYiDDILca	HERVK10
2903P	D B	RT	RT4	100	ihYmDDILca	HML1
2904P	D B	RT	RT4	100	ihYmDDILla	HML6
2905P	D B	RT	RT4	100	vhYmDDILla	MMTV
2906P	D C	RT	RT4	100	lqYvDDLlla	MLV
2907P	D D	RT	RT4	100	vsYmDDIlya	BLV
2908P	D L	RT	RT4	100	yqYmDDLyvg	HIV-1
2909P	D S	RT	RT4	100	qvYvDDIyls	HSRV
2910P	C G	RT	RT4	100	yvYvDDIyls	Gypsy
3001P	D A	RT	RT5	100	GlkineaKtQ	LPDV
3002P	D A	RT	RT5	100	GftispdKvQ	RSV
3003P	D B	RT	RT5	100	GlaiasdKiQ	HERVK10
3004P	D B	RT	RT5	100	GlviapdKiQ	HML1
3005P	D B	RT	RT5	100	GleiaseKvQ	IAPha
3006P	D B	RT	RT5	100	GlvvsteKiQ	MMTV
3007P	D B	RT	RT5	100	GlhiapeKvQ	MPMV
3008P	D C	RT	RT5	100	GyrasakKaQ	MLV
3009P	D D	RT	RT5	100	GfqvaseKtS	BLV
3010P	D L	RT	RT5	100	GlttpdkKhQ	HIV-1
3011P	D S	RT	RT5	100	GyvvslkKse	HSRV

```
       3012P    C G              RT     RT5  100       nmrvsqeKtrFFke                    Gypsy
       3101P    D B              RT     RT6  100       TLNDFQKLLgDInW                    MMTV
       3102P    D B              RT     RT6  100       TLwDvQKLvgslqW                    RSV
       3103P    C G              RT     RT6  100       DPeKVkAIQEYPeP                    Gypsy
       3200SPA  C C              RT     RT7  010       !DG(25-50)E(5-30)DS(35-70)N(3)D    RNaseH
       3301P    D B              SU     SU2  100       WmdnpteVYvndsvW                   HERVK10
       3302P    D B              SU     SU2  100       WsdalseIYhdqgaW                   HML-6.29
       3303P    D B              SU     SU2  100       WdreivpVYvndtsL                   JSRV
       3304P    D L              SU     SU2  025       YkGiflWR                          CAEV
       3401P    D B              SU     SU3  100       KRdfgItAAMIIAI                    IAPm
       3402P    D C              SU     SU3  100       RRalgmiifaiv                      HML6.17
       3403P    D C              SU     SU3  100       KRgiviGnWkdnew                    HERV-E
       3404P    C C              SU     SU3  100       KRepvsltlalllg                    MuLV
       3405P    C C              SU     SU3  100       Krviplitlmvglgl                   HERV-H RGH2 VIPLITLMVGLGL
       3406P    D D              SU     SU3  100       RRavpvAVwLvsAL                    HTLV-1
       3407P    D E              SU     SU3  100       KRdlGlHStLNSWWN                   WDSV
       3501P    D C              TM     TM1  050       SQMAWENKIAL                       HERV I
       3601P    D B              TM     TM2  100       rqtviwmgdrlmslehrfqlqC            HERV K10
       3602P    D B              TM     TM2  100       ydvvrvlgeqvqsinfrmkiqC            JSRV
       3603P    D B              TM     TM2  100       eevvlelgqdvanlktrmstrC            MMTV
       3701P    D A              TM     TM3  100       lqnraAIDFLllahghgC                ASLV
       3702P    D B              TM     TM3  100       LksmVlwlgEqV                      HML6.17
       3703P    D B              TM     TM3  100       LdlaeEqIGVLhqmaQLgC               IAPm
       3704P    D C              TM     TM3  100       yQNRLALDYLLA                      HERV E/ERV3
       3705P    D C              TM     TM3  100       lQNhRGLDlLTAekGGLCIfLE            HERV H/ERV9
       3706P    D C              TM     TM3  100       lQNrRGLDmLTAaqGGICLaLD            HERVES02A
       3707P    D E              TM     TM3  100       GCFiPkHpWsAG                      WDSV
       3708P    D L              TM     TM3  100       lQaRvLAVERYLKDQQL                 HIV-1 MN
       3709P    D S              TM     TM3  100       vnPLKNGSYLvlas                    HSRV
       3800SPA  C ABCDELSGO      TM     TM4  035       L(6)L(6)L(6)L(0-5)!C(4-15)C       Heptad repeat-CC
       3900HYF  C ABCDESLGO      TM     TM5  065       X                                 hydrophobic motif, changed from 60 050531
       4000LTR  D ABCDESLGO      5LTR   5LT  300       X                                 5'LTR
       4001LTR  D ABCDESLGO      3LTR   3LT  300       X                                 3'LTR
       4100PPT  C ABCDESLGO      PPT    PPT  050       X                                 PPT motif
::
LTR1Motifs::
       5001N    D B              ????   ???? 100       AATGGATTAAGGGCGGTGCAaGATGTGCTT    HERV-K, U3
       5002N    D B              ????   ???? 100       TTGGCAAAAGCCAAAGCCTAGGACAAATAC    JSRV, U3 (NF-1 like)
       5003N    D C              ????   ???? 100       CTTTGTTtcCTgcTTTCAAgGCcagACTTC    RTVL Ia, U3
       5004N    D C              ????   ???? 100       TAYACRtCCAGATGGCCagAAGTAACTGAA    RTVL-H consensus, U3
       5005N    D S              ????   ???? 100       CTTGATTGtATtGAAgGATGCAAAGcATTG    HERV-L, 5'and 3'versions of U3
       5006N    D S              ????   ???? 100       GCAGCCAGGCAGGCATAGGCTGAAGTAAAC    HERV-S U5
       5007N    D C              ????   ???? 100       TTCACCTTGATCAAAAACCACCAAAATCCA    HERV-F U3
       5008N    D C              ????   ???? 100       CCAgGAATgTCAGGTGACCATcAGgTGAT     HERV-FRD U3
       5009N    D B              ????   ???? 100       TTATAGAAAGAAGCAAACCTTCTTGGAATG    HML5 U3
       5010N    D C              ????   ???? 100       TAAGTACTGGCAGCCAGCCTGCGGATGTGA    HERV.ADP U3
::
LTR2Motifs::
       5101N    D B              ????   ???? 100       CCACCTtACGAGAAACACCCACAGGTGTGt    HERV-K, U5
       5102N    D B              ????   ???? 100       TTCTCTCCCTGTGCAGGTGCGACTCTTGTT    JSRV, U5
```

```
       5103N    D C              ????   ???? 100       yTTrTTGCTCACACAAAGCCTGTTTGGTCT     HERV-H consensus, U5
       5104N    D C              ????   ???? 100       TTCTGCAaAAGTAAATTTGCCTTGCTGAGA     RTVL Ia, U5
       5105N    D S              ????   ???? 100       TACTAGTTCTGTCCTTcTAGAGAACCTTGA     HERV-L, U5
       5106N    D S              ????   ???? 100       CATCTACCTGCTGTCTCTCTCAAGTGTTC      HERV-S U3
       5107N    D C              ????   ???? 100       TTTGACTGTTGAGCCATTTTTCATGTT        HERV-F U3
       5108N    D C              ????   ???? 100       ACTCTgCCTTATACAAgTaAGaTgAATTCT     HERV-FRD U5
       5109N    D B              ????   ???? 100       AACTCTTTACAGCACACTcTTTgGGTGT       HML5 U5
       5110N    D C              ????   ???? 100       GAGTCCATTCcTTGgGTTTGGttggGTGAA     HERV.ADP U5
::
KozakMotif::
       8000N    D ABCDESLGO      ????   ???? 100       gccaccATGg                         Kozak consensus
::
SpliceAcceptorMotif::
       8100SPL  C ABCDESLGO      ????   ???? 100       tttnctag                           Splice acceptor consensus
::
SpliceDonorMotif::
       8110SPD  C ABCDESLGO      ????   ???? 100       caggtaagt                          Splice donor consensus
::
SlipperyMotif::
       8200SLI  C ABCDESLGO      ????   ???? 100       xxxyyy                             Slippery sequence
::
ProteaseCleavageMotif::
       8300PCS  C ABCDESLGO      ????   ???? 100       wPf fPv yPi                        Protease cleavage site
::
PseudoKnotMotif::
       8400PKN  C ABCDESLGO      ????   ???? 100       X                                  Pseudoknot motif
::
ComponentMotifs::
       8900SLI  C ABCDESLGO      ????   ???? 100       xxxyyy                             Slippery sequence
       8901PKN  C ABCDESLGO      ????   ???? 100       X                                  Pseudoknot motif
::
FrameShifterMotif::
       8500COI  C ABCDESLGO      ????   ???? 100       8900 8901 12<16                    Coincidence of slippery and pseudoknot
::
RMotif::
       8600R    C ABCDESLGO      ????   ???? 100       X                                  Motif for R part of LTR
::
NotInUse::
       9100SSQ  C ABCDESLGO      SU     SU1  060       X                                  von Heijne weight matrix
::
```

Supplementary Information: RetroTector© Output

HERV-Fc1 (AL354685)

1. LTRID
2. RetroVID
3. ORFID
   - 3.1. ORFID-Gag
     - 3.1.1.Gag Putein
   - 3.2. ORFID-Pro
     - 3.2.1. Pro Putein
   - 3.3. ORFID-Pol
     - 3.3.1.Pol Putein
   - 3.4. ORFID-Env
     - 3.4.1. Env Putein
4. XonID
5. Chainview

1. LTRID

Executor: LTRID
DNAFile: HERV-Fc1-AL354685_1.txt
{ Created by SweepDNA with parameters
{ NewDNADirectory: NewDNA
{ ExecutorToUse: LTRID
{ ChunkOverlap: 15000
{ LINETolerance: 10
{ ALUTolerance: 10
{ ChunkSize: 115000
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

2. RetroVID

Executor: RetroVID
DNAFile: HERV-Fc1-AL354685_1.txt
Database: Ordinary
P1_LTRpair::
  PairFactor: 0.324
  RepeatLength: 5
  IntegrationSites: aaaat/tg<>ca/aaaat
::
P1_5LTR::
 ScoreFactor: 0.162
 VirusGenus: C
 First: 40963
 Last: 41338
 Hotspot: 1.0 1.0 41257 AATAAA
 U5NN: 0.4 2.0 41268
 GTModifier: 0.41 1.0
 U3NN: 0.9 1.0 41115
 TATAA: 1.0 1.0 41169
 MEME50: 0.55 1.0 41144
 Motifs1: 1.17 4.0 41122
 Motifs2: 1.43 4.0 41275
 Transsites: 0.2 1.0
 CpGModifier: 0.58 1.0
 Spl8Modifier: 0.21 1.0
 ShortDescription:
AATAAA(41257);U5NN:0.2(41268);GT:0.41;U3NN:0.9(41115);TATAA:1.0(41169);MEME50:0.55(41144);Mot1:0.29(41122);Mot2:0.36(41275);Trans:0.2;CpG:0.58;Spl8:0.21;
::
P1_3LTR::
 ScoreFactor: 0.162
 VirusGenus: CS
 First: 48536
 Last: 48905
 Hotspot: 1.0 1.0 48824 AATAAA
 U5NN: 0.44 2.0 48835
 GTModifier: 0.34 1.0
 U3NN: 0.88 1.0 48665
 TATAA: 1.0 1.0 48736
 MEME50: 0.58 1.0 48711
 Motifs1: 1.2 4.0 48642
 Motifs2: 1.16 4.0 48833
 Transsites: 0.3 1.0
 CpGModifier: 0.72 1.0
 Spl8Modifier: 0.24 1.0
 ShortDescription:
AATAAA(48824);U5NN:0.22(48835);GT:0.34;U3NN:0.88(48665);TATAA:1.0(48736);MEME50:0.58(48711);Mot1:0.3(48642);Mot2:0.29(48833);Trans:0.3;CpG:0.72;Spl8:0.24;
::
SingleLTR_S1::
 ScoreFactor: 0.181
 VirusGenus: CS
 First: 70136
 Last: 68338
 Hotspot: 1.0 1.0 68952 AGTAAA
 U5NN: 0.71 2.0 68758
 GTModifier: 0.95 1.0
 U3NN: 0.9 1.0 69820
 TATAA: 0.84 1.0 68997
 MEME50: 0.56 1.0 69895
 Motifs1: 1.43 4.0 69212
 Motifs2: 1.49 4.0 68966
 Transsites: 0.31 1.0
 CpGModifier: 0.18 1.0

Spl8Modifier: 0.5 1.0
ShortDescription:
AGTAAA(68952);U5NN:0.35(68758);GT:0.95;U3NN:0.9(69820);TATAA:0.84(68997);MEME50:0.56(69895);Mot1:0.36(69212);Mot2:0.37(68966);Trans:0.31;CpG:0.18;Spl8:0.5;_catgctg/tg<>ca/catgctg
::
SingleLTR_S2::
ScoreFactor: 0.185
VirusGenus: BC
First: 66266
Last: 64804
Hotspot: 1.0 1.0 65648 ATTAAA
U5NN: 0.63 2.0 65544
GTModifier: 0.67 1.0
U3NN: 0.9 1.0 66098
TATAA: 1.0 1.0 65847
MEME50: 0.56 1.0 64938
Motifs1: 1.8 4.0 66003
Motifs2: 1.56 4.0 65679
Transsites: 0.21 1.0
CpGModifier: 0.14 1.0
Spl8Modifier: 0.55 1.0
ShortDescription:
ATTAAA(65648);U5NN:0.31(65544);GT:0.67;U3NN:0.9(66098);TATAA:1.0(65847);MEME50:0.56(64938);Mot1:0.45(66003);Mot2:0.39(65679);Trans:0.21;CpG:0.14;Spl8:0.55;_caaataaa/tg<>ca/caaataaa
::
{ Created by LTRID with parameters
{ Motifs2Weight: 4.0
{ Motifs1Weight: 4.0
{ U5NetWeight: 2.0
{ SplitOctamerWeight: 1.0
{ LINELTRTolerance: 10.0
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ U3NetWeight: 1.0
{ DOSingleLTRs: Yes
{ InputFile:
{ Debugging: No
{ MEME50Weight: 1.0
{ CpGWeight: 1.0
{ TransSitesWeight: 1.0
{ MaxPairsOutput: 50
{ Database: Ordinary
{ TATAAWeight: 1.0
{ AATAAAWeight: 1.0
{ SingleLTRThreshold: 0.18
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\LTRID_001Script_.txt
{ ExponentStrength: 2.0
{ LTRepTolerance: 7.0
{ GTWeight: 1.0
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 50873 milliseconds
{ Latecoming LTRCandidates: 4
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3. ORFID

Executor: ORFID
DNAFile: HERV-Fc1-AL354685_1.txt
Strand: Primary
Database: Ordinary
VirusGenus: C
GeneScripts::
ORFID_P00001CGag_001Script.txt
ORFID_P00001CPro_001Script.txt
ORFID_P00001CPol_001Script.txt
ORFID_P00001CEnv_001Script.txt
::
{ Created by RetroVID with parameters
{ FrameFactor: 1.5
{ ImproveHitsMax: 5
{ MaxSubGeneSkip: 4
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ InputFile:
{ Debugging: No
{ FinalSelectionThreshold: 150.0
{ MakeChainsFiles: 0
{ BrokenPenalty: 0.9
{ Database: Ordinary
{ Strand: Both
{ SDFactor: 5.5
{ ORFIDMinScore: 200
{ KeepThreshold: 25.0
{ LengthBonus: 1.02
{ BrokenPasses: 1
{ ConservationFactor: 2
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\RetroVID_001Script_.txt
{ SubGeneHitsMax: 5
{ SelectionThreshold: 100.0
{ FitPuteins: Yes
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 49532 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3.1. ORFID-Gag

Executor: ORFID
DNAFile: HERV-Fc1-AL354685_1.txt
Strand: Primary
Database: Ordinary
VirusGenus: C
Gene: Gag
FirstDNAStart: 41122
LastDNAStart: 41892
ChainNumber: 00001
FirstDNAEnd: 43141
LastDNAEnd: 43510
HitInfo::
42029 14 CA0 ANN ABCDELSGO #CAStartNNData.txt_____ CA Start NN
42059 20 MA2 P D #iPPPYvepta HTLV, end of MA
42641 44 CA1 P C #teVVQGPGEXPGaFLECLqEAY S71
43025 54 NC1 P D #CyrClkeGHwarDC BLV
43094 42 NC2 P D #CplCqdptHwkrDC HTLV-I
::
{ Created by RetroVID with parameters
{ FrameFactor: 1.5
{ ImproveHitsMax: 5
{ MaxSubGeneSkip: 4
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ InputFile:
{ Debugging: No
{ FinalSelectionThreshold: 150.0
{ MakeChainsFiles: 0
{ BrokenPenalty: 0.9
{ Database: Ordinary
{ Strand: Both
{ SDFactor: 5.5
{ ORFIDMinScore: 200
{ KeepThreshold: 25.0
{ LengthBonus: 1.02
{ BrokenPasses: 1
{ ConservationFactor: 2
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\RetroVID_001Script_.txt
{ SubGeneHitsMax: 5
{ SelectionThreshold: 100.0
{ FitPuteins: Yes
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 49522 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3.1.1. Gag Putein

Executor: Puteinview
NumberOfHits: 4
LeadingInfo::

Leading path in Gag
Starts at=41648
Putein string of length=137 and total score=141.29758
MGGAQSKIDPKTPLGCLLAnFeALGLSMDLKRKRLIFFCLVAWPQYKLDNQSRWPPEGTFDFQILQDLDNLCRRQGKWSEVPYVQ
AFWDLRSRPDLCAKCSLGQVLLAKASPSNKEPDSSPLSEPPEALALPPLPAA
yielding average=1.0238955
Inside limits=1.0
For fit to alignment 0.7777778*3.0
Kozak score=0.625*0.5
Yielding path score = 3.6697292
::
TrailingInfo::

Trailing path in Gag
Ends at=43225
Putein string of length=67 and total score=49.243042
CFKCGNEGHWAZQCPYPKEPTRPcPNCHqMghwkseCPSvGASTVPLRCENSETTGGAFqLLSMDDD
yielding average=0.7241624
Inside limits=1.0
For fit to alignment 0.5*3.0
For stop codon 0.5*0.5
Yielding path score = 2.4741623
::
Gene: Gag
Genus: c
DNAFile: HERV-Fc1-AL354685_1.txt
EstimatedStartPosition: 41648
EstimatedLastPosition: 43225
LengthInside: 526
LengthTotal: 795
AlignedAcids: 495
AverageScoreInside: 0.8661258
AverageScoreTotal: 0.5730594
MostUsedRow: 2 (HERV-H_gag_rekonstr)
StopCodonsInside: 1(1, 11, 32)
StopCodonsTotal: 7
AmbiguousAcidsInside: 0
AmbiguousAcidsTotal: 0
ShiftsInside: 0
ShiftsTotal: 1
LongestRun: 526 at 41648
{ Starts at position 41122, ends at 43508
LongestORF: 41648
<<mggaqskidpktplgcllanfealglsmdlkrkrliffclvawpqykldnqsrwppegtfdfqilqdldnlcrrqgkwsevpyvqafwdlrsrpdlcakcslgqvllakaspsnkepdssplseppe
alalpplpaalpppypgssgptptapplpptppsspanppasalpppspvsahtrsktdllcllrevagaegvvrvhvpfsltdlskiekrpgsfsanptlyikqfrylcqaydltwrdlhiiltstlspeererv
qavarqhadqihltdpampvgtlavpaaepdwdyqagqtgrrrrdqmvqcllasmqaasnktvnfdklreiiqgsdenpavflnclteallqytrldptspagatvlathvisqsagdirkklkkveegp
qtpiqdlvkmafrvynsreetaeaqrqarlkqkvqfqtqalvaaprlagsgsqpkggsghrappgacfkcgneghwa
Putein::

1_____
3_____ 4_____ 5_____

typrkpkltkplpacsiktlycpsrarlpwpsllgpvnlarelhzzsrspltishlnsvrzfsldtclpkpdigaktreetplzprvgehllslparnqtrasafglcywvslpsvlzapgtsvcnrghsessphqf
pttgprtrrrpsssasairasrkvggcpsltlnrppqdMGgAQSKIdPKTPLgCLLAnFeALgLSmDLKrKRLIFFClVAWPQYKLDNqSrWPPEGTFDFqILQD
LDNLCRRQGKWSEVPYVQAFwDLRSrPdLCAKCSLGQVLLAKASPSNKEPDSSPLSEPPEALALPPLPAALPPPYPGSSGPTPTAPPLP
PTPPSSPAnPPAsAlPPPsPVsAhTrSktdLLcLlREvAGAeGVvRVHVPFSLtDLSkIeKrpGSFSaNPTlYIKqFRYLcQaYDLTWRDLHIILTSTL
sPEERERVQAvARQHADqiHLTDPAmPVGtlAVPaaEPDWDYqAgQTGRRRRDQMVQCLLAsmQAASnKtVNfDKLREIIQGSDENPAV
FLNCltEALiQYTrLDPTSPAGATVLAtHVISqSAgDIRKKLKKvEeGpQTPIQDLVKmAFRVYnSREETaEAQRQARLKQKvQfQTQALV
AAPRlAGSGSQPKGGsGhRApPGaCFKCGnEGHWAZqCPyPKEPTRPcPnCHqmghwksecPSvGaSTvPLrcenSetTGGafqllsmdddzrgpdsgtpl
tlaeprvtlqvagksisflvhmgatysvlpsfgvssfpspvtvvgidgtpsthrqtpppslscrlddtlishsfliipscpvllf

3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111

1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111
>! !
! ! ! !
! <
41122 41212 41302 41392 41482 41572 41663
41753 41843 41933 42023 42113 42203 42293
42383 42473 42563 42653 42743 42833 42923
43013 43103 43193 43283 43373 43463
::
Hits::
1 42029 14 CA0 ANN ABCDELSGO #CAStartNNData.txt_____________ CA Start NN
2 42059 20 MA2 P D #iPPPYvepta HTLV, end of MA
3 42641 44 CA1 P C #teVVQGPGEXPGaFLECLqEAY S71
4 43025 54 NC1 P D #CyrClkeGHwarDC BLV
5 43094 42 NC2 P D #CplCqdptHwkrDC HTLV-I
::
{ Created by ORFID with parameters
{ NonAlignedScore: 0.4
{ GlycosylationFactor: 0.2
{ StopCodonFactor: -0.4
{ LastDNAEnd: 43510
{ NonORFHexamerFactor: -0.1
{ MotifHitFactor: 0.2
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ VirusGenus: C
{ StopCodonValue: -15
{ MasterSkipPenalty: -0.2
{ InORFBonus: 0.1
{ OutputFile: ORFIDout_P00001CGag_001.txt
{ InputFile:
{ MinHitScore: 15
{ Debugging: No
{ Strand: Primary
{ Database: Ordinary
{ Gene: Gag
{ PuteinFile: Putein_P00001CGag_001.txt
{ ORFHexamerFactor: 0.2
{ ChainNumber: 00001
{ FirstDNAStart: 41122
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\ORFID_P00001CGag_001Script.txt
{ FirstDNAEnd: 43141
{ LastDNAStart: 41892
{ MinAverageScore: 20
{ FrameShiftFactor: -1.5
{ ScoreFactor: 0.1
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 8823 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3.2. ORFID-Pro

Executor: ORFID
DNAFile: HERV-Fc1-AL354685_1.txt
Strand: Primary
Database: Ordinary
VirusGenus: C
Gene: Pro
FirstDNAStart: 42545
LastDNAStart: 43283
ChainNumber: 00001
FirstDNAEnd: 43538
LastDNAEnd: 43817
HitInfo::
43310 22 PR2 P C #lvDTGAqhSv MuLV, motif A
43502 40 PR3 P C #llGRdllt MuLV, motif B
::
{ Created by RetroVID with parameters
{ FrameFactor: 1.5
{ ImproveHitsMax: 5
{ MaxSubGeneSkip: 4
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ InputFile:
{ Debugging: No
{ FinalSelectionThreshold: 150.0
{ MakeChainsFiles: 0
{ BrokenPenalty: 0.9
{ Database: Ordinary
{ Strand: Both
{ SDFactor: 5.5
{ ORFIDMinScore: 200
{ KeepThreshold: 25.0
{ LengthBonus: 1.02
{ BrokenPasses: 1
{ ConservationFactor: 2
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\RetroVID_001Script_.txt
{ SubGeneHitsMax: 5
{ SelectionThreshold: 100.0
{ FitPuteins: Yes
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 49522 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3.2.1. Pro Putein

Executor: Puteinview
NumberOfHits: 2
LeadingInfo::

Leading path in Pro
Starts at=43229
Putein string of length=27 and total score=21.414242
RGPDSGTPlTLAEPRVTLQVAGKSISF
yielding average=0.76479435
Inside limits=1.0
For fit to alignment 0.5*3.0
For cleavage site=0.17777778*0.5
Yielding path score = 2.3536832
::
TrailingInfo::

Trailing path in Pro
Ends at=43633
Putein string of length=44 and total score=35.494114
LfGRDLLSKLgASIrlhpSLpsSAiSlLpLLALSdDTpSPIpLL
yielding average=0.7887581
Inside limits=1.0
For fit to alignment 0.51111114*3.0
For cleavage site=0.5777778*0.5
Yielding path score = 2.6109805
::
Gene: Pro
Genus: c
DNAFile: HERV-Fc1-AL354685_1.txt
EstimatedStartPosition: 43229
EstimatedLastPosition: 43633
LengthInside: 135
LengthTotal: 423
AlignedAcids: 112
AverageScoreInside: 0.7049479
AverageScoreTotal: 0.22498336
MostUsedRow: 1 (HERV_E_4_1_Trolig_POL_M10976)
StopCodonsInside: 0(0, 4, 5)
StopCodonsTotal: 2
AmbiguousAcidsInside: 0
AmbiguousAcidsTotal: 0
ShiftsInside: 0
ShiftsTotal: 0
LongestRun: 135 at 43229
{ Starts at position 42545, ends at 43814
LongestORF: 43229
rgpdsgtpltlaeprvtlqvagksisflvhmgatysvlpsfgvssfpspvtvvgidgtpsthrqtppplscrlddtlishsfliipscpvllfgrdllsklgasirlhpsl
pssaisllpllalsddtpspipll>>
Putein::

1_____ 2_____

gqtgrrrrdqmvqcllasmqaasnktvnfdklreiiqgsdenpavflncltealiqytrldptspagatvlathvisqsagdirkklkkveegpqtpiqdlvkmaf
rvynsreetaeaqrqarlkqkvqfqtqalvaaprlagsgsqpkggsghrappgacfkcgneghwazqcpypkeptrpcpnchqmghwksecpsvgastv
plrcensettggafqllsmdddzrgpDSGtplTlaEPRVTLqVaGKsIsFLVhmGAtySVLpSfGVSsfpspVTVvGidgTpSthRQt
PppLSCrLddTlIsHSFLiiPsCPVlLfGRDLLSKLgASIrlhpSLpsSaiSlLpLLaLsddTpspIplLpvpvdpivwdistpsiarhha
pimiklkdptkfpsrpqfpisvehrqglkpiitrllqqh

1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111111111111111111111111111111111111
! ! !
! > <
42545 42635 42725 42815 42905
42995 43085 43175 43265 43355
43445 43535 43625 43715 43805
::
Hits::
1 43310 22 PR2 P C #lvDTGAqhSv MuLV, motif A
2 43502 40 PR3 P C #llGRdllt MuLV, motif B
::
{ Created by ORFID with parameters
{ NonAlignedScore: 0.4
{ GlycosylationFactor: 0.2
{ StopCodonFactor: -0.4
{ LastDNAEnd: 43817
{ NonORFHexamerFactor: -0.1
{ MotifHitFactor: 0.2
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ VirusGenus: C
{ StopCodonValue: -15
{ MasterSkipPenalty: -0.2
{ InORFBonus: 0.1
{ OutputFile: ORFIDout_P00001CPro_001.txt
{ InputFile:
{ MinHitScore: 15
{ Debugging: No
{ Strand: Primary
{ Database: Ordinary
{ Gene: Pro
{ PuteinFile: Putein_P00001CPro_001.txt
{ ORFHexamerFactor: 0.2
{ ChainNumber: 00001
{ FirstDNAStart: 42545
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\ORFID_P00001CPro_001Script.txt
{ FirstDNAEnd: 43538
{ LastDNAStart: 43283
{ MinAverageScore: 20
{ FrameShiftFactor: -1.5
{ ScoreFactor: 0.1
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 6710 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3.3. ORFID-Pol

Executor: ORFID
DNAFile: HERV-Fc1-AL354685_1.txt
Strand: Primary
Database: Ordinary
VirusGenus: C
Gene: Pol
FirstDNAStart: 43039
LastDNAStart: 43737
ChainNumber: 00001
FirstDNAEnd: 46645
LastDNAEnd: 46997
HitInfo::
43838 32 RT1 P C #wNtPllpVKK MLV
43994 67 RT2 P C #svlHLkDaFFtiPL HERV-H
44099 66 RT3 P C #qltWtrLPQGfknSP MLV
44207 60 RT4 P C #lqYvDDLlla MLV
44300 22 RT5 P C #GyrasakKaQ MLV
45831 18 IN3 P C #vtsaCkvCqqvnaga BAEV
45930 26 IN4 P C #hWeidfte MuLV, motif C
46110 63 IN5 P C #gSDNGPafvSQv BAEV
46194 60 IN6 P C #AYqPQSsgKVERmnrt HERV-E
46584 57 IN7 P C #eprWkGPyiVLlttpt BAEV
::
{ Created by RetroVID with parameters
{ FrameFactor: 1.5
{ ImproveHitsMax: 5
{ MaxSubGeneSkip: 4
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ InputFile:
{ Debugging: No
{ FinalSelectionThreshold: 150.0
{ MakeChainsFiles: 0
{ BrokenPenalty: 0.9
{ Database: Ordinary
{ Strand: Both
{ SDFactor: 5.5
{ ORFIDMinScore: 200
{ KeepThreshold: 25.0
{ LengthBonus: 1.02
{ BrokenPasses: 1
{ ConservationFactor: 2
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\RetroVID_001Script_.txt
{ SubGeneHitsMax: 5
{ SelectionThreshold: 100.0
{ FitPuteins: Yes
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 49532 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3.3.1. Pol Putein

Executor: Puteinview
NumberOfHits: 10
LeadingInfo::

Leading path in Pol
Starts at=43634
Putein string of length=68 and total score=51.805576
PVPVDPIVWDISTPSIARHHAPIMIKLKDPTkfPSRpQFPISVEHRQGLKPIITRLLQQHILIPVNSR
yielding average=0.75080544
Inside limits=1.0
For fit to alignment 0.44444445*3.0
For cleavage site=0.5777778*0.5
Yielding path score = 2.3730278
::
TrailingInfo::

Trailing path in Pol
Ends at=46751
Putein string of length=56 and total score=45.85149
tPRWSGPYTVILTTPRATKLlGLPSWYHlSQLKKAPTQhDWSSKLTPTRLRITHGQ
yielding average=0.8044121
Inside limits=1.0
For fit to alignment 0.6111111*3.0
Yielding path score = 2.6377454
::
Gene: Pol
Genus: c
DNAFile: HERV-Fc1-AL354685_1.txt
EstimatedStartPosition: 43634
EstimatedLastPosition: 46751
LengthInside: 1039
LengthTotal: 1318
AlignedAcids: 936
AverageScoreInside: 0.736851
AverageScoreTotal: 0.58087116
MostUsedRow: 1 (MuLV_pol_U92)
StopCodonsInside: 1(14, 23, 44)
StopCodonsTotal: 15
AmbiguousAcidsInside: 0
AmbiguousAcidsTotal: 0
ShiftsInside: 1
ShiftsTotal: 2
LongestRun: 568 at 43634
{ Starts at position 43039, ends at 46995
LongestORF: 43634
<<pvpvdpivwdistpsiarhhapimiklkdptkfpsrpqfpisvehrqglkpiitrllqqhilipvnsrcntpilpirkasgayrlvqdlriineavvpifpvvpn
pytllsripptthftvldlkddfftiplhpdcyflfaftwedpdthvssqfawtvlpqgfrdsphlfgqalakdlstctladstlllyvddlllcspslsvsqqdtatilnf
lgkqgyrvtphkvqlctptvtylgisltattkslttdrvslikdlqlpqdadkilsfvglvgffrhwipnfgvlakplyqaaketptsplsdpalvarhfhrlqqcllta
pvvslpnplrpfhlytdelqgvatgllgqpvgptyqvvaylsrqldpstrgwqpclralaaaaeltkealkltlshpltvysphrltdvlshkclahlapsriqlfhvl
fvenpditltaspplnpatllpieasepppvlshscpelltsnpnsrlglfdrplsnpdstlfvdgssvltpcgrrqaayavvthdktveaaalplgttsqkaellaltra
lllsqgqrvniytdskyayshcthafcslagarfpyyerdfnrqrasyp
Putein::

1_____ 2_____ 3_____
4_____ 5_____
6_____ 7_____ 8_____
9_____ 10_____

errplgltmpvpzgtdptmpzlppdgtlevzvpqrrsvhsastlzklrddrwrlpitqhgrrlkrprlgnpphscraqgnapgsrzvhifsrayggylfcfaflrrv
qfpvpghgsgdzrypfhpssdpppivlpagrhshlpflphypflsrpalrkglavzvrglhsvapqpplqcnlfassagtzrzhsfadpiaPVPvdPiVwDi
STPSIARHhAPimIkLKdpTkfPSRpQfPISvEhRQGlKPiItRLLqQhILiPvnSrcNTPiLPirKasGaYRlVQDLRiINea
VvpIfPvVPNPYtLLSriPPttThfTVLDLKdDFFtipLHPdcyfLFAFtWeDPdthvSsQfaWTvLPQGFrdSPhLFgqALak
DLstctLaDstLLlYVDDLLLcsPSLsvsQQdTAtiLnfLGKqGYRVtphKvQLCtptVTYLGiSLtattKsLtTdRvsliKdlql
PqdadkiLsFvGlvGFfRhWIPnFgvLAkPLYqaaKETPTsPlsDpaLvArhFhrlqQcLLTAPvvsLPnpLrPFhLYtDElQG
VAtGlLgQpvGPtyqvVAYLSrqLDPStrGWqPCLRAlAAaAeLTKeAlKLTLshPLTVysPHrLtdvLshKcLAHLAps
RiQlFhvlfvenPditlTASppLNPATLLPiEaSEppPVlsHsCPELLTsnpnsRlgLfDrPLsnpDsTlfvDGSSvLTpcGrRqA
AyAVVThdKTVeAAALPlGTtSQKAELlALTrALlLsqGqrvNIYTDSkYAysiAHtHsvlwQeRGfLTmkGtsIvNgpl
IhkpLnALqAPReVAIIHCksHQhsKDPVAqGNnLADsTAKslAlTSapaPapamfLsgSrTPaYsPqETfhlIsnLKGmT
DqdGziWVdnRialPEsQAqaiItdvHKTlLiGPKLLhQfLEPifLcPgLQSLihQVhqTCAvCstVNtqGglrRpGPhhQlRG
HqPGEdwqlDFThmPRhkhYrYLLtlVDTFtGWIEAFPTarETgevAVsVLLEhIiPRFGlPrsLqSDNGPAFVSKitQq
VsesLrVtWKLHipYRPQSSGkVERaNSllKEhLTKLTLETklSWVTLLPLALtRlRaaPrGPtGLsPFElLYGrPfLLP
gLpPTVSPppLAsyLpYLTllrDlLrKhAdAclPEpTPssPdaPvvlsPGDSVLlkelQsKtLtPRWsGPYTViLTTPrAtKllG
lpsWyHlSqLKkAPTQhDwSSKLTPtrLRIThGqtfptmpptppdppnphsaqzlppnztpvpmevlpazdldprqpalhchtgngglptsr
lsgpsnfzlhflzkcsaglvqsh

333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
333333333333333333111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111122222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
2222222222222222222222222222222222222222222222222222222222
!
>
! !
< ! !

43039	43129	43219	43309	43399
43489	43579	43670	43760	43850
43940	44030	44120	44210	44300
44390	44480	44570	44660	44750
44840	44930	45020	45110	45200
45290	45381	45471	45561	45651
45741	45831	45921	46011	46101
46191	46281	46371	46461	46551
46641	46731	46821	46911	

::
Hits::
1 43838 32 RT1 P C #wNtPllpVKK MLV
2 43994 67 RT2 P C #svlHLkDaFFtiPL HERV-H
3 44099 66 RT3 P C #qltWtrLPQGfknSP MLV
4 44207 60 RT4 P C #lqYvDDLlla MLV
5 44300 22 RT5 P C #GyrasakKaQ MLV
6 45831 18 IN3 P C #vtsaCkvCqqvnaga BAEV
7 45930 26 IN4 P C #hWeidfte MuLV, motif C
8 46110 63 IN5 P C #gSDNGPafvSQv BAEV
9 46194 60 IN6 P C #AYqPQSsgKVERmnrt HERV-E
10 46584 57 IN7 P C #eprWkGPyiVLlttpt BAEV
::
{ Created by ORFID with parameters

{ NonAlignedScore: 0.4
{ GlycosylationFactor: 0.2
{ StopCodonFactor: -0.4
{ LastDNAEnd: 46997
{ NonORFHexamerFactor: -0.1
{ MotifHitFactor: 0.2
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ VirusGenus: C
{ StopCodonValue: -15
{ MasterSkipPenalty: -0.2
{ InORFBonus: 0.1
{ OutputFile: ORFIDout_P00001CPol_001.txt
{ InputFile:
{ MinHitScore: 15
{ Debugging: No
{ Strand: Primary
{ Database: Ordinary
{ Gene: Pol
{ PuteinFile: Putein_P00001CPol_001.txt
{ ORFHexamerFactor: 0.2
{ ChainNumber: 00001
{ FirstDNAStart: 43039
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\ORFID_P00001CPol_001Script.txt
{ FirstDNAEnd: 46645
{ LastDNAStart: 43737
{ MinAverageScore: 20
{ FrameShiftFactor: -1.5
{ ScoreFactor: 0.1
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 34349 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3.4. ORFID-Env

Executor: ORFID
DNAFile: HERV-Fc1-AL354685_1.txt
Strand: Primary
Database: Ordinary
VirusGenus: C
Gene: Env
FirstDNAStart: 45471
LastDNAStart: 47271
ChainNumber: 00001
FirstDNAEnd: 48379
LastDNAEnd: 49153
HitInfo::
48088 62 TM3 P C #lQNhRGLDlLTAekGGLCIfLE HERV H/ERV9
48316 13 TM5 HYF ABCDELSGO #X_____ hydrophobic motif
::
{ Created by RetroVID with parameters
{ FrameFactor: 1.5
{ ImproveHitsMax: 5
{ MaxSubGeneSkip: 4
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ InputFile:
{ Debugging: No
{ FinalSelectionThreshold: 150.0
{ MakeChainsFiles: 0
{ BrokenPenalty: 0.9
{ Database: Ordinary
{ Strand: Both
{ SDFactor: 5.5
{ ORFIDMinScore: 200
{ KeepThreshold: 25.0
{ LengthBonus: 1.02
{ BrokenPasses: 1
{ ConservationFactor: 2
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\RetroVID_001Script_.txt
{ SubGeneHitsMax: 5
{ SelectionThreshold: 100.0
{ FitPuteins: Yes
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 49532 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

3.4.1. Env Putein

Executor: Puteinview
NumberOfHits: 2
LeadingInfo::

Leading path in Env
Starts at=46764
Putein string of length=441 and total score=433.02655
MLLLLLTLLTPIVPSNSLLTEPPFRWRfYLHETWTQGNRLSTVtLATVDCqPHGCQAQvTfNfTSfKSVLRg
wSnPTICFvYDqtHSnCRDYwvDTNGGCPYAYCRMHVTqLhtaKKLQhTYrltsDGRtTYFLTIPDPwdSrwVSG
VTGRlYRWpTdSYPVGKLrIFLTyIRVipQVLSNlKDQAdNIKHqEEVINTLVqSHPKADMVTYDDKAEAGP
FSWITLVRHGARLVNMAGLVNLSHCFLCtALSqPPlVAVPLPQAFNTSGNHTAHPSgVfSeQvPLFRdPLQP
qFPFcyTTpNSSwCNQTYSGSLSnLsAPAgGyFWCNfTLTkhlnIssNNTlSrNLClpISLVPRLTLYSEaELSSLVN
PPMRQKRAVFPPLVIGVSLTSSLVASGLGTGaIVHFlSSsQDLSIkLQmAIeASaeSLaSLQRQITSVAkvA
yielding average=0.97969806
Inside limits=1.0
For fit to alignment 0.6111111*3.0
Kozak score=0.71875*0.5
For SpliceAcceptorMotif at 46752 0.5714286*0.7
von Heijne score with LATVDCqPHGCQAQv =0.25985792*0.5
Majored:-1.0
Yielding path score = 3.7023354
::
TrailingInfo::

Trailing path in Env
Ends at=48393
Putein string of length=102 and total score=104.23819
MQNRRALDLLTADKGGTCmFLGEECCYYINeSGLVeTSlLTLDKIRDgLhRpSStPNYGGgWWQSPLTTWI
IPFISPILiICLLLLIAPCVLKFIKNRISEV
yielding average=1.0120213
Inside limits=1.0
For fit to alignment 0.6888889*3.0
Yielding path score = 3.0786877
::
Gene: Env
Genus: c
DNAFile: HERV-Fc1-AL354685_1.txt
EstimatedStartPosition: 46764
EstimatedLastPosition: 48393
LengthInside: 543
LengthTotal: 1226
AlignedAcids: 453
AverageScoreInside: 0.9125116
AverageScoreTotal: 0.4041548
MostUsedRow: 1 (GaLV_env)
StopCodonsInside: 0(27, 29, 0)
StopCodonsTotal: 10
AmbiguousAcidsInside: 0
AmbiguousAcidsTotal: 0
ShiftsInside: 1
ShiftsTotal: 1
LongestRun: 541 at 46768
{ Starts at position 45471, ends at 49150
LongestORF: 46768
<<llllltlltpivpsnslleppfrwrfylhetwtqgnrlstvtlatvdcqphgcqaqvtfnftsfksvlrgwsnpticfvydqthsncrdywvdtnggcpyaycr
mhvtqlhtakklqhtyrltsdgrttyfltipdpwdsrwvsgvtgrlyrwptdsypvgklrifltyirvipqvlsnlkdqadnikhqeevintlvqshpkadmvty
ddkaeagpfswitlvrhgarlvnmaglvnlshcflctalsqpplvavplpqafntsgnhtahpsgvfseqvplfrdplqpqfpfcyttpnsswcnqtysgslsnl
sapaggyfwcnftltkhlnissnntlsrnlclpislvprltlyseaelsslvnppmrqkravfpplvigvsltsslvasglgtgaivhfisssqdlsiklqmaieasaes

laslqrqitsvakvamqnrraldlltadkggtcmflgeeccyyinesglvetslltldkirdglhrpsstpnygggwwqspltttwiipfispiliicllllliapcvlkfi
knrisev>>
Putein::

1_____________________ 2___________

hckshqhskdpvaqgnnladstakslaltsapapapamflsgsrtpayspqetfhlisnlkgmtdqdgziwvdnrialpesqaqaiitdvhktlligpkllhqfle
piflcpglqslihqvhqtcavcstvntqgglrrpgphhqlrghqpgedwqldfthmprhkhyrylltlvdtftgwieafptaretgevavsvllehiiprfglprsl
qsdngpafvskitqqvseslrvtwklhipyrpqssgkveransllkehltkltletklswvtllplaltrlraaprgptglspfellygrpfllpglpptvspppplasylp
yltllrdllrkhadaclpeptpsspdapvvlspgdsvllkelqsktltprwsgpytvilttpratkllglpswyhlsqlkkaptqhdwsskltptrlrithgqtfptML
LLLLTLLTPiVPSNSLLTePPFrWRfYLHeTWTQGnRLSTVtLATVDCqPHGCQAQvTfNfTSfKSVLRgwSnPT
ICFvYDqtHSnCRDYwvDTNGGCPYAYCRMHVTqLhtaKKLQhTYrltsDGRtTYfLTIPDPwdSrwVSGVTGRl
YRWpTdSyPVGKLrIfLTyIrVipQVLSNlKDQAdNIKHqEEVINTLVqSHPKADmVTYDDKAEAGPfswITLV
RhGARLVNmAGLVNLShCFLCtALSqPPlVAVPLPQAfNTSGNhTAHPSgVfSeQvPLFRdPLQPqFPFcyTTpN
SSwCNQTYSGSLSnLsAPAgGyFWCNfTLTkhlnIssNNTlSrNLClpISLVPrLTLYSEaELSSLvnPPmRQKRAVF
pPLVIGvsLTSSLVASGLGTGaIVhFISSsQDLSIkLQmAIeASaeSLaSLQRQiTSvAkvAmQNRRALDLLTAdK
GGTCmFLGEECCYYINeSGLVeTSlLTLDKIRDgLhRpSStPNYGGgWWqSPLTTWIiPFiSPILiICLLLLIaPCv
LKFIKNRISEVsrvtvnqmllhpysrlptsedhyddaltqqeaarzlrrpfflqyevgmlgrspkmavppgpkmaartpssppppplgvshqifppdg
hfpmtaaprsrnlsqktetyzapphtlyknpllpqsgatslallvrtseprpraplikqvasdhzlpkfcaalligylssqagqkzikqnlfsnglvakatlvpetqn
aplydyvqvhtsvepqfthlzygikapilqgczedsirpsaskwlaqcsayzzvqzt

2222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
2222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
2222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
2222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
2222222222222222222222222222222222222222222222222222222222222222222222222222222233333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333333
33333333333333333333333333333333333333333333333333333333
! !
>!
! !
! ! < ! !

45471	45561	45651	45741	45831
45921	46011	46101	46191	46281
46371	46461	46551	46641	46731
46822	46912	47002	47092	47182
47272	47362	47452	47542	47632
47722	47812	47902	47992	48082
48172	48262	48352	48442	48532
48622	48712	48802	48892	48982
49072				

::
Hits::
1 48088 62 TM3 P C #lQNhRGLDlLTAekGGLCIfLE HERV H/ERV9
2 48316 13 TM5 HYF ABCDELSGO #X___________ hydrophobic motif
::
{ Created by ORFID with parameters
{ NonAlignedScore: 0.4
{ GlycosylationFactor: 0.2
{ StopCodonFactor: -0.4
{ LastDNAEnd: 49153
{ NonORFHexamerFactor: -0.1
{ MotifHitFactor: 0.2

{ DNAFile: HERV-Fc1-AL354685_1.txt
{ VirusGenus: C
{ StopCodonValue: -15
{ MasterSkipPenalty: -0.2
{ InORFBonus: 0.1
{ OutputFile: ORFIDout_P00001CEnv_001.txt
{ InputFile:
{ MinHitScore: 15
{ Debugging: No
{ Strand: Primary
{ Database: Ordinary
{ Gene: Env
{ PuteinFile: Putein_P00001CEnv_001.txt
{ ORFHexamerFactor: 0.2
{ ChainNumber: 00001
{ FirstDNAStart: 45471
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\ORFID_P00001CEnv_001Script.txt
{ FirstDNAEnd: 48379
{ LastDNAStart: 47271
{ MinAverageScore: 20
{ FrameShiftFactor: -1.5
{ ScoreFactor: 0.1
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Execution time was 32827 milliseconds
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

4. XonID

Executor: XonID
DNAFile: HERV-Fc1-AL354685_1.txt
Strand: Primary
Database: Ordinary
ChainNumber: 1
ChainStart: 41098
ChainEnd: 48819
{ Created by RetroVID with parameters
{ FrameFactor: 1.5
{ ImproveHitsMax: 5
{ MaxSubGeneSkip: 4
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ InputFile:
{ Debugging: No
{ FinalSelectionThreshold: 150.0
{ MakeChainsFiles: 0
{ BrokenPenalty: 0.9
{ Database: Ordinary
{ Strand: Both
{ SDFactor: 5.5
{ ORFIDMinScore: 200
{ KeepThreshold: 25.0
{ LengthBonus: 1.02
{ BrokenPasses: 1
{ ConservationFactor: 2
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\RetroVID_001Script_.txt
{ SubGeneHitsMax: 5
{ SelectionThreshold: 100.0
{ FitPuteins: Yes
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

5. Chainview

Executor: Chainview
DNAFile: HERV-Fc1-AL354685_1.txt
Selected: Yes
Database: Ordinary
PSingleLTRs::
::
SSingleLTRs::
68952 (70136-68338)
AGTAAA(68952);U5NN:0.35(68758);GT:0.95;U3NN:0.9(69820);TATAA:0.84(68997);MEME50:0.56(69895)
;Mot1:0.36(69212);Mot2:0.37(68966);Trans:0.31;CpG:0.18;Spl8:0.5;_catgctg/tg<>ca/catgctg
65648 (66266-64804)
ATTAAA(65648);U5NN:0.31(65544);GT:0.67;U3NN:0.9(66098);TATAA:1.0(65847);MEME50:0.56(64938);
Mot1:0.45(66003);Mot2:0.39(65679);Trans:0.21;CpG:0.14;Spl8:0.55;_caaataaa/tg<>ca/caaataaa
::
ChainP1::
A:0.35
B:0.35
C:0.99
D:0.69
E:0.35
L:0.1
S:0.4
G:0.06
O:0.06

41098
48819
Type C Score= 1568
SubGene 5LTR, type CS, score=91 , hotspot 41171
5LT:ABCDELSGO (5'LTR): Score=91 at 41171 frame 1 [41098-41252]
X scored against

tgactgcagccccgagaagtcgaaacctatcccagaaaaccgaaacttactaagcccctccccgcgtgctctataaaaaccctctactgccccagtcgggcgcgac
ttccctggccctccttgttaggaccagtgaacctcgcccgagagctcca (155 bases)
SubGene PBS, type C, score=107 , hotspot 41340
PBS:C (tRNAPhe-HERV.F): Score=107 at 41340 frame 2 [41340-41357]
TGGtgccgaaacccggaa scored against
tggtgccaaaacccggga (18 bases)
SubGene MA, type CD, score=165 , hotspot 41582
MA2:D (HTLV, end of MA): Score=20 at 42059 frame 1 [42059-42088]
iPPPYvepta scored against
lpppypgssg (30 bases)
MA3:ABCDELSGO (Exp Wills acid motif): Score=100 at 42062 frame 1 [41972-42073]
K(1-5)K(10-25)!PPPY scored against
kaspsnkepdssplseppealalpplpaalpppy (102 bases)
SubGene CA, type C, score=78 , hotspot 42641
CA0:ABCDELSGO (CA Start NN): Score=14 at 42029 frame 1 [42029-42118]
CAStartNNData.txt scored against
alalpplpaalpppypgssgptptapplpp (90 bases)
CA1:C (S71): Score=44 at 42641 frame 1 [42641-42706]
teVVQGPGEXPGaFLECLqEAY scored against
reiiqgsdenpavflnclteal (66 bases)
SubGene NC, type D, score=144 , hotspot 43025
NC1:D (BLV): Score=54 at 43025 frame 1 [43025-43066]
CyrClkeGHwarDC scored against
cfkcgneghwazqc (42 bases)
NC2:D (HTLV-I): Score=42 at 43094 frame 1 [43094-43135]

CplCqdptHwkrDC scored against
cpnchqmghwksec (42 bases)
SubGene Prot, type C, score=94 , hotspot 43310
PR2:C (MuLV, motif A): Score=22 at 43310 frame 1 [43310-43339]
lvDTGAqhSv scored against
lvhmgatysv (30 bases)
PR3:C (MuLV, motif B): Score=40 at 43502 frame 1 [43502-43525]
llGRdllt scored against
lfgrdlls (24 bases)
SubGene RT, type C, score=371 , hotspot 44099
RT1:C (MLV): Score=32 at 43838 frame 1 [43838-43867]
wNtPllpVKK scored against
cntpilpirk (30 bases)
RT2:C (HERV-H): Score=67 at 43994 frame 1 [43994-44035]
svlHLkDaFFtiPL scored against
tvldlkddfftipl (42 bases)
RT3:C (MLV): Score=66 at 44099 frame 1 [44099-44143]
qltWtrLPQGfknSP scored against
qfawtvlpqgfrdsp (45 bases)
RT4:C (MLV): Score=60 at 44207 frame 1 [44207-44236]
lqYvDDLlla scored against
llyvddlllc (30 bases)
RT5:C (MLV): Score=22 at 44300 frame 1 [44300-44329]
GyrasakKaQ scored against
gyrvtphkvq (30 bases)
SubGene IN, type C, score=335 , hotspot 45930
IN3:C (BAEV): Score=18 at 45831 frame 2 [45831-45875]
vtsaCkvCqqvnaga scored against
vhqtcavcstvntqg (45 bases)
IN4:C (MuLV, motif C): Score=26 at 45930 frame 2 [45930-45953]
hWeidfte scored against
dwqldfth (24 bases)
IN5:C (BAEV): Score=63 at 46110 frame 2 [46110-46145]
gSDNGPafvSQv scored against
qsdngpafvski (36 bases)
IN6:C (HERV-E): Score=60 at 46194 frame 2 [46194-46241]
AYqPQSsgKVERmnrt scored against
pyrpqssgkveransl (48 bases)
IN7:C (BAEV): Score=57 at 46584 frame 2 [46584-46631]
eprWkGPyiVLlttpt scored against
tprwsgpytvilttpr (48 bases)
SubGene TM, type C, score=103 , hotspot 48082
TM3:C (HERV H/ERV9): Score=62 at 48088 frame 3 [48088-48153]
lQNhRGLDlLTAekGGLCIfLE scored against
mqnrraldlltadkggtcmflg (66 bases)
TM5:ABCDELSGO (hydrophobic motif): Score=13 at 48316 frame 3 [48316-48351]
X scored against
piliiclllllia (36 bases)
SubGene 3LTR, type CS, score=91 , hotspot 48738
3LT:ABCDELSGO (3'LTR): Score=91 at 48738 frame 2 [48665-48819]
X scored against

tgacagcagccccgagaagtcgaaacctatcccagaaaaccgaaacttactaagcccctccccacacgctctataaaaaccctctactgccccagtcgggtgcga
cttccctggccctccttgttaggaccagtgaacctcgcccgagagctcca (155 bases)

Integration sites ttcaga/tg<>ca/ttaata

::
P1_5LTR::
ScoreFactor: 0.148

VirusGenus: CS
First: 41098
Last: 41252
Hotspot: 0.5 1.0 41171 tataaa
U5NN: 0.44 2.0 41177
GTModifier: 0.24 1.0
U3NN: 0.9 1.0 41115
TATAA: 1.0 1.0 41169
MEME50: 0.55 1.0 41144
Motifs1: 1.17 4.0 41122
Motifs2: 1.05 4.0 41204
Transsites: 0.24 1.0
CpGModifier: 0.74 1.0
Spl8Modifier: 0.25 1.0
ShortDescription:
tataaa(41171);U5NN:0.22(41177);GT:0.24;U3NN:0.9(41115);TATAA:1.0(41169);MEME50:0.55(41144);Mot1:0.29(41122);Mot2:0.26(41204);Trans:0.24;CpG:0.74;Spl8:0.25;
::
P1_3LTR::
ScoreFactor: 0.156
VirusGenus: CS
First: 48665
Last: 48819
Hotspot: 0.5 1.0 48738 tataaa
U5NN: 0.43 2.0 48812
GTModifier: 0.33 1.0
U3NN: 0.88 1.0 48665
TATAA: 1.0 1.0 48736
MEME50: 0.58 1.0 48711
Motifs1: 1.17 4.0 48689
Motifs2: 1.05 4.0 48771
Transsites: 0.47 1.0
CpGModifier: 0.83 1.0
Spl8Modifier: 0.29 1.0
ShortDescription:
tataaa(48738);U5NN:0.21(48812);GT:0.33;U3NN:0.88(48665);TATAA:1.0(48736);MEME50:0.58(48711);Mot1:0.29(48689);Mot2:0.26(48771);Trans:0.47;CpG:0.83;Spl8:0.29;
::
P1SlipperyMotifHits::
41835 41841 100
42797 42803 100
43727 43733 100
46960 46966 100
::
P1PseudoKnotMotifHits::
::
P1SpliceAcceptorMotifHits::
41175 41193 18
41207 41225 39
41361 41379 18
41394 41412 59
41458 41476 80
41512 41530 18
42036 42054 18
42060 42078 80
42402 42420 39
42726 42744 39
43195 43213 18
43345 43363 18
43403 43421 39

43552 43570 59
43585 43603 18
43743 43761 18
43923 43941 80
44141 44159 39
44221 44239 59
44243 44261 59
44274 44292 18
44417 44435 18
44445 44463 39
44556 44574 18
44582 44600 18
44649 44667 59
44732 44750 18
44824 44842 18
44838 44856 18
44862 44880 18
44905 44923 18
44997 45015 59
45106 45124 18
45151 45169 39
45256 45274 39
45275 45293 59
45281 45299 18
45351 45369 39
45429 45447 18
45556 45574 59
45606 45624 59
45699 45717 18
45769 45787 18
45798 45816 18
45813 45831 18
45979 45997 59
46054 46072 18
46525 46543 39
46784 46802 59
47159 47177 18
47231 47249 18
47499 47517 39
47533 47551 18
47617 47635 18
47623 47641 18
47708 47726 18
47826 47844 18
47998 48016 18
48040 48058 59
48104 48122 39
48332 48350 59
48505 48523 80
48742 48760 18
48774 48792 39
::
P1SpliceDonorMotifHits::
41449 41457 33
41595 41603 33
41871 41879 33
42905 42913 33
43290 43298 100
43872 43880 33
44010 44018 33

44439 44447 33
45142 45150 33
45535 45543 33
45922 45930 33
46033 46041 33
46540 46548 33
46936 46944 33
47199 47207 33
47723 47731 33
48384 48392 33
48492 48500 33
::
{ Execution time was 49491 milliseconds
{ Created by RetroVID with parameters
{ FrameFactor: 1.5
{ ImproveHitsMax: 5
{ MaxSubGeneSkip: 4
{ DNAFile: HERV-Fc1-AL354685_1.txt
{ InputFile:
{ Debugging: No
{ FinalSelectionThreshold: 150.0
{ MakeChainsFiles: 0
{ BrokenPenalty: 0.9
{ Database: Ordinary
{ Strand: Both
{ SDFactor: 5.5
{ ORFIDMinScore: 200
{ KeepThreshold: 25.0
{ LengthBonus: 1.02
{ BrokenPasses: 1
{ ConservationFactor: 2
{ ScriptPath: C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1\RetroVID_001Script_.txt
{ SubGeneHitsMax: 5
{ SelectionThreshold: 100.0
{ FitPuteins: Yes
{ and plugins
{ LTRAlign
{ LTRhistogram
{ ComplementDNA
{ LTRImprover
{ Belongs in C:\Retrotector\RetroTector10\Workplace\HERV-Fc1-AL354685\HERV-Fc1-AL354685_1
{ Created “DATE/TIME” under RetroTector version 1.0 “DATE”
{ using Database:Ordinary last modified “DATE/TIME”

Delineation of the four major retroviral genes in reference genomes.

Genus	Virus	RT_gag	Annot_gag	RT_pro	Annot_pro	RT_pol	Annot_pol	RT_env	Annot_env
Lenti	HIV1HXB2 NC_001802_RNA	336 - 1835	336 - 1838	1799 - 2092	?	2096 - 4639	2096 - 4639	5641 - 8338	5771 - 8341
Delta	HTLV1 NC_001436_RNA	780 - 1736	450 - 1739	1739 - 2344	1718 - 2245	2245 - 4833	2245 - 4836	4829 - 6292	4829 - 6295
	HTLV2 NC_001488	807 - 2105	807 - 2108	2108 - 2611	?	2593 - 5184	2239 - 5187	5180 - 6640	5180 - 6640
Alpha	ALV NC_001408 RNA	372 - 2093	372 - 2102	2094 - 2474	2103 - 2474	2477 - 5197	2495 - 5182	No env!	5069 - 6872
	RSV NC_001407 RNA	380 - 2101	380 - 2485?	2103 - 2482	?	2485 - 7187	2482 - 5190	No env!	5078 - 9392
Gamma	MoMLV NC_001501 RNA	357 - 1970	357 - 1973	1974 - 2291	1959 (1974) - 2336	2292 - 5570	2337? - 5573	5513 - 7507	5513 - 7510
	FeLV NC_001940	907 - 2424	907 - 2412	2428 - 2745	2413 - 2787	2746 - 6042	2788 - 6042	5988 - 7853	5988 - 7916
	Wooly monkey sarcoma virus (w/o sis oncogene) NC_1514 RNA	1614 - 2864	645 - 2867	2690 - 3637	3368 - 3640	3636 - 6281	3888 - 6284	6366 - 8345	6314 - 8041
Beta	JSRV NC_001494 RNA	263 - 2098	263 - 2101	1984 - 2859	1993 - 2862	2856 - 5438	3108 - 5441	5350 - 7194	5350 - 7197
	MPMV NC_001550 RNA	269 - 2239	269 - 2242	2116 - 3006	1844 - 2865	3003 - 5579	2092 - 5579	5671 - 7748	5621 - 7378
Epsilon	SnakeheadRV NC_001724_RNA	640 - 2140	337 - 2193	2149 - 2748	2194 - ?	2749 - 6387	2194 - 6390	6131 - 9532	6184 - 9535
	WDSV NC_001867 RNA	800 - 2545	800 - 2548	2549 - 2923	2545- ?	2924 - 6055	? - 6058	6115 - 8454 (EnvTrace)	5856 - 12708
	Xen1 AJ506107	2088 - 3842	2088 - 3845	3846 - 4346	3846 - ?	4347 - 7498	? - 7051	8103 - 9610	7498 - 9494

Comment to the table:

RT predicts the four major structural protein genes with precision in the simple retroviruses. The complex viruses (lenti, delta and epsilon retroviruses), which have one ore several additional regulatory protein genes, can cause problems. The genes most often affected are the *gag* and *env*. Additional reading frames are occasionally present before *gag* and around *env*. Similar problems are caused by the sarcoma viruses, where oncogenes disrupt the retroviral structure.

allrepsfam	erv3like	erv9like	herv19like	herv48like	hervadplik	hervelike	hervfblike	hervfclike	hervfrdlik	hervhlike	hervilike	hervl66lik	hervllike	hervrblike	hervslike	hervtlike	hervwlike	hml1	hml10	hml2	hml3	hml4	hml5	hml6	hml7	hml8	hml9	huersp3lik	mer41like	mer66like	rhervilike	allrepstot
BAEV_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	6
ERVL_A1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ERVL_A2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ERVL_B1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL_B2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL_B3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL_B4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL_C1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL_D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL_E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL_F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HERV15	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	9	18
HERV16	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	4
HERV16_B1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HERV17	1	3	0	0	0	0	0	2	0	0	0	0	0	1	0	0	75	0	0	0	3	0	0	0	0	0	0	6	0	0	0	91
HERV3	24	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	2	0	0	2	32
HERV30	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	6
HERV70	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	4
HERV9	0	176	0	1	0	0	0	1	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	10	1	0	1	198
HERVE	24	0	0	0	0	34	0	4	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	2	0	0	0	67
HERVFH19	0	0	5	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
HERVFH21	0	0	1	2	0	0	12	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
HERVH	1	2	1	16	0	6	5	2	0	763	0	0	1	0	0	1	1	1	0	2	0	1	0	1	0	0	0	0	1	1	0	806
HERVH48	0	0	1	22	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	28
HERVI	0	0	0	0	0	0	0	0	0	1	5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
HERVIP10F	0	0	1	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	1	2	0	1	0	0	0	0	0	2	1	0	12
HERVIP10FH	0	38	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	41
HERVK	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	15	0	22	0	0	0	2	0	0	2	0	0	0	0	42
HERVK11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	4	0	0	0	0	0	6	0	12	0	0	0	0	0	23
HERVK11D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	10	0	0	0	0	0	0	11
HERVK13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	6
HERVK14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	1	0	2	0	3	0	0	0	0	0	0	0	23
HERVK14C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	6
HERVK22	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	21	3	0	0	0	1	0	0	0	28
HERVK3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	22	0	0	1	0	0	0	0	24
HERVK9	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	43	0	0	67	43	0	8	0	0	1	1	0	0	0	166
HERVKC4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	0	0	0	0	0	0	0	0	0	0	0	0	6
HERVL	0	0	0	0	0	0	0	0	0	2	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
HERVL_A1	0	0	0	0	0	0	0	0	0	1	0	0	113	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	117
HERVL_A1_A	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
HERVL_A2	0	0	0	0	0	0	0	0	0	0	0	0	45	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	47
HERVL_A2_A	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
HERVL_A2_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HERVL_B1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
HERVL_B3	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
HERVL_B3_B	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
HERVL_B4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HERVL_B5_A	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
HERVL18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
HERVL18_A1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HERVL32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
HERVL40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HERVL66	0	0	0	0	0	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
HERVL74	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
HERVP71A	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
HERVS71	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	16
HUERS_P1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
HUERS_P2	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	5	0	0	0	8
HUERS_P3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	30	1	0	0	32
HUERS_P3B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
HARLEQUIN	2	3	0	0	0	10	0	1	0	3	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	21
LOR1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LOR1A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LOR1B	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LOR1B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	2
LTR10A	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR10A_INT	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR10B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR10B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR10B1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR10B1_IN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR10C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR10D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR10E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR10F	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
LTR10F_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR10G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR10G_INT	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR12	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	21
LTR12B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR12C	0	10	0	0	0	1	0	0	0	2	0	0	0	1	0	1	0	0	0	1	1	0	0	1	0	0	0	1	0	0	0	19
LTR12C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR12D	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
LTR13A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR14A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	3
LTR14B	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2
LTR14C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
LTR16A	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3
LTR16A1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	3
LTR16B	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	4
LTR16C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR16D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR17	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	6
LTR18A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR18B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
LTR19A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR19A_INT	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR19B	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR19B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR19C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2
LTR19C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR1B	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
LTR1B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2
LTR1C	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR1D	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	1	0	0	0	0	0	1	0	0	0	3	0	0	0	9
LTR1D_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	4
LTR2	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	5
LTR21A	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR21B_INT	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
LTR22A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR22B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
LTR23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR23_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
LTR24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR24_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR24B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR24C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR25_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR26B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR26E_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
LTR27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR27_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR27B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR27B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	4	0	0	0	6
LTR28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR28_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	3
LTR29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR29_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR2B	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR2C	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
LTR3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	4
LTR30	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR33	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	3	0	0	0	0	0	0	1	0	0	0	6
LTR33A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR37A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2
LTR37B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
LTR38_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR38B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR38C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR39	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR39_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR3A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
LTR3B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR3B1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR3B_	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR40A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
LTR40B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR40C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR42	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR43	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
LTR43_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR45_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR45B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR45C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR46	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR46_INT	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

LTR47A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR47B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR48_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR48B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR48B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR49_INT	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
LTR52_INT	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR54	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR54B	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
LTR56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR56_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR57_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR58_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR5B	1	0	0	1	0	3	0	0	0	0	0	0	0	0	0	0	0	7	0	1	0	0	0	0	0	0	0	0	1	0	0	14
LTR5_HS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR62	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR6A	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5
LTR6B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR7	0	0	0	0	0	0	0	1	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	51
LTR7_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR71A	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	4
LTR71B	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR73	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR75_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR76	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR76_INT	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
LTR77_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR7B	0	0	0	1	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
LTR8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	3
LTR8_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR8A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR8A_INT	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
LTR9	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	5
LTR9_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2
LTR9B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR9B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER101B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER110_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER110A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER110A_IN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER11A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
MER11B	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	3
MER11C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	3	0	0	0	0	0	6
MER11D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	4	0	0	0	0	0	0	5
MER21A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER21B	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
MER21B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER21C	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
MER21C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER31_INT	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
MER31A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER31B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
MER34	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER34_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER34B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER34B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
MER34D	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER34D_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER39	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
MER39B	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER4_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER41_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
MER41A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER41A_INT	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER41B	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER41B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER41C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	5
MER41C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	5	0	0	6
MER41D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
MER41D_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER41E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
MER41E_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER48	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
MER49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER4A	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3
MER4A_INT	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2
MER4A1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3
MER4A1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER4B	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	6
MER4B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER4C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER4C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER4D	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4
MER4D_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER4D1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER4E	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER4E_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER4E1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER4E1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER50	0	0	1	2	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
MER50_INT	0	0	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5
MER50B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER50B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER51_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER51A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	3
MER51A_INT	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	4
MER51B	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER51B_INT	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER51C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER51D_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER51E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER51E_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER52_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	6
MER52A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
MER52A_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	6
MER52C	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	3
MER52C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER52D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER54A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
MER54A_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER54B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER57_INT	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER57A	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER57A_INT	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
MER57B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER57B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER61A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER61A_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER61B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER61B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER61C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER61D_INT	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER61E	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	3
MER61E_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER61F	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
MER61F_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
MER65_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER65A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER65A_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER65B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER65C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER65C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER65D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER66_INT	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	4
MER66A	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER66A_INT	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2
MER66B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER66B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	2
MER66C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER66C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
MER67B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER67C	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
MER67D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2
MER68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER70A	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER70B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER70B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER72B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER73	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER74A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER74B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER74C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1

MER76_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER77	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
MER83	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER83_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER83A_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER83B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER83C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER84_INT	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
MER87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER89	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER89_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
MER9	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	1	0	0	0	0	0	0	6
MER90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER90_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
MER92B	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MER93B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER93B__2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1A	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2
MLT1A_INT	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT1A0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	2
MLT1A0_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1A1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
MLT1A1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT1B	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3
MLT1B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
MLT1C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT1D	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	6
MLT1D_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT1E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1E1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	5
MLT1E1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1E1A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
MLT1E2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	2
MLT1E2_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1E3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT1F	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
MLT1F_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
MLT1F1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	3
MLT1F2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
MLT1G	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	4
MLT1G1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2
MLT1G1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1H	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2
MLT1H1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
MLT1H1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
MLT1H2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	2
MLT1H2_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1I	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	3
MLT1J	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3
MLT1J_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT1J1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
MLT1J2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	3
MLT1K	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3
MLT1L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	3
MLT2A1	0	0	0	0	0	0	0	0	0	1	0	0	0	46	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	48
MLT2A2	0	0	0	0	0	0	0	0	0	1	0	0	0	27	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	30
MLT2B1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3
MLT2B2	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3
MLT2B3	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
MLT2B4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT2B5	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT2C1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT2C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MLT2D	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT2E	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MLT2F	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MST_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
MSTA	0	0	0	0	0	0	0	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	1	4	0	0	0	0	0	0	0	0	9
MSTA_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3
MSTB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	3
MSTB_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MSTB1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MSTB1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MSTB2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
MSTB2_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MSTC	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
MSTC_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MSTD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MSTD_INT	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
PABL_A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PABL_A_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PABL_B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PABL_B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRIMA4_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRIMA41_IN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRIMA4_LTR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRIMAX_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
THE1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
THE1A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2
THE1A_INT	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	3
THE1B	0	1	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	2	0	13
THE1B_INT	0	2	0	0	0	0	0	0	0	7	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	13
THE1C	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
THE1C_INT	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
THE1D	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	4
THE1D_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	3	0	0	0	0	1	0	0	0	6
RTTOTAL	75	277	17	51	9	65	24	28	4	911	11	4	322	6	7	39	96	119	5	36	180	54	27	59	14	16	4	114	29	7	15	2625

hervdfamil	erv3like	erv9like	herv19like	herv48like	hervadplik	hervelike	hervfblike	hervfclike	hervfrdlik	hervhlike	hervilike	hervl66lik	hervllike	hervrblike	hervslike	hervtlike	hervwlike	hml1	hml10	hml2	hml3	hml4	hml5	hml6	hml7	hml8	hml9	huersp3lik	mer41like	mer66like	rhervilike	hervdtotal
RTTOTAL	74	273	15	49	9	66	23	29	4	885	8	2	307	6	7	38	93	117	4	40	175	49	28	59	14	15	4	112	28	7	15	0
BAEV_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ERVL	0	1	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	8	0	0	1	0	0	0	0	0	0	0	5256
HERV15	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	11	50
HERV16	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2049
HERV17	1	3	0	2	0	0	0	3	0	1	0	0	0	1	0	0	76	0	0	0	2	0	0	0	0	0	0	3	0	0	0	315
HERV3	27	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	3	91
HERV30	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	30
HERV70	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	885
HERV9	0	202	0	1	0	0	1	1	0	0	0	0	0	0	0	1	11	0	0	0	0	0	0	0	0	0	0	11	1	0	0	418
HERVE	28	0	0	0	0	52	0	6	0	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	181
HERVFH19	0	0	6	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	55
HERVFH21	0	0	1	2	0	0	12	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	48
HERVH	0	0	1	17	0	0	7	4	0	857	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	1	0	1	0	1306
HERVH48	0	0	1	24	0	0	1	2	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	57
HERVI	0	0	0	0	0	0	1	1	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	111
HERVI_LIKE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HERVK	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	21	0	33	0	0	0	3	0	0	2	0	0	0	0	121
HERVK11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	5	0	0	1	0	0	8	14	14	0	0	0	0	0	145
HERVK13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	1	0	0	24
HERVK14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	0	1	0	2	0	5	0	0	0	0	0	0	0	96
HERVK14C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	28
HERVK22	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	26	3	0	0	0	0	0	0	0	163
HERVK3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	29	0	0	1	0	0	0	0	109
HERVK9	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	45	0	0	68	42	0	7	0	0	1	1	0	0	0	230
HERVKC4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	3	0	0	0	0	0	0	0	0	0	0	0	0	24
HERVL	0	0	0	0	0	0	0	0	0	1	0	0	288	0	0	1	0	1	0	0	5	0	0	0	0	0	0	1	0	0	0	2462
HERVL18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	165
HERVL32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45
HERVL40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	452
HERVL47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	222
HERVL66	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45
HERVL74	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	259
HERVS71	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	60
HUERS_P1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HUERS_P2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0
HUERS_P3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	29	1	0	0	0
HUERS_P3B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
HARLEQUIN	2	2	0	0	0	11	0	1	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	0	0	0	0	2	0	0	0	630
LOR1_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LOR1A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1021
LOR1B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1443
LTR1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1396
LTR10A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	232
LTR10B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	186
LTR10C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	426
LTR10D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	158
LTR10E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	206
LTR12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	4172
LTR13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	426
LTR14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	62
LTR14B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	233
LTR14C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	91
LTR15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	227
LTR16A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	4548
LTR16A1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	2738
LTR16B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	1	0	0	0	2316
LTR16C	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	6724
LTR16D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1128
LTR17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	441
LTR18B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	255
LTR19_INT	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0
LTR19A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	274
LTR19B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	197
LTR19C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	157
LTR1B	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2362
LTR2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	717
LTR21A	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	62
LTR22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	321
LTR22B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	165
LTR23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	818
LTR23_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	305
LTR24_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR24B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	469
LTR25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	271
LTR25_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	458
LTR26B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	304
LTR26E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	234
LTR27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	943
LTR28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	517
LTR29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	676
LTR2B	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	269
LTR2C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	143
LTR30	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	199
LTR32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	782
LTR33	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	6369
LTR33A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2618
LTR34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	353
LTR36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	473
LTR37A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1961
LTR37B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1982
LTR38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	308
LTR38B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	222
LTR38C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	132
LTR39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1182
LTR3B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	55
LTR4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	64
LTR40A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1508
LTR40B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	900
LTR40C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	446
LTR41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1315
LTR43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	655
LTR43_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	178
LTR45B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	425
LTR45C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	390
LTR46_INT	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR47A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	695
LTR47B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	318
LTR48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1226
LTR48B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1063
LTR49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1401
LTR49_INT	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
LTR50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1859
LTR51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	402
LTR52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1075
LTR52_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LTR53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	569
LTR54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1102
LTR54B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1069
LTR55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	404
LTR56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	953
LTR57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	169
LTR58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	81
LTR5B	1	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	504
LTR5_HS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	494
LTR60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	248
LTR61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	127
LTR62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	215
LTR64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	240
LTR65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	456
LTR66	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	278
LTR67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	4875
LTR68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	341
LTR6A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	205
LTR6B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95
LTR7	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1220
LTR71B	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	214
LTR72	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	253
LTR75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	416
LTR8	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3537
LTR8A	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	2682
LTR9	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1931
LTR9B	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	743
MER101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	377
MER101B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	244
MER110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	732
MER110_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER110A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	648
MER110A_IN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER11A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	806
MER11B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	482
MER11C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	750
MER11D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	185
MER21_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER21A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1842
MER21A_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER21B	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	3612
MER21B_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER21C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	4850
MER21C_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER31_INT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MER31A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1321
MER31B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1463
MER34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2384
MER34B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1304
MER39	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0	3141
MER39B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1153

Detailed comparison of ReTe with RepeatMasker (hg15)

This is given in the contingency table, which lists hg15 fidings (S7). To enable a qualitative comparison with the other approaches to retroviral sequence detection, ReTe chains were categorised into 31 groups (erv3like, erv9like, herv19like, herv48like, hervadplike, hervelike, hervfblike, hervfclike, hervfrdlike, hervhlike, hervilike, hervl66like, hervllike, hervrblike, hervslike, hervtlike, hervwlike, hml1-10, huersp3like, mer41like, mer66like and rhervilike). They are based on around 80% similarity of their Pol putein to the same number of reference HERV Pol proteins. HERV classification is the subject of a forthcoming paper (Blomberg et al, in preparation), in which 62% of ReTe chains in hg15 were classified in this manner. Repeatmasker entries are often fragments of longer transposon segments. Recognition of a sequence as "retroviral" is often difficult. Many ReTe chains are classified into similar groups as Repeatmasker entries, among them 176 erv9like/HERV9, 776 hervhlike/HERVH and 113 hervllike/HERVLA1 ReTe/Repeatmasker coincidences. The very old and mutated, remotely retroviruslike, MalR elements, 4% of the human genome [IHGSC, 2001], are entirely missed by ReTe. Neither are single LTRs (approximated by RepeatMasker to 0.7 % of the human genome) efficiently detected by ReTe.

In a more recent comparison of ReTe and RM (table S10), a small number of RM missed elements were identified in galGal3, canFam2 and hg18. galGal3 had the highest proportion of RM misses (84). Notable among them were 7 betaretroviruslike and one gammaretroviruslike chains with a ReTe score>400. In canFam2, 24 ReTe chains were missed by RM. They contained 3 betaretroviruslike and one errantiviruslike chain with a ReTe score above 400. The few RM-missed elements in hg18 (28) were mainly weakly ReTe-scoring with a limited similarity to *Errantivirus*. Repeatmasker output for the three genomes has become more complete from early versions (hg15, gagGal1 and canFam1)(our observations). The reasons for the few remaining RM-missed elements, and their properties, should be further investigated.

Detailed Comparison of ReTe with HERVd (hg15)

This is given in the contingency table (S8), also based on hg15. The nomenclature is highly concordant for some ReTe groups versus a HERVd family: erv9like/HERV9, herv19like/HERVFH1, hervhlike/HERVH, hml2/HERVK , hml5/HERVK22, hml6/HERVK3, hervwlike/HERV17, hervtlike/HERVS71, huersp3like /HUERSP3, hervllike/HERVL were unambiguous ReTe/HERVd concidences. In other cases assignments crosses between ReTe groups and HERVd families, like for hervelike and erv3like ReTe groups, where variously HERV3 or HERVE HERVd families were involved. Some HERVd families have no or just a few ReTe chain counterparts: 15 ervllike are recorded by ReTe vs 5256 in HERVd (in the following written as 15/5256), herv16like 4/ 2049, HERVL 297/2462, LOR1a 0/1024 and LOR1B 1/1443. The "Other" group (29 125 elements or element pieces) in HERVd contains 19 mer41like, 24 huersp3like and 50 erv9like ReTe chains.

ROC curves and sensitivity plots (hg18, canFam2 and galGal3)

To provide an updated comparison, the hg18 (March 2006) assembly was used. HERVd does not cover this version. Instead, we constructed likely elements from RepeatMasker output for hg18 (downloaded in April, 2007). As described below, overlapping elements, and elements with a proximity of <500 bp to elements of the same group and polarity, were connected to

form “RM chains”. These are the basis for the plots for the human, dog and chicken genomes, in supplementary material S11-S16.

There are a number of difficulties with the comparison of ReTe and RepeatMasker (RM):

1. It must be stressed that the RepeatMasker output is difficult to organise into proviruses. The “id” number is given, on unclear grounds, for a collection of like and probably connected repeats. However, this property only covers a minority of likely connected repeat fragments. Instead, we (JB) had to write a program which provisionally joins RM repeats into proviral chains.

2. A minority of allelic chromosome names are erratic, like “qbl.hap1”. It is easy to miss some overlaps due to possible minor chromosome name differences in RM and ReTe. However, this error was compensated for, if detected. It can only account for a few missed ReTe-RM chain overlaps.

3. A basic problem is secondary integration of proviruses into each other. This may create two incomplete halves and one complete provirus. RM lists all repeats, but may fail to knit them together. ReTe may or may not knit the outer halves together due to its broken chain function, possibly discarding the inner element. Old elements often have multiple secondary integrations. They may distort the proviral structure beyond ReTe recognition. It is a likely cause of ReTe misses in many of the old elements (MalR, ERVL, many of the MER elements), despite the ability of ReTe to detect highly mutated proviruses (this paper). In addition, the MalR and ERVL elements may have structural deviations from the current ReTe structural model. Coming versions of ReTe will aim at alleviating this restriction.

4. ReTe rarely detects chains less than 1000 bp long. Usually, they are more than 2000 bp long, whereas RM can detect much shorter repeats, and single LTRs.

Thus, the comparison can only be approximate.

In the data behind table S10, and figures S11-S16, ReTe sensitivity was calculated versus a more or less reduced set of RM hits.

5. For calculation of specificity, the abstract RM negative/ReTe negative value must be estimated. If it is a large number, its exact size will not affect calculations much. It was arbitrarily set to genome size / 10 000.

Enumeration of ReTe chains with a score of >=300, and approximated RM connected elements in three genomes.

Genome	ReTe+RM+	ReTe+RM-	ReTe-RM+	Sens%	Spec%
galGal3 large set	417	82	306	57.7	99.97
galGal3 small set	404	82	15	96.4	99.97
canFam2 large set	377	33	3200	10.5	99.97
canFam2 small set	376	23	35	91.5	99.99
hg18 large set	3511	33	11453	23.5	99.99
hg18 small set	3389	28	2130	61.4	99.98

ReTe chains were tested for coincidence with RM elements connected via “id” assignment, overlap, or proximity less than 500 bp. The RM elements considered were of LTR type. Each genome was studied in two ways:

The “large set” disregarded MalR and ERVL elements, single LTRs insofar as they could be identified programmatically, and elements shorter than 1000 bp.

The “small set” disregarded a larger number of elements. MalR and ERVL elements, single LTRs, MER and some other elements (see below), and elements shorter than 2000 bp were excluded. Many of the MER elements are common to the human and dog genomes, some also occur in the chicken genome. MER elements were encountered in some previous analyses by us and others [Jern, 2005; Jern et al., 2005; Jern et al., 2005; Jurka et al., 2005; Oja et al., 2005]. Most of them are evolutionarily old elements with multiple secondary integrations which disrupt proviral structure. As seen from the table, the removal of these elements resulted in significantly higher sensitivity values for ReTe. Despite drastic removal of LTR elements in the small set, very few ReTe+ elements were removed.

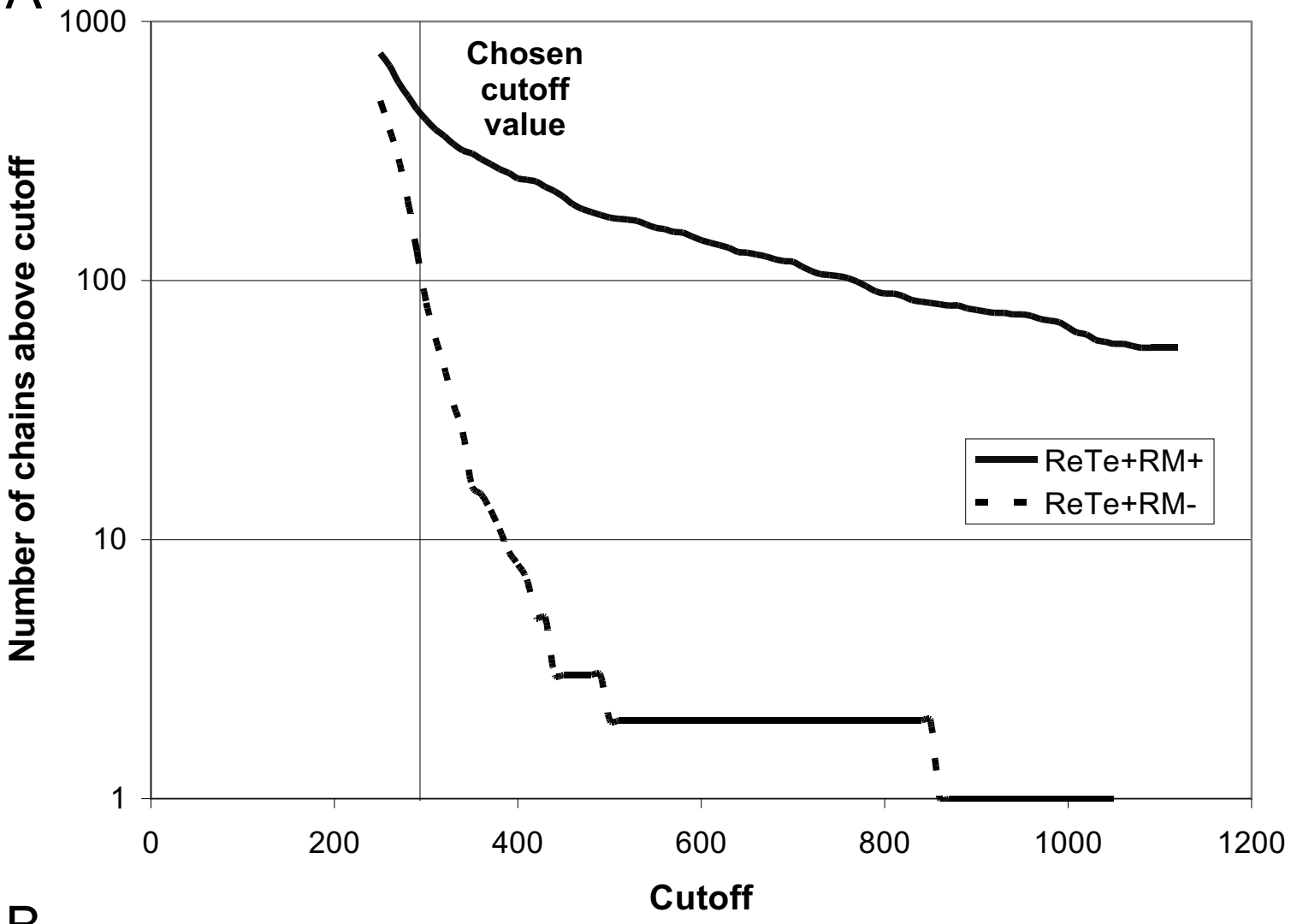
The following elements were disregarded in the “small set”:

LOR1; PABL; MER41, 66, 84, 51, 83, 50, 4, 65, 61, 57, 92, 52, 90, 34, 31, 67, 110, 89, 21, 39, 49, 101, 70, 72, 87, 90, 92, 93, 95; CARLTR1, 2, 4, 5, 6, 7, 8; CARERV4, 2; GGLTR1, LTR34, 75, 48, 28, 23, 24, 8, 49, 27, 37, 48, 54, 12, 26, 2, 35, 38, 44, 45, 61, 64, 6, 77, 7; HUERS-P2; HUERS-P3"

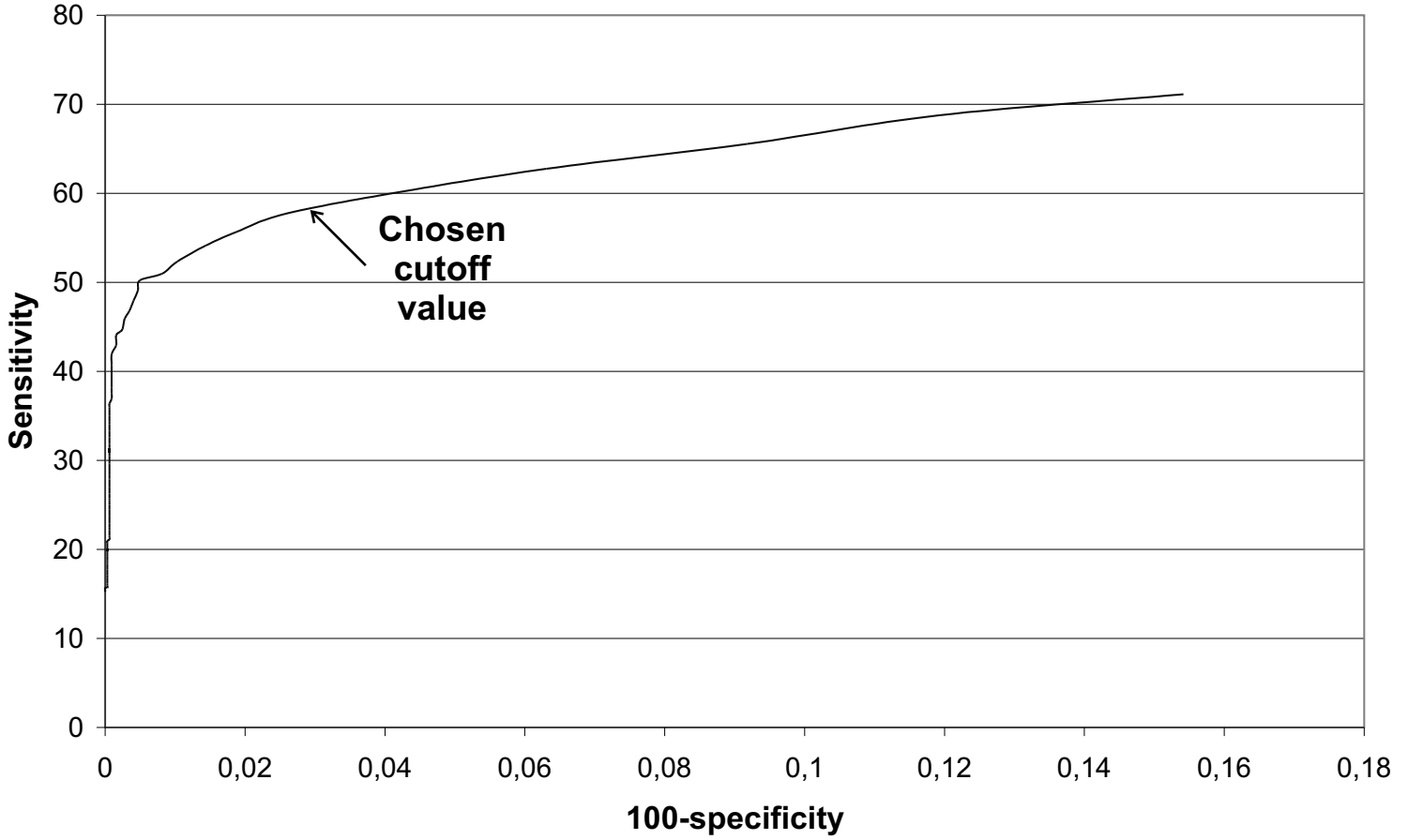
IHGSC. 2001 Feb 15. Initial sequencing and analysis of the human genome. Nature:860-921.
Jern P. 2005. Genomic Variation and Evolution of HERV-H and other Endogenous Retroviruses (ERVs). Uppsala: Thesis, Acta Universitatis Upsaliensis. 77 p.
Jern P, Sperber GO, Ahlsen G, Blomberg J. 2005. Sequence variability, gene structure, and expression of full-length human endogenous retrovirus h. J Virol 79(10):6325-6337.
Jern P, Sperber GO, Blomberg J. 2005. Use of endogenous retroviral sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy. Retrovirology 2:50.
Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. 2005. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110(1-4):462-467.
Oja M, Sperber GO, Blomberg J, Kaski S. 2005. Self-organizing map-based discovery and visualization of human endogenous retroviral sequence groups. Int J Neural Syst 15(3):163-179.

RM vs ReTe for galGal3

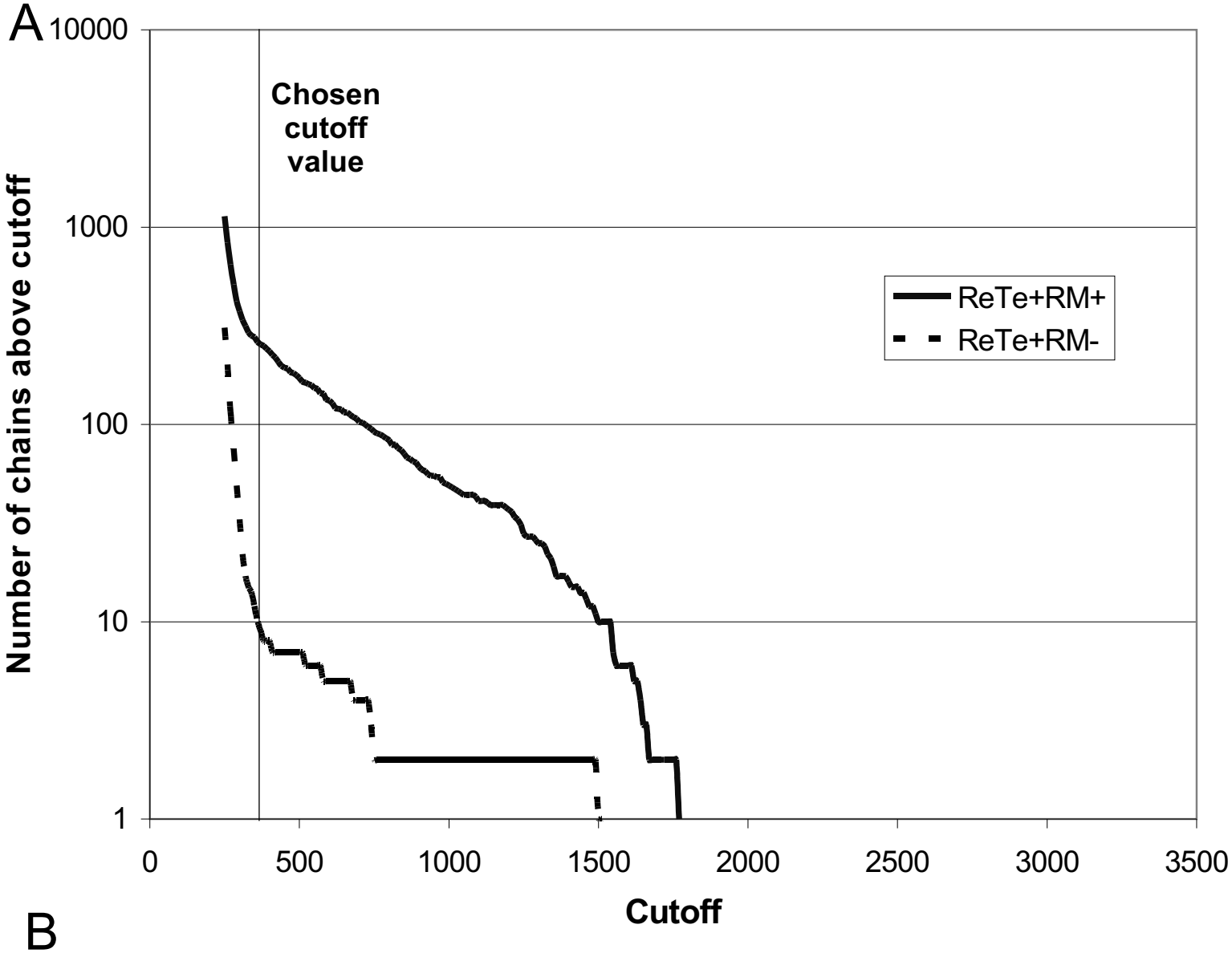
A

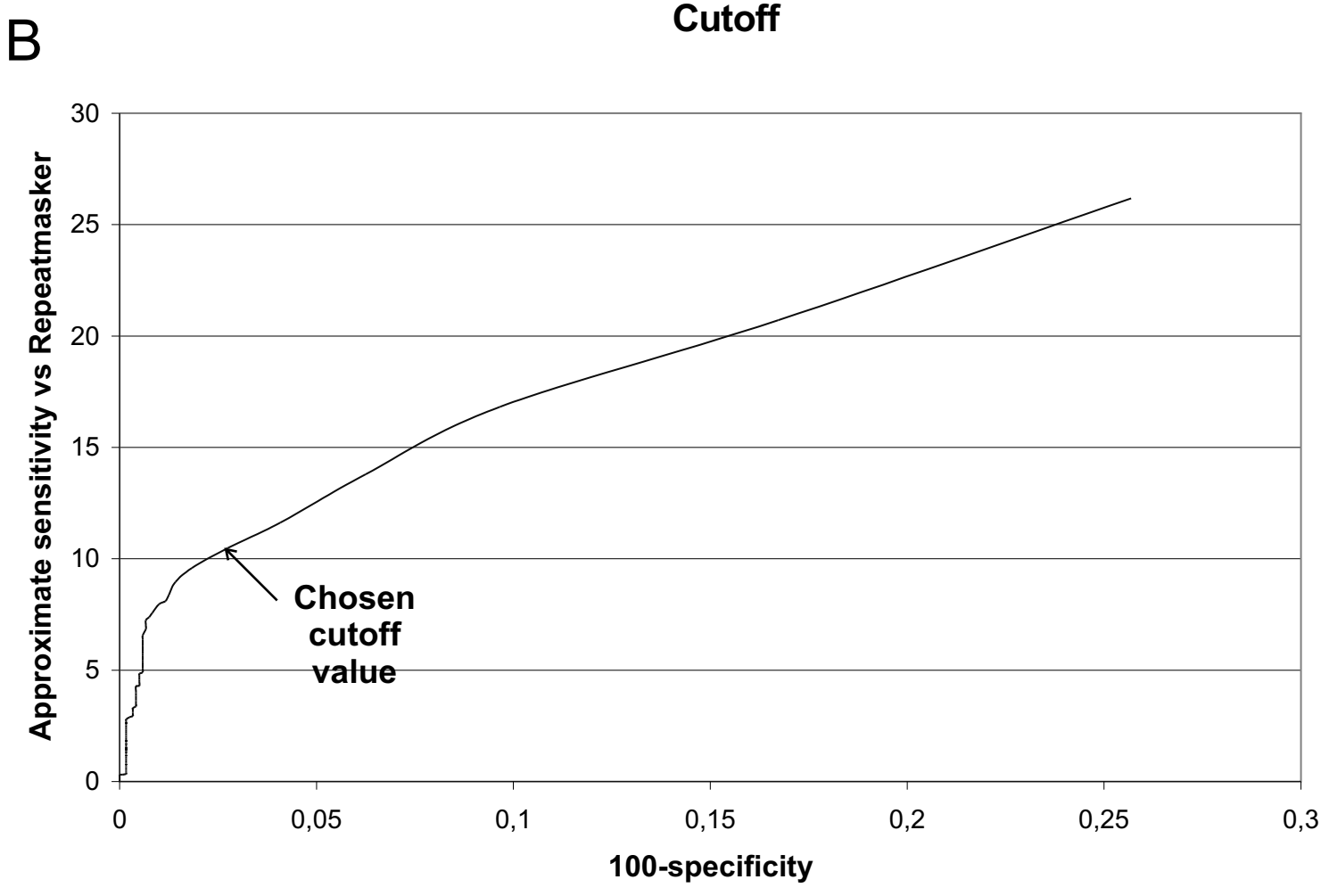


B



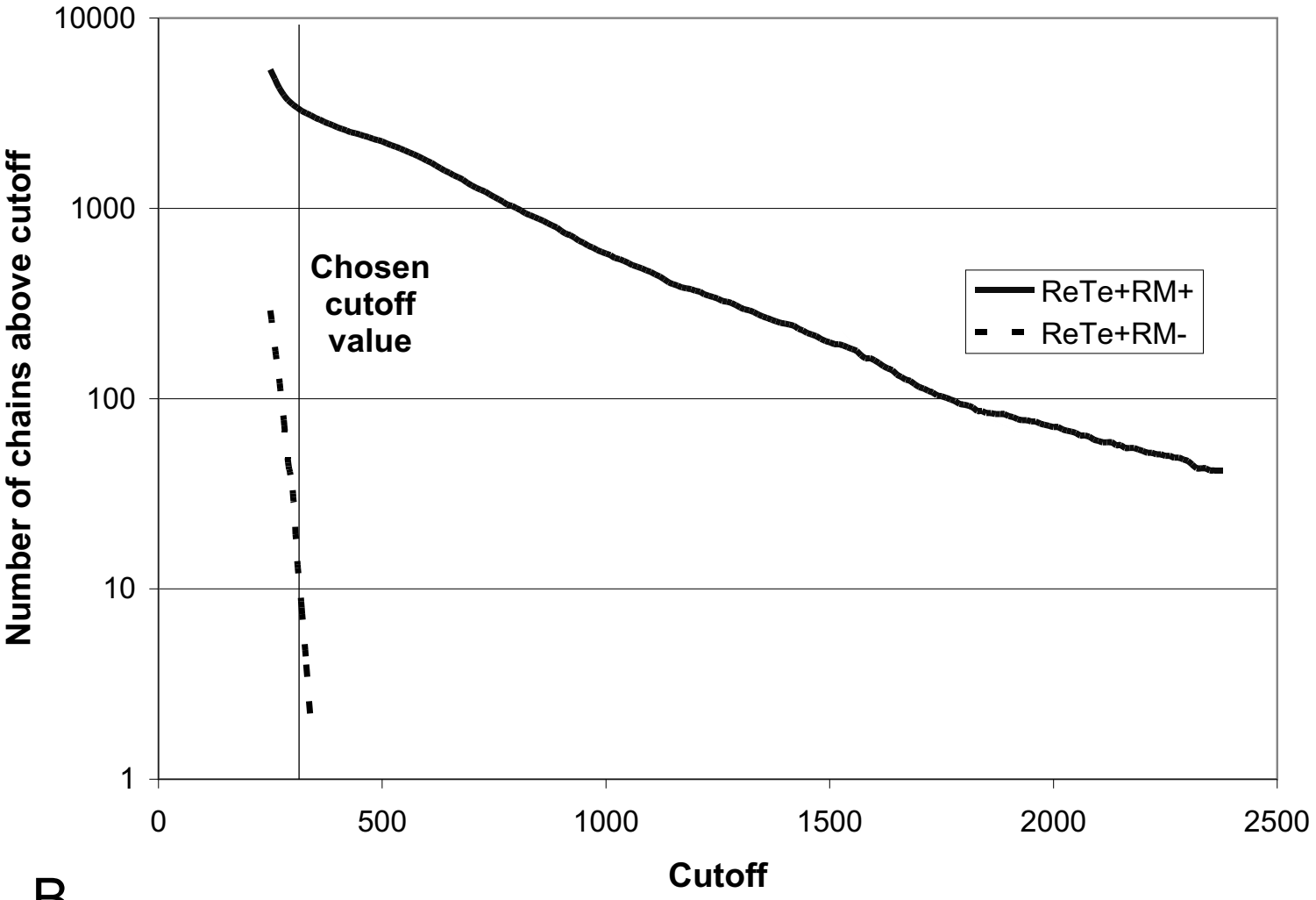
RM vs ReTe for canFam2



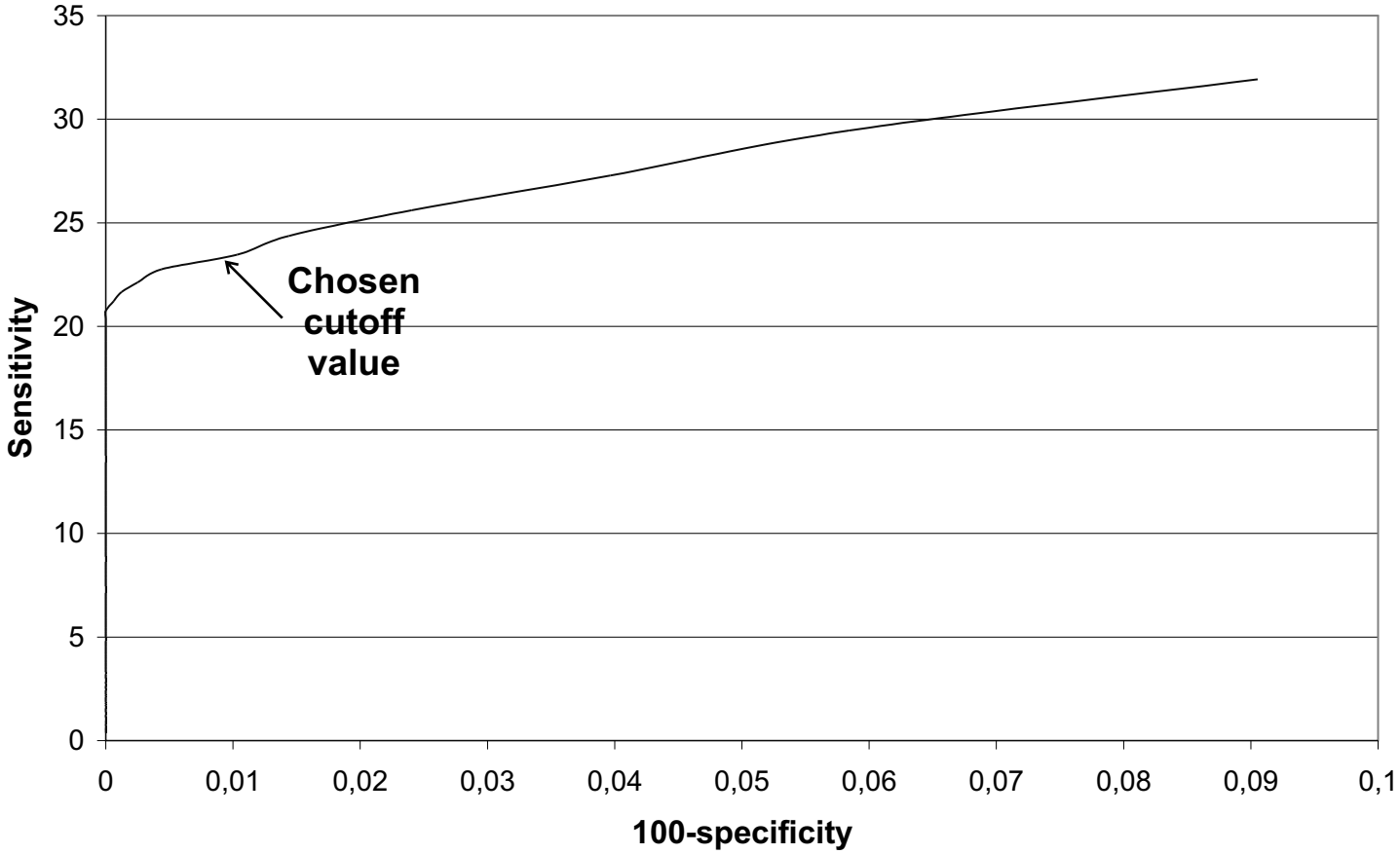


RM vs ReTe for hg18

A

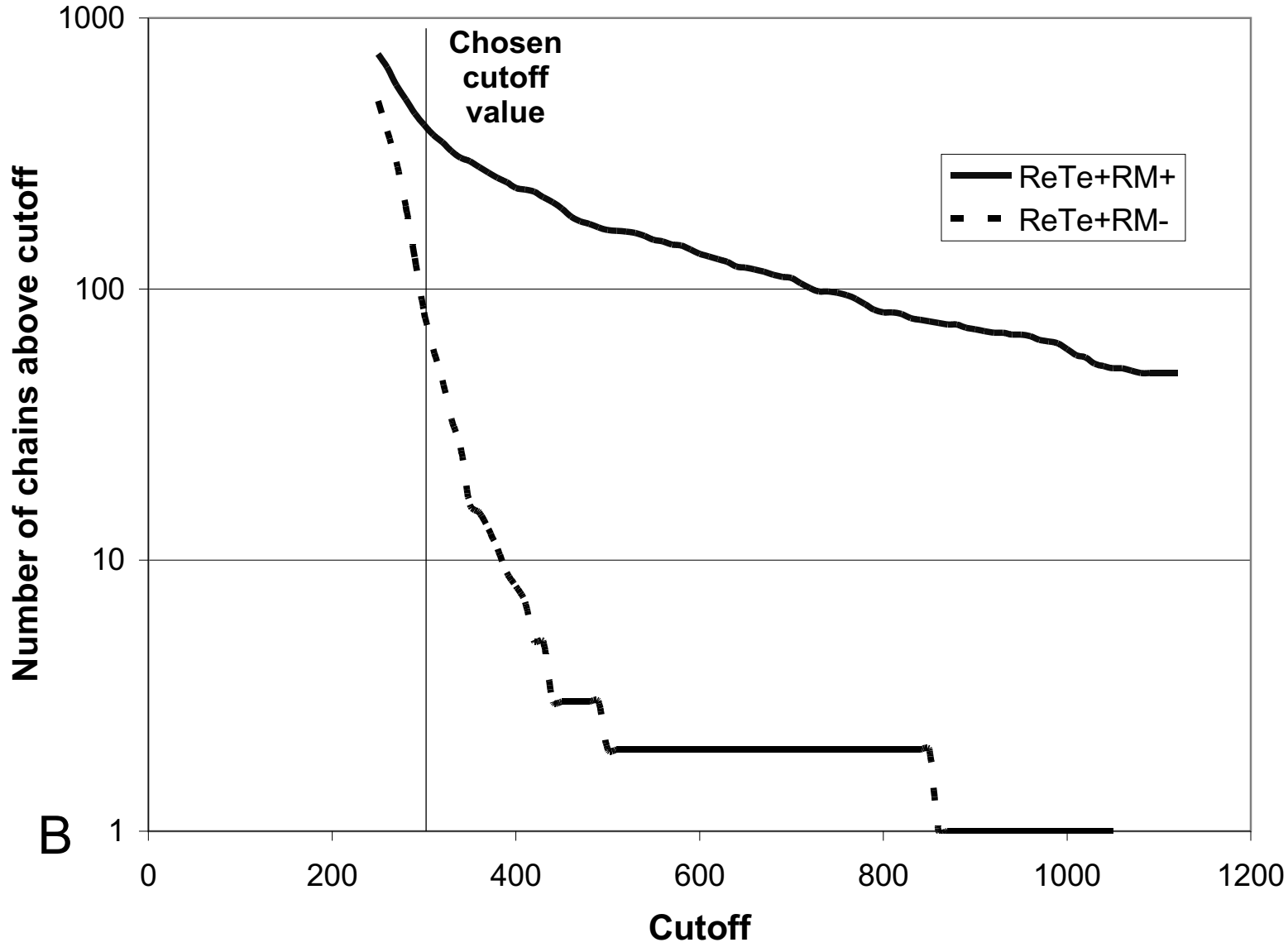


B

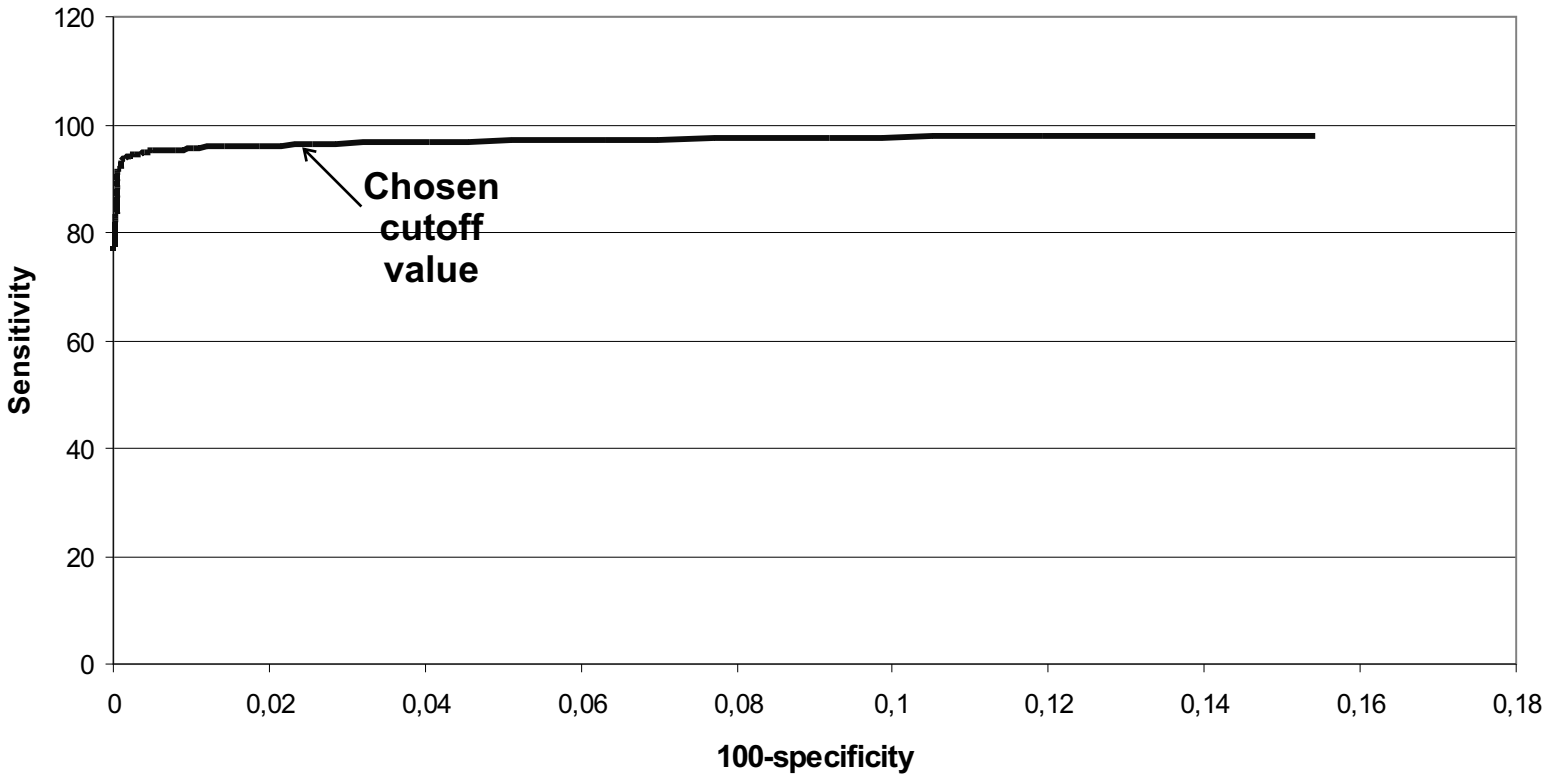


RM vs ReTe for galGal3, smaller set

A

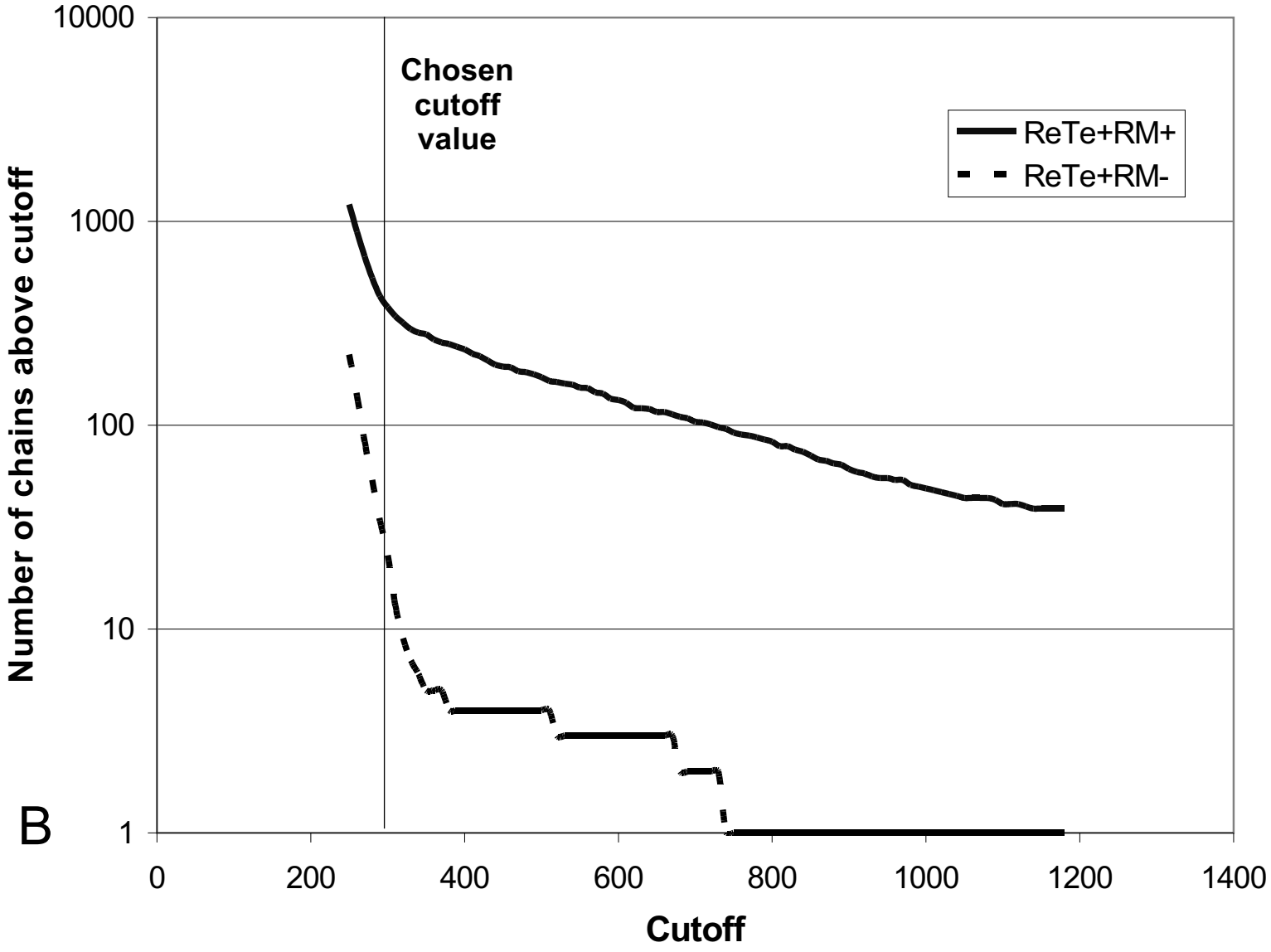


B

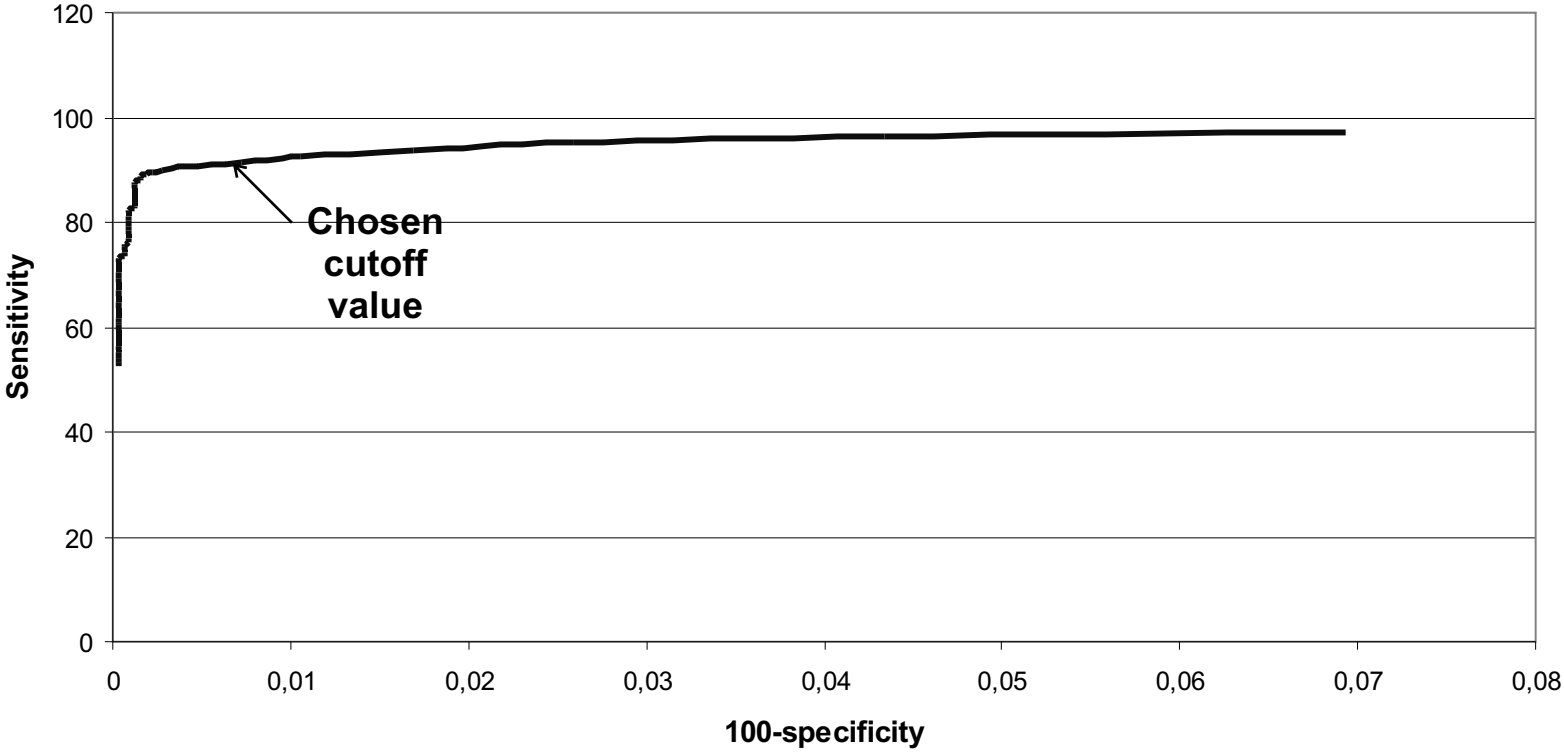


A

RM vs ReTe for canFam2, smaller set

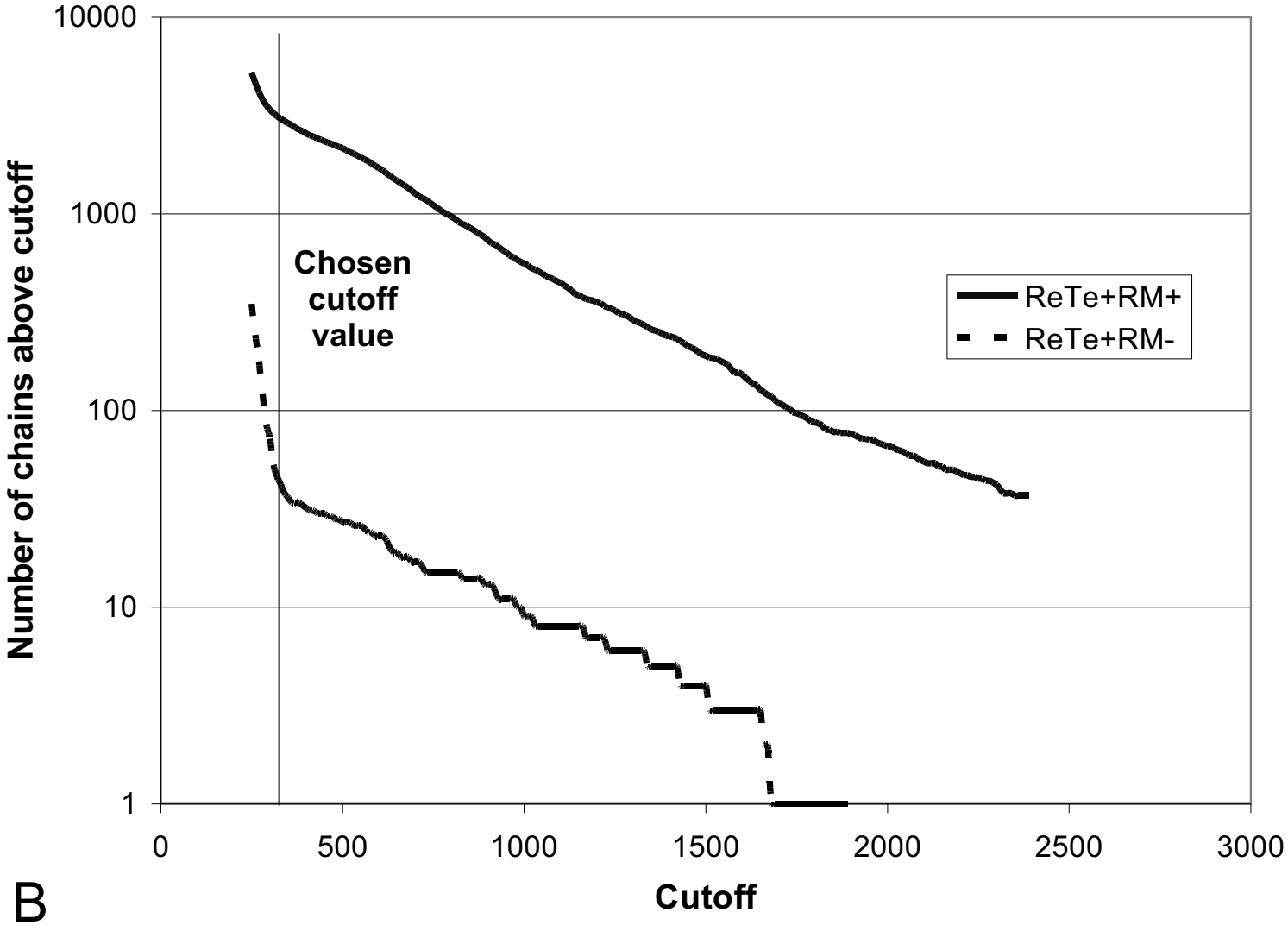


B

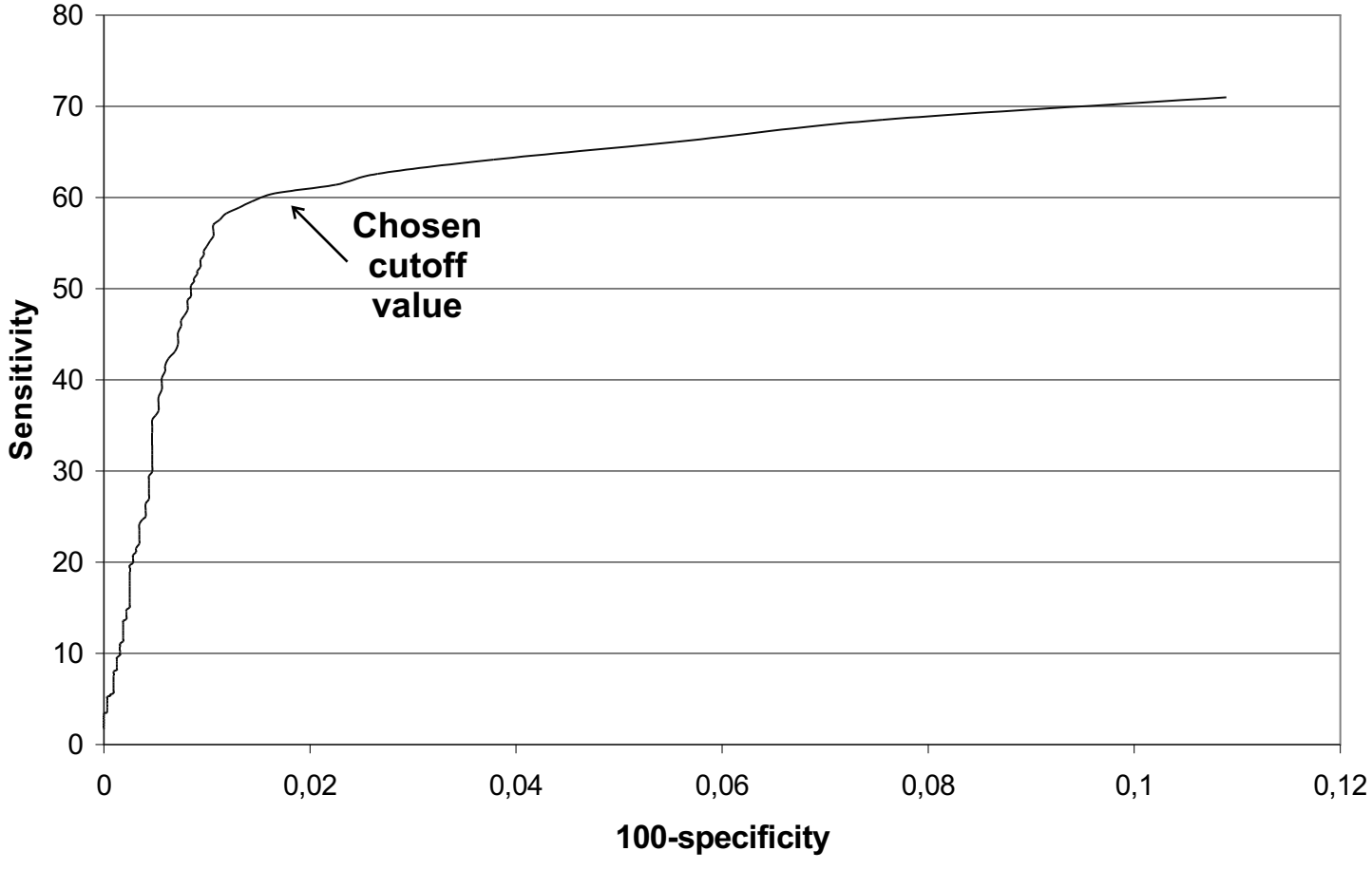


RM vs ReTe for hg18, smaller set

A



B



Description of some proviral chains missed in older or recent versions of RM.

In this section, only elements with a ReTe score above 500 are considered.

homo sapiens: The proviruses HERVFc1 (chr Xq21.33) and HERVFRD (chr 6p24.2) were missing in hg15. In later (hg18) versions, HERVFc1 is entirely covered by HERVFH21 RM repeats. HERVFRD is partially covered by MER93 and MER50 repeats in hg18.

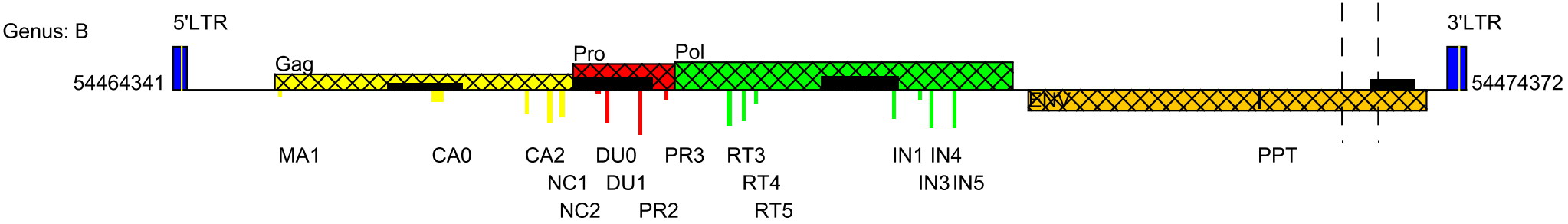
canis familiaris: A group of 7 HERVFc like repeats was not covered by RM in canFam1. In canFam2, they are all covered by repeats named CFERVF1.

gallus gallus: Several betaretroviruslike elements are still not covered in galGal3. One of them is shown in S17.

The retroviral genome diagrams (S17) produced from a retroviral genomic SQL database by ReTe have these conventions:

ReTe delineated gag, pro, pol and env genes are shown in yellow, red, green and orange, respectively. Hatching denotes a gene with more than one stop and frameshift. An Env which was reconstructed by EnvTracer is shown below the base line. A black bar shows the longest open reading frame within a gene. Motif hits are shown below the base line as bars whose length is proportional to the hit score.

Betaretroviruslike chain in galGal3, chromosome 2_54464341



Gammaretroviruslike chain in canFam2, chromosome 3_85023462