Readme file for the Supplementary information

Supp_1 RetroTector©010 interpretation of the ERV3 7q11 locus and a discussion of the XonID findings.doc

This describes features detected by RetroTector version 010 from this locus. Gag-, Pro- , Pol-, and Env puteins are defined. The putative XORF protein is discussed.

Supp_2 A detailed list of the ERV3-like elements.

This list contains all ERV3-like elements detected by RetroTector 010 in the human genome, version hg15. A larger subset of 76 ERV3-related sequences were first collected using similarity to the ERV3 7q11 Pol putein, and the BLASTlike protein similarity algorithm of RetroTector Shell version 01. This larger subset was further analysed by Clustal W 1.83 alignment of Pol puteins (Supp_3a), a cladogram using the neighbor-joining method, and a crosswise similarity table of the Pol puteins (Supp_4a). Using criteria described in the main text, 42 ERV3-like elements were defined.

Supp_3a Clustal W 1.83 Alignment of ERV3-related elements.

As described under Supp_2, 76 ERV3-related sequences were aligned together with gammaretroviral reference sequences.

Supp_3b Cladogram of ERV3-related elements with details.

The aligment of Supp_3a gave rise to a guide tree, based on the neighbor-joining method. Besides the chromosomal locations, the xx of each element are presented.

Supp_4a. Similarity cross table of Pol of ERV3-related elements.

The cross table shows that the 42 ERV3-like elements are separated from other similar elements, the HERV-E like and other gammaretroviruses. A similarity of 80% (shown in red; defined in Materials and Methods) delineates the ERV3-like sequences from the other. However, the division between ERV3-like and HERV-E like is not entirely clear. Some ERV3-like elements are more than 80% similar to the HERV-E like elements.

Supp_4b. Similarity cross table of Env of ERV3-related elements.

This was performed as described under Supp_4a. Env is more divisive than Pol. The ERV3-like elements appear in two Env-defined subgroups.

Supp_5 Alignment of the ERV3 5´and 3´LTRs.

The purpose of the alignment is to highlight differences between the two LTRs. Promoter function normally resides in the 5´LTR. Selection for Env expression should be accompanied by a functional 5´LTR.

Supp_6a.TFBS findings for the ERV3 LTRs from ConSite.

Transcription factor binding sites were predicted using default settings.

Supp_6b. A map of the ERV3 LTR, with TFBS predictions plotted.

Predicted transcription factor binding sites, and the degree of divergence between 5´and 3´LTRs, were plotted, together with U3, R and U5 borders. Likely TATAA and AATTAAA boxes are also indicated.

Supp_7. List of PBS sequences used by ERV3-like sequences.

RetroTector 010 interpretation of the ERV3 7q11 locus from hg15 (April 2003)

RETROVID output:

```
ChainS1::
   A:0.37
   B:0.37
   C:0.99
   D:0.68
   E:0.37
   L:0.06
   S:0.36
   G:0.06
   O:0.06

63858202
63848590
 Type C Score= 1754

SubGene 5LTR, type C, score=90 , hotspot 63857663
  5LT:ABCDELSGO (5'LTR): Score=90 at 63857663 frame 2 [63858202-63857610]
    X scored against

tgaggcaggaaatataaaaggaaaaacaagtaaagggaaaacaagtccttccctgatcagtctgactcactccaaagtcctgctggag
ctatgataacattatctgcaaggccaggcagggaccccaaaagaatgggctccaggagcagagatgagaaaaacaagttctccttatc
agtttccgccttgaaattctttccccataccattattctttgttctgctctcacaactatttttgtaactatttctgcaagtttgcaaagatttcataa
gttcctgtttttctttctgtagcacggcaaggtcacaagacatgcttaagtaagatagggtcatgttgcaaatcctgttgtaaaacctgtcac
ggtatgattaactgtctttgttctgcttctgtaagactgctttcctgtctcacaggtttcatgccaaaaacctgacccgcccctgttggttgcat
gtataaaagtcaagccctgtcattgttcagggctcagcctttggatgttcatcggctgggctggtggtcacctaaataaaatcctcctgttc
caccaagtggtctctccagcctcctgattcccacaaca (593 bases)

SubGene PBS, type C, score=141 , hotspot 63857606
  PBS:C (tRNAArg-ERV3, ver020308, Snakehead RV): Score=141 at 63857606 frame 2
[63857606-63857589]
    TGGtgagccagccaggag scored against
    tggcgagccagccaggag (18 bases)

SubGene MA, type CD, score=158 , hotspot 63856864
  MA1:E (WDSV): Score=1 at 63856864 frame 3 [63856864-63856841]
    MGNssstp scored against
    mggksskp (24 bases)
  MA2:D (HTLV, end of MA): Score=11 at 63856423 frame 3 [63856423-63856394]
    iPPPYvepta scored against
    tpppyvpiyp (30 bases)
  MA3:ABCDELSGO (Exp Wills acid motif): Score=100 at 63856420 frame 3 [63856478-
63856409]
    K(1-5)K(10-25)!PPPY scored against
    krkatrktsfagtsrgdrnppppy (70 bases)

SubGene CA, type C, score=61 , hotspot 63855756
```

```
  CA0:ABCDELSGO (CA Start NN): Score=14 at 63856095 frame 1 [63856095-63856006]
    CAStartNNData.txt scored against
    llhreapsphrsdavhfsdtksnvarlqta (90 bases)
  CA1:C (HERV E (4-1)): Score=33 at 63855756 frame 1 [63855756-63855691]
    seVIQGkEESPAkFHERLcEAY scored against
    levlqetnzspgqfyerlckaf (66 bases)

SubGene NC, type C, score=59 , hotspot 63855343
  NC1:C (S71): Score=59 at 63855343 frame 3 [63855343-63855302]
    CtyCkqiGHwkkEC scored against
    caqckkrghwkgec (42 bases)

SubGene Prot, type C, score=151 , hotspot 63855094
  PR2:C (MuLV, motif A): Score=45 at 63855094 frame 3 [63855094-63855065]
    lvDTGAqhSv scored against
    mvdtgaehlv (30 bases)
  PR3:C (MuLV, motif B): Score=56 at 63854905 frame 3 [63854905-63854882]
    llGRdllt scored against
    llgrdllq (24 bases)

SubGene RT, type C, score=525 , hotspot 63854225
  RT1:C (MLV): Score=58 at 63854501 frame 2 [63854501-63854472]
    wNtPllpVKK scored against
    wntplllvqk (30 bases)
  RT2:C (MLV): Score=87 at 63854330 frame 2 [63854330-63854289]
    tvlDLkDaFFclrL scored against
    tvldlkdaffclyl (42 bases)
  RT3:C (MLV): Score=85 at 63854225 frame 2 [63854225-63854181]
    qltWtrLPQGfknSP scored against
    qlawtslpqgfknsp (45 bases)
  RT4:C (MLV): Score=61 at 63854111 frame 2 [63854111-63854082]
    lqYvDDLlla scored against
    lqyidnllla (30 bases)
  RT5:C (MLV): Score=59 at 63854018 frame 2 [63854018-63853989]
    GyrasakKaQ scored against
    gyrvskkkaq (30 bases)

SubGene IN, type C, score=110 , hotspot 63852376
  IN5:C (MuLV, motif D): Score=14 at 63852215 frame 2 [63852215-63852192]
    gtDNgpaF scored against
    gsdnrlaf (24 bases)
  IN6:C (HERV-E): Score=45 at 63852131 frame 2 [63852131-63852084]
    AYqPQSsgKVERmnrt scored against
    ayrlqslgkvecmnwt (48 bases)
  IN7:C (BAEV): Score=15 at 63851738 frame 2 [63851738-63851691]
    eprWkGPyiVLlttpt scored against
    gpiwdglhtvilsipt (48 bases)

SubGene SU, type C, score=82 , hotspot 63849885
  SU3:C (HERV-E): Score=82 at 63849885 frame 1 [63849885-63849844]
```

```
    KRgiviGnWkdnew scored against
    krgitigdwkdnew (42 bases)

SubGene TM, type C, score=181 , hotspot 63849657
  TM3:C (HERV E/ERV3): Score=100 at 63849651 frame 1 [63849651-63849616]
    yQNRLALDYLLA scored against
    yqnrlaldylla (36 bases)
  TM5:ABCDELSGO (hydrophobic motif): Score=37 at 63849414 frame 1 [63849414-63849379]
    X scored against
    ivlailgvclil (36 bases)

SubGene 3LTR, type C, score=90 , hotspot 63848643
  3LT:ABCDELSGO (3'LTR): Score=90 at 63848643 frame 1 [63849181-63848590]
    X scored against

tgaagcaggaaatataaaaggaaaaacaagtaaagggaaaacaagtcctttcctgaccagtctgactcactccaaagtcctgctggag
ctatgataattatctgcaaggccaggcaggggctccgaaggagtgggctccaggagcagggatgagaaaaacaagttctccttatcag
tttccctgtttgaaattctctccccataacattattctttgttctgctctcacaactatttttgtaactatttctgcaagtctgtaaagattttgtaagt
tcttgtttttctttctgtagcatggcaaggtcacaagacatgtttaagtaaggtaggctcatgttgcaaatcctgttgtaaaacctgtcacggt
atgattaactgcctttgttctgcttctgtaagactgctttctcacctcgcaggttttgcgccaaaaacccgacttgcccctgcctgatgcatgt
ataaaagtcaagccctgtctttgttcagggctcagcctttggatgttaatccgctgggccagtggccacctaaataaaaccttcctgttgca
cccagtgatctctccggcctcctgattcccacaaca (592 bases)

Integration sites tata/tg<>ca/tata
```

ORFID output:

Gag defined by ORFID, 1338 [63856864]- 3068 [63855145]:

```
MgGkSSKPIPLeCMLrNFKzDytGDYeikLTPqRLRTLCEIEWPsFNVGWLaEGTIDREkIgrV
fKVVTGvGgQpGHPDQFPYIDSWLniVQtQPaWLQPCLASycKtLiArAKPkVKekSASLSA
TETKgKPqEKPVLQePAEEiETPPPYvPiyPPLPrTAPeQPDSdGDtLqATPqRGkSEPPPqEvk
kESqDdqaGRLRPGhTRvwQMPlRETRGPiYYDEQGQFkGGQPTFIYqPFSttDLLNWKhHT
PysSYTEKPQALTDLMQSIfQTqNPTWpDCkQLLLTLFNTEEcRRVtQAALhWLEHNAPE
gtLNvQaymLRQFPeaSlHWDpNDAAqlqyLQrYwEALLqGLKeGGKKAVNtgKilEVLQETn
zSPGQFYERLCKAFzLYTPFDTEAtENQcMvNAeFVGQAqgDIRRKLQKLEcftGmNATQ
LLevAtKVYvnRnQEKERLIRlKKKaDLLvavLTgReTSIASGcGhShGhGGvQARQRPEsQP
RLcrNQCAqCKKRGHWKgECPEGNEEndrdykTGKlPAKgyCVPRESDTyLirLaGTGeyedza
rlGsfSlP
```

gag DNA:

```
atgggagggaagtcatcaaaacccatcccattagaatgtatgttaaggaactttaagtaagattatacaggggattatgagatcaagttga
ccccccagaggttgagaactctctgtgaaatagaatggccctctttcaatgttggatggctggccgaaggaactatagatagggaaaaa
attggccgtgtatttaaggtggtgactggggtcggaggacagccagggcatccagaccaatttccttatattgactcatggctaaatatag
ttcaaactcaacctgcatggctgcagccttgcctggcttcttattgcaaaacgctcatagctcgagccaagcctaaagtaaaagaaaaat
cagcttcactgtcagctacagagacaaaaggaaagccacaagaaaaaccagttttgcaggaaccagcagaggagatagaaacccct
cctccctatgtgccaatctatccccctttaccaaggacggcccctgagcagccagactcagatggtgacacactccaggctacacctca
aagggggaaatctgagcccccacctcaggaggtcaagaaggaaagtcaggatgatcaagcaggccgccttcgacctggccacacc
agagtgtggcagatgcccctccgtgaaacacggggacctatctattatgatgaacaggggcaggttcaagggggggcaaccgactttt
atttaccagcctttttcaaccactgatctcctaaactggaaacaccatactccctactccctcctacacagagaagccccaagccctcaca
gatctgatgcagtccatttttcagacacaaaatccaacgtggccagattgcaaacagctcctcctgacactgtttaacaccgaggaatgc
cgaagggtgacccaggcagccctccactggctggaacacaatgcaccagaaggcacacttaatgtccaggcatatatgctcagggcc
aattcccagaagccagcctacactgggacccaaatgatgcagcccaattacagtacctacagaggtactgggaggcactcctgcaag
ggctaaaggaaggtgggaaaaaggcagtcaacacggggaaaatcttggaagtgctgcaggagactaattaaagtcctggtcagtttta
tgagagactctgtaaagcattctagctttacaccccgtttgacactgaggctactgagaaccagtgcatggtgaacgcagagtttgtagg
acaggcccagggagacatcaggcgaaagctgcaaaagctagagtgtttcacgggcatgaatgccactcagcttttagaagtggccac
caaggtgtatgttaaccgtaaccaggagacaaagagaggctgatcagacttaagaaaaaggccgatctgctagtggcagtcctcacag
gaagggaaactagcatcgcgagtgggtgtggacatagccatggacatggaggggttcaggccaggcagagacctgaaagtcaacc
aagactatgtaggaatcaatgtgcgcaatgcaaaaagaggggacactggaaaggtgaatgtccagagggcaatgaagaaaatgaca
gagactataaaactggaaaactgccagccaagggctattgtgtcccaagggagtcagatacctacttgatcaggctggcaggaactgg
agaatatgaagattaggccagactgggctccttctctttagccc
```

Detailed ORFID Gag result:

```
Executor: Puteinview
NumberOfHits: 5
LeadingInfo::

Leading path in Gag
Starts at=63856864
Putein string of length=368 and total score=312.7832
MgGKSSKPIPLECMLRNFKzDYTGDYEIKLTPQRLRTLCEIEWPSFNVGWLAEGTIDRE
KIGRVFKVVTGVGGQpGHPDQFPYIDSWLNIVQTQPaWLQPCLASycKtLiArAKPkVKe
kSASLSATETKgKPQEKPVLQePAEEiETPPPYVPIyPPLPRTAPEQPDSDGDtLqATPqRG
```

```
kSEPPPqEvkkESqDdqaGRLRPGhTRvWQMPlRETRGPiYYDEQGQFKGGQPTFIYQPFST
TDLLNWKHHTPysSYTEKPQALTDLMQSIFQTQNPTWPDCKQLLLTLFNTEEcRRVTQ
AALHWLEHNAPEgtLNvQAYmLRQFPEASLHWDPNDAAQLQYLQrYWEALLQGLKE
GGKKAVNTGKI
yielding average=0.84765095
Inside limits=1.0
Kozak score=0.53125*1
Comparing score=2.378901
::
TrailingInfo::

Trailing path in Gag
Ends at=63855144
Putein string of length=66 and total score=52.915493
CAqCKKRGHWKgECPEGNEEndrdykTGKlPAKgyCVPRESDTyLirLaGTGeyedzarlGsFSl
P
yielding average=0.7897835
Inside limits=1.0
For fit length (3-2)*0.5
For ending at cleavage site=0.44444445*0.5
Comparing score=2.5120058
::
EstimatedStartPosition: 63856864
EstimatedLastPosition: 63855145
LengthInside: 571
LengthTotal: 593
AlignedAcids: 494
AverageScoreInside: 0.8380022
AverageScoreTotal: 0.80218625
MostUsedRow: 10 (HERV-E_gag_AL391419_alg2_1fs)
StopCodonsInside: 4(20, 31, 15)
StopCodonsTotal: 4
AmbiguousAcidsInside: 0
AmbiguousAcidsTotal: 0
ShiftsInside: 5
ShiftsTotal: 6
{ Starts at position 63856913, ends at 63855122
Putein::
        1_______
2_________
3______________________________
4_____________________
5_____________

semkgrvsgfchnweeMgGkSSKPIPLeCMLrNFKzDytGDYeikLTPqRLRTLCEIEWPsFNVGW
LaEGTIDREkIgrVfKVVTGvGgQpGHPDQFPYIDSWLniVQtQPaWLQPCLASycKtLiArA
KPkVKekSASLSATETKgKPqEKPVLQePAEEiETPPPYvPiyPPLPrTAPeQPDSdGDtLqAT
PqRGkSEPPPqEvkkESqDdqaGRLRPGhTRvwQMPlRETRGPiYYDEQGQFkGGQPTFIYqP
FSttDLLNWKhHTPysSYTEKPQALTDLMQSIfQTqNPTWpDCkQLLLTLFNTEEcRRVtQ
AALhWLEHNAPEgtLNvQaymLRQFPeaSlHWDpNDAAqlqyLQrYwEALLqGLKeGGKKA
```

```
VNtgKilEVLQETnzSPGQFYERLCKAFzLYTPFDTEAtENQcMvNAeFVGQAqgDIRRKL
QKLEcftGmNATQLLevAtKVYvnRnQEKERLIRlKKKaDLLvavLTgReTSIASGcGhShGhG
GvQARQRPEsQPRLcrNQCAqCKKRGHWKgECPEGNEEndrdykTGKlPAKgyCVPRESDT
yLirLaGTGeyedzarlGsfSlPpepmvt

333333333333333333333333333333333333333333333333333333333333333333333333333
333333333333333333333333333333333333333333333333333333333333333333333333333
333333333333333333333333333333333333333333333333333333333333333333333333333
3333333333333333333111111111111111111111111111122222222222222222222222222222
22222222222222222222222222222222222222221111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111113333333333333333333333333333333333333333333333333333333333
33333333333333333333333333333333333333333333333333333333333331133333
                >
<
 63856912             63856822             63856732             63856642
63856552             63856462             63856372             63856282
63856192             63856101             63856010             63855920
63855828             63855738             63855648             63855558
63855466             63855376             63855286             63855196
::
Hits::
1 63856864 1 MA1 P E #MGNssstp WDSV
2 63856423 11 MA2 P D #iPPPYvepta HTLV, end of MA
3 63856095 14 CA0 ANN ABCDELSGO #CAStartNNData.txt_____________ CA Start NN
4 63855756 33 CA1 P C #seVIQGkEESPAkFHERLcEAY HERV E (4-1)
5 63855343 59 NC1 P C #CtyCkqiGHwkkEC S71
::
{ Created by ORFID with parameters
{ NonAlignedScore: 0.4
{ GlycosylationFactor: 0.2
{ StopCodonFactor: -0.4
{ LastDNAEnd: 63855121
{ NonORFHexamerFactor: -0.1
{ MotifHitFactor: 0.01
{ DNAFile: chr7_639.txt
{ StopCodonValue: -20
{ ScriptName: C:\rt010\Workplace\chr7\chr7_639\orfid_s00001cgag001Script_.txt
{ OutputFile: ORFIDout_S00001CGag_001.txt
{ InputFile:
{ MinHitScore: 15
{ Debugging: No
{ Database: Ordinary
{ Strand: Secondary
{ FrameGene: Gag
{ PuteinFile: Putein_S00001CGag_001.txt
{ ORFHexamerFactor: 0.2
{ ChainNumber: 00001
{ VirusGroup: C
{ FirstDNAStart: 63856913
```

```
{ FirstDNAEnd: 63855284
{ LastDNAStart: 63856864
{ MinAverageScore: 20
{ FrameShiftFactor: -0.9
{ ScoreFactor: 0.1
{ Execution time was 2140 milliseconds
{ Belongs in C:\rt010\Workplace\chr7\chr7_639
{ Created Sun May 11 18:13:04 GMT+01:00 2003 under RetroTector version 0.10 03 05 10
{ using Database:Ordinary last modified Fri Dec 20 13:49:56 GMT+01:00 2002
```

Pro defined by ORFID 3060 [63855142]- 3366 [63854836]:

```
QEPMVTLEVGGQlMnFMVDTGAEHLVVTQPIGPsSKdhTTIvGATGVpEKRPFCQPRRC
vfGGzEVQHkFLYLPNyPVPLLGRDLLQKLhAQITfGSqGIlL
```

pro DNA:

```
caggagcccatggtcacattagaagtaggaggccagctaatgaactttatggtagacactggggctgaacacttggtagtgacccagc
ccatagggccatcatccaaagatcatacaactattgttggggctacaggggtcccagagaagaggccattttgtcagccaaggaggtgt
gtctttggaggatgagaggtccaacataaattcctgtatctccccaattatccagtgcctctactaggaagagacctactccaaaaattgc
atgcacagattacctttgggtcacaaggagatattactt
```

Detailed ORFID Pro results:

```
Executor: Puteinview
NumberOfHits: 2
LeadingInfo::

Leading path in Pro
Starts at=63855142
Putein string of length=16 and total score=14.666667
QEPMVTLEVGGQLMNF
yielding average=0.8627451
Inside limits=1.0
Comparing score=1.862745
::
TrailingInfo::

Trailing path in Pro
Ends at=63854835
Putein string of length=23 and total score=17.838865
LLGRDLLQKLHAQITFGSQGILL
yielding average=0.7432861
Inside limits=1.0
For ending at stop codon=1.0*0.5
For ending at cleavage site=0.13333334*0.5
Comparing score=2.3099527
::
EstimatedStartPosition: 63855142
EstimatedLastPosition: 63854836
```

```
LengthInside: 102
LengthTotal: 225
AlignedAcids: 100
AverageScoreInside: 0.8950847
AverageScoreTotal: 0.6075172
MostUsedRow: 2 (Human_chromosome_18_AC022030_1)
StopCodonsInside: 1(9, 2, 1)
StopCodonsTotal: 3
AmbiguousAcidsInside: 0
AmbiguousAcidsTotal: 0
ShiftsInside: 1
ShiftsTotal: 3
{ Starts at position 63855284, ends at 63854603
Putein::
                                                                 1_________
2_______

ndrdyktgklpakgycvpresdtylirlagtgeyedzarlgsfslapQEPMVTLEVGGQlMnFMVDTGAEHLV
VTQPIGPsSKdhTTIvGATGVpEKRPFCQPRRCvfGGzEVQHkFLYLPNyPVPLLGRDLLQ
KLhAQITfGSqGIlLzltqpkamvltltipqaeewrlymkyrhqcslvrrrkrllfrpidkipgvwaedkppglavnqalvv
elkpgatpv

333333333333333333333333333333333333333333333333333333333333333333333333333
333333333333333333333333333333333333333333333333333333333333333333333331111
333333333333333333333333333333333333322222222222222222222222222222222222222
                                                 >    !
<
 63855283                    63855193                    63855103                    63855013
63854923                    63854830                    63854740                    63854648
::
Hits::
1 63855094 45 PR2 P C #lvDTGAqhSv MuLV, motif A
2 63854905 56 PR3 P C #llGRdllt MuLV, motif B
::
{ Created by ORFID with parameters
{ NonAlignedScore: 0.4
{ GlycosylationFactor: 0.2
{ StopCodonFactor: -0.4
{ LastDNAEnd: 63854602
{ NonORFHexamerFactor: -0.1
{ MotifHitFactor: 0.01
{ DNAFile: chr7_639.txt
{ StopCodonValue: -20
{ ScriptName: C:\rt010\Workplace\chr7\chr7_639\orfid_s00001cpro001Script_.txt
{ OutputFile: ORFIDout_S00001CPro_001.txt
{ InputFile:
{ MinHitScore: 15
{ Debugging: No
{ Database: Ordinary
{ Strand: Secondary
```

{ FrameGene: Pro
{ PuteinFile: Putein_S00001CPro_001.txt
{ ORFHexamerFactor: 0.2
{ ChainNumber: 00001
{ VirusGroup: C
{ FirstDNAStart: 63855284
{ FirstDNAEnd: 63854869
{ LastDNAStart: 63855121
{ MinAverageScore: 20
{ FrameShiftFactor: -0.9
{ ScoreFactor: 0.1
{ Execution time was 937 milliseconds
{ Belongs in C:\rt010\Workplace\chr7\chr7_639
{ Created Sun May 11 18:13:05 GMT+01:00 2003 under RetroTector version 0.10 03 05 10
{ using Database:Ordinary last modified Fri Dec 20 13:49:56 GMT+01:00 2002

Pol defined by ORFID 3449 [63854753]- 6638 [63851564]:

APMQPCTQEEEkLlFrPIDKiPgVWAEdkppGLAVnQALVVELKpGATPVqVcQYPlSQEdi
wGIyKHlKwLcDhGIiiqCQSPWNTPLLlVqKplPGpgsDeYiPVQgLhaVNqatVtIHPvVPNlYt
LmGLilaSaTWfTVLDLKDAFFCLyLaPvSQPiFAFzWdnsvtGTGgQLaWTsLPQGFKNSPTi
FgEALaSDLkAytPpndNcALLQYiDnLLLAAPTqEDCywGTQdLLhlLwKaGYRVSkKKAQI
ChekVKYLGfivsqGeRWLghgzKqAIcalPTPTTwRQigEFLGaAGFChiWIPnFSLiArPLYeaT
REGekePllzkaDQkKvFkQIKeALTqAPALGLPDtTKPfFLYVhElKGmAiGVLTQiiGsWhRP
VvYLSKqLTVvlGWPPCcKAlAtttLLTqeADKLTLGQqLTIqvPHAVitLmdQreqhWLSNptM
TzYQgLLcenpRitleTVnTLNPATLLPVkPGiTfRDCgnvdEvfssRgDLTDQPfkDpDveyfiDGS
SFvlEGvcwAGzAvVTldSVVeAzpqPAGTSAQKAELaALTrvLqLAEdKKIsVYTDSkYAFA
TlHVHGAIYKeRGLLTatGKEIKwKEzILqLLdAvwiPekVAamHCrGHQRaGtSEAKGNRkA
DrEAkwAamvtphfkKEAlAmplLpEpplQevSSySPnEkaVLPKKlEnILkEdGgnfDgRsaiPKmVa
PKFvKzLHHeLvGktaLeTLLgRhFYvPrLtAitQaVcEQyltCAQnNpzqwptRpPGiqeiGtmPcenll
mDTElPqagsYrYmLVlVyTFSGWAkAFPTrsKnszeVTKiLLrnIiPRFGlsVtLGSDNrlAFlAeii
QeLTrvLkIkWKLHtAYRlQSlGkVEcMNwTLKqLLkkycqETplRWdqVLPmVLLqVRcTPtk
qtGysPYEILfGkPsPiisqIkgnLrElGELTlRrqMQALgiAmzsVHgWvQKnAhKPDRPIHPFKPr
DSlWVkkznPttLgPiWdGlHTViLsiPTvVKVaGIvpWIHpsSqLKPAaQDkwTSqqDLdhaTQLI
LRGT

pol DNA:

gcaccaatgcagccttgtacgcaggaggaagagaaattattatttaggccaattgataaaattcctggagtatgggctgaagacaagcc
acctgggctggctgtaaatcaggcactggtagtagaattaaaaccaggagcaactccagttcaggtttgtcagtacccactttcccaaga
ggatatttggggcatttataagcatttaaagtggctctgtgaccatgggatcataatccaatgccagtcaccctggaacactccacttttgct
ggtacaaaaaccattgcctggaccaggatctgatgagtatataccggtgcagggcttgcatgctgtaaaccaagccacggtgaccata
catccagtagtaccaaacctgtatactttaatgggacttattctagcaagtgccacctggtttacagtcctggacttaaaggatgctttcttct
gtctctacctggcaccagttagtcagcccatctttgcattttaatgggacaattcagtcacaggcacagggggacagctcgcctggacta
gtctcccacaagggttcaagaattctcccacaatctttggggaagcactggcctcagacctcaaggcatacaccccaccaaatgacaac
tgcgccttgttgcagtacatagacaaccttcttttggcagccccaacccaagaggactgttactggggaacccaggacctcctccatctc
ttatggaaagcagggtatagagtatccaaaaagaaggcccaaatttgccatgaaaaggttaaatatttaggcttcatagtaagccaaggg
gaacgctggcttggccatggatgaaagcaagccatttgtgcacttccaactccaaccacctggcgccaaataggggaattcttagggg
cagcagggttctgccatatctggatcccaaatttctcacttatagccaggcccttatatgaagccacaagggagggtgaaaaggaaccc

```
ctcctctgaaaggctgaccagaagaaggtgtttaaacaaatcaaagaagccctaactcaggctccagccttaggactgccagatactac
taagcctttctttctatatgtccatgaattaaagggaatggctataggggtcctgactcaaatcataggatcatggcatcgcccagtggtgt
acttatccaaacagttgactctgtggtgctagggtggcctccttgctgtaaagcactagccactaccaccctattgacacaagaagctga
caaactgaccctggggcaacaattgaccatccaggtaccacatgcggttataactttaatggatcagagagagcagcattggttatcaaa
cccaacgatgacttgataccaagggctcctatgtgaaaatccacgcataactttagaaacagtgaacacccttaacccagctaccctgct
cccagtcaaaccgggaatcaccttccgtgactgtggcaacagtggacgaggtattctccagtcgaggagaccttacagaccagcctttc
aaggacccagatgtagaatacttcatagatggaagcagttttgtgctggaaggggtctgctgggctgggtaggcagtggtgaccttaga
ctcagtggtggaggcgtagcctcagcctgctggaacatcagcccagaaggcagagctagcagccctaacgagggttctccagctgg
cagaagacaaaaagatcagtgtttacacggattccaaatatgcttttgccacactgcatgttcatggagctatatataaagaaagaggact
cttaactgccacaggcaaagaaataaaatggaaagaatgaattctacagctcttagatgctgtatggatcccagagaaagtagcagctat
gcactgcaggggacaccagagggcagggacgtcagaggcaaaaggaaacagaaaggcagacagggaggcaaaatgggcagca
atggttacacctcattttaaaaaggaagccttagctatgcccctcctcccagagcctcccctccaggaggtctcaagttactctccgaatg
agaaggcctgttttgcccaaaaaactggaaaatatattgaaggaggatggtggaaattttctgatgggaggtcggccattcccaaaatgg
tggcccccaaatttgtgaagtaattgcatcatggaactcgtgggaaaaacagcactagaaacgctactaggacgccatttctatgtgccg
cgactcactgccatcacccaagctgtttgtgaacaatatttaacctgtgcccagaacaacccatgacaatggcccactcggcccccagg
aattcaggaaataggaaccatgccctgtgaaaacctgcttatggacttacagagttgcctcaggcagggagctatcggtacatgttagta
cttgtctacaccttttcaggatgggccaaggctttccctaccagaagcaagaactcatgagaagtgactaaaatactattaagaaacattat
ccccagatttggactgtctgtgactctagggtcagacaacaggctggcatttctagctgaaataattcaggaactaacacgagtgttaaaa
ataaaatggaaattacacacagcctaccggctgcagagcttaggaaaagtagagtgcatgaactggacactcaaacagctactgaaga
aatattgccaagaaacccctctgagatgggatcaagtcctgcccatggtcctccttcaagttaggtgcacccccaccaaacaaactggg
tattcgccctatgaaatactgttcggcaagccatccccaatcataagtcaaattaagggtaatctccgtgaactaggggaattaactttaag
gaggcaaatgcaggctttagggatagccatgtgaagtgtccatggctgggtacaggaaagaatgcccataagcctgatagatccaata
caccccttaaacccagggactctctttgggtcaaaaaataaaacccaaccactctgggacccatatgggatgggctccatactgtaatc
ttgtctattcccactgttgttaaagttgcaggaattgtgccttggatccatccatccagtcagctgaaaccagcagcccaggacaagtgga
ccagccaacaggacctagaccatgcaacccagctgatcctacgatggaacc
```

Detailed ORFID Pol result:

```
Executor: Puteinview
NumberOfHits: 8
LeadingInfo::

Leading path in Pol
Starts at=63854753
Putein string of length=84 and total score=64.92558
APMQPCTQEEEKLlFRPIDKIPGVWAEdKPPGLAVNQALVVELKPGATPVQVcQYPLS
QEdIWGIyKHLKWLcDHGIIIQCQSP
yielding average=0.76383036
Inside limits=1.0
Comparing score=1.7638304
::
TrailingInfo::

Trailing path in Pol
Ends at=63851563
Putein string of length=58 and total score=39.640514
GPIWDGLHTVILSIPTVVKVAGIVPWIHpSSQLKPAAQDkwTSQQDLDHaTQLILRGT
yielding average=0.67187315
Inside limits=1.0
For fit length (3-2)*0.5
```

```
Comparing score=2.171873
::
EstimatedStartPosition: 63854753
EstimatedLastPosition: 63851564
LengthInside: 1059
LengthTotal: 1175
AlignedAcids: 956
AverageScoreInside: 0.75426763
AverageScoreTotal: 0.68824154
MostUsedRow: 3 (GaLV_U26927)
StopCodonsInside: 12(39, 23, 72)
StopCodonsTotal: 14
AmbiguousAcidsInside: 0
AmbiguousAcidsTotal: 0
ShiftsInside: 9
ShiftsTotal: 12
{ Starts at position 63854869, ends at 63851326
Putein::
                                                                                1__________
2______________                3______________               4__________
5__________
6________              7_______________
8_______________

qitfgsqgditlsltqpkamvltltipqaeewrlymkyAPMQPCTQEEEkLlFrPIDKiPgVWAEdkppGLAVn
QALVVELKpGATPVqVcQYPlSQEdiwGIyKHlKwLcDhGIiiqCQSPWNTPLLlVqKplPGpg
sDeYiPVQgLhaVNqatVtIHPvVPNlYtLmGLilaSaTWfTVLDLKDAFFCLyLaPvSQPiFAFz
WdnsvtGTGgQLaWTsLPQGFKNSPTiFgEALaSDLkAytPpndNcALLQYiDnLLLAAPTqED
CywGTQdLLhlLwKaGYRVSkKKAQIChekVKYLGfivsqGeRWLghgzKqAIcalPTPTTwRQi
gEFLGaAGFChiWIPnFSLiArPLYeaTREGekePllzkaDQkKvFkQIKeALTqAPALGLPDtTK
PfFLYVhElKGmAiGVLTQiiGsWhRPVvYLSKqLTVvlGWPPCcKAlAtttLLTqeADKLTLG
QqLTIqvPHAVitLmdQreqhWLSNptMTzYQgLLcenpRitleTVnTLNPATLLPVkPGiTfRDCg
nvdEvfssRgDLTDQPfkDpDveyfiDGSSFvlEGvcwAGzAvVTldSVVeAzpqPAGTSAQKAEL
aALTrvLqLAEdKKIsVYTDSkYAFATlHVHGAIYKeRGLLTatGKEIKwKEzILqLLdAvwiP
ekVAamHCrGHQRaGtSEAKGNRkADrEAkwAamvtphfkKEAlAmplLpEpplQevSSySPnEka
VLPKKlEnILkEdGgnfDgRsaiPKmVaPKFvKzLHHeLvGktaLeTLLgRhFYvPrLtAitQaVcEQ
yltCAQnNpzqwptRpPGiqeiGtmPcenllmDTElPqagsYrYmLVlVyTFSGWAkAFPTrsKnszeV
TKiLLrnIiPRFGlsVtLGSDNrlAFlAeiiQeLTrvLkIkWKLHtAYRlQSlGkVEcMNwTLKqLLk
kycqETplRWdqVLPmVLLqVRcTPtkqtGysPYEILfGkPsPiisqIkgnLrElGELTlRrqMQALgiA
mzsVHgWvQKnAhKPDRPIHPFKPrDSlWVkkznPttLgPiWdGlHTViLsiPTvVKVaGIvpWIH
psSqLKPAaQDkwTSqqDLdhaTQLILRGTkasemttallwslrkltsphtaeazgdnspalvtqkltslctaeazvii
revnvvrnlkssnfpchinyfaiklslcl

333333333333333333333333333333333333332222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
```

```
111111222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222333333333333333322222222
222222222222233333333333333333333333333333333333333333333333333333333333333
332222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222222222222222222222222222222222222222222222
222222222222222222222222222222222221111111122222222222222222222222222222222
222222222222222222222222222222222222222222222332222222222222222222222222222
22222233333333333333333333333333333333333333333333
          !        > !
<
 63854869           63854779           63854687           63854597
63854507           63854417           63854327           63854237
63854147           63854057           63853967           63853877
63853787           63853697           63853607           63853515
63853425           63853335           63853244           63853154
63853064           63852974           63852884           63852794
63852704           63852611           63852520           63852430
63852338           63852248           63852158           63852068
63851978           63851888           63851795           63851705
63851615           63851522           63851431           63851341
::
Hits::
1 63854501 58 RT1 P C #wNtPllpVKK MLV
2 63854330 87 RT2 P C #tvlDLkDaFFclrL MLV
3 63854225 85 RT3 P C #qltWtrLPQGfknSP MLV
4 63854111 61 RT4 P C #lqYvDDLlla MLV
5 63854018 59 RT5 P C #GyrasakKaQ MLV
6 63852215 14 IN5 P C #gtDNgpaF MuLV, motif D
7 63852131 45 IN6 P C #AYqPQSsgKVERmnrt HERV-E
8 63851738 15 IN7 P C #eprWkGPyiVLlttpt BAEV
::
{ Created by ORFID with parameters
{ NonAlignedScore: 0.4
{ GlycosylationFactor: 0.2
{ StopCodonFactor: -0.4
{ LastDNAEnd: 63851325
{ NonORFHexamerFactor: -0.1
{ MotifHitFactor: 0.01
{ DNAFile: chr7_639.txt
{ StopCodonValue: -20
{ ScriptName: C:\rt010\Workplace\chr7\chr7_639\orfid_s00001cpol001Script_.txt
{ OutputFile: ORFIDout_S00001CPol_001.txt
{ InputFile:
{ MinHitScore: 15
{ Debugging: No
{ Database: Ordinary
{ Strand: Secondary
{ FrameGene: Pol
{ PuteinFile: Putein_S00001CPol_001.txt
```

{ ORFHexamerFactor: 0.2
{ ChainNumber: 00001
{ VirusGroup: C
{ FirstDNAStart: 63854869
{ FirstDNAEnd: 63851677
{ LastDNAStart: 63854602
{ MinAverageScore: 20
{ FrameShiftFactor: -0.9
{ ScoreFactor: 0.1
{ Execution time was 2453 milliseconds
{ Belongs in C:\rt010\Workplace\chr7\chr7_639
{ Created Sun May 11 18:13:07 GMT+01:00 2003 under RetroTector version 0.10 03 05 10
{ using Database:Ordinary last modified Fri Dec 20 13:49:56 GMT+01:00 2002

Env defined by ORFID 6910 [63851292]- 8721 [63849481]:

MLGMNMLLITLFLLLPLSMLKGEPWEGCLHCTHTTWSGNIMTKTLLYHTYYECAGT
CLGTCTHNQTTYSVCDPGRGQPYVCYDPKSSPGTWFEIHVGSKEGDLLNQTKVFPSG
KDVVSLYFDVCQIVSMGSLFPVIFSSMEYYSSCHKNRYAHPACSTDSPVTTCWDCTT
WSTNQQSLGPIMLTKIPLEPDCKTSTCNSVNLTILEPDQPIWTTGLKaPLGARVSGEEIG
PGAYVYLYIIKKTRTRSTQQfRVFESFYEHVNQKLPEPPPLASNLFAQLAENIASSLHV
ASCYVCGGMNMGDQWPWEARELMPQDNFTLTASSLEPAPSSQSIWFLKTSIIGKFCI
ARWGKAFTDPVGELTCLGQQYYNETLGKTLWRGKSNNSESPHPSPFSRFPSLNHSWY
QLEAPNTWQAPSGLYWICGPQAYRQLPAKWSGACVLGTIRPSFFLMPLKQGEALGYP
IYDETKRKSKRGITIGDWKDNEWPPERIIQYYGPATWAEDGMWGYRTPVYMLNRIIR
LQAVLEIITNETAGALNLLAQQATKMRNVIYQNRLALDYLLAQEEGVCGKFNLTNCC
LELDDEGKVIKEITAKIQKLAHIPVQTWKG

env DNA:

atgctgggtatgaacatgctactcatcactttgttcttgctactccccttatccatgttaaaaggagaaccctgggagggatgcctccactg
cacccacactacgtggtcggggaacatcatgactaaaaccctgttgtatcacacttattatgagtgtgctgggacctgcctaggaacttgt
actcacaaccagacaacctactcagtctgtgacccaggaaggggccagccttatgtgtgttatgaccctaagtcttcacctgggacctg
gtttgaaattcatgtcgggtcaaaggaaggggatcttctaaaccaaaccaaggtatttccctctggcaaggatgtcgtatccttatactttg
atgtttgccagatagtatccatgggctcactctttcccgtaatcttcagttccatggagtactatagtagctgccataaaaataggtatgcac
accctgcttgttccaccgattccccagtaacaacttgctgggactgcacaacgtggtccactaaccaacaatcactagggccaattatgc
ttaccaaaataccattagaaccagattgtaaaacaagcacttgcaattctgtaaatcttaccatcttagagccagatcagcccatatggaca
acaggtttaaaagcaccgctaggggcacgagtcagcggtgaagaaattggcccaggagcctatgtctatctatatatcataaagaaaac
tcggacccgctcaacccaacagttccgagtttttgagtcattctatgagcatgttaaccagaaattgcctgagccccctcccttggccagt
aatttattcgcccaactggctgaaaacatagccagcagcctgcacgttgcttcatgttatgtctgtgggggaatgaacatgggagaccaa
tggccatgggaagcaagggaactaatgccccaagataatttcacactaaccgcctcttccctcgaacctgcaccatcaagtcagagcat
ctggttcttaaaaacctccattattggaaaattctgtattgctcgctggggaaaggcctttacagacccagtaggagagttaacttgcctag
gacaacaatattacaacgagacactaggaaagactttatggaggggcaaaagcaataattctgaatcaccacacccaagcccattctct
cgtttcccatctttaaaccattcttggtaccaacttgaagctccaaatacctggcaggcaccctctggcctctactggatctgtgggccaca
agcatatcgacaactgccagctaaatggtcaggggcctgtgtactggggacaattaggccgtccttcttcctaatgcccctaaaacagg
gagaagccttaggataccccatctatgatgaaactaaaaggaaaagcaaaagaggcataactataggagattggaaggacaatgaat
ggcctcctgaaagaataattcaatattatggcccagccacctgggcagaagatggaatgtggggataccgcaccccagtttacatgctt
aaccgcattataagattgcaggcagtactagaaatcattaccaatgaaactgcaggggccttgaatctgcttgcccagcaagccacaaa
aatgagaaatgtcatttatcaaaatagactggccttagactacctcctagcccaggaagagggagtatgcggaaagttcaaccttactaa

ctgctgcctggaacttgatgacgaaggaaaggtcatcaaagaaataactgctaaaatccaaaagttagctcacatcccagttcagacttg
gaaagga

Detailed ORFID Env result:

```
Executor: Puteinview
NumberOfHits: 3
LeadingInfo::

Leading path in Env
Starts at=63851292
Putein string of length=469 and total score=468.1137
MLGMNMLLITLFLLLPLSMLKGEPWEGCLHCTHTTWSGNIMTKTLLYHTYYECAGT
CLGTCTHNQTTYSVCDPGRGQPYVCYDPKSSPGTWFEIHVGSKEGDLLNQTKVFPSG
KDVVSLYFDVCQIVSMGSLFPVIFSSMEYYSSCHKNRYAHPACSTDSPVTTCWDCTT
WSTNQQSLGPIMLTKIPLEPDCKTSTCNSVNLTILEPDQPIWTTGLKaPLGARVSGEEIG
PGAYVYLYIIKKTRTRSTQQFRVFESFYEHVNQKLPEPPPLASNLFAQLAENIASSLHV
ASCYVCGGMNMGDQWPWEARELMPQDNFTLTASSLEPAPSSQSIWFLKTSIIGKFCI
ARWGKAFTDPVGELTCLGQQYYNETLGKTLWRGKSNNSESPHPSPFSRFPSLNHSWY
QLEAPNTWQAPSGLYWICGPQAYRQLPAKWSGACVLGTIRPSFFLMPLKQGEALGYP
IYDETKRKS
yielding average=0.99598664
Inside limits=1.0
For fit length (89-2)*0.5
Kozak score=0.53125*1
For SpliceAcceptorMotif hit at 63851311 0.85714287*1.5
von Heijne score with WEGCLHCTHTTWSGN =0.025809368*0.5
Comparing score=47.325855
::
TrailingInfo::

Trailing path in Env
Ends at=63849480
Putein string of length=57 and total score=57.0
YQNRLALDYLLAQEEGVCGKFNLTNCCLELDDEGKVIKEITAKIQKLAHIPVQTWKG
yielding average=0.98275864
Not inside limits=0
For fit length (35-2)*0.5
For ending at stop codon=1.0*0.5
Comparing score=17.98276
::
EstimatedStartPosition: 63851292
EstimatedLastPosition: 63849481
LengthInside: 604
LengthTotal: 1026
AlignedAcids: 602
AverageScoreInside: 0.996877
AverageScoreTotal: 0.5567527
MostUsedRow: 3 (ERV3SU+TM)
StopCodonsInside: 0(0, 39, 38)
```

```
StopCodonsTotal: 7
AmbiguousAcidsInside: 0
AmbiguousAcidsTotal: 0
ShiftsInside: 0
ShiftsTotal: 6
{ Starts at position 63851677, ends at 63848587
Putein::

1_____________                                             2___________
3___________

givpwihpssqlkpaaqdkwtsqqdldhatqlilrwnqgasemttallwslrkltsphtaeazgdnspalvtqkltslctaeazviire
vnvvrnlkssnfpchinyfaiklslclafspqekafsvMLGMNMLLITLFLLLPLSMLKGEPWEGCLHCT
HTTWSGNIMTKTLLYHTYYECAGTCLGTCTHNQTTYSVCDPGRGQPYVCYDPKSSP
GTWFEIHVGSKEGDLLNQTKVFPSGKDVVSLYFDVCQIVSMGSLFPVIFSSMEYYSSC
HKNRYAHPACSTDSPVTTCWDCTTWSTNQQSLGPIMLTKIPLEPDCKTSTCNSVNLTI
LEPDQPIWTTGLKaPLGARVSGEEIGPGAYVYLYIIKKTRTRSTQQfRVFESFYEHVNQ
KLPEPPPLASNLFAQLAENIASSLHVASCYVCGGMNMGDQWPWEARELMPQDNFTL
TASSLEPAPSSQSIWFLKTSIIGKFCIARWGKAFTDPVGELTCLGQQYYNETLGKTLW
RGKSNNSESPHPSPFSRFPSLNHSWYQLEAPNTWQAPSGLYWICGPQAYRQLPAKWS
GACVLGTIRPSFFLMPLKQGEALGYPIYDETKRKSKRGITIGDWKDNEWPPERIIQYY
GPATWAEDGMWGYRTPVYMLNRIIRLQAVLEIITNETAGALNLLAQQATKMRNVIY
QNRLALDYLLAQEEGVCGKFNLTNCCLELDDEGKVIKEITAKIQKLAHIPVQTWKGzs
pdslfrgwfsslggfktlvqivlailgvclilpcllplivkniqtaiealvdrrtttrlmaltkyzplprkellplevnedsdafsztlliksik
ggneagnikgktskgktsfpdqsdslqspagamiiickaagapkewapgagmrktsspyqfpclkfsphniilccshnyfcnyfc
ksvkilzvlvflsvawqghktclskvgsccksccktchgminclcsasvrllshlagfapktrlappdacikvkpclcsglslwmlir
wasghlnktfllhpvispaszfpqh

222222222222222222222222222222222222222222222222222222222222222222222222333
333333333333333333333333333333333333333333333333333311111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111111111111111111111111111111111111111111111111111111111111
111111111111111111113333333333333333333333222222222222222222222222222222222
221111111111111111111111111111111111111111111111111111111111111111111111111
111122222222222222222222222222222222222222222222222
                                                               >
!                                           !     !
!
<
 63851675          63851585          63851495          63851404
63851314          63851223          63851133          63851043
63850953          63850863          63850773          63850683
63850593          63850503          63850413          63850323
```

```
63850233        63850143        63850053        63849963
63849873        63849783        63849693        63849603
63849513        63849423        63849333        63849243
63849153        63849059        63848969        63848877
63848787        63848696        63848606
::
Hits::
1 63849885 82 SU3 P C #KRgiviGnWkdnew HERV-E
2 63849651 100 TM3 P C #yQNRLALDYLLA HERV E/ERV3
3 63849414 37 TM5 HYF ABCDELSGO #X___________ hydrophobic motif
::
{ Created by ORFID with parameters
{ NonAlignedScore: 0.4
{ GlycosylationFactor: 0.2
{ StopCodonFactor: -0.4
{ LastDNAEnd: 63848586
{ NonORFHexamerFactor: -0.1
{ MotifHitFactor: 0.01
{ DNAFile: chr7_639.txt
{ StopCodonValue: -20
{ ScriptName: C:\rt010\Workplace\chr7\chr7_639\orfid_s00001cenv001Script_.txt
{ OutputFile: ORFIDout_S00001CEnv_001.txt
{ InputFile:
{ MinHitScore: 15
{ Debugging: No
{ Database: Ordinary
{ Strand: Secondary
{ FrameGene: Env
{ PuteinFile: Putein_S00001CEnv_001.txt
{ ORFHexamerFactor: 0.2
{ ChainNumber: 00001
{ VirusGroup: C
{ FirstDNAStart: 63851677
{ FirstDNAEnd: 63849351
{ LastDNAStart: 63850501
{ MinAverageScore: 20
{ FrameShiftFactor: -0.9
{ ScoreFactor: 0.1
{ Execution time was 3657 milliseconds
{ Belongs in C:\rt010\Workplace\chr7\chr7_639
{ Created Sun May 11 18:13:11 GMT+01:00 2003 under RetroTector version 0.10 03 05 10
{ using Database:Ordinary last modified Fri Dec 20 13:49:56 GMT+01:00 2002
```

ERV3 XORF predicted by XONID at 6637 [63851565] - 6909 [63851293]:

The only suitably positioned methionine for a putative XORF protein is at 6623 [63851579] in frame 1. It requires a shift to frame 3 approximately at 6738 [63851464]. Alternatively, the putative XORF protein could have an internal ribosomal entry start (IRES) after this frameshift. Using the mentioned methionine, the reconstructed putative ERV3 XORF protein (134 aa) becomes

MEPRCQZDDNSPALVTPEADQSTHGZSLRRQQPCSSPQKLTSLCTAEAZVIIREVNVVRNLKSSNFPCHINYFAIKLSLCLAFSPQEKAFSVHAGYEHATHHFVLATPLIHVKRRTLGGMPPLHPHYVVGEHHD.

The underlined portion is from frame 1. The rest is from frame 3. It contains three stops (shown as “Z”). Therefore, unless there is an internal ribosomal entry sequence (IRES) past the last stop codon, the putative XORF protein should not be active in ERV3 7q11. Its C-terminal 20 amino acids have a moderate similarity to a histidine-rich domain positioned close to the carboxy terminus of the latrophilin G-protein-coupled seven-transmembrane brain-specific toxin receptors (see supplementary data). The XORF sequence is present in several ESTs >98% identical to ERV3 7q11 (Fig 1). An XORF reminiscent of the one in ERV3 7q11 was seen in 10 (chr 6, Y, Y, Y, Y, 11, 19, 7, 14 and 14) of the 41 ERV3like sequences defined here. The consensus XORF for them became

GACAACAGCCCTGCTCTGGTCACTCCAGAAGCTGACCAGTCTACGCATCGCTGAAGCTTGAGGAGACAACAAGCCATGCTCTAGTCACCCCAGAAGCTTACTAGTCTACGCACGGCCGAAGCTTGAGGAAGTCAACATCAGATAAGTAAATGTGGATAGAAATCTTAAGTCAAGTAATTTTCTTTGTAATATTAATTGTTTTACTATTGTTCTGTCATTTTCCTCAATCTCCTCCACCGGGTAAAAACCTCTTTTGTCCTTGCTGGATATAAACATGCTACTCTTTACTTTGTTCTTACTACTCCCCATA This is translatable to QQPCSGHSRSZPVYASLKLEETTSHALVTPEAYZSTHGRSLRKSTSDKZMWIEILSQVIFFVILIVLLLFCHFPQSPPPGKNLFCPCWIZTCYSLLCSYYSP or TTALLWSLQKLTSLRIAEAZGDNKPCSSHPRSLLVYARPKLEEVNIRZVNVDRNLKSSNFLCNINCFTIVLSFSSISSTGZKPLLSLLDINMLLFTLFLLLPI.

The latter contains a similarity to the C terminus of the Pol protein suggested by Cohen, O´Connel and Kato (Virology 1987) AA88026 (M12140.1):

1 gyspyerlfg kpspiisqik gnlrelgelt lrrqmqalgi amxsvhgwvq ermpislidp 61 ihpfkprdsl wvkkxnpttl gpiwdglhtv ilsiptvvkv agivpwihps sqlkpaaqdk 121 wtsqqdldha tqlilrwnqg asemttallw slrkltsprt aea, where the overlapping similarity is underlined.

A search for similar nucleic acids and proteins, other than human and chimpanzee genomic ones, in GenBank, did not reveal related nonretroviral sequences. Thus, the review of the ten ERV3 7q11 related XORFs did neither reveal a consensus with all open reading frame, nor a protein of obvious function.

JB 040510

Identification of the loci of 41 ERV3-like sequences

RetroTector-defined chains were located in the hg15 (April 2003) human genome using Genome Browser at the UCSC web site. EST data were as shown there. It is not a complete EST or cDNA list.

List of 23 tree defined ERV3:

Chr7 63858202 1754 po CR	7q11.21	EST: M55422, AK074464,	In Env tree&DM
Chr6 110218096 843 po CX	|Tandem HERV3 on 6q21.	No ESTs	
Chr6 110214535 786 po CX	|		
ChrY 3018841 881 po CR	Yp11.2 LTR76+LTR76int	No ESTs	
ChrX 86876098 639 po CX	Xq21.31 LTR76+LTR76int	No ESTs (similar to previous!)	
Chr14 68899917 965 po CX	14q24.2 LTR4+HERV3	No ESTs	In Env tree&DM
ChrY 23874437 661 po CX	Yq11.223 HERV3	No ESTs	
ChrY 27518418 649 po CX	Yq11.23 HERV3	No ESTs	
ChrY 25016892 657 po CX	Yq11.223 HERV3	No ESTs	
ChrY 19932937 520 po CX	Yq11.222 LTR4+HERV3	No ESTs (contains MER11A)	In Env tree&DM
ChrY 20104614 518 po CX	Yq11.222 LTR4+HERV3	No ESTs	In Env tree&DM
Chr19 58660155 598 po CX	19q13.42 HERV3	No ESTs	In Env tree&DM
Chr19 58621153 500 po CX	19q13.42 HERV3	No ESTs (two tandem HERV3)	In Env tree&DM
Chr7 139686987 710 po CX	7q34 HERV3	No ESTs	
Chr7 137651301 1097 po CP	7q34 HERV3	No ESTs	
Chr11 29579197 930 po CP	11p14.1 HERV3	No ESTs	In Env tree&DM
Chr4 117329205 488 po CX	4q26 HERV3+LTR76	No ESTs	
Chr21 38522185 383 po CX	21q22 HERV3+LTR76	CA432258 enchondroma +AceView Feebaw?	
Chr4 70173370 1049 po CX	4q13.2 Harlequin+LTR2B		
Chr4 70022513 788 po CK	4q13.2 HERV3+LTR76	BX095654 fetal liver skawkaw	
Chr14 105225894 1027 po CX	14q32.33 HERV3+LTR76	No ESTs	In Env tree&DM
Chr7 119241536 761 po CX	7q31.31 HERV3+LTR4	No ESTs	In Env tree&DM
ChrY 22564970 567 po CX	Yq11.223 HERV3+LTR15	AK097136 normal spleen rohiru	

List of 18 tree defined RRHERVI-HERV15:

Chr6 64985396 1041 po CX	6q12 HERV3+LTR30	No ESTs	In Env tree&DM
ChrY 20355209 659 po CX	Yq11.222	BC022384 skel muscle BI020800 adult marrow Aceview sleydaw	In Env tree&DM
Chr3 34567166 523 po CX	3p23 HERV3+LTR30	No ESTs	
Chr3 122072468 245 po CX	3q13.33 HERV3+LTR30	No ESTs	
Chr11 58788475 541 po CX	11q12.1 HERV3+LTR30	No ESTs	In Env tree&DM
Chr6 137421042 373 po CX	6q23.3 HERV3,HERV15+LTR30	No ESTs	In Env tree&DM
Chr4 115910608 550 po CP	4q26 HERV15+LTR30	No ESTs	In Env tree&DM
Chr21 14409346 1477 po CP	21q11.2 HERV15+LTR15	No ESTs	In Env tree&DM
ChrX 51259700 1396 po CP	Xp11.22 HERV15+LTR15 cont MER11C	No ESTs	In Env tree&DM
ChrY 13743185 977 po CI	Yq11.21 HERV15+LTR15	No ESTs	
ChrX 117198092 983 po CX	Xq24 HERV15 + LTR15	No ESTs	In Env tree&DM
Chr19 21166030 1412 po CX	19p13.11 HERV15 + LTR15	No ESTs	In Env tree&DM
Chr19 21213743 913 po CX	19p13.11 HERV15 + LTR15	No ESTs	
Chr19 15957774 569 po CX	19p13.12 HERV15 Contains LTR5/61	No ESTs	
Chr11 77508681 297 po CX	11q13.5 HERV3+LTR61	AK093267 spliced env testis AK096213 spliced env adult ovary	
Chr13 98250769 738 po CX	13q32.3 HERV3+LTR61 Cont:s LTR5(B)	No ESTs	
Chr6 39968183 513 po CP	6p21.2 HERV3+LTR15	No ESTs	
Chr15 66867263 645 po CX	15q23 HERV3+LTR76	No ESTs	

(Total of 41 ERV3 like)

The RRHERVI element at ChrY 14533115 (AF290423, "HERV15Yq2") was not identified in hg15. It may thus be polymorphic.

```
CLUSTAL W (1.81) multiple sequence alignment


HERVI_polputein                     --------------------------------------------------
HERVIP10F_polputein                 --------------------------------------------------
HERVadp_AC005741                    --------------------------------------------------
HERV-PT47D                          --------------------------------------------------
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      --------------------------------------------------
HERV_Y_chr12_51022911_pol           --------------------------------------------------
HERV17_polputein                    --------------------------------------------------
HERV_W_chr7_9105739_syncytin_p      --------------------------------------------------
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  --------------------------------------------------
ERVfrd_AC004022                     --------------------------------------------------
PRIMA41_polputein                   --------------------------------------------------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 --------------------------------------------------
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           --------------------------------------------------
ChrY_13790029_693_po_C_Xz12/11      --------------------------------------------------
HERVS71_polputein                   --------------------------------------------------
BaEV_M7_pol                         --------------------------------------------------
HERVR_polputein                     --------------------------------------------------
AKR_MLV_MLOCG_RT                    --------------------------------------------------
GaLV_pol                            --------------------------------------------------
Chr3_158587414_1662_po_C_Ez10/      --------------------------------------------------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      --------------------------------------------------
Chr5_111890459_1626_po_C_Ez10/      --------------------------------------------------
ChrX_48766518_1600_po_C_Ez6/3       --------------------------------------------------
ChrX_152580713_1572_po_C_Ez10/      --------------------------------------------------
ChrX_139287533_1154_po_C_Xz8/4      --------------------------------------------------
Chr2_52615786_912_po_C_Xz11/5       --------------------------------------------------
HERV-E_4_1                          --------------------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      --------------------------------------------------
ChrY_8634032_950_po_C_Xz14/10       --------------------------------------------------
Chr3_126735360_997_po_C_Xz5/10      --------------------------------------------------
Chr3_126713203_958_po_C_Xz7/8       --------------------------------------------------
Chr3_75226118_977_po_C_Xz8/9        --------------------------------------------------
Chr7_6629919_827_po_C_Xz8/7         --------------------------------------------------
Chr4_9166812_811_po_C_Xz8/10        --------------------------------------------------
Chr12_8478642_834_po_C_Xz12/7       --------------------------------------------------
Chr10_15187175_787_po_C_Xz11/5      --------------------------------------------------
Chr11_71819740_854_po_C_Xz11/9      --------------------------------------------------
Chr11_71519974_886_po_C_Xz13/9      --------------------------------------------------
Chr8_7271441_969_po_C_Xz10/6        --------------------------------------------------
Chr8_11770066_826_po_C_Xz9/8        --------------------------------------------------
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      --------------------------------------------------
Chr14_50205291_837_po_C_Xz10/8      --------------------------------------------------
Chr7_6597120_842_po_C_Xz7/7         --------------------------------------------------
Chr11_71544658_783_po_C_Xz8/13      --------------------------------------------------
Chr11_67999520_760_po_C_Xz8/8       --------------------------------------------------
Chr3_75461975_796_po_C_Xz7/9        --------------------------------------------------
Chr12_52228939_692_po_C_Xz3/10      --------------------------------------------------
Chr10_15211234_286_po_C_Xz11/1      --------------------------------------------------
Chr12_100973302_945_po_C_Xz12/      --------------------------------------------------
Chr19_9685420_817_po_C_Xz6/10       --------------------------------------------------
Chr13_40936820_719_po_C_Xz9/10      --------------------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      --------------------------------------------------
ERV3_chr7q21_AC073210_polputei      --------------------------------------------------
Chr6_110218096_843_po_C_Xz16/8      --------------------------------------------------
Chr6_110214535_786_po_C_Xz16/8      --------------------------------------------------
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        --------------------------------------------------
Chr14_68899917_965_po_C_Xz6/8       --------------------------------------------------
ChrY_23874437_661_po_C_Xz15/9       --------------------------------------------------
ChrY_27518418_649_po_C_Xz15/9       --------------------------------------------------
ChrY_25016892_657_po_C_Xz13/12      --------------------------------------------------
ChrY_19932937_520_po_C_Xz10/14      --------------------------------------------------
ChrY_20104614_518_po_C_Xz10/14      --------------------------------------------------
Chr19_58660155_598_po_C_Xz21/1      --------------------------------------------------
Chr19_58621153_500_po_C_Xz21/1      --------------------------------------------------
```

```
Chr7_139686987_710_po_C_Xz11/7      --------------------------------------------------
Chr7_137651301_1097_po_C_Pz13/      --------------------------------------------------
Chr11_29579197_930_po_C_Pz8/12      --------------------------------------------------
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      --------------------------------------------------
Chr4_70022513_788_po_C_Kz10/12      --------------------------------------------------
Chr14_105225894_1027_po_C_Xz13      --------------------------------------------------
Chr7_119241536_761_po_C_Xz11/1      --------------------------------------------------
ChrY_22564970_567_po_C_Xz9/10       --------------------------------------------------
Chr21_14409346_1477_po_C_Pz5/6      --------------------------------------------------
ChrX_51259700_1396_po_C_Pz7/8       --------------------------------------------------
Chr6_64985396_1041_po_C_Xz11/1      --------------------------------------------------
ChrY_20355209_659_po_C_Xz7/11       --------------------------------------------------
Chr3_34567166_523_po_C_Xz8/6        --------------------------------------------------
Chr3_122072468_245_po_C_Xz9/9       --------------------------------------------------
Chr11_58788475_541_po_C_Xz8/22      ANVTGIGTLLAKQCLNLSQQDPVWWMQSESESKACHEGGNGRKSIKAYST
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------------------
ChrY_13743185_977_po_C_Xz14/8       --------------------------------------------------
HERV15                              --------------------------------------------------
ChrX_117198092_983_po_C_Xz8/8       --------------------------------------------------
Chr19_21166030_1412_po_C_Xz9/1      --------------------------------------------------
Chr19_21213743_913_po_C_Xz14/1      --------------------------------------------------
Chr19_15957774_569_po_C_Xz8/10      --------------------------------------------------
Chr11_77508681_297_po_C_Xz19/1      --------------------------------------------------
Chr13_98250769_738_po_C_Xz7/13      --------------------------------------------------
Chr6_39968183_513_po_C_Pz10/9       --------------------------------------------------
Chr15_66867263_645_po_C_Xz11/1      --------------------------------------------------
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                --------------------------------------------------
MER66_polputein                     --------------------------------------------------


HERVI_polputein                     --------------------------------------------------
HERVIP10F_polputein                 --------------------------------------------------
HERVadp_AC005741                    --------------------------------------------------
HERV-PT47D                          -------------------------AQLTFSSCZSLKQKLIYWGGLMLKL
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      --------------------------------------------------
HERV_Y_chr12_51022911_pol           --------------------------------------------------
HERV17_polputein                    --------------------------------------------------
HERV_W_chr7_9105739_syncytin_p      --------------------------------------------------
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  --------------------------------------------------
ERVfrd_AC004022                     --------------------------------------------------
PRIMA41_polputein                   --------------------------------------------------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 --------------------------------------------------
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           --------------------------------------------------
ChrY_13790029_693_po_C_Xz12/11      --------------------------------------------------
HERVS71_polputein                   --------------------------------------------------
BaEV_M7_pol                         --------------------------------------------------
HERVR_polputein                     --------------------------------------------------
AKR_MLV_MLOCG_RT                    ----------------------------AQVMGPMGQPLQVLTLNIEDEH
GaLV_pol                            --------------------------------------------------
Chr3_158587414_1662_po_C_Ez10/      --------------------------------AKVTGNRSIMALTVPREE
Chr3_79627623_1138_po_C_Ez4/7       ----------------------------------------------PWEE
Chr19_20477244_1350_po_C_Ez10/      ----------------------ISFTKHGSLQLKLLGTGVIIVLTVPWEE
Chr5_111890459_1626_po_C_Ez10/      --------------------------------QKLPRTGVIMALTVPQEE
ChrX_48766518_1600_po_C_Ez6/3       -----------------------------------PETGVIMALMVTREK
ChrX_152580713_1572_po_C_Ez10/      --------------------------------------------------
ChrX_139287533_1154_po_C_Xz8/4      -----------------------------------PGTGVIMALIVPQEE
Chr2_52615786_912_po_C_Xz11/5       ------------------------------------------ALTVPQEE
HERV-E_4_1                          --------------------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      -----------------------------------PGMGVIVALSLPZ-E
ChrY_8634032_950_po_C_Xz14/10       --------------------------------------------------
Chr3_126735360_997_po_C_Xz5/10      -------------------------TDHGSLLLKLPGTGVIMTLMVPGDE
Chr3_126713203_958_po_C_Xz7/8       --------------------------------AKVTGTGVIMTLMVPGDE
Chr3_75226118_977_po_C_Xz8/9        -----------------------------------PGTGVIMTLMVPZEE
Chr7_6629919_827_po_C_Xz8/7         ---------------------------------------VIMTLMVPPEE
Chr4_9166812_811_po_C_Xz8/10        ------------------------------------ATGIIMTLMVPGEE
Chr12_8478642_834_po_C_Xz12/7       -----------------------------------PRTGVIMTLIVPREE
```

```
Chr10_15187175_787_po_C_Xz11/5      -----------------------------------PRTGVIMTLMVPQEE
Chr11_71819740_854_po_C_Xz11/9      -----------------------------------PGTGVIMTLVVPREE
Chr11_71519974_886_po_C_Xz13/9      -----------------------------------PRTGVIMTLMVPREE
Chr8_7271441_969_po_C_Xz10/6        --------------------------------AKVTGMGVIMTLMVPREE
Chr8_11770066_826_po_C_Xz9/8        -----------------------------------PGTGVIMTIIVPREL
Chr4_4128362_651_po_C_Xz7/11        ------------------------------------ATGVIMTLTVPREE
Chr14_50203955_850_po_C_Xz10/8      ------------------------------------ATGVIMTLTVPREE
Chr14_50205291_837_po_C_Xz10/8      ------------------------------------ATGVIMTLTVPREE
Chr7_6597120_842_po_C_Xz7/7         --------------------------------------------AGPRED
Chr11_71544658_783_po_C_Xz8/13      ------------------------------------ATGVIMTLTVPREE
Chr11_67999520_760_po_C_Xz8/8       ------------------------------------ATGVIMTLWSPERR
Chr3_75461975_796_po_C_Xz7/9        --------------------------------------------------
Chr12_52228939_692_po_C_Xz3/10      --------------------------------------------------
Chr10_15211234_286_po_C_Xz11/1      --------------------------------------------------
Chr12_100973302_945_po_C_Xz12/      ---------------------------QGSLQLKLPGTGVIMALTVPWEE
Chr19_9685420_817_po_C_Xz6/10       --------------------------------------------------
Chr13_40936820_719_po_C_Xz9/10      -----------------------------------PGTGVIMTLMVPREE
Chr7_63858202_1754_po_C_Rz12/9      --------------------------------------------------
ERV3_chr7q21_AC073210_polputei      --------------------------------------PQAEEWRLYMKY
Chr6_110218096_843_po_C_Xz16/8      --------------------------------------------------
Chr6_110214535_786_po_C_Xz16/8      --------------------------------------------------
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        --------------------------------------------------
Chr14_68899917_965_po_C_Xz6/8       --------------------------------------------------
ChrY_23874437_661_po_C_Xz15/9       --------------------------------------------------
ChrY_27518418_649_po_C_Xz15/9       --------------------------------------------------
ChrY_25016892_657_po_C_Xz13/12      --------------------------------------------------
ChrY_19932937_520_po_C_Xz10/14      --------------------------------------------------
ChrY_20104614_518_po_C_Xz10/14      --------------------------------------------------
Chr19_58660155_598_po_C_Xz21/1      --------------------------------------------------
Chr19_58621153_500_po_C_Xz21/1      --------------------------------------------PTPLFQ
Chr7_139686987_710_po_C_Xz11/7      -----------------------------MVLTFTVLQAEEWRLYKKRFQ
Chr7_137651301_1097_po_C_Pz13/      -------------------------NLTDPKAMVLTLTVLQTEEWRLYTK
Chr11_29579197_930_po_C_Pz8/12      ------------------PQVAMTLNLTLPKAMVIILTIPQTDKWRLYTK
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       ---------------------------TPLQCMLKNFKKSYAECCG----
Chr4_70173370_1049_po_C_Xz11/8      ------------------------------------------------AK
Chr4_70022513_788_po_C_Kz10/12      ------------------------------------------------AK
Chr14_105225894_1027_po_C_Xz13      ------------AQITFGPQGDVTLNLAHSEAMVLTLTVPRAEEWRLYAG
Chr7_119241536_761_po_C_Xz11/1      ------------------PQGDMTLKLTYPKAMVLTLTIPQTEEWRLYTK
ChrY_22564970_567_po_C_Xz9/10       -------------------------------------------------A
Chr21_14409346_1477_po_C_Pz5/6      ------------------------QRKVMVLTLTIPKTDEWRLHESSCQE
ChrX_51259700_1396_po_C_Pz7/8       ----------------------PRTERAMVLTLTIQKTEEWRLYESNCQE
Chr6_64985396_1041_po_C_Xz11/1      ---------QISFTLEGDMTLNLGQRKAIIMTLTVPTTEEWRLYER-CKI
ChrY_20355209_659_po_C_Xz7/11       --------------------------------------------------
Chr3_34567166_523_po_C_Xz8/6        --------------------------------------------------
Chr3_122072468_245_po_C_Xz9/9       --------------------------------------------------
Chr11_58788475_541_po_C_Xz8/22      GVHAEKLLKGFNGDYGVKLTPQKLRTLCEIDWPSFNVGWPHEELES-PPP
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------------------
ChrY_13743185_977_po_C_Xz14/8       --------------------------------------------------
HERV15                              --------------------------------------------------
ChrX_117198092_983_po_C_Xz8/8       --------------------------------------------------
Chr19_21166030_1412_po_C_Xz9/1      ----------------------NMTLEVGKPKAMVLTLALLIAEEWRLYE
Chr19_21213743_913_po_C_Xz14/1      --------------------------------------------------
Chr19_15957774_569_po_C_Xz8/10      ----------------------NLTHPKAIVLTLTIPQAEEWRLHRKKLQ
Chr11_77508681_297_po_C_Xz19/1      --------------------------------------------------
Chr13_98250769_738_po_C_Xz7/13      --------------------------------------------------
Chr6_39968183_513_po_C_Pz10/9       ------------------------------------------PRSGGYMQ
Chr15_66867263_645_po_C_Xz11/1      -----------------AQITYGPQGNVTLNLAHLETMVLTLTVPQAEEW
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                --------------------------------------------------
MER66_polputein                     --------------------------------------------------


HERVI_polputein                     --------------------------------GRKLRKASNSP-----NP
HERVIP10F_polputein                 --------------------------------------------------
HERVadp_AC005741                    --------------------------NPLVWLREGNRGGLTIT-----PL
HERV-PT47D                          GIGLQVSPRGFLTSLNLLTTADEKYINPSVWSKEGNZGKLRVP-----PI
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         -----------------NMGNKLPICCPLLEEGINPEVWALEGQFGRAKN
Chr13_54564403_572_po_C_Yz7/14      -----------------NIGEGTPVCCPLLEEELILKSRQQKDNIDEQR-
HERV_Y_chr12_51022911_pol           --------------------------------------------------
HERV17_polputein                    --------------------------------------------------
HERV_W_chr7_9105739_syncytin_p      -----------------NIGEGTPVCCPLLEEGINPEVWATEGQYGRAKN
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  --------------------------------------------------
```

```
ERVfrd_AC004022                     --------------------------------------------------
PRIMA41_polputein                   ---------------GRKVNRDLSSLPIEVTSQINPVVWDIEVP-DKALN
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 --------PTLKHATFPYPPS------------VVNPAVZDTSTPSVAEH
HERV_H_RTVLH2_pol                   ---ASLTIPRLZHTHCHLFPSSKPPSHPPLVSPLLNPQVZDTSIPSLVTD
HERVFH21_pol                        ------ASPAPDPSPQHPLPTS-----------LVNPVVWDTTTPSIAAH
HERVRblike_chr4_109047953           --PPEHGLQLQALLVSTEAPHPEVEAIPQEVFDKVKPEVWAZDQPEREIN
ChrY_13790029_693_po_C_Xz12/11      ----PPENKTNPLLLELQTLFP------ZVWAEPIPPGRAKHH----PPV
HERVS71_polputein                   -----------------------------------------------LPV
BaEV_M7_pol                         --------------------------------------------------
HERVR_polputein                     --------------------------------------------------
AKR_MLV_MLOCG_RT                    RLHETSKEPDVSLGSTWLSDFP------QAWAETGGMGLAVRQ----APL
GaLV_pol                            --------------------------------------------------
Chr3_158587414_1662_po_C_Ez10/      EWRLFLIZPGRETKPALAKZWP------RVWTEDNPPGLVVNQ----APV
Chr3_79627623_1138_po_C_Ez4/7       EWRLFLTEPGQEIKPALAKQWS------RVWAEDNPPGLVVNQ----APV
Chr19_20477244_1350_po_C_Ez10/      EW--------EEIQPALAERWS------ZEWTEDNPPGLAANQ----API
Chr5_111890459_1626_po_C_Ez10/      EWRLFLTEPGQGIKPALAKQWP------ZIWAEDNPPGLVVNQ----APV
ChrX_48766518_1600_po_C_Ez6/3       EWRLFLTEPGQEIKPALAKRWP------RIWADDNPPGLVVNQ----DPV
ChrX_152580713_1572_po_C_Ez10/      ---------------------P------QVWAEDNPPGLVVNQ----APV
ChrX_139287533_1154_po_C_Xz8/4      EZRLFLTEPGQEIKPALAKRWP------ZVWAEDNPPGLAINQ----APV
Chr2_52615786_912_po_C_Xz11/5       EWRLFLTKPGQEIGPALAKRZP------KVWAEDNPT----------GPI
HERV-E_4_1                          --------------------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      EGRLYITEPGQEIGPAVAKWWP------RVWVEDNLPGLAINQ----APV
ChrY_8634032_950_po_C_Xz14/10       ---------------------A------QVWAEDNPLGLMVNK----ASI
Chr3_126735360_997_po_C_Xz5/10      EWRVFLTEPGQERRPALAKRWP------RVRAEENPPGL-----------
Chr3_126713203_958_po_C_Xz7/8       EWRVFLTEPGQERRPALAKRWP------RVRAEENPPGL-----------
Chr3_75226118_977_po_C_Xz8/9        EWRLFLNEPGQERRPALAKPWP------KVRAEDNLPGL-----------
Chr7_6629919_827_po_C_Xz8/7         EWRLFLTEPGQERRPALAKRWP------RARAEDNPPGL-----------
Chr4_9166812_811_po_C_Xz8/10        EWRLFLTELGQERRPALAKRWP------RVRAEDNPPGL-----------
Chr12_8478642_834_po_C_Xz12/7       EWRLFZTEPDQERRPALAKRWP------RVWQKTTLRDW-----------
Chr10_15187175_787_po_C_Xz11/5      EWRLFZTEPSQERRPALAKRWP------RVRAETTLRDL-----------
Chr11_71819740_854_po_C_Xz11/9      ERRLFZTEPGQERRPALAKRWP------RVWAEDNPPGL-----------
Chr11_71519974_886_po_C_Xz13/9      EWRLFZTELGQERRPALAKRWP------RVRVEDNPPGL-----------
Chr8_7271441_969_po_C_Xz10/6        EWRLFZTDPGKEKRPALAKRWP------RVRAEENPPGL-----------
Chr8_11770066_826_po_C_Xz9/8        EWRLFZTEPGQE-RPALAER-A------KSMAEDNPGAL-----------
Chr4_4128362_651_po_C_Xz7/11        EWRLLZTEPGQERRPALAKRWP------RVWAEDNPPGL-----------
Chr14_50203955_850_po_C_Xz10/8      EWRLFZTEPGQEIRPALAKRWP------RVWAEDNPPGL-----------
Chr14_50205291_837_po_C_Xz10/8      EWRLFZTEPGQEIRPALAKRWP------RVWAEDNPPGL-----------
Chr7_6597120_842_po_C_Xz7/7         QLWLS---GGQEYR-----------------QKTTLQGW-----------
Chr11_71544658_783_po_C_Xz8/13      EWRVFELRRAKREDQLWLSGGQ------EYGQKTTHQGW-----------
Chr11_67999520_760_po_C_Xz8/8       NGDFCELNRAKREDQLWLSRGQ------EYGQKTTLQGW-----------
Chr3_75461975_796_po_C_Xz7/9        ---------AKREDQLWLSGGP------EYRQKTTLWDW-----------
Chr12_52228939_692_po_C_Xz3/10      ---AQVTDGGLDWLTPLS--------------------------------
Chr10_15211234_286_po_C_Xz11/1      --RTILTSLNDIPEPSLCN-----------------FSL-----------
Chr12_100973302_945_po_C_Xz12/      EWRLFLTEPGQEIKPALAKRWPRVWASPGLTVNQAPILI-----------
Chr19_9685420_817_po_C_Xz6/10       -------------PQVZAEDNP-----LELAVNQTPILI-----------
Chr13_40936820_719_po_C_Xz9/10      ERRLFELSQAKREDQLWLSGGQ------EYGQETSLRDW-----------
Chr7_63858202_1754_po_C_Rz12/9      APMQPCTQEEEKLLFRPIDKIP------GVWAEDKP-PGLAV----NQAL
ERV3_chr7q21_AC073210_polputei      RHQCSLTQEEEKLLFRPIDKIP------GVWAEDKP-PGLAV----NQAL
Chr6_110218096_843_po_C_Xz16/8      ----------------------------------------AV----NQAL
Chr6_110214535_786_po_C_Xz16/8      ----------------------------------------AV----NQAL
ChrY_3018841_881_po_C_Rz9/12        --AKGLLKPELHELYGLPSKIP------GVGZRZAP-WTGWN----QAPV
ChrX_86876098_639_po_C_Xz7/8        --AKGLLKPELDELYGLPSKIP------GVGZRZAP-WTGWN----QAPV
Chr14_68899917_965_po_C_Xz6/8       -APDLPHMLGIEELFGLLTKIP------EYGLKITL-WDGCK----LALV
ChrY_23874437_661_po_C_Xz15/9       QSPELSHKPGLEKVFTLLGKIF------GVWANNNP-LSLAVKAVNQALI
ChrY_27518418_649_po_C_Xz15/9       --------PGLEKVFTLLGKIF------GVWANNNP-LSLAVKAVNQARI
ChrY_25016892_657_po_C_Xz13/12      QSPELSHKPGLEKVFTLLGKIF------GVWANNNP-LSLAVKAVNQALI
ChrY_19932937_520_po_C_Xz10/14      QSPELSQKPGLEKLFMLPGKIF------GLWAKNNP-LGLAL---NQAPV
ChrY_20104614_518_po_C_Xz10/14      QSPELSQKPGLEKLFMLPGKIF------GLWAKNNP-LGLAL---NQAPV
Chr19_58660155_598_po_C_Xz21/1      APMQPHTQVEEKLLFQLVKDIP------GVQAEDNP-RGLAVN---HAPV
Chr19_58621153_500_po_C_Xz21/1      VPVQPHMQEEQKSLFQLAKEIP------GVZAEDNP-PGLAVN---HAPV
Chr7_139686987_710_po_C_Xz11/7      APMQPHTQEEEKLLFQLVKEIT------GVWAEDNP-PGLAVN---HAPV
Chr7_137651301_1097_po_C_Pz13/      KPPELSHKPGILELZRLLNKIP------GVWAEDNP-PGLAKD---ZMPV
Chr11_29579197_930_po_C_Pz8/12      ESP----EPGINELYGLLNKIP------GVWAEDNP-PGLAKN---QAPV
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      KTL----EPRVNEMPGLLTKIP------RVWADSNP-PGLAVS---QAPV
Chr4_70022513_788_po_C_Kz10/12      KTP----EPGVNEMHGLLTKIP------EFGLKVITWTGASVN---QAPV
Chr14_105225894_1027_po_C_Xz13      KTP----EPGVNZMRGLLTKIT------GVWAESNL-PELAVN---PAPI
Chr7_119241536_761_po_C_Xz11/1      ESP----ELGINELHELFYRIP------GVZAEDNP-PGLPEN---QASV
ChrY_22564970_567_po_C_Xz9/10       QVPLKQGLLGPLEKYGLVTKIP------RVWAENNP-PGLAVN---QAPV
Chr21_14409346_1477_po_C_Pz5/6      YGRRYSVAEREKLFTDLLLKLS------GVWAEDSPPGLVVNQ----APV
ChrX_51259700_1396_po_C_Pz7/8       CGKRYTVAKGEKLFTDLLLKLP------GVWAEDNPTGLAVNR----APA
Chr6_64985396_1041_po_C_Xz11/1      CKDAFCQWENEAMYKELFLKLP------GVLVEDNPPGLAINQ----APV
ChrY_20355209_659_po_C_Xz7/11       -----------AMYKELFLNLA------GVWVEDSLPELAINQ----APV
Chr3_34567166_523_po_C_Xz8/6        ------QWGNEAIYEELFLKLP------GVWAEDNPLGLAINQ----ASV
```

```
Chr3_122072468_245_po_C_Xz9/9       -----ASGKMRQCIRNYFSSCQ------GSGWKTVPR-LAINQ----AHV
Chr11_58788475_541_po_C_Xz8/22      YAPIYSSLARLRQEAALAASGG------SDSEESTPQAIPCRE----DPE
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------------------
ChrY_13743185_977_po_C_Xz14/8       --ARRIPELDLHSMZGMLFKVP------GVWAEDNPPGLAVNR----HPV
HERV15                              --ARRIPELDLHSMZGMLFKVP------GVWAEDNPPGLAVNR----HPV
ChrX_117198092_983_po_C_Xz8/8       -----VTRTDLHNIWGMLFKVP------GVWAEDYP-GLAANR----PPV
Chr19_21166030_1412_po_C_Xz9/1      LCTRRLLELDLHNTWGMLFKVP------GVWAEDNPPGLVLNR----PPV
Chr19_21213743_913_po_C_Xz14/1      ------------------------------------------------PQ
Chr19_15957774_569_po_C_Xz8/10      APVQLHTRKEEKLLFGLVNKIP------KVZAEDNPPRLAVNH----PPV
Chr11_77508681_297_po_C_Xz19/1      -------------------------------AWGSSVKSISKF----SPV
Chr13_98250769_738_po_C_Xz7/13      -PVQPRMQDEEKZLFQLVREIP------RVWAEDNPPGLAVNH----APV
Chr6_39968183_513_po_C_Pz10/9       KYHLSPRLSEPLKIYELIVKIP------GVWAKDNPPGLAVNQ----ALV
Chr15_66867263_645_po_C_Xz11/1      RLYKKTLEPGVNEMHGVLTKIP------GVWAESNPLGLAVNQ----APV
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                ------AASGVRNQPTQCEEZP---ELSATRKETGQQPSLKDPHILWRRL
MER66_polputein                     --------------------------------------------------


HERVI_polputein                     HQANTPH-------------WEVVRRKQFPIPLEGMLGLKPIIES-----
HERVIP10F_polputein                 --------------------------------------------------
HERVadp_AC005741                    QIELKQA-------------EEVVCRKQYPISLEGRKCLQPVIEE-----
HERV-PT47D                          HIKLKTP-------------GEVVRRKQZLIPLQGRIGLKRIIEG-----
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         AHPVQIR---------LKDPT-TSLSKAISLRPEAHKGLQNIVKH-----
Chr13_54564403_572_po_C_Yz7/14      AHPVQVK---------LKD----SASFPYPLRPKAQQGLQKIVKD-----
HERV_Y_chr12_51022911_pol           ----PQV---------TKGFCLLSLQRQYPLRPKAKQRLQKIVKD-----
HERV17_polputein                    --------------------------------------------------
HERV_W_chr7_9105739_syncytin_p      ARPVQVK---------LKDSTSFPYQRQYPLRPKAQQGLQKIVKD-----
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  --------------------------------------------------
ERVfrd_AC004022                     --------------------------------------------------
PRIMA41_polputein                   VPLVCIZ---------LKPDALYSWKRQDPLKPEAZRGIHPLRTK-----
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 HTPVRIT---------LKEPTQFLSQKQYPIPQAALIGLKPIISR-----
HERV_H_RTVLH2_pol                   HAPLTIS---------LKPNHPYPRQCQYPIPQHALKRLKPVITR-----
HERVFH21_pol                        QDPIKIQ---------LKDPSKFPNIPQYPISLTHQKGLQPIINK-----
HERVRblike_chr4_109047953           VSPIKNK---------LKEGPQPIRKKAIPLKEGSLGSIQPALIW-----
ChrY_13790029_693_po_C_Xz12/11      VVELLAT-------------TLPVQVKQZSVSQQARERINPHIQQ-----
HERVS71_polputein                   VVELLAT-------------ALPVQVKQYPMSQWAREGINPHIQ------
BaEV_M7_pol                         --------------------------------------------------
HERVR_polputein                     --------------------------------------------------
AKR_MLV_MLOCG_RT                    IIPLKAT-------------STPVSIKQYPMSQEARLGIKPHIQR-----
GaLV_pol                            -------------------------------------GIRPHIQK-----
Chr3_158587414_1662_po_C_Ez10/      LIEVKPG-------------AQPIRQKQYPLPREALEGIQVHLRR-----
Chr3_79627623_1138_po_C_Ez4/7       LVEVKPG-------------AQPIRQKQYPFPREALKGIQFHLRH-----
Chr19_20477244_1350_po_C_Ez10/      LIEVKPG-------------PNZLDESSIPS-QEALEGIRVHLRC-----
Chr5_111890459_1626_po_C_Ez10/      LIEVKPG-------------AQPIRQKQYPVPREALEGIQAQLRH-----
ChrX_48766518_1600_po_C_Ez6/3       LIEVKPG-------------AQPIRQKQYPVPREALEGIQVHLRH-----
ChrX_152580713_1572_po_C_Ez10/      HIEVKPG-------------AQPIRQKQYPVPREALKGIQVHLRR-----
ChrX_139287533_1154_po_C_Xz8/4      LIEVKPG-------------GQPIRQKQYPVPREALEGIQVHLKR-----
Chr2_52615786_912_po_C_Xz11/5       ITEIKPG-------------TQPVRQKZYPVPREALEGIQVHLKC-----
HERV-E_4_1                          --------------------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      LIEVEPA-------------AQLARQKZYPVSREDLEGIQVYLKC-----
ChrY_8634032_950_po_C_Xz14/10       LIKV----------------AQPILQKQYLVPREALKGIQVHLRY-----
Chr3_126735360_997_po_C_Xz5/10      --PVKTG-------------AQPVRRKQDLVPREALQGIQVRLKH-----
Chr3_126713203_958_po_C_Xz7/8       --PVKTG-------------AQPVRQKQDPVPREALQGIQVRLKH-----
Chr3_75226118_977_po_C_Xz8/9        --PVKTG-------------AQPVRQKQEPVPREALQGIQVCLKH-----
Chr7_6629919_827_po_C_Xz8/7         --PVKTG-------------AQPVRQKQDPVPREALQGIQVRLKH-----
Chr4_9166812_811_po_C_Xz8/10        --PVKTG-------------AQPVRQKQDPVPREALQGIQVRLKH-----
Chr12_8478642_834_po_C_Xz12/7       --PVKTG-------------AQPVRQKQEPVPKEALQGIQVPLKH-----
Chr10_15187175_787_po_C_Xz11/5      --PVNTG-------------AQPVRQKQEPVPREALQGIQVRLKH-----
Chr11_71819740_854_po_C_Xz11/9      --PVKTG-------------ALPVRQKQGPVPREALQGIQVHLKH-----
Chr11_71519974_886_po_C_Xz13/9      --PVKTG-------------AQPVRQKZDLVPREALQGIQVRLRH-----
Chr8_7271441_969_po_C_Xz10/6        --PVKTG-------------AQPVRQKQDSVPREALQGIQVHLKH-----
Chr8_11770066_826_po_C_Xz9/8        --PVKTG-------------AQPVRQKEEPVPREALQGIQVHLKH-----
Chr4_4128362_651_po_C_Xz7/11        --PVKTG-------------AQPLRQKQDPVPREALQGIQVHLKH-----
Chr14_50203955_850_po_C_Xz10/8      --PVKTG-------------AQPLRQKQDPVPREALQGIQVHLKH-----
Chr14_50205291_837_po_C_Xz10/8      --PVKTG-------------AQPLRQKQDPVPREALQGIQVHLKH-----
Chr7_6597120_842_po_C_Xz7/7         --PVKTG-------------AQPLKQKQDPVPREALQGIQVHLKH-----
Chr11_71544658_783_po_C_Xz8/13      --PVKTG-------------SQPLRQKQDPVPREALQGIQVHLKH-----
Chr11_67999520_760_po_C_Xz8/8       --LVKTG-------------AQPLRQKQDLVPREALQGIQVHLKH-----
Chr3_75461975_796_po_C_Xz7/9        --PIKTG-------------AQPLRQKQDPVPREALQGIQAPLKH-----
Chr12_52228939_692_po_C_Xz3/10      --PEVLG-------------LQALRLASSYPP----------LLF-----
Chr10_15211234_286_po_C_Xz11/1      --PSVMA-------------AILTSQ------------------------
Chr12_100973302_945_po_C_Xz12/      --EVNPG-------------AQPIRQKQYPVPREALERIQVHLRC-----
```

```
Chr19_9685420_817_po_C_Xz6/10       --EVKPG-------------AQPVRZKQYPVPREALEGIQVHLMC-----
Chr13_40936820_719_po_C_Xz9/10      --PVKTG-------------AQPVRQKQDPVPREALQGIQVRLKH-----
Chr7_63858202_1754_po_C_Rz12/9      VVELKPG-------------ATPVQVCQYPLSQEDIWGIYKHLKW---LC
ERV3_chr7q21_AC073210_polputei      VVELKPG-------------ATPVQVCQYPLSQEDIWGIYKHLKW---LC
Chr6_110218096_843_po_C_Xz16/8      VIVELKS-------------SNTGQVHQYPLPQEAIQGIQEHFKW---LL
Chr6_110214535_786_po_C_Xz16/8      VIVELKS-------------SNTGQVHQYPLPQEAIQGIQEHFKW---LL
ChrY_3018841_881_po_C_Rz9/12        TVELKPG-------------EIPVRVCQYPLSQEAVQGIQGHLKR---LL
ChrX_86876098_639_po_C_Xz7/8        TVELKPG-------------EIPVRVCQYPLSQEAVQGIQGHLKR---LL
Chr14_68899917_965_po_C_Xz6/8       VVELKPG-------------ATPVRVHQYPLPZEAIWGIYNHLKW---LY
ChrY_23874437_661_po_C_Xz15/9       IVKLKPG-------------ATPVQVHQ-PISLRRHAGYSQTFKV---TL
ChrY_27518418_649_po_C_Xz15/9       IVKLKPG-------------ATPVQVHQ-PISLRRHAGYSQTFKV---TL
ChrY_25016892_657_po_C_Xz13/12      IVKLKPG-------------ATPVQVHQ-PISLRRHAGYSQTFKV---TL
ChrY_19932937_520_po_C_Xz10/14      IVKLKPG-------------ATTVEVLQYPFPZEDTQGIHKHLKZ---LF
ChrY_20104614_518_po_C_Xz10/14      IVKLKPG-------------ATTVEVLQYPFPZEDTQGIHKHLKZ---LF
Chr19_58660155_598_po_C_Xz21/1      VAESKPG-------------ATPVSVRQYPVPQKSIQGICKRSER---LY
Chr19_58621153_500_po_C_Xz21/1      VAELKPE-------------ATPVRVHQYPVPQKSVQGICKHLER---LY
Chr7_139686987_710_po_C_Xz11/7      VVELKPG-------------ADPVQVHQYPVTQKAVWGICKHLQP---LY
Chr7_137651301_1097_po_C_Pz13/      VVE-----LKPGATP---------VVCQYLVFPEAIZGVHKHLEZATLLT
Chr11_29579197_930_po_C_Pz8/12      VVENRVVELKPGAVSPVQFSFTPVQFHQYPFFQEVIWDIRKHLE---QLF
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      VVELKS-----------------SSVQQYLLLPEAILGVRKYLE----LH
Chr4_70022513_788_po_C_Kz10/12      VVELKLG-------------ETPVRVRQYSLLPEDIRGVHKHLEW---LH
Chr14_105225894_1027_po_C_Xz13      VVELKLG-------------ATLVHVRQYPLLPEAY-GSPQTLRL---TS
Chr7_119241536_761_po_C_Xz11/1      VVELKPG-------------ATPVPLHQYPLFQEDIWDIHKHLEW---LF
ChrY_22564970_567_po_C_Xz9/10       LVELKPG--------------ATLVSLSVTDFHISHQGVQKHIDW---LF
Chr21_14409346_1477_po_C_Pz5/6      IVELLQG-------------TYLVQICQYPIPTEVNQGIAKHLKW-----
ChrX_51259700_1396_po_C_Pz7/8       VVELLRR-------------TYPVRIRQYLIPMEAYWGIEKHLKR-----
Chr6_64985396_1041_po_C_Xz11/1      VVELLWG-------------TYPVZICHNPIPAEATHGITKHIDQ-----
ChrY_20355209_659_po_C_Xz7/11       VVELLRD-------------TYPVQIRQYSIPVKATHGITKHIVR-----
Chr3_34567166_523_po_C_Xz8/6        VVELLRG-------------NYPVZICQYHMPVEATHGIIKHIDW-----
Chr3_122072468_245_po_C_Xz9/9       VVELLRG-------------TYPVZIRQYPIPAEATHGITKHIDW-----
Chr11_58788475_541_po_C_Xz8/22      PKSQGRN-------------SRGMRSATLAQAMPEQCRRPSETWG-----
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------------------
ChrY_13743185_977_po_C_Xz14/8       VIELNTH-------------AALVZVCQYPLPKEVIEGITQHLNR-----
HERV15                              VIELNTH-------------AALVZVCQYPLPKEVIEGITQHLNR-----
ChrX_117198092_983_po_C_Xz8/8       VAELNPH-------------AAPVZVRQYPVPREAIDGIQKHLNR-----
Chr19_21166030_1412_po_C_Xz9/1      VVELNPH-------------ATLVZVRZYPLPREATEGITKHVNQ-----
Chr19_21213743_913_po_C_Xz14/1      VVELNPH-------------AAPVQVHQYPLPZEAIDRITKHLNW-----
Chr19_15957774_569_po_C_Xz8/10      VVKLKPG-------------ATLVRVRQYPVPZEAIZGICKHLDW-----
Chr11_77508681_297_po_C_Xz19/1      VAELKPG-------------ATPVWVRQYPVPQKAVQGVCKHLER-----
Chr13_98250769_738_po_C_Xz7/13      VVELKPG-------------ATPIQVRQYSVPQKAVRGICKHLER-----
Chr6_39968183_513_po_C_Pz10/9       LVELKSG-------------ASPVZVRQYPVFIDGIZGVQKHTDQ-----
Chr15_66867263_645_po_C_Xz11/1      VVELKPG-------------ATPVQVHQYPLLPDAIRGVHKHLEW-----
HERV18_polputein                    ---------------------------------------TETIKK-----
HERVIP10FH_polputein                CAQTKEG--------------EAAERVKYFLGGRDZGKSRGAVKHS----
MER66_polputein                     -------------------WSKHLLKVKMLNTCEWNRCLNSELXVR---F


HERVI_polputein                     -LINDGLLEPCMSPYNTPILPVKK-----SDGSYRLVKDLRAIN------
HERVIP10F_polputein                 -----------------PILPVKK-----SDGSYRLVQDLRAIN------
HERVadp_AC005741                    -LIKDGLLEPCMSPYNTPILPVK------SPGLYRLVQDLRTIN------
HERV-PT47D                          -LIKDGLLKSRMFPITPPILPVKK-----SDGSYQLVQDLRAIN------
HERV_W_chr6_141432567_ERV9_lik      -------------PCNTTILGVQKLN-----GQWRLVQDLRLIN------
ERV9_PH1_RT                         -LKAQGLVRKCSSPCNTPILEVQRP---------RLVQDLRLIN------
Chr13_54564403_572_po_C_Yz7/14      -LKAQGLVKACNSHCSTPILGVQKPN-----RQZRLVQDLRIIS------
HERV_Y_chr12_51022911_pol           -LKAQGLVKPCSSPCNTPILGVQKPN-----GQZRLVQNLRIFS------
HERV17_polputein                    ---AQGLVKPCSSPCNTPILGVQKPN-----GQWRLVQDLRIIN------
HERV_W_chr7_9105739_syncytin_p      -LKAQGLVKPCSNPCSTPILGVQKPN-----RQWRLVQDLRIIN------
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  ---------------NTPILGVQKPN-----GEWRLVQDLCLIN------
ERVfrd_AC004022                     ---THGLIS-CNSPYNTPILTVKKTN-----GEYRLVQNLZVIN------
PRIMA41_polputein                   -FLQFGLLRPFKFPCNTLILPIKKPN-----GDYIFVQDLZAVN------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 -LLASHLLRPTNSPFNTPVLPVKKPD-----GTYRLVQDLRLIN------
HERV_H_RTVLH2_pol                   -LLQHGLLKPINSPYHSPILPVLKPD-----KAYRLVQNLHLIN------
HERVFH21_pol                        -LCSRRLLRPTRSPYNTSILPVKKSD-----GSYQLVQDLRAIN------
HERVRblike_chr4_109047953           -FLQYGLIRPCQSSYDTPFLPLKKPHS----HKYRFVQDHRAIN------
ChrY_13790029_693_po_C_Xz12/11      -LLQTGILTPCQSIWNTPFFPVQK----PRTNDYWPVQDLTEVN------
HERVS71_polputein                   ------------SAWNTPFLLVQK----PGTNDYRPVQDLQEVN------
BaEV_M7_pol                         --------------WNTPLLPVKK----PGTQDYRPVQDLREIN------
HERVR_polputein                     --------------WNTPLLPVKK----PGTQDYRPVQDLREIN------
AKR_MLV_MLOCG_RT                    -LLDQGILVPCQSPWNTPLLPVKK----PGTNDYRPVQDLREVN------
GaLV_pol                            -FLDLGVLVPCRSPWNTPLLPVKK----PGTNDYRPVQDLREIN------
Chr3_158587414_1662_po_C_Ez10/      -LKAYGIIVPCQALWNTPLLPVHK----PGTKDYHPVQDLCLVN------
```

```
Chr3_79627623_1138_po_C_Ez4/7       -LKAFGIIVPCQSPWNTPLLPVPK----PGTKDYRPVQDLRLVN------
Chr19_20477244_1350_po_C_Ez10/      -FKAYGIIVLCQCPWNTPLLPVPK----PGTKDYQPVQDLRLVN------
Chr5_111890459_1626_po_C_Ez10/      -LKAYGIIVPCQSPWNTPLLPVPK----PGTKDYZPVQDLHLVN------
ChrX_48766518_1600_po_C_Ez6/3       -LKAFGIIVPCQSPWNLPFLPVPK----PGTKDYQPVQDLCLVN------
ChrX_152580713_1572_po_C_Ez10/      -LKAFGIIVPCQSPWEHPLLPVPK----PGTKDYR-VQDLRLVN------
ChrX_139287533_1154_po_C_Xz8/4      -LKAFGIIVPCQSPWN-PLLPVPK----PVTKDYRPVQDLHLVN------
Chr2_52615786_912_po_C_Xz11/5       -LRAFGITVPCQSPWNTPLLPVPK----SGTKDYRPVQDLCLVN------
HERV-E_4_1                          -LRTFRIIVPCQSPWNTPLLPVPK----PGTKDYRPVQDLRLVN------
ChrY_21962399_1137_po_C_Xz13/9      -LKAFGIIVPCQSPWNTPLLLVPK----PGTKYYRPVQDLHLVN------
ChrY_8634032_950_po_C_Xz14/10       -LIAYGIIVPCQSPWNKPLLPVP---------DCQPVQHLRLVN------
Chr3_126735360_997_po_C_Xz5/10      -LRTFGIIVPCQSPWNTPLLPVPK----PRTKDYRPVQDLRLLH------
Chr3_126713203_958_po_C_Xz7/8       -LRTFGIIVPCQSPWNTLLLPVPK----PRTKDYRPVQDLPLLH------
Chr3_75226118_977_po_C_Xz8/9        -LRTYGIIVPCQCPWNTPLLPVPK----PQTKDYRPVQDLRFLL------
Chr7_6629919_827_po_C_Xz8/7         -LRTFGIIVPCQSPWNTPFLPVPK----PPTKDYRQVQDLCLLH------
Chr4_9166812_811_po_C_Xz8/10        -LRTFGIIVLCQSPRNTPLLPVPK----PRTKDYRTVQDLRLLH------
Chr12_8478642_834_po_C_Xz12/7       -LRTFGMIVPCQSPWNTPLLPVSK----PRTKDYRPVQDLHLLH------
Chr10_15187175_787_po_C_Xz11/5      -LRTFGMIVPCQSPWNTPLLPVPK----PRTKDYQLVQDLRLLH------
Chr11_71819740_854_po_C_Xz11/9      -LRTFGIIVPCRSPWNTPLLSVPK----SRTKNYQPVQDLRLLH------
Chr11_71519974_886_po_C_Xz13/9      -LRTFGIIVPCQSPWNTPLLPVPK----PRTKDYRPVQDLRLLH------
Chr8_7271441_969_po_C_Xz10/6        -LRTFGMIVPCQSPWNTPLLPVPK----PRTKDYRPVQDLRLLH------
Chr8_11770066_826_po_C_Xz9/8        -LRTFGMIVPCQSPWNTPLLPVPK----PRTKDYRPVQDLRLLH------
Chr4_4128362_651_po_C_Xz7/11        -LRTFGMIVPCQSPRNTPLLPVPK----PWTKDYSPGQDSRLLS------
Chr14_50203955_850_po_C_Xz10/8      -LRTFVMIVPCRSPRNTPLLPVPK----LWTKDYRPGQDLRLLS------
Chr14_50205291_837_po_C_Xz10/8      -LRTFVMIVPCRSPRNTPLLPVPK----LWTKDYRPGQDLRLLS------
Chr7_6597120_842_po_C_Xz7/7         -LRTFGMIVPCQSPWNTPLLPVPK----PWAZDYRPGZDSRLLS------
Chr11_71544658_783_po_C_Xz8/13      -LRTFGMIVPCKSPRNTTLLPIPK----PCTKDR-----IRTCL------
Chr11_67999520_760_po_C_Xz8/8       -LRTFGMIVPCQSPQNTPLLPVPK----PWTKDYRLGQDLRLFS------
Chr3_75461975_796_po_C_Xz7/9        -LRTFGIIVPCQSPWNTPLLPVPK----PRTKDYSPGQDLRLLN------
Chr12_52228939_692_po_C_Xz3/10      -LKLSFSICCTSCQVYTPLLPVPK----PGTKDYRPVQDLRLVN------
Chr10_15211234_286_po_C_Xz11/1      -LVPVDMLAGYQSALKZ-----PK----TAFQERREGEFTNWVF------
Chr12_100973302_945_po_C_Xz12/      -LKAYGIIVPCQS-HGTPLLPVPK----PGTKDYZPVQDLRLVN------
Chr19_9685420_817_po_C_Xz6/10       -VKACGVIVPCQSPRSTTLLPVPK----PGTKDYRPAQDLRLVN------
Chr13_40936820_719_po_C_Xz9/10      -LRTFGIIVPCQSPWNTPLLPVPK----PRTKDYRLVPDLRLLN------
Chr7_63858202_1754_po_C_Rz12/9      DHGII---IQCQSPWNTPLLLVQKPLPGPGSDEYIPVQGLHAVN------
ERV3_chr7q21_AC073210_polputei      DHGII---IQCQSPWNTPLLLVQKPLPGPGSDEYIPVQGLHAVN------
Chr6_110218096_843_po_C_Xz16/8      EHGIL---AZCQSSWNTPLLLVQK----PESNQYRPVQDLHAVTSAGYRP
Chr6_110214535_786_po_C_Xz16/8      EHGIL---AZCQSSWNTPLLLVQK----PESNQYRPVQDLHAVTSAGYRP
ChrY_3018841_881_po_C_Rz9/12        EHRVL---AQCWSPWNTLLLPVRK----PGTNEYRLVQNLHAVN------
ChrX_86876098_639_po_C_Xz7/8        EHRVL---AQCWSPWNTLLLPVRK----PGTNEYRLVQNLHAVN------
Chr14_68899917_965_po_C_Xz6/8       DPGIL---VQYQSPWNTP-LPVQK----PGSGEYRLAQDLHAVN------
ChrY_23874437_661_po_C_Xz15/9       QTQNLSLVVWCESPWNTPLLPKQN----TESGKSKPVQNLHEMN------
ChrY_27518418_649_po_C_Xz15/9       QTQNLSLVVWCESPWNTPLLPKQN----TESGKSKPVQNLHEMN------
ChrY_25016892_657_po_C_Xz13/12      QTQNLSLVVWCESPWNTPLLPKQN----TESGKSKPVQNLHEMN------
ChrY_19932937_520_po_C_Xz10/14      KHRNF---VWRQSTWNTLLLPIQN----PGSGKYKPVQNLHDVN------
ChrY_20104614_518_po_C_Xz10/14      KHRNF---VWRQSTWNTLLLPIQN----PGSGKYKPVQNLHDVN------
Chr19_58660155_598_po_C_Xz21/1      KHKIL---VQCZSPZNTPLLTVQKPLPGPGSNEYRLVQDLHVVN------
Chr19_58621153_500_po_C_Xz21/1      KHKIL---VQCQSPZNTLLLPVQKPLPGPGSNEYRPVQYLHAIN------
Chr7_139686987_710_po_C_Xz11/7      KQGIL---VQCQLPWNTVLLPVQKPLPGPGSDEYRPVQDLHAVN------
Chr7_137651301_1097_po_C_Pz13/      WNHVV---VKFQSPWNIPLLPVQK-----PSGEYRPVQDLCAVNH-----
Chr11_29579197_930_po_C_Pz8/12      KHGII---VKCQSPWNTPLLPVQK-----PSGEZRLVQDLCAVNE-----
Chr4_117329205_488_po_C_Xz3/16      -----------NSKRLRTLCEIEW-----PSFNVRRLAEETIDRE-----
Chr21_38522185_383_po_C_Xz7/9       --------IKLTPQRLRTLCEIEW-----TSFGVRWLVEGTIDRE-----
Chr4_70173370_1049_po_C_Xz11/8      KHGII---VKCQSPWNTPLLPVRK-----PSGEYRPVQNLCVENZ-----
Chr4_70022513_788_po_C_Kz10/12      KHGII---AKCQSPWNNILLPVRK-----PSGEYRPVQDLHSENQ-----
Chr14_105225894_1027_po_C_Xz13      QVGTI---VKCQSLWNTPLLPVWK-----PSGEYRPVQDLCAVNQ-----
Chr7_119241536_761_po_C_Xz11/1      KHGIL---VKCQSPWNTPLLLVQK-----PSGEYRPVQDLRAIT------
ChrY_22564970_567_po_C_Xz9/10       KNRII---VKCRZVWTTPLLPVZK-----PSGEYGPVQDLRAIN------
Chr21_14409346_1477_po_C_Pz5/6      -LLEFGIIERCVSSWNAPLLLVLK-----LSGDYWPIQALKAIN------
ChrX_51259700_1396_po_C_Pz7/8       -LLEFWIIEECTSSWNTPLLPVLK-----PSGDYQPVQDZKAVN------
Chr6_64985396_1041_po_C_Xz11/1      -FLKFGIIERCASSWNTPLLPVLK-----PPGDYQPVQDLWAVN------
ChrY_20355209_659_po_C_Xz7/11       -LFKFGIIERCASPWNTPLLLVLK-----PIGDYWLAQDLQAIH------
Chr3_34567166_523_po_C_Xz8/6        -LLKFG-IERCASSWNTPLLLVLK-----PSGDYRPLQDLQAVN------
Chr3_122072468_245_po_C_Xz9/9       -LLKFGIMQRCALSWNTPLVPVLK-----PSRDYQPVKDLRAIN------
Chr11_58788475_541_po_C_Xz8/22      -QIYLNAQNEVEGGEWLFFYQPFS-----TTDLLNWRQHTPSYA------
Chr6_137421042_373_po_C_Xz6/9       -------------------------------------QHTPSYR------
Chr4_115910608_550_po_C_Pz8/12      ----------------------------------------PQVS------
ChrY_13743185_977_po_C_Xz14/8       -LYEQGIIVKCKSSWNTPLQPVHK-----PNGEYRPVQDFWVAN------
HERV15                              -LYEQGIIVKCKSSWNTPLQPVHK-----PNGEYRPVQDFWVAN------
ChrX_117198092_983_po_C_Xz8/8       -LYTHGIIVKCKSSWNTPLLPMHK-----SNGEYRPVQDLRVVN------
Chr19_21166030_1412_po_C_Xz9/1      -LYEHEIIVKYKSSWNTPLLPVLK-----PNGEYRPVZDLQVVN------
Chr19_21213743_913_po_C_Xz14/1      -LYMHEIIAKCKSSWNTPLLPVRK-----PNGEYRPVQDLWAVN------
Chr19_15957774_569_po_C_Xz8/10      -LYKHRILVRCQSPRNTPLLPAZKPTPGPGSNVYRPVQDLRARN------
Chr11_77508681_297_po_C_Xz19/1      -LYKYGILIRCQS-RNTPLLPVZKPLLGPRSDEZIPVQDLHAVN------
Chr13_98250769_738_po_C_Xz7/13      -LYKHRILVQCQSPWNTPLLPIRKPLPGPESDEYRPVQDLRAVN------
Chr6_39968183_513_po_C_Pz10/9       -LFKHGIIIKCRSPWNIPLLPVQKP-----SGKYKPVQNLHAIN------
Chr15_66867263_645_po_C_Xz11/1      -LHKHGIIVKCEZSWNTPLLPVQKA-----SGEYRPVQDLHAVN------
HERV18_polputein                    -LKEVQIVRGTHSPYNFLVWPVRK-----PDGTWRMTVDYRELN------
```

```
HERVIP10FH_polputein                -LVRTYQGERNRRAVKHSLVG---------TRERKPQGVVKYLS------
MER66_polputein                     KILNCKIVLIYRXPMSGGQFRISCSLX-----LYKMCQRNVFFIG-----
                                                               .

HERVI_polputein                     --------------QTVQTTNPVVPNPYTILSK--IPYNHQWFTVIDLKD
HERVIP10F_polputein                 --------------QIVQTTHPVVPNPYTILSK--IPYNHQWFTVIDLKD
HERVadp_AC005741                    --------------QIIQTRHPVMPNPYTFLRK--IPYEHKWFRVVDLKD
HERV-PT47D                          --------------QIVQTTHPVVPCPYTILSK--IPYNHQWFTVIDLK-
HERV_W_chr6_141432567_ERV9_lik      --------------EAVIPLYPVVPNPYILLSQI--PEEAERFMVLDLKD
ERV9_PH1_RT                         --------------EAVISLYPVVPNPYTLLSQI--PEEAEWFTVLDLKD
Chr13_54564403_572_po_C_Yz7/14      --------------EAVVPLYPAVPSPHTLLSQI--PEEAEWFIVLDLKD
HERV_Y_chr12_51022911_pol           --------------EVVVPLDPAVPNPYTLLSQI--PEEAEWFTVLDPKD
HERV17_polputein                    --------------EAVVPLYPAVPNPYTLLSQI--PEE-QWFTVLDLKD
HERV_W_chr7_9105739_syncytin_p      --------------EAVVPLZPAVPSPYTLLSQI--PEEAEWFTVLDLQD
HERV9_polputein                     -----------------------------------------------PQG
HUERS-P3_polputein                  --------------EAVVPIHPVVPNPYTLLTQI--PEGTKWFTVLDLKD
ERVfrd_AC004022                     --------------EAVVRIHPMPQP--LCNSRR--FPDAQWFSVLDFKN
PRIMA41_polputein                   --------------SAVIPKHPVVLNPYMLLPG---LGDASWFTVLDLRD
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------SI--PPSTTHYSVLDLKH
HERV_H48I_polputein                 --------------QAVLPVCPVVPNPYTLLSAI--PSNTTHFSVLNLKD
HERV_H_RTVLH2_pol                   --------------QIVLPIHPMVPNPYTLLSSI--PPSTTHYSVLDLKH
HERVFH21_pol                        --------------QAVLPIHPVVPNPYTLLSLI--PSNTTHYTAIDLKD
HERVRblike_chr4_109047953           --------------DTVEDIHPIVANPYTMFASL--PKDHEWFTVLDLKD
ChrY_13790029_693_po_C_Xz12/11      --------------RQIVTVHPTVPNPYTLLRL--LPPEHTICTVLDLKD
HERVS71_polputein                   --------------KWTVTVHPTVPNPYILLGL--LPPEHTAYTVLDLKD
BaEV_M7_pol                         --------------KRTVDIHPTVPNPYNLLST--LKPDYSWYTVLDLKD
HERVR_polputein                     --------------KRTVDIHPTVPNPYNLLST--LKPDYSWYTVLDLKD
AKR_MLV_MLOCG_RT                    --------------KRVEDIHPTVPNPYNLLSG--LPPSHQWYTVLDLKD
GaLV_pol                            --------------KRVQDIHPTVPNPYNLLSS--LPPSYTWYSVLDLKD
Chr3_158587414_1662_po_C_Ez10/      --------------QATVTLHPTVPNPYTLSG--LLPAEDCWFTCLDLKD
Chr3_79627623_1138_po_C_Ez4/7       --------------QATVTLHPTVPNPYTLLG--LL--------------
Chr19_20477244_1350_po_C_Ez10/      --------------QATVTLHPTVPNPYTLLG--LLPAEDIWFTHLDLKD
Chr5_111890459_1626_po_C_Ez10/      --------------QATVTLHPTVPNPYTLLG--LLLAEDSCFTCLDLKD
ChrX_48766518_1600_po_C_Ez6/3       --------------QATVTLHPTVPNLYTLLG--LLQAEDSWFTCLDLKD
ChrX_152580713_1572_po_C_Ez10/      --------------QATVTLHPTVPNPYTLLG--LLPAEDSWFTCLDLKD
ChrX_139287533_1154_po_C_Xz8/4      --------------QATVTLHPTVPNPYTLLG--LLPAEDSWFTCLELKD
Chr2_52615786_912_po_C_Xz11/5       --------------QPTVTLHPTVPNPYTLLG--LLPAEDSWFTCLDLKD
HERV-E_4_1                          --------------QATVTLHPTVPNLYTLLG--LLPAEDSWFTCLDLKD
ChrY_21962399_1137_po_C_Xz13/9      --------------QATVTLHPMVPNLYTLLE--FLPAEDNLFTSLDLKD
ChrY_8634032_950_po_C_Xz14/10       --------------QATVTLHPTVPNLYTFLG--LLLAEDSWFTCLELKD
Chr3_126735360_997_po_C_Xz5/10      --------------QATLTLHPTVPNPSTLLV--LPPAEDSWFTFWNLKD
Chr3_126713203_958_po_C_Xz7/8       --------------QATLTLHPTVPNPSTLLG--LLPAEDSWFTCLDLKD
Chr3_75226118_977_po_C_Xz8/9        --------------QATLTLHPTVPNPSTLLG--LPPAENSWFTCLDLKD
Chr7_6629919_827_po_C_Xz8/7         --------------QATLTFHPTVPNPSTLLG--LLPAEDSWFTFLDLKD
Chr4_9166812_811_po_C_Xz8/10        --------------QATETFHPTVPNPSTLLG--LLPAEDSWFSCLDLKD
Chr12_8478642_834_po_C_Xz12/7       --------------QATLTFHPTVPNPSTLLG--LPPAEDSWFTGLDLKD
Chr10_15187175_787_po_C_Xz11/5      --------------QATLTFHPTVPNPSTLLG--LPPAEDSWFTCLDLKD
Chr11_71819740_854_po_C_Xz11/9      --------------QAKLTFHPTVPNPSTLLG--FPPAEDSWFTCLDLKD
Chr11_71519974_886_po_C_Xz13/9      --------------QATLTFHPTVPNPSILLG--LPPAEDSWFTCLDLKD
Chr8_7271441_969_po_C_Xz10/6        --------------QATLTLHPTVPNPSTLLG--LPPAEDSWFTCLDLKD
Chr8_11770066_826_po_C_Xz9/8        --------------QATLTLHPTVPNAYTLLG--FLPAEDSWFTCLDLKD
Chr4_4128362_651_po_C_Xz7/11        --------------LATLTFHPTVPSPSTLVG--LLPAEDSWFTCLDLRD
Chr14_50203955_850_po_C_Xz10/8      --------------QATLTFHPTVPSPSTLLG--LLPAEDSWFTCLDLKD
Chr14_50205291_837_po_C_Xz10/8      --------------QATLTFHPTVPSPSTLLG--LLPAEDSWFTCLDLKD
Chr7_6597120_842_po_C_Xz7/7         --------------QATLTFHSTVPSPSTLLG--LLPAEDSWFTCLDPRD
Chr11_71544658_783_po_C_Xz8/13      --------------VKTLTFHPTVPSPSTLLG--LLPAEDGWFTCLDLRD
Chr11_67999520_760_po_C_Xz8/8       --------------QATLTLHPTVPSPSMLLG--LLPAEDSWFTCLDLRD
Chr3_75461975_796_po_C_Xz7/9        --------------QATLTLLPIVPNPSTLLG--LLPAEDSWFTCLDLKD
Chr12_52228939_692_po_C_Xz3/10      --------------QATVTLHPTVPNPYTLLG--LLPAEDSWFTCLDLKD
Chr10_15211234_286_po_C_Xz11/1      --------------CPVINLEQFIGNTVQIIC--PY-----ILFSLDSCQ
Chr12_100973302_945_po_C_Xz12/      --------------QAAVTRHPTVPNPYTLLG--LLPAEDSWFTCLDLKD
Chr19_9685420_817_po_C_Xz6/10       --------------QASDS-APTVPNPYTLLE--LLPAEDSWFTCLDLKD
Chr13_40936820_719_po_C_Xz9/10      --------------QATLTFHPTVPNPSTLLG--LLPAEDSWFTCLDLKD
Chr7_63858202_1754_po_C_Rz12/9      --------------QATVTIHPVVPNLYTLMGL--ILASATWFTVLDLKD
ERV3_chr7q21_AC073210_polputei      --------------QATVTIHPVVPNLYTLMGL--ILASATWFTVLDLKD
Chr6_110218096_843_po_C_Xz16/8      VQDZMNECRLVQDLQSTVTIHPMVPNPYTLMGL--IPASAAWFTCLDLKD
Chr6_110214535_786_po_C_Xz16/8      VQDZMNECRLVQDLQSTVTIHPMVPNPYTLMGL--IPASAAWFTCLDLKD
ChrY_3018841_881_po_C_Rz9/12        --------------QATVTIHPVVPNLYTLMGL--IPANAAWFTCLDLKD
ChrX_86876098_639_po_C_Xz7/8        --------------QATVTIHPVVPNLYILMGL--IPANAAWFTCLDLKD
Chr14_68899917_965_po_C_Xz6/8       --------------QTTVTIPPVVPNLYTLMGL--IAASAAWFTCLDLKD
ChrY_23874437_661_po_C_Xz15/9       --------------QATVTFHLVIPNPYMLIRF--ILASAAZFTZLDLKN
ChrY_27518418_649_po_C_Xz15/9       --------------QATVTFHLVIPNPYMLIRF--ILASAAZFTZLDLKN
ChrY_25016892_657_po_C_Xz13/12      --------------QATVTFHLVIPNPYMLIRF--ILASAAZFTZLDLKN
ChrY_19932937_520_po_C_Xz10/14      --------------QATVTIHLVVPNPYTLIGL--ILASAAZFTSLDLKN
ChrY_20104614_518_po_C_Xz10/14      --------------QATVTIHLVVPNPYTLIGL--ILASAAZFTSLDLKN
Chr19_58660155_598_po_C_Xz21/1      --------------QATVTIHPVVPNPYTVITP--ASATZ--FTVLNLKD
```

```
Chr19_58621153_500_po_C_Xz21/1      --------------QATVTIHPVVPNPYTLIAL--IPASA--STVLDLKD
Chr7_139686987_710_po_C_Xz11/7      --------------WATVTIHPVVPNLYNLMEL--ISASATWFTVZDLKD
Chr7_137651301_1097_po_C_Pz13/      ---------------TTVNIHPIVPNPYTLMGL--IPASAAWFTCLDLKD
Chr11_29579197_930_po_C_Pz8/12      ---------------ATVTIHPVVPNPYSLTGL--ISASAVWFACLDLKD
Chr4_117329205_488_po_C_Xz3/16      ---------------TIGHVFKVVTGVGGQPGH--PDQFPYIDSWLNIVK
Chr21_38522185_383_po_C_Xz7/9       ---------------TIGHVFKVVTGLGGQPGH--PDQFPYIDSWLHIVQ
Chr4_70173370_1049_po_C_Xz11/8      ---------------ATVTIHPVVPNPYTLMGH--IPASATZFTVLDLKD
Chr4_70022513_788_po_C_Kz10/12      ---------------ATVTIHPVVSNLHTLMGH--IPASATZFTVLDLKN
Chr14_105225894_1027_po_C_Xz13      ---------------ATVTIHPVLPNLYTLMGH--IPVSATWFTVLDLKD
Chr7_119241536_761_po_C_Xz11/1      ---------------RTVTIYPVVPNSYTLMGF--IPASATWFTYLDLKD
ChrY_22564970_567_po_C_Xz9/10       --------------QVTETIYPVVPNPYTLLGQ--IPASAAWFIFFYLKD
Chr21_14409346_1477_po_C_Pz5/6      --------------EVVATLHAIVPNLYAMLAQ--IPASAAWFTCLHIKD
ChrX_51259700_1396_po_C_Pz7/8       --------------KVAATQHAVVLNPYTMLGQ--MPASAAWFTCLDIKD
Chr6_64985396_1041_po_C_Xz11/1      --------------KVAATLYAIVSNPYTMLGR--IPADGAWFTCLYIKD
ChrY_20355209_659_po_C_Xz7/11       --------------KVAATLYAIVPNPYTMLGZ--IPADVAWFTCLML-D
Chr3_34567166_523_po_C_Xz8/6        --------------KVAATLYAIVPNPYTMLGQ--IPADAAGFMCLYIKD
Chr3_122072468_245_po_C_Xz9/9       --------------KVAATLYAIVPNPYTMLGZ--IPVDAAWFTCLDIKD
Chr11_58788475_541_po_C_Xz8/22      --------------EKPQALIDLMQSIFLTHNPTWADCKELLLSLFNTEE
Chr6_137421042_373_po_C_Xz6/9       --------------EKPQALIDLMQSIFLTHNPTWVDCKHFFLSLFNMEE
Chr4_115910608_550_po_C_Pz8/12      --------------DT--QLTKVATKVYINR-----DQEAKKEANWKLKE
ChrY_13743185_977_po_C_Xz14/8       --------------KATVTIYAIVPNPYTMLGQ--IPAEAMWFMCLDLKD
HERV15                              --------------KATVTIYAIVPNPYTMLGQ--IPAG-HVVHVLDLKD
ChrX_117198092_983_po_C_Xz8/8       --------------KATVAIHATVPNPYTMLGQ--IPADATWITCLDLKD
Chr19_21166030_1412_po_C_Xz9/1      --------------KTTATIHATVPNPYTMLZH--IPAEAAGFTCLDLKD
Chr19_21213743_913_po_C_Xz14/1      --------------KAAVTIHAIVPNPYTMLGQ--IHADATWFMCLDLKD
Chr19_15957774_569_po_C_Xz8/10      --------------RATVAIHLVVPNLYTLMGL--IPASCTWFTVLDLKD
Chr11_77508681_297_po_C_Xz19/1      --------------RAMVTIHPVAPNPYTLMVL--IPTSATRFTVLDLKD
Chr13_98250769_738_po_C_Xz7/13      --------------RAMVTIHPVVPNPYTLIGL--IPASATWLTVLDLKD
Chr6_39968183_513_po_C_Pz10/9       --------------RVMETIHPVVPNPYNLMGQ--IPASAAWLICLDLKN
Chr15_66867263_645_po_C_Xz11/1      --------------QAKVTMHPVVPNLYSLMGH--ILASAAWFTVLDLKD
HERV18_polputein                    --------------KVTPPLHAAVPSIMDLMDR-LTTELGQYHYVVDLAN
HERVIP10FH_polputein                --------------RDVLEVKKGEKSPLRGAES-QKEVRNPHGGGLPQKE
MER66_polputein                     -------------KKXSSNLEVIKREVQNIRNX---XVEKRDIKNVMDRK


HERVI_polputein                     AFWACP-WLE-SRDTFAFEWEDPQLGZKQWYQ----------------WT
HERVIP10F_polputein                 AFWACP-LAEDSRDIFAFEWEDPHSGRKQQYR----------------WT
HERVadp_AC005741                    VFWACP-LDFRSRDLFAFEWENPITGRKQZYC----------------WT
HERV-PT47D                          VDFACP-LAKDSQDIFAFERRSP-LRAETTIS----------------MD
HERV_W_chr6_141432567_ERV9_lik      A-FFYIPLYSDSQFLFTFE---DPTDHTSQLT----------------WM
ERV9_PH1_RT                         A-FFCVPLHSDSQFLFAFE---DPTNHTSQLT----------------WM
Chr13_54564403_572_po_C_Yz7/14      A-FFCIPVHPDSQFLFAFE---DPK-PNVQLT----------------WT
HERV_Y_chr12_51022911_pol           A-FFCIPVHPDSQFLFAFR---DPSNLTSQLP----------------WT
HERV17_polputein                    A-FFCI-LHPDSQFLFAFE---DPSNPTSQLT----------------WT
HERV_W_chr7_9105739_syncytin_p      A-FFCIPVHPDSQFLFAFE---DTSNPTSQLT----------------WT
HERV9_polputein                     C-LLLFPLHSDSQFLFAFE---DPTDHTSQLT----------------WM
HUERS-P3_polputein                  A-FFCIPLHPDSQYPFAFK---DPSGQTAQLT----------------WT
ERVfrd_AC004022                     --FFCIPLDPSSQFIFVLELE-NEKGRSZQLT----------------WT
PRIMA41_polputein                   FFFFCIZVYPDSKFIFAFGWTDPDSHLVSQLT----------------WM
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       A-FFTIPLHPSTQPLFVFTWTDPDTHQSQQLT----------------WA
HERV_H48I_polputein                 A-FFSIPLHPDSQNLFAFTWENPDTHLSRQLT----------------WC
HERV_H_RTVLH2_pol                   A-FFTIPLHPSSQRLFAFTR-LTLTPIRLQIT----------------WA
HERVFH21_pol                        A-FFTIPLHPDSQNLFAFTRTDPDTLQSQQLT----------------WT
HERVRblike_chr4_109047953           A-FFCIPVEIESQLLFAFEWTDPETAAQFQYC----------------WT
ChrY_13790029_693_po_C_Xz12/11      AFFAIP-LAPKSQPIFAFEWTDPGSGNTTQLT----------------WT
HERVS71_polputein                   AFFAIP-LAPKSQPIFASEWMDPGSGDTTQLT----------------WT
BaEV_M7_pol                         AFFCLP-LAPQSQELFAFEWKDPERGISGQLT----------------WT
HERVR_polputein                     AFFCLP-LAPQSQELFAFEWKDPERGISGQLT----------------WT
AKR_MLV_MLOCG_RT                    AFFCLR-LHPTSQPLFAFEWRDPEMGISGQLT----------------WT
GaLV_pol                            AFFCLR-LHPNSQPLFAFEWKDPEKGNTGQLT----------------WT
Chr3_158587414_1662_po_C_Ez10/      AFFSIR-LAPDSQKLFAFQWEDPESGVTTZYT----------------WT
Chr3_79627623_1138_po_C_Ez4/7       -------LAPESQKLFAFQWEAPESGVTSQYT----------------WT
Chr19_20477244_1350_po_C_Ez10/      AICSIR-IAPESQKLFAFRWEDLQSGVTTQYT----------------WN
Chr5_111890459_1626_po_C_Ez10/      AIFSIR-ISPESQKLFAFQWEDPESGVTTQYT----------------WT
ChrX_48766518_1600_po_C_Ez6/3       AFFSIR-LAPESQKLFAFQWEDPESGVTTQYP----------------WT
ChrX_152580713_1572_po_C_Ez10/      AIFSIK-IAPESQKLFAFQWEDPESGVTTQYT----------------WT
ChrX_139287533_1154_po_C_Xz8/4      AFFSIR-LALESQKLFAFQZEDPGSGVTTQYT----------------WT
Chr2_52615786_912_po_C_Xz11/5       AFFSIR-LAPESQKLFAFQWEDLGSGVTTQYT----------------WT
HERV-E_4_1                          AFFSIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
ChrY_21962399_1137_po_C_Xz13/9      AFFSIR-LAPEGPKLFAFQGEDPGSGVITQYT----------------WT
ChrY_8634032_950_po_C_Xz14/10       VFFSIR-LDPESQNLFAFQZEYLEAGVTTZYT----------------WT
Chr3_126735360_997_po_C_Xz5/10      AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr3_126713203_958_po_C_Xz7/8       AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr3_75226118_977_po_C_Xz8/9        AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr7_6629919_827_po_C_Xz8/7         AFFPIR-LAPESQKLFAFQWEDPESGVTTQYT----------------WT
Chr4_9166812_811_po_C_Xz8/10        AFFPIR-LTPERKKLFAFQWEDPESGVSTQYT----------------WT
```

```
Chr12_8478642_834_po_C_Xz12/7       AFFPIR-LAPERQKLFAFQWADPESGVTTQYT----------------WT
Chr10_15187175_787_po_C_Xz11/5      VFFAIR-LAPERQKLFAFQWADPESGVTTQYT----------------WT
Chr11_71819740_854_po_C_Xz11/9      AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr11_71519974_886_po_C_Xz13/9      AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr8_7271441_969_po_C_Xz10/6        AFLPIR-LAPERQKLFAFQWEDPESGVTTZYT----------------WT
Chr8_11770066_826_po_C_Xz9/8        AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr4_4128362_651_po_C_Xz7/11        AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr14_50203955_850_po_C_Xz10/8      AFFPIR-SAPESQKLFAFQWEDPESGVTTQYT----------------WT
Chr14_50205291_837_po_C_Xz10/8      AFFPIR-SAPESQKLFAFQWEDPESGVTTQYT----------------WT
Chr7_6597120_842_po_C_Xz7/7         AFFPIR-LAPESQKLFAFQWEDPESGVTTQYT----------------WT
Chr11_71544658_783_po_C_Xz8/13      AFFPIR-LP-EIQKLFAFQWEDPESGVTTQYT----------------WT
Chr11_67999520_760_po_C_Xz8/8       AFFSIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr3_75461975_796_po_C_Xz7/9        AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr12_52228939_692_po_C_Xz3/10      AFFRIR-LAPESQKLFAFQWEDPESGITTQYT----------------WT
Chr10_15211234_286_po_C_Xz11/1      KLDMFR-NMPEILHHLTTDRN------TQIHA----------------WW
Chr12_100973302_945_po_C_Xz12/      AFFSNR-IAPESQKLYAFQWEDPESGITTQYT----------------WT
Chr19_9685420_817_po_C_Xz6/10       TFFSIR-LAPESQKLFAFQWEDLGSGVTTQYT----------------WT
Chr13_40936820_719_po_C_Xz9/10      AFFPIR-LAPERQKLFAFQWEDPESGVTTQYT----------------WT
Chr7_63858202_1754_po_C_Rz12/9      AFFCLY-LAPVSQPIFAFZWDNSVTGTG----------------GQLAWT
ERV3_chr7q21_AC073210_polputei      AFFCLY-LAPVSQPIFAFZWDNSVTGTG----------------GQLAWT
Chr6_110218096_843_po_C_Xz16/8      AFFCLH-LPPISQPIFAFQWEDPVTGTK----------------EQLTWT
Chr6_110214535_786_po_C_Xz16/8      AFFCLH-LPPISQPIFAFQWEDPVTGTK----------------EQLTWT
ChrY_3018841_881_po_C_Rz9/12        AFLFPC-VAPSSQPIFAFQZDDSVTGTG----------------EQLTWT
ChrX_86876098_639_po_C_Xz7/8        AFLFPC-LAPSSQPIFAFQZDDSVTGTG----------------EQLTWT
Chr14_68899917_965_po_C_Xz6/8       GFFCLC-LAPSSQPLFAFQWDDSVTGTG----------------EQLTWT
ChrY_23874437_661_po_C_Xz15/9       IFFYLC-LAPISQPIFAFQWDDZVTGTG----------------EQLTWT
ChrY_27518418_649_po_C_Xz15/9       IFFYLC-LAPISQPIFAFQWDDZVTGTG----------------EQLTWT
ChrY_25016892_657_po_C_Xz13/12      IFFYLC-LAPISQPIFAFQWDDZVTGTG----------------EQLTWT
ChrY_19932937_520_po_C_Xz10/14      IFFCLH-LAPVSQPILAFQWHDSVTGMG----------------EELTWA
ChrY_20104614_518_po_C_Xz10/14      IFFCLH-LAPVSQPILAFQWHDSVTGMG----------------EELTWA
Chr19_58660155_598_po_C_Xz21/1      AFFYIC-LAPVSQPIFAFHKN-EASQFT--------------------WT
Chr19_58621153_500_po_C_Xz21/1      AFFLLLSLAPVSQPIFAFHKRKEASQSPGLDSHKGLKTTTIWGPVVPPIG
Chr7_139686987_710_po_C_Xz11/7      AFLHIC-MVPVSQPIFAFQWDKEATQLT--------------------WT
Chr7_137651301_1097_po_C_Pz13/      AFFCLH-LHQLAH-FESRKWGESQFT----------------------WT
Chr11_29579197_930_po_C_Pz8/12      AFFCLR-LAPISP-FAFQ-WSELQFT----------------------WT
Chr4_117329205_488_po_C_Xz3/16      PNLCHC-SPAWQL-IAKRSWFKLSL-----------------------VK
Chr21_38522185_383_po_C_Xz7/9       NZPAWL-QLCLTA-YGKTLVAQAKLK----------------------VK
Chr4_70173370_1049_po_C_Xz11/8      TFFRLW-LAPINEPIFAFQWGKSQYT----------------------WT
Chr4_70022513_788_po_C_Kz10/12      AFFCLQ-FAPISQASFVFQWGKSQYT----------------------WT
Chr14_105225894_1027_po_C_Xz13      TFFCLQ-LAPISQPVFALQWGESQYT----------------------ZT
Chr7_119241536_761_po_C_Xz11/1      AFFCLH-LAPVSP--LFISLGDSQFT----------------------ZT
ChrY_22564970_567_po_C_Xz9/10       AFFCLQ-LVPVSZPIFAFHWGETQYT----------------------WT
Chr21_14409346_1477_po_C_Pz5/6      AFFCIQ-LAPVSRDIFAFDW-PSQYT----------------------WT
ChrX_51259700_1396_po_C_Pz7/8       AFFCIQ-LAPMSZDIFAFEWGPPQYT----------------------WT
Chr6_64985396_1041_po_C_Xz11/1      AFFCIQ-LAPES--IFAFEWGLSQYT----------------------WT
ChrY_20355209_659_po_C_Xz7/11       AFFCIQ-LAPESQGIFAFEWSSSQYT----------------------WT
Chr3_34567166_523_po_C_Xz8/6        ALFCIZ-LALESQGIFAFEWSITVY-----------------------LD
Chr3_122072468_245_po_C_Xz9/9       AFFCIQ-RAPESQGIFVFEWGSSQYT----------------------WT
Chr11_58788475_541_po_C_Xz8/22      HHRVIQ-AALQWLEDNEPAGTGDIRH----------------------YA
Chr6_137421042_373_po_C_Xz6/9       HRRVIQ-AALQWLEKNASSSTGDIRQ----------------------YT
Chr4_115910608_550_po_C_Pz8/12      RQIYZQ-QPSREGKLTLQGGVGAVMK----------------------EA
ChrY_13743185_977_po_C_Xz14/8       -FLCLR-LAPQSQPIFAFQWGQLZYT----------------------WT
HERV15                              -FLCLR-LAPQSQPIFAFQWGQLZYT----------------------WT
ChrX_117198092_983_po_C_Xz8/8       AFFCLK-LAPQSQ------RGQSQYT----------------------WT
Chr19_21166030_1412_po_C_Xz9/1      AFFCLR-LVPQSQPIFAFQWGQSQYT----------------------RR
Chr19_21213743_913_po_C_Xz14/1      AFFCLR-LAPQSQPVFAIQWGKSQYT----------------------WT
Chr19_15957774_569_po_C_Xz8/10      AFFCIZ-LAPVSQPIFAFQZDKEATQLT--------------------WT
Chr11_77508681_297_po_C_Xz19/1      AFFYIR-LARVSP--FAFQWDKEAMQLM--------------------WT
Chr13_98250769_738_po_C_Xz7/13      TFFCIC-LAPVSQPIFAFNRTKKLTQLT--------------------WT
Chr6_39968183_513_po_C_Pz10/9       TFFCLQ-LAPISQPIFAFHWD--ETQHT--------------------WT
Chr15_66867263_645_po_C_Xz11/1      AFFCLW-LAPISQPIFAFQZG--ESQYI--------------------WT
HERV18_polputein                    AFFSLT--LPESQEQFAFTWEGRQWT----------------------FT
HERVIP10FH_polputein                VRNPHGGRGVEPQKEVRNPHGGGZTSKRGEKS---------------PWR
MER66_polputein                     CIFCKEYIKKSNFXXRGILYSKFLSZ--------------------SKMT


HERVI_polputein                     VLPQG-------------------------------FMDSPNLFGQIL--
HERVIP10F_polputein                 VLPQG-------------------------------FTDSPNLFGQIL--
HERVadp_AC005741                    VLPQG-------------------------------FK-----FVQIL--
HERV-PT47D                          SLAQG-------------------------------FTDSPNIFGEIL--
HERV_W_chr6_141432567_ERV9_lik      ALP-G-------------------------------FRDSAHLFDQ----
ERV9_PH1_RT                         VLPQG-------------------------------FRDSPHLFGQ----
Chr13_54564403_572_po_C_Yz7/14      VLQ-G-------------------------------FKDSPHLFDQ----
HERV_Y_chr12_51022911_pol           VLPQE-------------------------------FRDSPHLFGQ----
HERV17_polputein                    VLPQG-------------------------------FRDSPHLFGQ----
HERV_W_chr7_9105739_syncytin_p      ILPQG-------------------------------FRDSPHLFGQ----
HERV9_polputein                     VLPQG-------------------------------FRDSPHLFGQ----
```

```
HUERS-P3_polputein                  VLPQG-------------------------------FZDSPHLFGQ----
ERVfrd_AC004022                     VLPQG-------------------------------FRDSPHLFGQ----
PRIMA41_polputein                   VFPQE-------------------------------CRDSPHLFGN----
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       VMLQG-------------------------------FRGSPYYFSQ----
HERV_H48I_polputein                 VLPQG-------------------------------FRDSPHLFGQ----
HERV_H_RTVLH2_pol                   VLPQG-------------------------------FTDRPHYFS-----
HERVFH21_pol                        VLPQG-------------------------------FQDSLLFFGQ----
HERVRblike_chr4_109047953           VLPQG-------------------------------FKNSPSIFGE----
ChrY_13790029_693_po_C_Xz12/11      QLPQG-------------------------------FKNSPTLFK-----
HERVS71_polputein                   WLPQG-------------------------------LKNSPTLFG-----
BaEV_M7_pol                         RLPQG-------------------------------FKNSPTLFD-----
HERVR_polputein                     RLPQG-------------------------------FKNSPTLFD-----
AKR_MLV_MLOCG_RT                    RLPQG-------------------------------FKNSPTLFD-----
GaLV_pol                            RLPQG-------------------------------FKNSPTLFD-----
Chr3_158587414_1662_po_C_Ez10/      WLPQG-------------------------------FKNSPTIFR-----
Chr3_79627623_1138_po_C_Ez4/7       RLPQG-------------------------------FKNSPTIFG-----
Chr19_20477244_1350_po_C_Ez10/      WLPQG-------------------------------FKNSLTIFG-----
Chr5_111890459_1626_po_C_Ez10/      QLTQG-------------------------------FENSPTIFG-----
ChrX_48766518_1600_po_C_Ez6/3       QLPQG-------------------------------FKNSLTIFG-----
ChrX_152580713_1572_po_C_Ez10/      WLPQG-------------------------------FKISPTIFG-----
ChrX_139287533_1154_po_C_Xz8/4      WLPQG-------------------------------FKNSPIIFGE----
Chr2_52615786_912_po_C_Xz11/5       WLPQG-------------------------------FKNSPTIFG-----
HERV-E_4_1                          QLPQR-------------------------------FKNSPTIFG-----
ChrY_21962399_1137_po_C_Xz13/9      RLPQG-------------------------------FKYSPTIFG-----
ChrY_8634032_950_po_C_Xz14/10       LLPQG-------------------------------FKNSTTIFG-----
Chr3_126735360_997_po_C_Xz5/10      GLPQG-------------------------------FKNSPTIFR-----
Chr3_126713203_958_po_C_Xz7/8       GLPQG-------------------------------FKNSPTIFG-----
Chr3_75226118_977_po_C_Xz8/9        GLPQG-------------------------------FKNFPTIFG-----
Chr7_6629919_827_po_C_Xz8/7         GLPQG-------------------------------FKNSPTIFG-----
Chr4_9166812_811_po_C_Xz8/10        GLPKG-------------------------------FKNSPTIFG-----
Chr12_8478642_834_po_C_Xz12/7       GLPQG-------------------------------FKNSPTIFR-----
Chr10_15187175_787_po_C_Xz11/5      GLPHE-------------------------------FKNFPTILG-----
Chr11_71819740_854_po_C_Xz11/9      GLPQG-------------------------------VKNSPTIFG-----
Chr11_71519974_886_po_C_Xz13/9      RLPQG-------------------------------FKNSPTIFG-----
Chr8_7271441_969_po_C_Xz10/6        GLPQR-------------------------------FKNSPTIFG-----
Chr8_11770066_826_po_C_Xz9/8        RLPQG-------------------------------FKNSPTIFG-----
Chr4_4128362_651_po_C_Xz7/11        GLPQG-------------------------------FKNSPTIFG-----
Chr14_50203955_850_po_C_Xz10/8      GLPQE-------------------------------FKNSPTTFG-----
Chr14_50205291_837_po_C_Xz10/8      GLPQE-------------------------------FKNSPTTFG-----
Chr7_6597120_842_po_C_Xz7/7         GLPQG-------------------------------FKNSPTIFG-----
Chr11_71544658_783_po_C_Xz8/13      GLPQG-------------------------------FKDSPTIFG-----
Chr11_67999520_760_po_C_Xz8/8       GLPQG-------------------------------FKNSPTIFG-----
Chr3_75461975_796_po_C_Xz7/9        GLPQG-------------------------------FKNSPTIFG-----
Chr12_52228939_692_po_C_Xz3/10      RLPQG-------------------------------FQSSPTIFR-----
Chr10_15211234_286_po_C_Xz11/1      HPKVH-------------------------------VZ------------
Chr12_100973302_945_po_C_Xz12/      RLPQG-------------------------------FKNSPTIFR-----
Chr19_9685420_817_po_C_Xz6/10       WLPQG-------------------------------FKNSPTIFR-----
Chr13_40936820_719_po_C_Xz9/10      GLPQG-------------------------------FKNSPTIFG-----
Chr7_63858202_1754_po_C_Rz12/9      SLPQG-------------------------------FKNSPTIFG-----
ERV3_chr7q21_AC073210_polputei      SLPQG-------------------------------FKNSPTIFG-----
Chr6_110218096_843_po_C_Xz16/8      RLPQG-------------------------------ZKNSPTIFA-----
Chr6_110214535_786_po_C_Xz16/8      RLPQG-------------------------------ZKNSPTIFA-----
ChrY_3018841_881_po_C_Rz9/12        RLPQG-------------------------------FKKSPTIFG-----
ChrX_86876098_639_po_C_Xz7/8        RLPQG-------------------------------FKKSPTILG-----
Chr14_68899917_965_po_C_Xz6/8       KLPQG-------------------------------FKNSPTIFE-----
ChrY_23874437_661_po_C_Xz15/9       RLPQR-------------------------------FKNTPTIFG-----
ChrY_27518418_649_po_C_Xz15/9       RLPQR-------------------------------FKNTPTIFG-----
ChrY_25016892_657_po_C_Xz13/12      RLPQR-------------------------------FKNTPTIFG-----
ChrY_19932937_520_po_C_Xz10/14      RLPQ---------------------------------------FG-----
ChrY_20104614_518_po_C_Xz10/14      RLPQ---------------------------------------FG-----
Chr19_58660155_598_po_C_Xz21/1      RLPQR-------------------------------FKNSPTLSHNLPQS
Chr19_58621153_500_po_C_Xz21/1      RLRQENGVNQGGRAYSEPRPRQCTPLGDRARLHLKNNNNNKNNYHNLPQF
Chr7_139686987_710_po_C_Xz11/7      RFLQG-------------------------------FKNFPTIFG-----
Chr7_137651301_1097_po_C_Pz13/      RLPQG-------------------------------FKKSPTIFN-----
Chr11_29579197_930_po_C_Pz8/12      RLPQA-------------------------------LKNSLTIFEE----
Chr4_117329205_488_po_C_Xz3/16      EKSAL-------------------------------LSATKTKGK-----
Chr21_38522185_383_po_C_Xz7/9       EKSAS-------------------------------LSATETKGK-----
Chr4_70173370_1049_po_C_Xz11/8      RLPQG-------------------------------FKNSPTIFEE----
Chr4_70022513_788_po_C_Kz10/12      RFPQG-------------------------------FKNSTTIFEE----
Chr14_105225894_1027_po_C_Xz13      RLPQG-------------------------------FK-SPTIFEE----
Chr7_119241536_761_po_C_Xz11/1      RLPRG-------------------------------FKNSLTIFEE----
ChrY_22564970_567_po_C_Xz9/10       RLPRG-------------------------------FKNSPTIFG-----
Chr21_14409346_1477_po_C_Pz5/6      RLLQG-------------------------------FKNSPT------IF
ChrX_51259700_1396_po_C_Pz7/8       RLPQG-------------------------------FKNSP-------IF
Chr6_64985396_1041_po_C_Xz11/1      RLHQG-------------------------------LKNSP------TIF
ChrY_20355209_659_po_C_Xz7/11       KLPQG-------------------------------FKNSP------TIF
```

```
Chr3_34567166_523_po_C_Xz8/6        QASQG-------------------------------FKCSP------TIF
Chr3_122072468_245_po_C_Xz9/9       RLPQG-------------------------------FKNPP------IIF
Chr11_58788475_541_po_C_Xz8/22      QQALP-------------------------------IEANPG----WDLN
Chr6_137421042_373_po_C_Xz6/9       QQALL-------------------------------IRLTQA----GTLT
Chr4_115910608_550_po_C_Pz8/12      SLDRD-------------------------------LRAGQG----ZKKI
ChrY_13743185_977_po_C_Xz14/8       RLPQG-------------------------------FKNSP------IIF
HERV15                              RLPQG-------------------------------FKNSP------IIF
ChrX_117198092_983_po_C_Xz8/8       RLPZG-------------------------------FKNS-------TIF
Chr19_21166030_1412_po_C_Xz9/1      RLPQG-------------------------------FKNSP------TIF
Chr19_21213743_913_po_C_Xz14/1      RLPQG-------------------------------FENS-------TIF
Chr19_15957774_569_po_C_Xz8/10      KLPQG-------------------------------FKNSLTIFE-----
Chr11_77508681_297_po_C_Xz19/1      RLPQG-------------------------------FKNSPTIFG-----
Chr13_98250769_738_po_C_Xz7/13      RVPKR-------------------------------FKNSPTIFK-----
Chr6_39968183_513_po_C_Pz10/9       RLPQG-------------------------------FKNSPMIFE-----
Chr15_66867263_645_po_C_Xz11/1      RLPQE-------------------------------FKNSPTIFE-----
HERV18_polputein                    VLPQG-------------------------------YMHSPTICHG----
HERVIP10FH_polputein                GLNSH---------------------------------------------
MER66_polputein                     GYLRKRZ-----------------------------YRTSKKXQACCRWS


HERVI_polputein                     ------EQVLDKVSVPKQLCLLQYVDDILISGEDIEKEAGFSTHIFDHLQ
HERVIP10F_polputein                 ------EQVLEKVVIPEQICLLQYVDDILISGEDIEKVTDFSTHILNHLQ
HERVadp_AC005741                    ------ENVLEESQTSRGTQLLQYVDYLLISGEKRARVSETTTNLLNFLG
HERV-PT47D                          ------EQALEKVFIPEQICLLQYMDDLLISGKDIEKVTNFSTHILNHLQ
HERV_W_chr6_141432567_ERV9_lik      ----ALAQDLGHFS-SPGTLVLQYVVDLLLATSSEALWQQATLGLLNFLA
ERV9_PH1_RT                         ----ALAKDLGHFS-SPGTLVLQYVDDLLLATSZEALCQQATLDLLNFLA
Chr13_54564403_572_po_C_Yz7/14      ----ALAQDLSQFS-YMDTLVLRYGDDLILATHSETLCHQATQALLNFLA
HERV_Y_chr12_51022911_pol           ----ALAQDLSQFS-YLDTLVLQYMDDLLLATRSETVCHQATQELLNFLA
HERV17_polputein                    ----ALAQDLSQFS-YLDTLVLRYVDDLLLAARSETLCHQATQALLNFLA
HERV_W_chr7_9105739_syncytin_p      ----ALPR-LEPIL-IPGHLSFG-VDDLLLAAHSETLCHQATQALFNFLA
HERV9_polputein                     ----ALAQDLGHFS-SPGTLVLQYVDDLLLATSSEASCQQATLDLLNFLP
HUERS-P3_polputein                  ----ALSKDLSEFS-HPQVKVLQYVDDILLCAPTEEASQEGTEALLNFLA
ERVfrd_AC004022                     ----ALARYLQDLSLYMGGHLLQYMDDLLICSPTRELG----IQHLNFLA
PRIMA41_polputein                   ----TLTRELKMLK-LNRGTSIWYVNDLLVASPTKRGSNKNAIKLLNFLD
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       ----ALSHDLFLH---PSLLLIQYIGDVLLVAPP--ELLNKTHFLLQHLF
HERV_H48I_polputein                 ----ALARDLCTLS-LKPSTLLQCVNDLLLCSPSQRDCNAHTLSLLNFLA
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        ----ALAQDLASLD-LSPSRLLQYLDDLLLCSPSLKNSQTHTATLLNFLT
HERVRblike_chr4_109047953           ----ALIQDLRGLQ-LENGVLLQYVDDLLISSTSKQECQENTVKTLNHLA
ChrY_13790029_693_po_C_Xz12/11      -EALQQDLIPFQ-ART-LIVLLQYTDDLLLGTETTDSCLQHTKDLFYLLQ
HERVS71_polputein                   -EALQQDLIPFZ-ASHPNCTLLQYIDDLFLATETTDSCLQHTRDLLYLLQ
BaEV_M7_pol                         -EALHRDLTDFR-TQHPEVTLLQYVDDLLLAAPTKKACTQGTRHLLQELG
HERVR_polputein                     -EALHRDLTDFR-TQHPEVTLLQYVDDLLLAAPTKKACTQGTRHLLQELG
AKR_MLV_MLOCG_RT                    -EALHRDLADFR-IQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLG
GaLV_pol                            -EALHRDLAPFR-ALNPQVVLLQYVDDLLVAAPTYEDCKKGTQKLLQELS
Chr3_158587414_1662_po_C_Ez10/      -EALARDLQKFP-AKHLGCVLLLYVDDLLLGHSTAVWCAKGMDALLRHLE
Chr3_79627623_1138_po_C_Ez4/7       -EALARDLQKFP-AKDLGCVLLMYVDDLLLGHSMAAGCAKGTDALFRHLQ
Chr19_20477244_1350_po_C_Ez10/      -EALTZDLQKFP-AKDLSYVLLLYVDDLLLGHSMAVGCSKG-DALFQHLE
Chr5_111890459_1626_po_C_Ez10/      -EALARDLQKFP-AKDLGCILLLYMDDLLLGHSTAVRCAKGMDALLQHLE
ChrX_48766518_1600_po_C_Ez6/3       -EALAZDLQKFP-AKDLG--LLQYTDDLLLGHSTAVGCTKGTDALLWHLE
ChrX_152580713_1572_po_C_Ez10/      -EALAZDLQKFP-AKDLGCVLLLYVDYLLLGHSMAVGCAKGTDALLQHLE
ChrX_139287533_1154_po_C_Xz8/4      -QALAQDLQKFP-AKDLGCVLLQYVDDLLLGHSMADRCAKGMVALLRHLE
Chr2_52615786_912_po_C_Xz11/5       -EALAQDLQKFP-TRDLGCVLLQYVNNLLLGHPTAVGCAMRKDTLIQHLE
HERV-E_4_1                          -EALARDLQKFP-TRDLGCVLLQYVDDLLLGHPTSR-VAKGTDALLRHLE
ChrY_21962399_1137_po_C_Xz13/9      -EALAQDPQKFP-SRDLGCVLLQYVDDRLLGQPMAVWCTKG-ITLLQHLE
ChrY_8634032_950_po_C_Xz14/10       -EALVZNLQKFP-PEDLGCILILYVDNHLLGHSMESGVQKG-RVLLWHLZ
Chr3_126735360_997_po_C_Xz5/10      -EALARDPQKFP-TRDLGCVLLQYVDDLLLGHPTAVGCAKGTDALHRHLG
Chr3_126713203_958_po_C_Xz7/8       -EALARDPQKFP-SRDLGCVLLQZVDDLLLGHPTAVGCAKGTDALHRHLE
Chr3_75226118_977_po_C_Xz8/9        -EAWARDLQNFP-SRGLGCVLLQZVDDLLLGHPTAFGCAKGTDALHQHLE
Chr7_6629919_827_po_C_Xz8/7         -EAWARDLQKFP-SRDLGCVLLQZVDDLLLGHPTAVGCAKGTDALHRHLE
Chr4_9166812_811_po_C_Xz8/10        -EAWARDIQKFP-SGDIGCMLLQZVDDLLLGHPTAVGCAKGTDALHRHLE
Chr12_8478642_834_po_C_Xz12/7       -EAWARELQKFP-SRDLGCVLLQZVDDLLLGHPTAVGCAKGTDALHWHLE
Chr10_15187175_787_po_C_Xz11/5      -DAWARDLQKFP-SRDLGCVLLQZVDDLLLGHPTAVGCAKGIDALHRHLE
Chr11_71819740_854_po_C_Xz11/9      -EAWARDLQKFP-SRDLGCVLLZZVDDLLLGHPTAVGCAKGTDALHRHLE
Chr11_71519974_886_po_C_Xz13/9      -EAWARDLQKFP-SRDLGCVLLQZVDDLLLEHPTAVRCAKGTVVLHWHLE
Chr8_7271441_969_po_C_Xz10/6        -EAWDRDLQKFP-SRDLGCLLLQZVDDLLLGHPTAVGCAKGTDALHRTWR
Chr8_11770066_826_po_C_Xz9/8        -EAWAQDLQKFP-SRDLGCVLLQZADDLLLGHPTAVGCAKGKDALHRHLE
Chr4_4128362_651_po_C_Xz7/11        -EALARDLQKFP-TRDLGCVLLRZVDELLLGHPTAVGCAKGTDALRRHLD
Chr14_50203955_850_po_C_Xz10/8      -EALARDLQKFP-TGDLGCALLRZVDELLLGHPTAVGCAKGRDCPTLTPG
Chr14_50205291_837_po_C_Xz10/8      -EALARDLQKFP-TGDLGCALLRZVDELLLGHPTAVGCAKGRDCPTLTPG
Chr7_6597120_842_po_C_Xz7/7         -EAWARDLQKFP-TRDLGCALLRZVDELLLGHPTAVGCAKGMDCPMPTPG
Chr11_71544658_783_po_C_Xz8/13      -GALARDLQKFP-SRDLGCALLRZVDELLLGHPTAVGCAKGTDALRRHLE
Chr11_67999520_760_po_C_Xz8/8       -EALAZDLQKFP-TRDLGCALLQZVDELLLGHLMAVGCTKGTDALRRHLE
Chr3_75461975_796_po_C_Xz7/9        -EALARDLQKFP-TRDLGCMLLQZVDDLLLGHPTAVRCAKGTDALHWHLE
Chr12_52228939_692_po_C_Xz3/10      -EALATDLQKLP-TRDLGCVLLQYVDDLLLRHPTAVRCAKGTDALLQRLE
Chr10_15211234_286_po_C_Xz11/1      ---VARDLQKFP-TRDLGCVLLWZVDELWLGHPTAVGCTKGTDALRPHLE
```

```
Chr12_100973302_945_po_C_Xz12/      -EALAGDLQKFH-AKDPGCIZLLYVDDLLLEHSTAVGCAKGMDALLRDLV
Chr19_9685420_817_po_C_Xz6/10       -EALARDLQKLP-AEDLGCVLLLYMDDVLLGHSTAVGCTKGTDALLWHLE
Chr13_40936820_719_po_C_Xz9/10      -EALARDLQKFP-TRDLGCVLLQZVDDLLLGHPKAVRCAKGTDALHRHLE
Chr7_63858202_1754_po_C_Rz12/9      -EALASDLKAYTPPNDN-CALLQYIDNLLLAAPTQEDCYWGTQDLLHLLW
ERV3_chr7q21_AC073210_polputei      -EALASDLKAYTPPNDN-CALLQYIDNLLLAAPTQEDCYWGTQDLLHLLW
Chr6_110218096_843_po_C_Xz16/8      -EALAADLKAYTPPDDN-CTLLQYVDDLLLMAPTRKDCYWGTQGLPHLLZ
Chr6_110214535_786_po_C_Xz16/8      -EALAADLKAYTPPDDN-CTLLQYVDDLLLMAPTRKDCYWGTQGLPHLLZ
ChrY_3018841_881_po_C_Rz9/12        -EAVP-RTQGLHPANNN-CALLNYVDDLLLVAPTHEDCYQRTQDLLHFVC
ChrX_86876098_639_po_C_Xz7/8        -EALP-RTQGLHPANNN-CALLNYVDDLLLVAPTREDCYQRTQDLLHFVW
Chr14_68899917_965_po_C_Xz6/8       -EALASDLKAYTLPNNN-CALLQYIDDILLAAPTHEDCYQGTQDFLYLLW
ChrY_23874437_661_po_C_Xz15/9       -DTLASDIKAFTTPNDN-YTLLQYINDLLLAAPTPGGRYQGIQDFLHLLY
ChrY_27518418_649_po_C_Xz15/9       -DTLASDIKAFTTPNDN-YTLLQYINDLLLAAPTPGGRYQGIQDFLHLLY
ChrY_25016892_657_po_C_Xz13/12      -DTLASDIKAFTTPNDN-YTLLQYINDLLLAAPTPGGRYQGIQDFLHLLY
ChrY_19932937_520_po_C_Xz10/14      -NTLESDLKAHPHQMT--TVLLQYVNNLLLADPTPGDCYQGTQVFLHHVY
ChrY_20104614_518_po_C_Xz10/14      -NTLESDLKAHPHQMT--TVLLQYVNNLLLADPTPGDCYQGTQVFLHHVY
Chr19_58660155_598_po_C_Xz21/1      LEKHWSDLMAYTLPNSN-CTLLQYADELLLPAPIZEDCYRGTQDLFHFLZ
Chr19_58621153_500_po_C_Xz21/1      RETLASDLMAYTLPSNN-CALLQYIDELSLPAPTZEECYRGTQDLLHLLZ
Chr7_139686987_710_po_C_Xz11/7      -EALASDLKAYTLPNDN-CASLQYVDDLLLAAPTQEDCYQGTQDLLHLLZ
Chr7_137651301_1097_po_C_Pz13/      -------LKAYTPPNDN-CTLLQYIDDLFLAAPTQEDCFQGTQDLLHLL-
Chr11_29579197_930_po_C_Pz8/12      --ALASDLKAZ--PNDN-CALLEYIDDLLPADPTQEDCLQGIQDLLHLL-
Chr4_117329205_488_po_C_Xz3/16      -------------PQEK-PVLQELSEETEILPPYAPIYPLYQGRLPR-S-
Chr21_38522185_383_po_C_Xz7/9       -------------PQEK-PVLQESPEEIGILP--TVIYPPLLRPAPEES-
Chr4_70173370_1049_po_C_Xz11/8      --AQASDLKAYTPPNSD-CVLLQYIDDLLLAAPTZEDCFEGTQDLLHLL-
Chr4_70022513_788_po_C_Kz10/12      --ALASDLKVYTPPNSY-CVLLQYIDDLLLAAPTQEDCFQGTQDLLHLL-
Chr14_105225894_1027_po_C_Xz13      --ALASDLKAYTPPSSD-CVLLZYIDDLLLAAPTWKDCFQETQDLLHLLG
Chr7_119241536_761_po_C_Xz11/1      --ALASDLKAYTLPNYN-CALLEYIDDLLLAVPTQKDCFQGTQDLLYLL-
ChrY_22564970_567_po_C_Xz9/10       -EVLASDLGAFIPPNDK-CFZLQFIDNLLFMVPTQVDCLQGTEKLLHLLW
Chr21_14409346_1477_po_C_Pz5/6      EEALASDLKAFTPPSDP-CILLPYIDDLLLATPTRKECMQGTESLLZALW
ChrX_51259700_1396_po_C_Pz7/8       EEGLVSDLKAFTPPSDR-CILLQYIDYLLLAAPTRKECIEGTENLLZVLW
Chr6_64985396_1041_po_C_Xz11/1      KEALP-DLKAFMPPSDR-RVLLQYIDDLLLAAPTKEECFQGTESFLRVLW
ChrY_20355209_659_po_C_Xz7/11       EEALASDLKAFMPPSGH-CVLLQYIDDL---APTKEECFQGTEALH-VLW
Chr3_34567166_523_po_C_Xz8/6        EEALASDLKAFVPPSDR-CVLFQYMDDLLLAAPTKEECFQGTESLLRVLW
Chr3_122072468_245_po_C_Xz9/9       EEALT-DLKAFVPPSDP-CVLLQYMDDVLLAAPTKEECFQGTESLLRVLW
Chr11_58788475_541_po_C_Xz8/22      QAQGLQSLQQYZEVLTW-NKGWREKGHQYQKCLKRSARSQESPSEFCERL
Chr6_137421042_373_po_C_Xz6/9       RLKGYKFAAVLRGTPNG-INAGEKKATNIGKVSEVRQKPDKSPGEFYERL
Chr4_115910608_550_po_C_Pz8/12      NVNGAKGKDTERMNQEN-KENGQD--HGIRKPLQGLLHPGKTQNLLHLLW
ChrY_13743185_977_po_C_Xz14/8       EDALATNLEAFAPFSDN-SVVLQYIHDLLFAAPRREEYLQGIERLLHLLC
HERV15                              EDALATNLEAFAPFSDN-SVVLQYIHDLLFAAPRREEYLQGIERLLHLLC
ChrX_117198092_983_po_C_Xz8/8       EEALATDLEAFAPPSDN-YVLLQHIDDFLFAAPTREEZLQGTERLLHVLC
Chr19_21166030_1412_po_C_Xz9/1      KEALATDLEAIAPSSDN-SVLLQYIDDLLLTAPKREECLQGTER-LDLLC
Chr19_21213743_913_po_C_Xz14/1      EEALATDLEAFLPPSDN-CVLLZYIDDLLFTAPMREECFQGTERLLHLLR
Chr19_15957774_569_po_C_Xz8/10      -EALASNLKAYTLPNDN-CALFQYVDDLLLAAPTREDYYQGTQDLLHLLW
Chr11_77508681_297_po_C_Xz19/1      -ETLASDLKAYTPPND--LCLLQYTGDLLLAAPTQEDYYQRTQDLLHLLW
Chr13_98250769_738_po_C_Xz7/13      -EALASDLRPIPHQMITNGALLWYMDDLLLVAPTZEDCYQGTRDLLHLLG
Chr6_39968183_513_po_C_Pz10/9       -EVLASDL-AYVPPNDK-CILLQCIENLLFAAPTZVDCLQVTKKLLHLLW
Chr15_66867263_645_po_C_Xz11/1      -EALASDLKAYTAPNSD-CVLLHYIDDLLLAAPTREDCFQGTQDLLHLLW
HERV18_polputein                    --LVATDLAAWKCPKG--VHLFHYIDDIMLTSDSLADLEAAAPLLRQHLA
HERVIP10FH_polputein                ----KPPVVRKIFRTPLSFLLREERVATPAGPSPRGRGRRGENSSISGWQ
MER66_polputein                     VXVMIRFMKGNLZKEFLCMIKLAIIKRKLFTIDFLKNGYVXKPNFLKVLI


HERVI_polputein                     FE-------GLRVSK------GKLQCMEPKVKYLGRLISAG-KRRIGPEW
HERVIP10F_polputein                 FE-------GLRVSK------RKLQYVEPEVKYLGHLISAG-KRRIGPER
HERVadp_AC005741                    ER-------GLZVSK------NKMQFVEKEVKYLGHLINEG-KWRINPER
HERV-PT47D                          FE-------ELRVSK------KKLQYVEPEVKYLGHLISAG-KZRIGPER
HERV_W_chr6_141432567_ERV9_lik      NQ-------GYKVSK------SKAQFCPQQVKYLGLILARG-TRALSKER
ERV9_PH1_RT                         NQ-------GYKVSM------SKAQLCLQQVKYLGLILAKG-TRALSKEZ
Chr13_54564403_572_po_C_Yz7/14      TC-------GYKVSK------LKAQLCSQQVKYLGLKLSKG-TRALSEER
HERV_Y_chr12_51022911_pol           TN-------GYMVSK------TKSQLCSQQVKYLGLKLSKG-NRALSEEC
HERV17_polputein                    TC-------GYKVSK------PKAQLCSQQVKYLGLKLSKG-TRALSEER
HERV_W_chr7_9105739_syncytin_p      TC-------GYMVSK------PKAQLCSQQ--YLGLKLSKG-TRALSEEH
HERV9_polputein                     NQ-------GYKASR------SKVQLCLQQAKYLGLILARG-TRALSKEZ
HUERS-P3_polputein                  DR-------GYKVSK------SKAQLCXTSVKYLGLVLSEG-TRALGEER
ERVfrd_AC004022                     DR-------RYKVSK------AKTQLLRQAIQYLGIIMSLK-EHKLSAEZ
PRIMA41_polputein                   TN-------GYRVSL------HKAHVSTQEVKYLRYVLTPG-TWAIAPRK
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       SK-------GYPS---------KAQMSSPSVTYLGIILHKN-TRALPADS
HERV_H48I_polputein                 ER-------GYQVSP------KKAQICTPSVTYLGLALTPZ-TRGLTTDR
HERV_H_RTVLH2_pol                   ----------------------QAQVSSSSVTYLSIILIK--THVLSLLI
HERVFH21_pol                        NK-------GYRVSS-------KEQLSTSMVTYLGIQLSPG-ARAMTPAR
HERVRblike_chr4_109047953           AC-------EYKVSS------KKAQICKQTIEYLGFLLQE--QRALTVER
ChrY_13790029_693_po_C_Xz12/11      EL-------RYQVSA------KSQLCLP-RVSYLGYEINRG-KRALISAC
HERVS71_polputein                   EL-------GYWVSA------KKAQLCL-QTFYLGYKINKG-ERALATAR
BaEV_M7_pol                         EK-------GYRASA------KKAQICQTKVTYLGYILSEG-KRWLTPGR
HERVR_polputein                     EK-------GYRASA------KKAQICQTKVTYLGYILSEG-KRWLTPGR
AKR_MLV_MLOCG_RT                    NL-------GYRASA------KKAQICQKQVKYLGYLLKEG-QRWLTEAR
GaLV_pol                            KL-------GYRVSA------KKAQLCQREVTYLGYLLKEG-KRWLTPAR
```

```
Chr3_158587414_1662_po_C_Ez10/      DY-------GYKVSK------KKAHVCRQQVRYLGFTIRKG-ERSLGSER
Chr3_79627623_1138_po_C_Ez4/7       DC-------EYKVSK------KKAQICRQQVCYLGFTIWKG-GHSLGSEK
Chr19_20477244_1350_po_C_Ez10/      DC-------GYKVPK------KKSQICRQQVRYLGFTIWKG-EHSLWSER
Chr5_111890459_1626_po_C_Ez10/      DC-------GYKVSK------KKAQICRQQVRYLGFTIQKG-ECSLGSZR
ChrX_48766518_1600_po_C_Ez6/3       DC-------GYKVSK------KKAQICRQQVHYLGFTIQKG-ERSLESER
ChrX_152580713_1572_po_C_Ez10/      DC-------GYKVSK------KKAQICRQQVSYLGFTVQKG-ECSLGSER
ChrX_139287533_1154_po_C_Xz8/4      DC-------RYKVSK------KKAQICRQZVRYLGFTIRKG-ERSLGSER
Chr2_52615786_912_po_C_Xz11/5       DC-------GYKVSK------KRAQNCIQQVRYLGFTIGHR-EHSLGSER
HERV-E_4_1                          DV-------G-KVSK------KKAQICRQQVCYLGFTIQQG-EHSLGSER
ChrY_21962399_1137_po_C_Xz13/9      NC-------GHKVSK------KKAQLSRQQVRQLGYYLTRG-AQ---PER
ChrY_8634032_950_po_C_Xz14/10       DH-------GYKVSK------KKSQICRQQVCYLGFTIQKE-ERGL-VRK
Chr3_126735360_997_po_C_Xz5/10      LW-------VEGVQK------K-AQICRWQVSYLGLTIQQGSEHSPGSER
Chr3_126713203_958_po_C_Xz7/8       DC-------GZKVSK------KKAQICRRQVSYLGLTIRQGSERSPGSER
Chr3_75226118_977_po_C_Xz8/9        DG-------GZKVSK------KKAQICRQQVRYLGFTIRQGSERSPGSEZ
Chr7_6629919_827_po_C_Xz8/7         DC-------GZKVSK------KKAQICRQQVRSLGLTILQGSERSLGSEZ
Chr4_9166812_811_po_C_Xz8/10        DC-------GEGVQE------ESSDLL-AAVTFLGIDYLTGVGTQPGIRM
Chr12_8478642_834_po_C_Xz12/7       DC-------GZKVSK------KKAQICZQQVGYLGLTIQQGSERSPGSER
Chr10_15187175_787_po_C_Xz11/5      DC-------GZKVSK------KKAQICRQQVGYLGLTIRQGSERNLGSER
Chr11_71819740_854_po_C_Xz11/9      DC-------G-EVSK------RKAQICZQKVCYLGLTIRQGWEGSPGSER
Chr11_71519974_886_po_C_Xz13/9      DC-------GYQLSK------KKAQICRQQVSFWGLTIZQGSERSPGSER
Chr8_7271441_969_po_C_Xz10/6        TV-------G-RVSK------KKAQICRQQVRYLGFTIRQGSERSPRSER
Chr8_11770066_826_po_C_Xz9/8        DC-------GZKVSK------KKAQICZQQVRYLGFTIZQGSEHSPGSER
Chr4_4128362_651_po_C_Xz7/11        DC-------GCKVSK------KKAQICRQQVRSLGYTIRQG-ERSPGSER
Chr14_50203955_850_po_C_Xz10/8      GL-------GYKLSK------KKAQICRQQVRSLGFTIZQG-ERSLGSER
Chr14_50205291_837_po_C_Xz10/8      GL-------GYKLSK------KKAQICRQQVRSLGFTIZQG-ERSLGSER
Chr7_6597120_842_po_C_Xz7/7         GL-------GCKLSK------KKAQICRQQVRSLGFTIRQG-ECSPGSER
Chr11_71544658_783_po_C_Xz8/13      DC-------GYKVSK------KKAQICRQQVRSLGFTIRQG-ERSPGSER
Chr11_67999520_760_po_C_Xz8/8       DC-------GCKVSK------KKAQICZQQVCSLGFSIRQG-ERSPGSER
Chr3_75461975_796_po_C_Xz7/9        GC-------GZKVSK------KKAQICRQQVRYLRFTIQQGSERSPGSER
Chr12_52228939_692_po_C_Xz3/10      DC-------GYKVSK------KKAEICRQQVRYLGFTIRHG-ECSLGSER
Chr10_15211234_286_po_C_Xz11/1      DC-------GYKVSK------KKAQICZQQVCSLGYTIZQG-ECSLVSET
Chr12_100973302_945_po_C_Xz12/      D-------------K------KKAQICRQQVRYLGFTIQKG-EHSLGSER
Chr19_9685420_817_po_C_Xz6/10       DC-------GYKVSK------KKAKTCRQHVRYLGFTIRKG-EHTL-AER
Chr13_40936820_719_po_C_Xz9/10      DC-------GYKVSK------KKAQICRQQLRYLGFTIRQG-ERSPGSER
Chr7_63858202_1754_po_C_Rz12/9      -------KAGYRVSK------KKAQICHEKVKYLGFIVSQG-ERWLGHGZ
ERV3_chr7q21_AC073210_polputei      -------KAGYRVSK------KKAQICHEKVKYLGFIVSQG-ERWLGHGZ
Chr6_110218096_843_po_C_Xz16/8      -------KARYKVSK------RKAQICQQKVKYLGFIVTQG-EHQLGIEZ
Chr6_110214535_786_po_C_Xz16/8      -------KARYKVSK------RKAQICQQKVKYLGFIVTQG-EHQLGIEZ
ChrY_3018841_881_po_C_Rz9/12        -------KAGYKVSR------KKA-ICQERVKYVGFIVSQG-ERWLGNEZ
ChrX_86876098_639_po_C_Xz7/8        -------KAGYKVSR------KKA-ICQERVKYVGFIVSQG-ERWLGNEZ
Chr14_68899917_965_po_C_Xz6/8       -------KAGYKVSR------KKAQICQERVKYLGFLVSQG-ECQLSNEQ
ChrY_23874437_661_po_C_Xz15/9       -------GKLVIRIQ------KEGQICQESVKYLDFIVSQW-EHQLSSEQ
ChrY_27518418_649_po_C_Xz15/9       -------GKLVIRIQ------KEGQICQESVKYLDFIVSQW-EHQLSSEQ
ChrY_25016892_657_po_C_Xz13/12      -------GKLVIRIQ------KEGQICQESVKYLDFIVSQW-EHQLSSEQ
ChrY_19932937_520_po_C_Xz10/14      -------RKGYKVFR------KKAQFCQESIKNLGFIVSQG-EHWLNSEQ
ChrY_20104614_518_po_C_Xz10/14      -------RKGYKVFR------KKAQFCQESIKNLGFIVSQG-EHWLNSEQ
Chr19_58660155_598_po_C_Xz21/1      -------KISYKISK------KQAQICLEEVPYLGVIVSHK-EHWLGYEW
Chr19_58621153_500_po_C_Xz21/1      -------KASYKISK------KKAQIFHEEVTYLGVILSHE-EHWFGYER
Chr7_139686987_710_po_C_Xz11/7      SRLQSTZKESSNFKT------KQKPICHEEVKYLGFIVSQG-ERRLGDER
Chr7_137651301_1097_po_C_Pz13/      -------YK---VSR------KKAQICSRRVQYLGFEVSQG-EQWLGSER
Chr11_29579197_930_po_C_Pz8/12      -------GKPDIRVQ------KEAQICSKRVQYLGFSESRG--MIARQZA
Chr4_117329205_488_po_C_Xz3/16      -------QTQD--NQ------TRASPQKEKSKPLPQEVKEE-----SQDD
Chr21_38522185_383_po_C_Xz7/9       -------NSNG--NE------PWALPQKEKSKPLP-QVKEE-----VRMI
Chr4_70173370_1049_po_C_Xz11/8      ------WKAGYKVSG------NKSQICSESDQYLGFYISQR-KRWLGSEQ
Chr4_70022513_788_po_C_Kz10/12      ------WKAGYKVSG------KKAQICFESFQYLGFYISZG-KRWLGCEW
Chr14_105225894_1027_po_C_Xz13      RQD---IKAGYKVSG------KKDQICSESVQYLGFYISEG-KRLLGS-M
Chr7_119241536_761_po_C_Xz11/1      ------WKAGYKVFR------KKAQLR--KVQYLGFIVSQR-ERWLSLEQ
ChrY_22564970_567_po_C_Xz9/10       -------KAKYKVSK------KKAHICSEGVIYLGFEISEG-KCELGSER
Chr21_14409346_1477_po_C_Pz5/6      KA-------GYKVSK------EKAQICGQGVQYLGFYISQG-QHELGWEQ
ChrX_51259700_1396_po_C_Pz7/8       EA-------GYKVSK------KKAHICGQGVQYLGFPIPQG-RRELVRER
Chr6_64985396_1041_po_C_Xz11/1      EA-------GYKVSK------EKAQICGQRARYLGFNISQG-QRELGHEZ
ChrY_20355209_659_po_C_Xz7/11       EA-------GYKMSK------EKAQICAKEQIILASTFLKG-SVKGGCER
Chr3_34567166_523_po_C_Xz8/6        KA-------GYKVSK------EKAQICGQRARYLGFNISQG-QYELGLEQ
Chr3_122072468_245_po_C_Xz9/9       EA-------GYKVSR------KKAQICGQRARYLGFNISQG-QHELGHEZ
Chr11_58788475_541_po_C_Xz8/22      CK-------AYQLYMPFDPE-AAEKQCLVNRAFVGQAQGNI-RHKLQKLE
Chr6_137421042_373_po_C_Xz6/9       CE-------AYQLYMPLDPE-AAGNQCMVNVAFVSQVQE---DRKLEGFE
Chr4_115910608_550_po_C_Pz8/12      KA-------GYKVSG------KKAQICFKSVQYLGFCISQG-KRWLGGEZ
ChrY_13743185_977_po_C_Xz14/8       EA-------GYKVSK------DKAKVCFLEVGYLGFMVSQS-LHRLGSAC
HERV15                              EA-------GYKVSK------DKAKVCFLEVGYLGFMVSKP-AQAW-KAC
ChrX_117198092_983_po_C_Xz8/8       EA-------GYKVSK------DKAKVCFQEVGYLGFMVSQG-QRRL-KAR
Chr19_21166030_1412_po_C_Xz9/1      EA-------GYKVFK------DKAEVCFWEVAYLGFIVSQG-QCRLGSAR
Chr19_21213743_913_po_C_Xz14/1      EA-------GCKVSK------DKAQVCFZEVGYLGFMVSQG-QCRLGSAC
Chr19_15957774_569_po_C_Xz8/10      KT-------SYKVSKKK------AQICHEKVKYLSFIVSHR-EHQLGYKQ
Chr11_77508681_297_po_C_Xz19/1      KA-------GYKVSKKK------AZICHKEVKYLGFVVSHG-ECWLGYEZ
Chr13_98250769_738_po_C_Xz7/13      KQ-------VTKYCREKKEIRMLRCLCRKGRTRDSILKKSC-TLNYCFAE
Chr6_39968183_513_po_C_Pz10/9       KA-------EYKVSKKK------AQICLAGVKYLGFEISHG-KRELGSEZ
Chr15_66867263_645_po_C_Xz11/1      KE-------GYKMSEKA-------QICFESVQYLGFYISEG-ERWLGSEQ
```

```
HERV18_polputein                    AC-------GWAVNE-------SKVKGLDCLKFLGVIWSGKTKAIPEAII
HERVIP10FH_polputein                RQ-------GKTSRE------EREKRERERGRERERKRERGRERDKDGSK
MER66_polputein                     XZ----------ITR---------NFVFQXCXLVLLKASQRXHVTXSFSF


HERVI_polputein                     VEG-----IVSLPL--------PQTKQELRKFLGLVGYCRLRINSHALNS
HERVIP10F_polputein                 IEG-----IVSLPL--------PQTKQELRKFLGLVGYCRLWIDSYALHS
HERVadp_AC005741                    LLE-----IMGLPL--------PKTKRELZNFFGLTHYYRLWMS-YAQKT
HERV-PT47D                          IDG-----IMSLPL--------LQTKQELRKFLGLVTYCHLRIDSYALHS
HERV_W_chr6_141432567_ERV9_lik      I-----QPILAYPH--------PKTLKQLWGFLGITGFCZLWIPRYSEIA
ERV9_PH1_RT                         I-----QPILAYPR--------PKTLKQLREFLGITSFCRLWIPGYSETA
Chr13_54564403_572_po_C_Yz7/14      I-----QPKLAYPY--------PKTIKQLRGFLGITGFCRIWIPRYSKIA
HERV_Y_chr12_51022911_pol           I-----ZPILAYPH--------PKTLKQLTGFLGITGFCQIWMSRHGKIA
HERV17_polputein                    I-----QPILAYPH--------PKTLKQLRGFLGITGFCRIWIPRYGEIA
HERV_W_chr7_9105739_syncytin_p      I-----QPILAYPH--------PKTLKQLRGFLGVIGFCRKWIPRYGEIA
HERV9_polputein                     I-----QPILAYPC--------PKTLKQLQGFLGITGFCRLWIPGYSEIV
HUERS-P3_polputein                  I-----KXISSFPL--------PKTLKQLRGFWGITGFCRLWIPGYGEIX
ERVfrd_AC004022                     I-----QAILRIT---------PTTQKQFZAFLGITRYCRLWILECGGII
PRIMA41_polputein                   K-----EVILGILE--------PQTRKQLWDFLRMGRFC-LMVPGFGHMA
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       VZ----LISQ-TPT--------PSTKQQLFSILGMVGYFRVRIPGFAILT
HERV_H48I_polputein                 IS----LLQS-LLP--------PQTKQEILSFLGLAGYFRLWVPSFALLA
HERV_H_RTVLH2_pol                   MS----NZSP-KPQ--------FFTKQQLLSFLGMVS-------------
HERVFH21_pol                        AA----LIDKZSTP--------PSSKSEILSFLGLAGFFRIWIPNFALLA
HERVRblike_chr4_109047953           L-----SAIASIVM--------PTTRKQLKGFLGMAGFCKIWSPNYGNSK
ChrY_13790029_693_po_C_Xz12/11      KE-----AIPQIPT--------P-TKRQVREFLGAIRYCRLWILRSAEIA
HERVS71_polputein                   KE-----AILQIPT--------PTTRRWVHEFLGAVGYCRLWILGFAEIT
BaEV_M7_pol                         IE-----TVARIPP--------PRNPREVREFLGTAGFCRLWIPGFAELA
HERVR_polputein                     IE-----TVARIPP--------PRNPREVREFLGTAGFCRLWIPGFAELA
AKR_MLV_MLOCG_RT                    KE-----TVMGQPT--------PKTPRQLREFLGTAGFCRLWIPGFAEMA
GaLV_pol                            KA-----TVMKIPV--------PTTPRQVREFLGTAGFCRLWIPGFASLA
Chr3_158587414_1662_po_C_Ez10/      KQ-----VICSLLEZZIE---QPRTRRQVREFLGAVGFCRLWIPNFAVLA
Chr3_79627623_1138_po_C_Ez4/7       KQ-----VICSLPE--------PKTRRQVREFLGAG-VYRLWIPNFAILA
Chr19_20477244_1350_po_C_Ez10/      KQ-----VICSLPE--------PKTRRQVREFLGAVGFCTLWIPNFAVLA
Chr5_111890459_1626_po_C_Ez10/      KQ-----VTCSLPE--------PKTRRQVREFLGAVGFCRLGISNFAMLA
ChrX_48766518_1600_po_C_Ez6/3       KQ-----VICSLPE--------PKTRRQVREFLGAVGFCRLWIPNFAVLA
ChrX_152580713_1572_po_C_Ez10/      KQ-----VICSLLE--------PKTRRQVREFLGAVGFCRLWIPNFAVLA
ChrX_139287533_1154_po_C_Xz8/4      KQ-----VICSLPE--------PKTKRQVREFLGAVGFCRLRIPNFVVLA
Chr2_52615786_912_po_C_Xz11/5       KQ-----VICNLPE--------PNTRRQVREFLGAVGFCRLWIPNFAVLA
HERV-E_4_1                          KQ-----VICNLPE--------PKTRRQVREFLGAVGFCRLWIPNFAVLA
ChrY_21962399_1137_po_C_Xz13/9      KQ-----VICNLLE--------PKTRMQVREFLGAVEFCRLWIPNFAVLA
ChrY_8634032_950_po_C_Xz14/10       KQ-----VIYSLPE--------PRTRRQVKEFLGAVGFCRLWIPNFAVLE
Chr3_126735360_997_po_C_Xz5/10      KQ-----VICHLAE--------PKSRKQVREFLGAVGFCRLWIPNFAVLA
Chr3_126713203_958_po_C_Xz7/8       KQ-----VICNLAE--------PKSRRQVREFLGAVGFCRLWIPNFAVLA
Chr3_75226118_977_po_C_Xz8/9        KQ-----VICNLAE--------PKSRRQVREFLGAVGFCRLWIPNFAVLA
Chr7_6629919_827_po_C_Xz8/7         KQ-----VICKLGE--------PKSRRQVRGFLGAVGFCRRWIPNFAVLA
Chr4_9166812_811_po_C_Xz8/10        KA-----GHLNLVE--------PKSRRQVREFL-TVGFCRLWIPNFAVLA
Chr12_8478642_834_po_C_Xz12/7       KQ-----VICHLAE--------PKGRRQVREFLGAVGFCRLWIPNFAEVA
Chr10_15187175_787_po_C_Xz11/5      KQ-----IICHLAE--------PKGRRZVREFLGAVGFCRLWIPKFAELA
Chr11_71819740_854_po_C_Xz11/9      KQ-----VICHLAE--------PKSZRQVREFLGAVGFCRLWIPNFAVLA
Chr11_71519974_886_po_C_Xz13/9      KQ-----VICNLAE--------PKSRRQVREFLGAVGFCRLWIPNFAVLA
Chr8_7271441_969_po_C_Xz10/6        KQ-----VICHLAE--------PKSRRQVGEFLGAVGFCRLWIPNFAVLA
Chr8_11770066_826_po_C_Xz9/8        KQ-----VICHLAE--------PKIZRQVREFLGAVGFCRLWIPNFAVLA
Chr4_4128362_651_po_C_Xz7/11        KQ-----VICNLAE--------PKSRRQVREFLGAAGFCRLQIPNFAVLA
Chr14_50203955_850_po_C_Xz10/8      KQ-----VICHLVE--------PKSRRQVREFLGAVGFCRLRIPNFAVZA
Chr14_50205291_837_po_C_Xz10/8      KQ-----VICHLVE--------PKSRRQVREFLGAVGFCRLRIPNFAVZA
Chr7_6597120_842_po_C_Xz7/7         KQ-----VICNLAQ--------PKSRRQVREFLGAVGFCRLRIPNFAVLA
Chr11_71544658_783_po_C_Xz8/13      KQ-----VICNLAE--------PKSRRQVREFLGAVGFCRLWIPNFAVLA
Chr11_67999520_760_po_C_Xz8/8       KP-----VICNLAE--------PKSRRQVREFLGAVGFCRLRIPNFAVLA
Chr3_75461975_796_po_C_Xz7/9        KQ-----VICNLAE--------PKSRRQVREFLGAVGFCRLWIPNFAVLA
Chr12_52228939_692_po_C_Xz3/10      KQ-----VICNLLE--------PKSRRQVREFLGAVGFCRLWIPNFAVLA
Chr10_15211234_286_po_C_Xz11/1      KQ-----VICHLAE--------PKSRRQVREFLGAVRFCRLWIPNFAVLA
Chr12_100973302_945_po_C_Xz12/      KQ-----VICSLPA--------PKTRRQVREFLGAVRFCRLWIPNFAVLA
Chr19_9685420_817_po_C_Xz6/10       KQ-----VICNLPE--------PKTRRQVREFLGAVGFCRLZIPNFAVLA
Chr13_40936820_719_po_C_Xz9/10      KQ-----VICNLAE--------PKSRRQVREFLGAVGFCRLWIPNFAVLA
Chr7_63858202_1754_po_C_Rz12/9      KQ-----AICALP--------TPTTWRQIGEFLGAAGFCHIWIPNFSLIA
ERV3_chr7q21_AC073210_polputei      KQ-----AICALP--------TPTTWRQIGEFLGAAGFCHIWIPNFSLIA
Chr6_110218096_843_po_C_Xz16/8      KQ-----AVZTLP--------IPTTRHQIREFLVVAGFCHIWIPNFSLMA
Chr6_110214535_786_po_C_Xz16/8      KQ-----AVZTLP--------IPTTRHQIREFLVVAGFCHIWIPNFSLMA
ChrY_3018841_881_po_C_Rz9/12        KQ-----AVCALP--------TPTNRRQIIAFVISG------SQTSHLWL
ChrX_86876098_639_po_C_Xz7/8        KQ-----AVCALP--------TPTNRRQIIEFVISG------TQTSHLWL
Chr14_68899917_965_po_C_Xz6/8       KQ-----PVCALP--------TPTTWHQIREFFGAAGLCRIWIPNFSLMA
ChrY_23874437_661_po_C_Xz15/9       KP-----CLYALS--------TQTTQCQMGKFLGNAGFCCIWVPNFSLIA
ChrY_27518418_649_po_C_Xz15/9       KP-----CLYALS--------TQTTQCQMGKFLGNAGFCCIWVPNFSLIA
ChrY_25016892_657_po_C_Xz13/12      KP-----CLYALS--------TQTTQCQMGKFLGNAGFCCIWVPNFSLIA
ChrY_19932937_520_po_C_Xz10/14      KQ-----RVYALP--------TQATQCQIRKFLGKAWFCCIWIPDLLLMA
ChrY_20104614_518_po_C_Xz10/14      KQ-----RVYALP--------TQATQCQIRKFLGKAWFCCIWIPDLLLMA
```

```
Chr19_58660155_598_po_C_Xz21/1      KQ-----AICALP--------TPATQRQI-EFLEAARFR---IPNCSLMA
Chr19_58621153_500_po_C_Xz21/1      KQ-----AICALL--------TPTTRZQI-EFLDAATFCHIWIPSFSLMA
Chr7_139686987_710_po_C_Xz11/7      KQ-----AICALP--------TPTTZRQMREFLGAAGFCCIWIPNFSLMA
Chr7_137651301_1097_po_C_Pz13/      KK-----AVCALP--------TPNTQRQIREFLGAAGFCCIWIPNFSFMA
Chr11_29579197_930_po_C_Pz8/12      KL-----AVCALP--------TPTTQHQIRZFWGTAWFCHIWIPNYSLMA
Chr4_117329205_488_po_C_Xz3/16      RP-----AASGLA--------TP---RLCRCLSGRLRDPFIMIN----MA
Chr21_38522185_383_po_C_Xz7/9       KP-----AASSLA--------MP---GLCRCLPGKLRDPSIILN----MA
Chr4_70173370_1049_po_C_Xz11/8      KQ-----TICTLP--------TPTTR--TREFLGPAGSCCIWIPNFSLMA
Chr4_70022513_788_po_C_Kz10/12      KQ-----VVCALP--------TPTTWZKIREFLGVAGLCHIWIPNFSLMA
Chr14_105225894_1027_po_C_Xz13      KA-----GFCVLP--------TPTTWZQMRELLKAAGFCHIWIPCFSLMG
Chr7_119241536_761_po_C_Xz11/1      KQ-----AVCVLP--------TPTT-WCQIEFLGAAGFCCIWIPNFAVMA
ChrY_22564970_567_po_C_Xz9/10       KR-----TVCAVP--------PTTG-CQIREFLGAAGFCHIYIPNFLLMA
Chr21_14409346_1477_po_C_Pz5/6      KE-----TVCSIP--------QPDTRRQVREFLGAAGFSRIWIPNYSLLA
ChrX_51259700_1396_po_C_Pz7/8       KE-----TVCSIP--------WPDTTWHVREFLGVAGFCRIWIPNYLLLA
Chr6_64985396_1041_po_C_Xz11/1      KE-----TVCGIP--------QPGTGQQVWEFLGAAGLSHIWIPNYTSAK
ChrY_20355209_659_po_C_Xz7/11       KE-----TVHGIL--------RPDSRQQVRDFIEAAGFCHIZIPNYLFLA
Chr3_34567166_523_po_C_Xz8/6        KG-----TVCGIP--------ZPDTRQQVWEFIGAAGFCCIWIPNYSLLA
Chr3_122072468_245_po_C_Xz9/9       KE-----NMCGIP--------LPDTRRQVQELLGADGFCHIWIPNYSLLA
Chr11_58788475_541_po_C_Xz8/22      ALGVSATQLIEVA--------TKVYINQDQETKKEAN--QKLKKADLLAA
Chr6_137421042_373_po_C_Xz6/9       G--VNITQLIQVA--------TKVFVNQKEEAKRKARC-RAKEKADLLAV
Chr4_115910608_550_po_C_Pz8/12      KZ-----AVCAFP--------TPTTRCQIRKFLEAAGF-RIWTLNFLLMA
ChrY_13743185_977_po_C_Xz14/8       KE-----ALCALP--------TSVTRQQVREFLGAVRLCQIWIPNFSLIA
HERV15                              KE-----ALCALP--------TSVTRQQVREFLGAVGLCRIWIPNFSLIA
ChrX_117198092_983_po_C_Xz8/8       KE-----AVCALP--------TPVTR---REFLGVAGFCRICIPNFSLTA
Chr19_21166030_1412_po_C_Xz9/1      KE-----AMCIVH--------PVTRRQQVREFLGVVRFCRIWISNFSLIA
Chr19_21213743_913_po_C_Xz14/1      KK-----AVCALS--------TPVTRQ--REFLGMAGFCZIWIPNFSLIA
Chr19_15957774_569_po_C_Xz8/10      KQ-----AICALP--------TPTTRCQIRZFLEAAGFCHIWIPNFSLMA
Chr11_77508681_297_po_C_Xz19/1      KQ-----DVCALP--------IPTTRCQIREFLGAAGICPIWIPNLSLMA
Chr13_98250769_738_po_C_Xz7/13      ML-----LICSFAPATLTQPGAHKNMCCMKQGLRDLGLFRMCLVNKMFTG
Chr6_39968183_513_po_C_Pz10/9       KQ-----AVVLP-----------TPTTQRQVFLRVAGFCCIWIPNFSLMV
Chr15_66867263_645_po_C_Xz11/1      KQ-----AVWTLS--------TPITWZQIREFLGAAGFCHICLPNFSLMA
HERV18_polputein                    DK------IQAYPR--------PTMVRQLQTFVGLLGYWRAFVPHLAQMI
HERVIP10FH_polputein                RE--------------------RGRERETERKEGVKERERQKVKERKKER
MER66_polputein                     FVS-----SCEFFP-------LVLLLWPDAKVFCLRSVGAVFSPSIVSID


HERVI_polputein                     KLLYQKLAQGKP-------------EHLLWTSKEVD------QVKELKGI
HERVIP10F_polputein                 KLLYQKLAQEKP-------------NRLLWTSEEVD------QVEKLKER
HERVadp_AC005741                    KILYLKLLEEEP-------------NPLZWSPEEIQ------AVKELKQA
HERV-PT47D                          KLLHQKLAQEKP-------------NPLLWTSEEVD------QVKELTER
HERV_W_chr6_141432567_ERV9_lik      RPLYTLIKETQRANTH----------LVEWEPEAET------AFKTLKQA
ERV9_PH1_RT                         AP---LIKETQRANTH----------LVEWEPEAET------TFKTLKQA
Chr13_54564403_572_po_C_Yz7/14      RPLYTLIKETQKANTH----------LVRWTPEAEA------AFQALKZA
HERV_Y_chr12_51022911_pol           RPIYTLIKETQKANTY----------LVRWTPEAEV------AFQALKKA
HERV17_polputein                    RPLYTLIKETQKANTH----------LVRWTPEAEA------AFQALKKA
HERV_W_chr7_9105739_syncytin_p      RSLNTLIKETQKANTH----------LVRWTTEVEV------AFQAL---
HERV9_polputein                     RPLHTLIKETQRANTY----------LVEVT-RGRN------SLQNLKAG
HUERS-P3_polputein                  CPLYHLIKETQAAKTH----------LLTWEPEAQK------AFNQLKQA
ERVfrd_AC004022                     KSLYQALKEGTNRDT-------------IWEKNQEQ------AFRZLKTA
PRIMA41_polputein                   KSLCKALKG---AHVI----------FLNAVSIVKH------TFNTLKEK
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       KPLYKLTKGNLVDPI-------------DPKSFPHS------SFHSLKTA
HERV_H48I_polputein                 KPLYQAAKGPLHEPL-------------NPAQPITQ------PFRLLRKA
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        HPLYKAAKGPLNEPL-------------NXSHNILP------SFRELQTA
HERVRblike_chr4_109047953           AAISKRSRP--------------------WEIEHQI------TFEQLKHK
ChrY_13790029_693_po_C_Xz12/11      KPLYTATGRNG---------------PLVWTDTKEZ------AFKNLKKA
HERVS71_polputein                   KPLYTTTRGNG---------------PHVWTD-KEQ------AFQNLKKA
BaEV_M7_pol                         APLYALTKEST---------------PFTWQTEHQL------AFEALKKA
HERVR_polputein                     APLYALTKEST---------------PFTWQTEHQL------AFEALKKA
AKR_MLV_MLOCG_RT                    APLYPLTKTGT---------------LFNWGPDQQK------AYQEIKQA
GaLV_pol                            APLYPLTKESI---------------PFIWTEEHQQ------AFDHIKKA
Chr3_158587414_1662_po_C_Ez10/      KPLYRVTKGGD-------------REPFEWGPLQQQ------AFCKLKEK
Chr3_79627623_1138_po_C_Ez4/7       KPLYGVT------------------EPSEWGPLRQQ------AFCELKEK
Chr19_20477244_1350_po_C_Ez10/      KPLYGITKGGN-------------WEPFEWGPLQQQ------AFCKLKEK
Chr5_111890459_1626_po_C_Ez10/      KPLYRVTKWGD-------------WERFEWEPLQQQ------AFCKLKEK
ChrX_48766518_1600_po_C_Ez6/3       KPLYRVTEEAT--------------EPFEWGPLQQQ------AFCKLKEK
ChrX_152580713_1572_po_C_Ez10/      KPFYRVTNGGD-------------REPFEWGPLQQQ------AFCKLKEK
ChrX_139287533_1154_po_C_Xz8/4      KSLYGVTKGGD-------------REPLEWGPLQKQ------AFCKLKEK
Chr2_52615786_912_po_C_Xz11/5       KPLZGLTKGGT-------------RNLSNGGPKQQQ------AFHVLKDK
HERV-E_4_1                          KPLYEVTKAGD-------------QEPFEWGSQQQQ------AFHELKER
ChrY_21962399_1137_po_C_Xz13/9      NPVYGVTKRGE-------------KEHFKWGSQQQZ------ASHZLKEN
ChrY_8634032_950_po_C_Xz14/10       KPLYRVTQGGD-------------WVLFEWRPLQZQ------DFCKLKEK
Chr3_126735360_997_po_C_Xz5/10      KPLYEVTKGAG-------------TGNLWNGSQQQQ------LFHELKEK
Chr3_126713203_958_po_C_Xz7/8       KPLYEVTKGAG-------------TGNLWNGSQQQQ------VFHELKEK
Chr3_75226118_977_po_C_Xz8/9        KPLZEVTR-AG-------------TGNLWNADQQQQ------VFRELKEK
Chr7_6629919_827_po_C_Xz8/7         KPLYEVTKGVG-------------TGNLWNGSQQQQ------GFHELKEK
```

```
Chr4_9166812_811_po_C_Xz8/10        KTLYEVTKGAG-------------TGPLECGSQQQQ------VFHELKEK
Chr12_8478642_834_po_C_Xz12/7       -PFYEVTKGAE-------------TGNFWNGSQQQQ------VFHELZEK
Chr10_15187175_787_po_C_Xz11/5      -PLYEVTKGAG-------------TGLLECGSQQQQ------VFHELKEK
Chr11_71819740_854_po_C_Xz11/9      KPLYAVTKGAG-------------TGKLWNGCQQQQ------GFHEL-KK
Chr11_71519974_886_po_C_Xz13/9      KPLYEVTKGAG-------------TGNLWNVDQQQZ------VFHELKEK
Chr8_7271441_969_po_C_Xz10/6        KFLYEVTRGTG-------------TGNLWNEDQQQQ------VFHELKEK
Chr8_11770066_826_po_C_Xz9/8        KPLYAVTNGAG-------------TGNLWNADQQQQ------VFHELKEK
Chr4_4128362_651_po_C_Xz7/11        KTLYEVTKGAG-------------TGLFEWGFQQZQ------VFHELKEK
Chr14_50203955_850_po_C_Xz10/8      KPLYEVTKGTG-------------TEPFEWGSQQQQ------VFHELKEK
Chr14_50205291_837_po_C_Xz10/8      KPLYEVTKGTG-------------TEPFEWGSQQQQ------VFHELKEK
Chr7_6597120_842_po_C_Xz7/7         KLLYEVTKGQGGG-----------TGHFEWGSQQQQ------VFHELKEK
Chr11_71544658_783_po_C_Xz8/13      KPLYEVTKGAG-------------TGLFEWGSQQQQ------VFHELKEK
Chr11_67999520_760_po_C_Xz8/8       KTLYKVTRGRG-------------WEIFEWGSQQQQ------AFRELKEK
Chr3_75461975_796_po_C_Xz7/9        KPLYEITKGAG-------------TGPFELGSQQQQ------AFHELKEK
Chr12_52228939_692_po_C_Xz3/10      KPLYEVTKWGT--------------GPFEWGSQQQQ------AFHELKKK
Chr10_15211234_286_po_C_Xz11/1      KPLFEVTKEAG-------------TGPFEWGSQQQQ------VFHELKEK
Chr12_100973302_945_po_C_Xz12/      KPLYGVTKGGD-------------WEPCEWGPLQQQ------AFCKLKEK
Chr19_9685420_817_po_C_Xz6/10       KPLYRVTKGGD-------------QVSFERGPLQQQ------AFCKLKEK
Chr13_40936820_719_po_C_Xz9/10      KPFYEATKGVG-------------TANLWNGSQQQQ------VFHELKEK
Chr7_63858202_1754_po_C_Rz12/9      RPLYEATREGEK-------------EPLLZKADQKK------VFKQIKEA
ERV3_chr7q21_AC073210_polputei      RPLYEATREGEK-------------EPLLZKADQKK------VFKQIKEA
Chr6_110218096_843_po_C_Xz16/8      KPLYEDTKGGER-------------EPVLWEADQZR------AFKQIKEA
Chr6_110214535_786_po_C_Xz16/8      KPLYEDTKGGER-------------EPVLWEADQZR------AFKQIKEA
ChrY_3018841_881_po_C_Rz9/12        SPYMKPQKVES--------------EPLLWETDQEK------AFKQIKEA
ChrX_86876098_639_po_C_Xz7/8        SPYMKPQKVES--------------EPLLWETDQEK------AFKQIKEA
Chr14_68899917_965_po_C_Xz6/8       KPLYEATKLGGK-------------EPLLWEADQEK------AFKQIKKV
ChrY_23874437_661_po_C_Xz15/9       KPLYKGTKRGER-------------KPLLZEAVQEKK-----VFKQAKEA
ChrY_27518418_649_po_C_Xz15/9       KPLYKGTKRGER-------------KPLLZEAVQEKK-----VFKQAKEA
ChrY_25016892_657_po_C_Xz13/12      KPLYKGTKRGER-------------KPLLZEAVQEKK-----VFKQAKEA
ChrY_19932937_520_po_C_Xz10/14      KPLHKGTKRGER-------------EPLLLEVVQEKK-----VFKQPKQA
ChrY_20104614_518_po_C_Xz10/14      KPLHKGTKRGER-------------EPLLLEVVQEKK-----VFKQPKQA
Chr19_58660155_598_po_C_Xz21/1      KSLIGSHKKRRK-------------GAPPLKPNQEI------DFKQVKEA
Chr19_58621153_500_po_C_Xz21/1      KLLYDATKGGEK-------------EPLFZKTNQEI------AFKQIKEA
Chr7_139686987_710_po_C_Xz11/7      KPLYEATKWGEK-------------EPLLWETNQEI------AFKQIKEA
Chr7_137651301_1097_po_C_Pz13/      KPLYEATKEGEK-------------EPLLWEADQEK------AFKQIKEA
Chr11_29579197_930_po_C_Pz8/12      RPLYEATKEGE--------------EPLLWEADQDR------AFKQIKEA
Chr4_117329205_488_po_C_Xz3/16      R-----FKRGEK-------------EPLLWEVDQEK------VFKQIKEA
Chr21_38522185_383_po_C_Xz7/9       R-----FKGREK-------------EPLLWEADQEK------AFKKIKEA
Chr4_70173370_1049_po_C_Xz11/8      KPLYEATKRGK--------------ELLLWEADQEK------AFKEIKEA
Chr4_70022513_788_po_C_Kz10/12      KLLYETTKWGK--------------ETILWEANQEK------AFKEIKEA
Chr14_105225894_1027_po_C_Xz13      KPLYEATKRGKK-------------EPLLWEATQEK------AFZZIKGD
Chr7_119241536_761_po_C_Xz11/1      KPLPEATKQGER-------------EPLLWEADQER------TFKQIKEG
ChrY_22564970_567_po_C_Xz9/10       KLLYEATEWGKR-------------APSLGSQSGKS-------LQGIKRA
Chr21_14409346_1477_po_C_Pz5/6      KPLYEAMKGQG-------------KEPLLWGKEQD------MAFKEIKKA
ChrX_51259700_1396_po_C_Pz7/8       KPLYEATKVGI-------------KEPLLWGKEQD------MAFKEIKKA
Chr6_64985396_1041_po_C_Xz11/1      ----PLYETTK----------GRGERTPLWGKEQD------TAFKEIKKA
ChrY_20355209_659_po_C_Xz7/11       SVRLALCEATK----------GWGKRTPLVGKRAA------HGLQRNQKA
Chr3_34567166_523_po_C_Xz8/6        KTLYEATKVGE----------KEP---LLWRKEQD------MAFKEIKKV
Chr3_122072468_245_po_C_Xz9/9       KPLYKAPKLGR----------GQGGRIPLWEKEQD------MAFKDIKKA
Chr11_58788475_541_po_C_Xz8/22      ALMGREADFSKRHGWGGERGQYGQGFKSWLRLEKDZCVWCKRKGHZEDEC
Chr6_137421042_373_po_C_Xz6/9       ALVGRETGFVR-------------GVVVVAVMIED------KLGKTRKLS
Chr4_115910608_550_po_C_Pz8/12      KPLYEATKREK--------------EPLLWGKQQD------MAFKEIKKA
ChrY_13743185_977_po_C_Xz14/8       RPLFEASKGKE-------------REPLLWEKEZE------KAFKDIKEA
HERV15                              RPLFEASKGKE-------------REPLLWEKEQE------KAFKDIKEA
ChrX_117198092_983_po_C_Xz8/8       RPLYEAIKGKE-------------REPLLWQKKQE------KAFKDIKEA
Chr19_21166030_1412_po_C_Xz9/1      RPLHEATRGRE-------------REPLLWEKEQE------KTFKDIKEA
Chr19_21213743_913_po_C_Xz14/1      RPLYEATKGGZ-------------REPLLWEKEKE------KAFKGIKEA
Chr19_15957774_569_po_C_Xz8/10      KPLYEATKIG-------------EKQPLLSKTNQEI------AFEQIKEA
Chr11_77508681_297_po_C_Xz19/1      VPLYEATKGG-------------QKNPPLGDZ-SEN------AFQQIKEA
Chr13_98250769_738_po_C_Xz7/13      SILGKGHRHSPASI--------NQGHNALRKATGTS------ALESGVLS
Chr6_39968183_513_po_C_Pz10/9       KPLYEATNGG-------------IGEPLLWEADQKR------LFKKIKKA
Chr15_66867263_645_po_C_Xz11/1      KPLYEATKWG-------------EKEPLLLEANQEK------AFKEIKEA
HERV18_polputein                    KPLYQLTKKGA---------------AWDWDDEAET------AFLAAKRA
HERVIP10FH_polputein                DRSSKEKTVYP---------------------------------IPLKAR
MER66_polputein                     IILFLAFSLTLT---------------LLRRVLRAN------ECLPTSIP


HERVI_polputein                     LITALALALPSLKY-TSLFVSEK-GVALGVLIQEHR----GCWQPMAFLS
HERVIP10F_polputein                 LITAPVLALPSLEKPFHLFVNVDSGVALGVLTQEHR----GRRQPVAFLS
HERVadp_AC005741                    LITAPVLAFPPLEKPFHLFVTVDQGVALGLLTQTLG----GKRQPVAFVS
HERV-PT47D                          LITSPVLALPSLEKPSHIFFFFFFEMESCSVAQAG------------VQW
HERV_W_chr6_141432567_ERV9_lik      LVQAAALTLPTG-QNFSL--YVTDRAEIALGVLIQT--CGTTPQPVAYLS
ERV9_PH1_RT                         LVQAPALSLPTG-QNFSL--YVRERARIALGVLTQT--HGTTPQPVAYLS
Chr13_54564403_572_po_C_Yz7/14      LTQAPVLSLPTG-QDFSL--YVTEETGIALRVLTQV--RGMSLQPVAYLS
HERV_Y_chr12_51022911_pol           LTQVPVLSLPMG-QEFSL--YVTEK-GIALGVLTQV--RGMSLQPVAYLS
HERV17_polputein                    LTQAPVLSLPTG-QDFSL--YVTEKTGIALGVLTQV--RGTSLQPVAYLS
HERV_W_chr7_9105739_syncytin_p      -TQAPVLSLPTG-QDFSS--YVTEKTGIALGVLTQI--RGMSLQPVAYLT
```

```
HERV9_polputein                     PVQVPALSLPT---NFSL--YVTERAGIALGVLTQT--LGTTPQPVAYLS
HUERS-P3_polputein                  LLKAPALSLPVG-KAFNL--YVSERKGMALGVLTQA--RGPXQQPVGYLS
ERVfrd_AC004022                     LSQAPALELPILTKPFQF--FITEKAGIALGVLTET--FGPVKHAVGYFS
PRIMA41_polputein                   LGTTPALGIPNL-DKPFL--YLAEK-QGALEVLVQK--LGG-IPPVAYFS
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       LETVSTLALPDIPTLF-I--HTAEVQGCAVRILTQ----GPGSHPVAFLS
HERV_H48I_polputein                 LTSAPVLTLPDLTKPFSL--YTDERRGVTLGVLTQSK--GPTLQVVAYLS
HERV_H_RTVLH2_pol                   ----------------------------AVRILTQ----EPGPHPVAFLS
HERVFH21_pol                        LVTAPALSLPNISQPFTL--YTAKNRGIALGVLGQQKENPPSFAPVAYLS
HERVRblike_chr4_109047953           LISALAPGLPNPHKLFQP--YMHERLGLALWVLTQK--LGEILQPVAYFS
ChrY_13790029_693_po_C_Xz12/11      LTEAPTLALPNISKPFH---FIHESQGVAKGVLTQT--LGPWRCPVAYLS
HERVS71_polputein                   LTEVPALALPNISEPFHL--FVHESQGVTKGVLTQT--LGPWZCPVAYLS
BaEV_M7_pol                         LLSAPALGLPDTSKPFTL--FLDERQGIAKGVLTQK--LGPWKRPVAYLS
HERVR_polputein                     LLSAPALGLPDTSKPFTL--FLDERQGIAKGVLTQK--LGPWKRPVAYLS
AKR_MLV_MLOCG_RT                    LLTAPALGLPDLTKPFEL--FVDEKQGYAKGVLTQK--LGPWRRPVAYLS
GaLV_pol                            LLSAPALALPDLTKPFTL--YIDERAGVARGVLTQT--LGPWRRPVAYLS
Chr3_158587414_1662_po_C_Ez10/      FMLAPALGLSDVTKPFT--LYVSERETMAVGVLTQT--VGHWPRPAAYLS
Chr3_79627623_1138_po_C_Ez4/7       LMSAPALGLPDLAKPFT--LYTSERKKMAAGVLTQT--VGPWLRPVAYLS
Chr19_20477244_1350_po_C_Ez10/      LTLVPALGLPDLTKPFA--LYVSERETMAVGVLTQT--VGSWPRPVAYLS
Chr5_111890459_1626_po_C_Ez10/      FMSAPALGLPDLTKPFT--LYVSEREKTAVGVLIQT--VGPWPTPVAYLS
ChrX_48766518_1600_po_C_Ez6/3       LMSAPSPGLPDLTKPFT--LYVSEREKMAVGVLTQT--VGPWPRPVVYLS
ChrX_152580713_1572_po_C_Ez10/      LMSAPVLGLPDLTKPFT--LYVSEREKMAVGVLTQT--VGPQPRPVAYLS
ChrX_139287533_1154_po_C_Xz8/4      LMSAPALGLSDLTKPFT--LYASEREKIAVGVLTQT--VGPWPRPVAYLS
Chr2_52615786_912_po_C_Xz11/5       LMTALALGLPDLTKPFT--LYVSEREKMAVGVLTQT--VGPWLRPVAYLS
HERV-E_4_1                          LMSVPALGLPDLTKPFT--LYVSESEKMAVGVLTQT--VGPWPRPVTYLS
ChrY_21962399_1137_po_C_Xz13/9      LMSAPALGLPDQTKPFT--LYVSEGEZMAVGVLTQS--MGPWPRLMAZLS
ChrY_8634032_950_po_C_Xz14/10       FMSTPAIGLPDLTKPFA--FYVSEREKMAVGVLTQT--VGPWPRPVVYLS
Chr3_126735360_997_po_C_Xz5/10      LLAAPALGLPDLTKPFP--LYASEREKMAAGLZTQT--VGPWLRPVTYLS
Chr3_126713203_958_po_C_Xz7/8       LLAAPALGLPELRKPFP--LCASEREKMAAGLZTQT--VGPWLRLVAYLS
Chr3_75226118_977_po_C_Xz8/9        LLAAPALGLPDLTKPFP--LYASEREKMAAGLZTQT--VGPWPRPVAYLS
Chr7_6629919_827_po_C_Xz8/7         LMSAPALGLPDLTKPFP--LYASEREKMAAGLZTQT--VGPRPRLGAYLS
Chr4_9166812_811_po_C_Xz8/10        LMSAPALGLPDLTKPFP--LYASEREKMASGLZTQT--VGPRPRPGAYLP
Chr12_8478642_834_po_C_Xz12/7       LQAGPALGLPDLTKLFP--LYASEREKLAAGLZTQT--VGPRPRLVSYLS
Chr10_15187175_787_po_C_Xz11/5      LLAAPALGLPDLTKPFP--LYASEREKMAAGLZSZT--VGPWPRPVAYLS
Chr11_71819740_854_po_C_Xz11/9      PLAAPALGLPDLTKPFP--LYASEREKMAAGLZTQT--VGPRPRLGAYLS
Chr11_71519974_886_po_C_Xz13/9      LLAAPALGLPDLTKPFP--SYN-EREKMAAGLZTQT--VGPRPRPGSYLS
Chr8_7271441_969_po_C_Xz10/6        LLAAPALRLPDLTKPFP--LYASERERMAAGLZTQT--VGPWLRLVAYLC
Chr8_11770066_826_po_C_Xz9/8        LLAAPALGLPDLTKPFP--LYIKERKDGSWTL-TQT--VGPRPRPVAYLS
Chr4_4128362_651_po_C_Xz7/11        LRSAPALGLPDMTKPSP--ZYASEREKMAAGLZTQT--VWPRERPVAYLS
Chr14_50203955_850_po_C_Xz10/8      LMSAPALGLPDLTKPFS--SYVSERHKMAAGLZTQT--VGPWPRPVAYLS
Chr14_50205291_837_po_C_Xz10/8      LMSAPALGLPDLTKPFS--SYVSERHKMAAGLZTQT--VGPWPRPVAYLS
Chr7_6597120_842_po_C_Xz7/7         LMSAPALGLPDLTKPFT--SYVSERGKMAAGLZTQT--VGPWPRPVAYLS
Chr11_71544658_783_po_C_Xz8/13      LMSAPALGLPDLTKPFS--SYVPEREKRAAGLZTQT--VGPWPRPVAYLS
Chr11_67999520_760_po_C_Xz8/8       LMSAPALGLPDLTKSFT--LCVREREKMAAGLZTQT--VGPWPRPVAYLS
Chr3_75461975_796_po_C_Xz7/9        RLAAPALGLPDLTKPFT--TYVSEREKRAAGLZTQT--VGPWQRPVAYLS
Chr12_52228939_692_po_C_Xz3/10      LMSAPALGLPDLTKPFT--LYVSEREKMAVGVFIQT--VGPWPRPVAYLS
Chr10_15211234_286_po_C_Xz11/1      LMSAPALGLPDLTKPFT--SYVSEREKMAAGLZTQT--VGPWPRLVAYVS
Chr12_100973302_945_po_C_Xz12/      LMSAPALGLPDLTKPFT--LCVSEREKMAVGVSTQT--VGYWPRPVAYLS
Chr19_9685420_817_po_C_Xz6/10       LTSAPALGLPDFTKPFT--LYVSEREKMAVGVLTQS--VGPWPRPVAYLS
Chr13_40936820_719_po_C_Xz9/10      LLAAPALGLPDLTKPFP--LYASEREKMAAGLZTQT--VGPWPRPVAYLS
Chr7_63858202_1754_po_C_Rz12/9      LTQAPALGLPDTTKPFFL--YVHELKGMAIGVLTQI--IGSWHRPVVYLS
ERV3_chr7q21_AC073210_polputei      LTQAPALGLPDTTKPFFL--YVHELKGMAIGVLTQI--IGSWHRPVVYLS
Chr6_110218096_843_po_C_Xz16/8      ITZAPALGLPDITKPFFL--YVHEZKGMAIGVLTZV--IESLHRPMTYSS
Chr6_110214535_786_po_C_Xz16/8      ITZAPALGLPDITKPFFL--YVHEZKGMAIGVLTZV--IESLHRPMTYSS
ChrY_3018841_881_po_C_Rz9/12        LTQVPALGLPDITKP-FL--YVHERKGMAIGVLTQV--LGSWHCLVAYLS
ChrX_86876098_639_po_C_Xz7/8        LTQVPALGLPDITKP-FL--YVHERKGMAIGVLTQV--LGSWHCLVAYLS
Chr14_68899917_965_po_C_Xz6/8       LTQVPALGLPDITKTFFL--YVHEZKGIAIGVLTQV--MGSWHRLVAHLS
ChrY_23874437_661_po_C_Xz15/9       LTQAQAIGLPNISKPFFP--HVHEZKGLAIGVLTQV--VKSZIRPVMYLS
ChrY_27518418_649_po_C_Xz15/9       LTQAQAIGLPNISKPFFP--HVHEZKGLAIGVLTQV--VKSZIRPVMYLS
ChrY_25016892_657_po_C_Xz13/12      LTQAQAIGLPNISKPFFP--HVHEZKGLAIGVLTQV--VKSZIRPVMYLS
ChrY_19932937_520_po_C_Xz10/14      LTKALAIGLPNIIKPLF------EQKGLAIGVLSQV--MKSZHHPGVYLS
ChrY_20104614_518_po_C_Xz10/14      LTKALAIGLPNIIKPLF------EQKGLAIGVLSQV--MKSZHHPGVYLS
Chr19_58660155_598_po_C_Xz21/1      LTQVPALRLZDITKS-FY--MFMNKKTMTIGALTQL--IGSWHRPVAHLS
Chr19_58621153_500_po_C_Xz21/1      LTQAPALRLZDITK--FL--YVHEQKTMTIGALTQL--IGSWRHPVAYLS
Chr7_139686987_710_po_C_Xz11/7      LAQSPALGLPDVTKPSFL--YVHEQKGVAIRVLTQA--TGSWHHPVAYLS
Chr7_137651301_1097_po_C_Pz13/      LTQAPAVGLPDLTKPFFL--YVHEZKGIAIGVLIQV--IGSWHGPVAZLS
Chr11_29579197_930_po_C_Pz8/12      LTQAPALGLPDLTKPFFL--YVHEQKRMAIGVLTQV--IGSWYHPVEYLS
Chr4_117329205_488_po_C_Xz3/16      L------GLPDITKHSFL--YVHN-EMMTIGVLTQV--IGSWHCLVANLS
Chr21_38522185_383_po_C_Xz7/9       LTQAPVLGLPDITEPFFL--YVHEQKGMAVGVLTQF--IGLWHRLVAYLS
Chr4_70173370_1049_po_C_Xz11/8      LTQAPTLGLPDLTKPFFL--YVHEQKGMAVRVLTQA--IGLWHHPVAYLS
Chr4_70022513_788_po_C_Kz10/12      LTLVPALGLTDLTKPFFL--YVHEQKGMAIGVLTKS--IGSWHCPATYLS
Chr14_105225894_1027_po_C_Xz13      LTEAPALGLPDLTKLFFL--YAHEZKGMVIGVLTQV--IGSWHHPVAYLS
Chr7_119241536_761_po_C_Xz11/1      LTQSPALGLPDITKPFFS--IYPZMKRNGYTGPDSS--HGIMASPVAYSS
ChrY_22564970_567_po_C_Xz9/10       LIQALALGLPDQTKPFFL--YVYEWKEVAVGILTQM--MGSWHIPAVYLS
Chr21_14409346_1477_po_C_Pz5/6      LIQAPALGLPDMTKAFY--LYVHERKEMATVVLVQM--LGSWYWPVAYLS
ChrX_51259700_1396_po_C_Pz7/8       LIQALALGLPDLTKPFY--LYVHERKEMATGVLVQM--LGSWYLPVAYLS
Chr6_64985396_1041_po_C_Xz11/1      LIQVPALELPDMTRPFY--LYVHERKGIATGVLVQK--LGSZYQPMAYLF
```

```
ChrY_20355209_659_po_C_Xz7/11       LIQSPALGLPDMTKPFY--LYVHKRKGIATGALVQT--LGSWYRPMTYLP
Chr3_34567166_523_po_C_Xz8/6        LIQIPALGLPDMIKPFY--LYVHERKGIVTGVLVQT--LGSWYRPMAYLS
Chr3_122072468_245_po_C_Xz9/9       LIQAPALGLPDMTKPFY--LYVHERKGIATGILVQT--LGSWYWPTAYFS
Chr11_58788475_541_po_C_Xz8/22      QETKENGQNHRIRKPLAKGYCTWEKRKPSAPAVEGK--LQSVRKKAQICF
Chr6_137421042_373_po_C_Xz6/9       QDKRAN---LGLREINVZDASRWDTRRMNA-KFSGA--LVRKRFPPEIAS
Chr4_115910608_550_po_C_Pz8/12      LIQAPALGLPDMKKPFY--LYVHESZGIATGVLVQI--LGSWYWPMAYZS
ChrY_13743185_977_po_C_Xz14/8       LIQAPALGLPDVKKPFF--LYVDEZKGMLFGVLTQL--LGSWYQPVAYLS
HERV15                              LIQAPALGLPDVKKPFF--LYVDEZKGMVFGVLTQL--LGSWHQPVAYLS
ChrX_117198092_983_po_C_Xz8/8       LIQAPALGLPNVKKP-F--LYVDECQGMAVGVLTQF--LGSWHRPVAYLS
Chr19_21166030_1412_po_C_Xz9/1      LIQASVCGLPDVRKTLL--LYVDERKGMAVGVLTQL--LGSLHWLVAYLS
Chr19_21213743_913_po_C_Xz14/1      LIQAPALGLPDIKKPFF--LYVDEZKGMAVGVLTQL--LGSWHWPVPYLS
Chr19_15957774_569_po_C_Xz8/10      LTWAPALGLPDITKP--FFLYVHEQKEMAIGVLTQV--IGSZHCPVAYLS
Chr11_77508681_297_po_C_Xz19/1      LTZALALGLPDITK----FLYVHEQKGMAIGVLTQV--VRSZHHPVTYZS
Chr13_98250769_738_po_C_Xz7/13      KVSPHVIVZNMASWZ--EKDLSVPQSDTRKGSVLRW--ISNRGKPLAVEI
Chr6_39968183_513_po_C_Pz10/9       LIQSPGLGLPDQTKH--FFLYVHEWKGMAIGVLTQI--MRSSHRPVAYLT
Chr15_66867263_645_po_C_Xz11/1      LTQALGRGLSVHVGV--HLHLFRWVEVFPTRTQKAQ--EVTKVLLKDIIL
HERV18_polputein                    IQQAQALRVVDQGRPFEL-DVHVTTDGFGWGLWQRT---ERLRTPVGFWS
HERVIP10FH_polputein                VNLKPILIIEGLLHDPIT-LQYHLVVSVNKGVARKLR----PLQPIAFLS
MER66_polputein                     SGLERLIGYKSFDSKSLX-QRKFQRNLKLTQAMTGNRGLDIPLCPLXNLA


HERVI_polputein                     KILDLVTCGWP-QCIQSIAATAVLVEESRKLTFG----------------
HERVIP10F_polputein                 KVLDPVTCGWP-QCIQSIAATAILVEESRKLTFG----------------
HERVadp_AC005741                    KLLDSVSCRWP-KCMKAVAATALLVEESQKLTFG----------------
HERV-PT47D                          HDLGSLKLHLP-GSRHPPASASRVAGLQRPPPCP----------------
HERV_W_chr6_141432567_ERV9_lik      KEMDVVSKGWP-QCLRVVAVVAVLVLEAIKIIQG----------------
ERV9_PH1_RT                         KEIDVVAKGWP-HCLRVVAAVAVLASEAIKIIQG----------------
Chr13_54564403_572_po_C_Yz7/14      KENDVVAKGWP-HCLWVVAAVAVLVSAAVKITG-----------------
HERV_Y_chr12_51022911_pol           KEIDVVAG-CP-HCLWVVTAAAVLVSEAVKMIQG----------------
HERV17_polputein                    KEIDVVAKGWP-HCLRVVAAVAVLVSEAVKIIQG----------------
HERV_W_chr7_9105739_syncytin_p      KEIDVVAKGZP-HCLRVVVAVAVLVSEAVKIIQG----------------
HERV9_polputein                     KEIDVVAKGWP-HCLRVVAALAIFVRGYQNNTR-----------------
HUERS-P3_polputein                  KELDLVAKGWP-ACLRAVAXVALLVPEASKLTLG----------------
ERVfrd_AC004022                     KNLDPAAQWWP-HCLKVVGTAALILEEDFNITMG----------------
PRIMA41_polputein                   KZLDHVALGWP-RCLRAVAAMALLVYEANKLALG----------------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       KQLDLTVLGWP-SCLHAASAATLILLEALKITNY----------------
HERV_H48I_polputein                 KQLEATVLGWP-ACLRALVAAAVLTLESLKLSLH----------------
HERV_H_RTVLH2_pol                   KZLNLTVLAZP-SCLRAAAATALILLETLKITNY----------------
HERVFH21_pol                        KQLDNTVKGWP-TCLKVLAAAAVLALESRKLTFS----------------
HERVRblike_chr4_109047953           KQLHTVARGWP-PCLRAVATTCLLLKEAEKLTLG----------------
ChrY_13790029_693_po_C_Xz12/11      KRLDPVAAGWP-TCLZAVVATASLVQEANKLTLG----------------
HERVS71_polputein                   KILDPVAPG-P-SCLZAIAAKASLVQEADKLTLG----------------
BaEV_M7_pol                         KKLDPVAAGWP-PCLRIMAATAMLVKDSAKLTLG----------------
HERVR_polputein                     KKLDPVAAGWP-PCLRIMAATAMLVKDSAKLTLG----------------
AKR_MLV_MLOCG_RT                    KKLDPVAAGWP-PCLRMVAAIAVLTKDAGKLTMG----------------
GaLV_pol                            KKLDPVASGWP-TCLKAVAAVALLLKDADKLTLG----------------
Chr3_158587414_1662_po_C_Ez10/      KQLDGVSKGWP-PRLRALAATALLAQEADKLTLG----------------
Chr3_79627623_1138_po_C_Ez4/7       KQLDGVSKGWP-PCLRALAATALLVQEANKLTLG----------------
Chr19_20477244_1350_po_C_Ez10/      KQLDGVSKGWP-PCVRALAATALLAQEADKLTLG----------------
Chr5_111890459_1626_po_C_Ez10/      KQLDGVSKGWP-PCLRALAATALLAQEADKLTLV----------------
ChrX_48766518_1600_po_C_Ez6/3       KQLDGVSKVWP-PCLRALAATALLAQEADKLTLG----------------
ChrX_152580713_1572_po_C_Ez10/      KQLDEVSKVWP-PCLRALAATALZAQEADKLTLG----------------
ChrX_139287533_1154_po_C_Xz8/4      KQLDGVSKGWP-PCLRALAATALLAQDADMLILE----------------
Chr2_52615786_912_po_C_Xz11/5       KQLDGVSKGWP-PCLRALAATALLAQEAEKLTLG----------------
HERV-E_4_1                          KQLDGVSKGWP-PCLRALAATALLVQEADKLILG----------------
ChrY_21962399_1137_po_C_Xz13/9      KQLKGVS-GLA-PYLRALAATALLAKEVDKLTLG----------------
ChrY_8634032_950_po_C_Xz14/10       KQLDGVSKGWP-PCLRALAATALLAQEANKIAVG----------------
Chr3_126735360_997_po_C_Xz5/10      KQLDGVSKGWP-PCLRALAATALPLQEANKLTLG----------------
Chr3_126713203_958_po_C_Xz7/8       KQLDGVSKGWP-PCLRALAATALLVQEANKLTLG----------------
Chr3_75226118_977_po_C_Xz8/9        KQLDGVSKGLS-PCLRALAATALPVQEANKLTLG----------------
Chr7_6629919_827_po_C_Xz8/7         QQLDGDSKGWP-PCWRALAATALRVQEANKLTLG----------------
Chr4_9166812_811_po_C_Xz8/10        KQLDGVSKGWP-PCWR-LGATALLVQEANKLTLG----------------
Chr12_8478642_834_po_C_Xz12/7       QGLDGVSKGWP-PCWRALAATALPVQEANKLTLG----------------
Chr10_15187175_787_po_C_Xz11/5      QQLDGVSKGWP-PCWKALAATALPVQEANKLTLG----------------
Chr11_71819740_854_po_C_Xz11/9      QQLDGVSKGWP-PCWRALAATALLVQEANKLTLG----------------
Chr11_71519974_886_po_C_Xz13/9      QQLDWVSKGWP-PCWRALAATALPVQEANKLTLG----------------
Chr8_7271441_969_po_C_Xz10/6        KPLDGVSKGWP-PCWRALAATALLVQEANKLTLG----------------
Chr8_11770066_826_po_C_Xz9/8        QQLDGVSKRWP-PCWRALAASALLVQEANKLTLG----------------
Chr4_4128362_651_po_C_Xz7/11        QQLDRVSKGWP-PCWRALAATALRVQEANKLTLG----------------
Chr14_50203955_850_po_C_Xz10/8      KQLDGVSKGWP-LCLRALAATALLVQEANKLTLG----------------
Chr14_50205291_837_po_C_Xz10/8      KQLDGVSKGWP-LCLRALAATALLVQEANKLTLG----------------
Chr7_6597120_842_po_C_Xz7/7         KQLDGDSKGWP-LCLRALAATALLVQEAKKLTLG----------------
Chr11_71544658_783_po_C_Xz8/13      KQLDWVSKGWP-LCLGALAATALLVQEAKKLTLG----------------
Chr11_67999520_760_po_C_Xz8/8       KQLDGVSKGWP-LCWRALAATALLVQEAKKLTLG----------------
Chr3_75461975_796_po_C_Xz7/9        KQLDGVSKGWP-LCLRALAATALLVQEVNKLTLG----------------
Chr12_52228939_692_po_C_Xz3/10      KQLDGVPKGWP-LCLRALAATALLVQEANKLTLG----------------
```

```
Chr10_15211234_286_po_C_Xz11/1      KQVDGVSKGWP-PCLRALAATALLVQEANKLTLG----------------
Chr12_100973302_945_po_C_Xz12/      KQLDRVSKGWP-PCLRALAATALLAQEADKLTLG----------------
Chr19_9685420_817_po_C_Xz6/10       KQLDGVSKGWP-PCLRALAATALLAQEADKLTLG----------------
Chr13_40936820_719_po_C_Xz9/10      KQLDGVSKGWP-PCLRALAATALLVQEANKLTLG----------------
Chr7_63858202_1754_po_C_Rz12/9      KQLT-VVLGWP-PCCKALATTTLLTQEADKLTLG----------------
ERV3_chr7q21_AC073210_polputei      KQLT-VVLGWP-PCCKALATTTLLTQEADKLTLG----------------
Chr6_110218096_843_po_C_Xz16/8      KQQDSVVLGZL-PCLRALAATTILTQEADKLTLG----------------
Chr6_110214535_786_po_C_Xz16/8      KQQDSVVLGZL-PCLRALAATTILTQEADKLTLG----------------
ChrY_3018841_881_po_C_Rz9/12        KQLDSVAWGWP-PCLRALAATALLTQDADKLTLG----------------
ChrX_86876098_639_po_C_Xz7/8        KQLDSVAWGWP-PCLRALAATALLTQDADKLTLG----------------
Chr14_68899917_965_po_C_Xz6/8       KQLDSVVLEWP-PCLRALAATALLAQEADKMTLG----------------
ChrY_23874437_661_po_C_Xz15/9       TQLNSLALAZP-PCLRALADTTLLEQQDRKLALR----------------
ChrY_27518418_649_po_C_Xz15/9       TQLNSLALAZP-PCLRALADTTLLEQQDRKLALR----------------
ChrY_25016892_657_po_C_Xz13/12      TQLNSLALAZP-PCLRALADTTLLEQQDRKLALR----------------
ChrY_19932937_520_po_C_Xz10/14      TQLNSVALGZP-PWLRALTATSLLEQEAKKLTLR----------------
ChrY_20104614_518_po_C_Xz10/14      TQLNSVALGZP-PWLRALTATSLLEQEAKKLTLR----------------
Chr19_58660155_598_po_C_Xz21/1      KQPDSVVLRWP-LCLKALAATILLTQEASKLTLG----------------
Chr19_58621153_500_po_C_Xz21/1      QQLDSMVLRWP-LCLKALAATILLTQEASKLTLG----------------
Chr7_139686987_710_po_C_Xz11/7      KQLDSAAR--M-ACFNTLATSVLLAQEASKLTLG----------------
Chr7_137651301_1097_po_C_Pz13/      KQLDSVALGWP-LCLKALAATALLAQEADQLTLG----------------
Chr11_29579197_930_po_C_Pz8/12      KELDLVALGWP-LCLKPLAATALLAQKADKLTLG----------------
Chr4_117329205_488_po_C_Xz3/16      KQLDSMALRWP-PWLRALAA-------FIKLTLG----------------
Chr21_38522185_383_po_C_Xz7/9       KQLDSVALRZP-PCFKTLAVTALLAQEANKLTLG----------------
Chr4_70173370_1049_po_C_Xz11/8      RQLNSVAPGWS-LCFKTLAITARLAQEANKLTLG----------------
Chr4_70022513_788_po_C_Kz10/12      RQLDSVAFGWP-PSFKALAATALLGQKANKLTLG----------------
Chr14_105225894_1027_po_C_Xz13      RQLDTVALAWP--LFKALAATASLAQEANKLTLG----------------
Chr7_119241536_761_po_C_Xz11/1      KQLDTIALGWP-PCLRALTDFALLAQEADELTVG----------------
ChrY_22564970_567_po_C_Xz9/10       KQLDPLVLSWP-PCFKTSAATTLLSQEANKLTLG----------------
Chr21_14409346_1477_po_C_Pz5/6      KQLDFVAKGWP-PCFKELAATALLAEDANKLTFG----------------
ChrX_51259700_1396_po_C_Pz7/8       KQLDWVATGWP-PRFKVLAVTALLAEDANKLTFG----------------
Chr6_64985396_1041_po_C_Xz11/1      K----TRLGGY-G--MALAATALLAQDANKLTFG----------------
ChrY_20355209_659_po_C_Xz7/11       KZLDFVAMGWP-LGFKALAATALLAENANKLTFG----------------
Chr3_34567166_523_po_C_Xz8/6        KRLDLVAMGWP-PCFKALAATVLLVEDANKLTFK----------------
Chr3_122072468_245_po_C_Xz9/9       KCLDLVAMGWP-PCFKALAATALLVEDANKLTFG----------------
Chr11_58788475_541_po_C_Xz8/22      KSFQYLGFYVS-QGKRWLGGEZKQAVCALPTPTTQHQIRKFLETAGFCHI
Chr6_137421042_373_po_C_Xz6/9       TNLLYTRRAHD-SKSRSKKSHDKDPYCPNNR-------------------
Chr4_115910608_550_po_C_Pz8/12      KGLDLVAMGWP-LCFKALAATVLLAEDANKLKLG----------------
ChrY_13743185_977_po_C_Xz14/8       KRLDLVALGWP--CLRALAATAILIEDANKLALG----------------
HERV15                              KRLDLVALGWP-HCLRALAATAILIEDANKLALG----------------
ChrX_117198092_983_po_C_Xz8/8       KRLDLVALGWP-HCLRALAATAILIEDANKLALG----------------
Chr19_21166030_1412_po_C_Xz9/1      KRLDZVALGWH-HCLRALAATAILIEDANKLALG----------------
Chr19_21213743_913_po_C_Xz14/1      KRLDLVALDWP-YCLRALAATTTLIEDANKLALG----------------
Chr19_15957774_569_po_C_Xz8/10      KELDS-MHGAR-MASFSFLFFFFLRRSLAVLPRL----------------
Chr11_77508681_297_po_C_Xz19/1      KQLDSVMLGWP-PCVKALAATTLLTQEASKLTLR----------------
Chr13_98250769_738_po_C_Xz7/13      EEGHCLLPALGTECLGIKPDVPLFNSEIGEKPPY----------------
Chr6_39968183_513_po_C_Pz10/9       KQLDLMALGWP-PCFKALATTALLAQEANKLALG----------------
Chr15_66867263_645_po_C_Xz11/1      RFGLLLTLGSDSGPAFVAEIVQDLTQLLNKTEAT----------------
HERV18_polputein                    QLWKGAELRYS--LIEKQLAAAYAALQARESVAG----------------
HERVIP10FH_polputein                KILNPVTRGWP-KCIQSVAATALLTEESRKTTFR----------------
MER66_polputein                     RFTRPFKNGNK-------------VLPLTRLSLT----------------


HERVI_polputein                     --------------------------------------------------
HERVIP10F_polputein                 --------------------------------------------------
HERVadp_AC005741                    --------------------------------------------------
HERV-PT47D                          --------------------------------------------------
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      --------------------------------------------------
HERV_Y_chr12_51022911_pol           --------------------------------------------------
HERV17_polputein                    --------------------------------------------------
HERV_W_chr7_9105739_syncytin_p      --------------------------------------------------
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  --------------------------------------------------
ERVfrd_AC004022                     --------------------------------------------------
PRIMA41_polputein                   --------------------------------------------------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 --------------------------------------------------
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           --------------------------------------------------
ChrY_13790029_693_po_C_Xz12/11      --------------------------------------------------
HERVS71_polputein                   --------------------------------------------------
BaEV_M7_pol                         --------------------------------------------------
HERVR_polputein                     --------------------------------------------------
AKR_MLV_MLOCG_RT                    --------------------------------------------------
```

```
GaLV_pol                            --------------------------------------------------
Chr3_158587414_1662_po_C_Ez10/      --------------------------------------------------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      --------------------------------------------------
Chr5_111890459_1626_po_C_Ez10/      --------------------------------------------------
ChrX_48766518_1600_po_C_Ez6/3       --------------------------------------------------
ChrX_152580713_1572_po_C_Ez10/      --------------------------------------------------
ChrX_139287533_1154_po_C_Xz8/4      --------------------------------------------------
Chr2_52615786_912_po_C_Xz11/5       --------------------------------------------------
HERV-E_4_1                          --------------------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      --------------------------------------------------
ChrY_8634032_950_po_C_Xz14/10       --------------------------------------------------
Chr3_126735360_997_po_C_Xz5/10      --------------------------------------------------
Chr3_126713203_958_po_C_Xz7/8       --------------------------------------------------
Chr3_75226118_977_po_C_Xz8/9        --------------------------------------------------
Chr7_6629919_827_po_C_Xz8/7         --------------------------------------------------
Chr4_9166812_811_po_C_Xz8/10        --------------------------------------------------
Chr12_8478642_834_po_C_Xz12/7       --------------------------------------------------
Chr10_15187175_787_po_C_Xz11/5      --------------------------------------------------
Chr11_71819740_854_po_C_Xz11/9      --------------------------------------------------
Chr11_71519974_886_po_C_Xz13/9      --------------------------------------------------
Chr8_7271441_969_po_C_Xz10/6        --------------------------------------------------
Chr8_11770066_826_po_C_Xz9/8        --------------------------------------------------
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      --------------------------------------------------
Chr14_50205291_837_po_C_Xz10/8      --------------------------------------------------
Chr7_6597120_842_po_C_Xz7/7         --------------------------------------------------
Chr11_71544658_783_po_C_Xz8/13      --------------------------------------------------
Chr11_67999520_760_po_C_Xz8/8       --------------------------------------------------
Chr3_75461975_796_po_C_Xz7/9        --------------------------------------------------
Chr12_52228939_692_po_C_Xz3/10      --------------------------------------------------
Chr10_15211234_286_po_C_Xz11/1      --------------------------------------------------
Chr12_100973302_945_po_C_Xz12/      --------------------------------------------------
Chr19_9685420_817_po_C_Xz6/10       --------------------------------------------------
Chr13_40936820_719_po_C_Xz9/10      --------------------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      --------------------------------------------------
ERV3_chr7q21_AC073210_polputei      --------------------------------------------------
Chr6_110218096_843_po_C_Xz16/8      --------------------------------------------------
Chr6_110214535_786_po_C_Xz16/8      --------------------------------------------------
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        --------------------------------------------------
Chr14_68899917_965_po_C_Xz6/8       --------------------------------------------------
ChrY_23874437_661_po_C_Xz15/9       --------------------------------------------------
ChrY_27518418_649_po_C_Xz15/9       --------------------------------------------------
ChrY_25016892_657_po_C_Xz13/12      --------------------------------------------------
ChrY_19932937_520_po_C_Xz10/14      --------------------------------------------------
ChrY_20104614_518_po_C_Xz10/14      --------------------------------------------------
Chr19_58660155_598_po_C_Xz21/1      --------------------------------------------------
Chr19_58621153_500_po_C_Xz21/1      --------------------------------------------------
Chr7_139686987_710_po_C_Xz11/7      --------------------------------------------------
Chr7_137651301_1097_po_C_Pz13/      --------------------------------------------------
Chr11_29579197_930_po_C_Pz8/12      --------------------------------------------------
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      --------------------------------------------------
Chr4_70022513_788_po_C_Kz10/12      --------------------------------------------------
Chr14_105225894_1027_po_C_Xz13      --------------------------------------------------
Chr7_119241536_761_po_C_Xz11/1      --------------------------------------------------
ChrY_22564970_567_po_C_Xz9/10       --------------------------------------------------
Chr21_14409346_1477_po_C_Pz5/6      --------------------------------------------------
ChrX_51259700_1396_po_C_Pz7/8       --------------------------------------------------
Chr6_64985396_1041_po_C_Xz11/1      --------------------------------------------------
ChrY_20355209_659_po_C_Xz7/11       --------------------------------------------------
Chr3_34567166_523_po_C_Xz8/6        --------------------------------------------------
Chr3_122072468_245_po_C_Xz9/9       --------------------------------------------------
Chr11_58788475_541_po_C_Xz8/22      WTLSFLLMAKPLYEATKWGEGTPLWGKEQVMAFKEIKKALIQALALGLPD
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------------------
ChrY_13743185_977_po_C_Xz14/8       --------------------------------------------------
HERV15                              --------------------------------------------------
ChrX_117198092_983_po_C_Xz8/8       --------------------------------------------------
Chr19_21166030_1412_po_C_Xz9/1      --------------------------------------------------
Chr19_21213743_913_po_C_Xz14/1      --------------------------------------------------
Chr19_15957774_569_po_C_Xz8/10      --------------------------------------------------
Chr11_77508681_297_po_C_Xz19/1      --------------------------------------------------
Chr13_98250769_738_po_C_Xz7/13      --------------------------------------------------
Chr6_39968183_513_po_C_Pz10/9       --------------------------------------------------
```

```
Chr15_66867263_645_po_C_Xz11/1      --------------------------------------------------
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                --------------------------------------------------
MER66_polputein                     --------------------------------------------------


HERVI_polputein                     --------------------------------------------------
HERVIP10F_polputein                 --------------------------------------------------
HERVadp_AC005741                    --------------------------------------------------
HERV-PT47D                          --------------------------------------------------
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      --------------------------------------------------
HERV_Y_chr12_51022911_pol           --------------------------------------------------
HERV17_polputein                    --------------------------------------------------
HERV_W_chr7_9105739_syncytin_p      --------------------------------------------------
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  --------------------------------------------------
ERVfrd_AC004022                     --------------------------------------------------
PRIMA41_polputein                   --------------------------------------------------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 --------------------------------------------------
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           --------------------------------------------------
ChrY_13790029_693_po_C_Xz12/11      --------------------------------------------------
HERVS71_polputein                   --------------------------------------------------
BaEV_M7_pol                         --------------------------------------------------
HERVR_polputein                     --------------------------------------------------
AKR_MLV_MLOCG_RT                    --------------------------------------------------
GaLV_pol                            --------------------------------------------------
Chr3_158587414_1662_po_C_Ez10/      --------------------------------------------------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      --------------------------------------------------
Chr5_111890459_1626_po_C_Ez10/      --------------------------------------------------
ChrX_48766518_1600_po_C_Ez6/3       --------------------------------------------------
ChrX_152580713_1572_po_C_Ez10/      --------------------------------------------------
ChrX_139287533_1154_po_C_Xz8/4      --------------------------------------------------
Chr2_52615786_912_po_C_Xz11/5       --------------------------------------------------
HERV-E_4_1                          --------------------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      --------------------------------------------------
ChrY_8634032_950_po_C_Xz14/10       --------------------------------------------------
Chr3_126735360_997_po_C_Xz5/10      --------------------------------------------------
Chr3_126713203_958_po_C_Xz7/8       --------------------------------------------------
Chr3_75226118_977_po_C_Xz8/9        --------------------------------------------------
Chr7_6629919_827_po_C_Xz8/7         --------------------------------------------------
Chr4_9166812_811_po_C_Xz8/10        --------------------------------------------------
Chr12_8478642_834_po_C_Xz12/7       --------------------------------------------------
Chr10_15187175_787_po_C_Xz11/5      --------------------------------------------------
Chr11_71819740_854_po_C_Xz11/9      --------------------------------------------------
Chr11_71519974_886_po_C_Xz13/9      --------------------------------------------------
Chr8_7271441_969_po_C_Xz10/6        --------------------------------------------------
Chr8_11770066_826_po_C_Xz9/8        --------------------------------------------------
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      --------------------------------------------------
Chr14_50205291_837_po_C_Xz10/8      --------------------------------------------------
Chr7_6597120_842_po_C_Xz7/7         --------------------------------------------------
Chr11_71544658_783_po_C_Xz8/13      --------------------------------------------------
Chr11_67999520_760_po_C_Xz8/8       --------------------------------------------------
Chr3_75461975_796_po_C_Xz7/9        --------------------------------------------------
Chr12_52228939_692_po_C_Xz3/10      --------------------------------------------------
Chr10_15211234_286_po_C_Xz11/1      --------------------------------------------------
Chr12_100973302_945_po_C_Xz12/      --------------------------------------------------
Chr19_9685420_817_po_C_Xz6/10       --------------------------------------------------
Chr13_40936820_719_po_C_Xz9/10      --------------------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      --------------------------------------------------
ERV3_chr7q21_AC073210_polputei      --------------------------------------------------
Chr6_110218096_843_po_C_Xz16/8      --------------------------------------------------
Chr6_110214535_786_po_C_Xz16/8      --------------------------------------------------
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        --------------------------------------------------
Chr14_68899917_965_po_C_Xz6/8       --------------------------------------------------
ChrY_23874437_661_po_C_Xz15/9       --------------------------------------------------
ChrY_27518418_649_po_C_Xz15/9       --------------------------------------------------
ChrY_25016892_657_po_C_Xz13/12      --------------------------------------------------
ChrY_19932937_520_po_C_Xz10/14      --------------------------------------------------
```

```
ChrY_20104614_518_po_C_Xz10/14      --------------------------------------------------
Chr19_58660155_598_po_C_Xz21/1      --------------------------------------------------
Chr19_58621153_500_po_C_Xz21/1      --------------------------------------------------
Chr7_139686987_710_po_C_Xz11/7      --------------------------------------------------
Chr7_137651301_1097_po_C_Pz13/      --------------------------------------------------
Chr11_29579197_930_po_C_Pz8/12      --------------------------------------------------
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      --------------------------------------------------
Chr4_70022513_788_po_C_Kz10/12      --------------------------------------------------
Chr14_105225894_1027_po_C_Xz13      --------------------------------------------------
Chr7_119241536_761_po_C_Xz11/1      --------------------------------------------------
ChrY_22564970_567_po_C_Xz9/10       --------------------------------------------------
Chr21_14409346_1477_po_C_Pz5/6      --------------------------------------------------
ChrX_51259700_1396_po_C_Pz7/8       --------------------------------------------------
Chr6_64985396_1041_po_C_Xz11/1      --------------------------------------------------
ChrY_20355209_659_po_C_Xz7/11       --------------------------------------------------
Chr3_34567166_523_po_C_Xz8/6        --------------------------------------------------
Chr3_122072468_245_po_C_Xz9/9       --------------------------------------------------
Chr11_58788475_541_po_C_Xz8/22      TKAFLLYVHARKGIATGLLVQTLGSWYLPVAYLSKRQDLVAMGWPPVSGI
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------------------
ChrY_13743185_977_po_C_Xz14/8       --------------------------------------------------
HERV15                              --------------------------------------------------
ChrX_117198092_983_po_C_Xz8/8       --------------------------------------------------
Chr19_21166030_1412_po_C_Xz9/1      --------------------------------------------------
Chr19_21213743_913_po_C_Xz14/1      --------------------------------------------------
Chr19_15957774_569_po_C_Xz8/10      --------------------------------------------------
Chr11_77508681_297_po_C_Xz19/1      --------------------------------------------------
Chr13_98250769_738_po_C_Xz7/13      --------------------------------------------------
Chr6_39968183_513_po_C_Pz10/9       --------------------------------------------------
Chr15_66867263_645_po_C_Xz11/1      --------------------------------------------------
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                --------------------------------------------------
MER66_polputein                     --------------------------------------------------


HERVI_polputein                     -----------------WRLTVSTPHQ-----------VRAILNKKA---
HERVIP10F_polputein                 -----------------GKLTVSTPHQ-----------VRTILNQRA---
HERVadp_AC005741                    -----------------GALIVSTPHQ-----------VRNILNZKA---
HERV-PT47D                          -----------------ANFLCFLVET-----------GFHRVSQDG---
HERV_W_chr6_141432567_ERV9_lik      -----------------KDLTGWTTHDVNG---ILGA-------------
ERV9_PH1_RT                         -----------------KDLT-STTYDVNG---ILGA-------------
Chr13_54564403_572_po_C_Yz7/14      ------------------DLTVWTSHDVNS---IHSA-------------
HERV_Y_chr12_51022911_pol           -----------------RDLTVWTSHDMNG---ILTA-------------
HERV17_polputein                    -----------------RDLTVWTSHDVNG---ILTA-------------
HERV_W_chr7_9105739_syncytin_p      -----------------RDLTVWTSHDVNG---ILTA-------------
HERV9_polputein                     -----------------KDLTVWTTHVVNG---ILGA-------------
HUERS-P3_polputein                  -----------------NDLTVYTPHNVAG---LLSS-------------
ERVfrd_AC004022                     -----------------QSIWLLMSHQIDP---LLNL-------------
PRIMA41_polputein                   -----------------QHLEVLTPHQVQG---VLEA-------------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       -----------------AQLILYSSHNFQN---LFSSS------------
HERV_H48I_polputein                 -----------------ANLTVYSTHNIKD---MLAHR------------
HERV_H_RTVLH2_pol                   -----------------VQLTLYISHNFQN---LFSSS------------
HERVFH21_pol                        -----------------QNTTVYSPHNLQD---LLSSR------------
HERVRblike_chr4_109047953           -----------------QPTMISVPLSLG-----VVRT------------
ChrY_13790029_693_po_C_Xz12/11      -----------------QNFTLMAPHA-----------IEILLPSAS---
HERVS71_polputein                   -----------------QNLTLMAPHA-----------IETLLQRRS---
BaEV_M7_pol                         -----------------QPLTVITPHT-----------LEAIVRQPP---
HERVR_polputein                     -----------------QPLTVITPHT-----------LEAIVRQPP---
AKR_MLV_MLOCG_RT                    -----------------QPLVILAPHA-----------VEALVKQPP---
GaLV_pol                            -----------------QNVTVIASHS-----------LESIVRQPP---
Chr3_158587414_1662_po_C_Ez10/      -----------------QNLNIKVPHA-----------VVTLMNTKG---
Chr3_79627623_1138_po_C_Ez4/7       -----------------QNLSIKAPHA-----------MVTLMNTKG---
Chr19_20477244_1350_po_C_Ez10/      -----------------PNLNIS-PHA-----------VVTLMNTKG---
Chr5_111890459_1626_po_C_Ez10/      -----------------QNLNIKAPHA-----------VVTLMNTKG---
ChrX_48766518_1600_po_C_Ez6/3       -----------------QNLNIKAPHA-----------VVTLMKTKG---
ChrX_152580713_1572_po_C_Ez10/      -----------------QNLNIKALHA-----------VVTLMNSKG---
ChrX_139287533_1154_po_C_Xz8/4      -----------------QNLSIKAPHA-----------VVTLMNTKG---
Chr2_52615786_912_po_C_Xz11/5       -----------------QNLNIKAPHA-----------VLTLTNTKG---
HERV-E_4_1                          -----------------QNLNIKDPHA-----------VVTLMNTRG---
ChrY_21962399_1137_po_C_Xz13/9      -----------------QNLNIKAPHA-----------VVTLMNTKK---
ChrY_8634032_950_po_C_Xz14/10       -----------------QNLNEKAPSV-----------VT-LMNTKV---
Chr3_126735360_997_po_C_Xz5/10      -----------------QNLNIKASRA-----------VVTLMNTKG---
Chr3_126713203_958_po_C_Xz7/8       -----------------QNLNIKASRA-----------VVTLMNTKG---
Chr3_75226118_977_po_C_Xz8/9        -----------------QNLDIKSSRA-----------VVTLMNTKG---
```

```
Chr7_6629919_827_po_C_Xz8/7         -----------------QNLNRKASRA-----------VVTLMNTKG---
Chr4_9166812_811_po_C_Xz8/10        -----------------QNLNIKASHA-----------VVTLMNAKG---
Chr12_8478642_834_po_C_Xz12/7       -----------------QNLNIKASRA-----------VVTLMNTKG---
Chr10_15187175_787_po_C_Xz11/5      -----------------QNLNIKASRA-----------VVTLMNTKG---
Chr11_71819740_854_po_C_Xz11/9      -----------------QKLNIKASRA-----------VVTLMNTKG---
Chr11_71519974_886_po_C_Xz13/9      -----------------QNLNIKASHA-----------VVTLMNTKG---
Chr8_7271441_969_po_C_Xz10/6        -----------------RNLNIKASRA-----------VV--MNTKG---
Chr8_11770066_826_po_C_Xz9/8        -----------------QSLNINASHS-----------VVTLMSSKG---
Chr4_4128362_651_po_C_Xz7/11        -----------------QNLNRKAPHA-----------VVTLMNTEG---
Chr14_50203955_850_po_C_Xz10/8      -----------------QNLNTKAPHA-----------VVTLMNTKG---
Chr14_50205291_837_po_C_Xz10/8      -----------------QNLNTKAPHA-----------VVTLMNTKG---
Chr7_6597120_842_po_C_Xz7/7         -----------------QNLNRKAPHA-----------VVTLMNTKG---
Chr11_71544658_783_po_C_Xz8/13      -----------------QNLNRKAPHT-----------VVTLMNTKG---
Chr11_67999520_760_po_C_Xz8/8       -----------------QNLNRKAPHA-----------VVTLMNIKG---
Chr3_75461975_796_po_C_Xz7/9        -----------------QNLNIKASHA-----------VVTLMNTKG---
Chr12_52228939_692_po_C_Xz3/10      -----------------QNLNIKAPHA-----------VLTLMNTKG---
Chr10_15211234_286_po_C_Xz11/1      -----------------ZNLTEKAPHA-----------VVTLMNTKG---
Chr12_100973302_945_po_C_Xz12/      -----------------QNLNIKAPHA-----------VVTLVNTKG---
Chr19_9685420_817_po_C_Xz6/10       -----------------QNLNINVPHT-----------VVTLMNTKG---
Chr13_40936820_719_po_C_Xz9/10      -----------------QNLNIKASHA-----------VVTLMDTKG---
Chr7_63858202_1754_po_C_Rz12/9      -----------------QQLTIQVPHA-----------VITLMDQRE---
ERV3_chr7q21_AC073210_polputei      -----------------QQLTIQVPHA-----------VITLMDQRE---
Chr6_110218096_843_po_C_Xz16/8      ------------------QLTIWVLHS-----------IITLMDQRG---
Chr6_110214535_786_po_C_Xz16/8      ------------------QLTIWVLHS-----------IITLMDQRG---
ChrY_3018841_881_po_C_Rz9/12        -----------------QQLTPHS--------------VITLMDQRG---
ChrX_86876098_639_po_C_Xz7/8        ------------------QLTPHS--------------VMTLMDQRG---
Chr14_68899917_965_po_C_Xz6/8       -----------------QQLTIQVPHL-----------VITLMDQRR---
ChrY_23874437_661_po_C_Xz15/9       -----------------QQLTIGIPQS-----------VLTLMDZTA---
ChrY_27518418_649_po_C_Xz15/9       -----------------QQLTIGIPQS-----------VLTLMDZTA---
ChrY_25016892_657_po_C_Xz13/12      -----------------QQLTIGIPQS-----------VLTLMDZTA---
ChrY_19932937_520_po_C_Xz10/14      -----------------QQLTIQVPHS-----------VLTLMDDRA---
ChrY_20104614_518_po_C_Xz10/14      -----------------QQLTIQVPHS-----------VLTLMDDRA---
Chr19_58660155_598_po_C_Xz21/1      -----------------QQLTVWVPHZ-----------VITLMDQRR---
Chr19_58621153_500_po_C_Xz21/1      -----------------QQLTVPVPHS-----------VITLMDQRR---
Chr7_139686987_710_po_C_Xz11/7      -----------------QQLTIWVLHSVQQLSIWVLHSVITSIDQRR---
Chr7_137651301_1097_po_C_Pz13/      -----------------QQLTIQVLHS-----------VITLMDQRW---
Chr11_29579197_930_po_C_Pz8/12      -----------------QQLTI-VPYL-----------VIILMDQRG---
Chr4_117329205_488_po_C_Xz3/16      -----------------QQLTIQVPHS-----------VTTLMDQRG---
Chr21_38522185_383_po_C_Xz7/9       -----------------QQLTIQLPHS-----------VITLMNQRG---
Chr4_70173370_1049_po_C_Xz11/8      -----------------QQVTIQLLHS-----------VITLMDQRG---
Chr4_70022513_788_po_C_Kz10/12      -----------------QQMTIQVPHS-----------VVTLMDQRG---
Chr14_105225894_1027_po_C_Xz13      -----------------RQLIIQVPYS-----------VITLMDKRG---
Chr7_119241536_761_po_C_Xz11/1      -----------------QQLTIWVPHL-----------VITLMDQRG---
ChrY_22564970_567_po_C_Xz9/10       -----------------QKLIIRVPHT-----------VIILMDQRG---
Chr21_14409346_1477_po_C_Pz5/6      -----------------QKLIIQVPHT-----------VVTLTEQRG---
ChrX_51259700_1396_po_C_Pz7/8       -----------------QKLIIQVPHT-----------VVTLMEQRG---
Chr6_64985396_1041_po_C_Xz11/1      -----------------QRLIICVPHM-----------VITLLEQRG---
ChrY_20355209_659_po_C_Xz7/11       -----------------QRLIIQVPHT-----------VITLMEQNG---
Chr3_34567166_523_po_C_Xz8/6        -----------------QSLIIWMPHA-----------VVTLMEQKG---
Chr3_122072468_245_po_C_Xz9/9       -----------------QRLIIGAPHG-----------RVTLMEQRG---
Chr11_58788475_541_po_C_Xz8/22      GSHYHMLAEDANKLTFGQRLIIWVPHT-----------AITLMEQRG---
Chr6_137421042_373_po_C_Xz6/9       -----------------GMVIIWVPHT-----------VVTLMEQRG---
Chr4_115910608_550_po_C_Pz8/12      -----------------QRLIIRVPHT-----------VITLMEQSG---
ChrY_13743185_977_po_C_Xz14/8       -----------------QKIIFRVPHT-----------VVTLMGQRG---
HERV15                              -----------------QKIIFRVPHT-----------VVTLMEQRG---
ChrX_117198092_983_po_C_Xz8/8       -----------------QKLIV-VPHA-----------VVTLMEQRG---
Chr19_21166030_1412_po_C_Xz9/1      -----------------QKLKIGVPHA-----------VVSLMDQRG---
Chr19_21213743_913_po_C_Xz14/1      -----------------QKLIIQVPHT-----------VVTLMEQRE---
Chr19_15957774_569_po_C_Xz8/10      ------------ECSSVISAHC----------------NLRLLGSRDSCV
Chr11_77508681_297_po_C_Xz19/1      ------------QQLTIREPHS----------------VITLMDRPSSCV
Chr13_98250769_738_po_C_Xz7/13      ------------GGRRDVCSNA----------------ALLFFTPLRCLG
Chr6_39968183_513_po_C_Pz10/9       ------------QKLIIQVLHT----------------VITLMDQREHHZ
Chr15_66867263_645_po_C_Xz11/1      ------------HSLPAAELRK----------------SGAHEPDTQAAA
HERV18_polputein                    ----------------WATVIVRMTYPIAG-------WVRSWVDPPD---
HERVIP10FH_polputein                -----------------GNLIVSTPHQVR--------TILSPKKKKK---
MER66_polputein                     -----------------KNLKEGSPES-----------------------


HERVI_polputein                     ---GRRLTDSRILK-YGAILLKKD--------------------------
HERVIP10F_polputein                 ---GRWLTDSRILK-YEAILLEKDDLTLTTDNSLNPAGF-LTGNPN----
HERVadp_AC005741                    ---RTWLMDSQILK-YEAILVEZDDLVITDTCLNLASFLZKGEENK----
HERV-PT47D                          ---LDLLTSZ----------------------------------------
HERV_W_chr6_141432567_ERV9_lik      -KZSLWLSDNHLLG-YQVLLLEGLVLQICTCV-ALNPATFLPE-DGKPI-
ERV9_PH1_RT                         -KGSLWLSDNCLLR-YQALLLEGPVLQICTCM-ALNPATFLPE-DGEPI-
Chr13_54564403_572_po_C_Yz7/14      -KGDLWLSDNHLLK-YQ-LLLEGPVLQLRTCA-TLNPATFLPD-NEEKI-
HERV_Y_chr12_51022911_pol           -KGDLWLSDNDLLK-YQALLLEGPVLRLCTCA-TLNZATCLPD-NEEKI-
HERV17_polputein                    -KGDLWLSDNRLLK-YQALLLEGPVLRLRTCA-TLNPATFLPD-NEEKI-
```

```
HERV_W_chr7_9105739_syncytin_p      -KGDLWLSDNCLLK-CQALLLEGPVLRLCTCA-TLNPATFLPD-NEEKI-
HERV9_polputein                     -KGILWLSDYHLLR-YQALLLEGPVLQISTCV-ALNPATFLPE-NGEPI-
HUERS-P3_polputein                  -RGSLWLTDSXLLK-YQALLLEXSXIQLKTCS-HLNPDTFLXE-ETGEP-
ERVfrd_AC004022                     -KGPZWLTDNTLIK-YQVLLLENPQVTVEQCS-TINPALLLPLPGDDNS-
PRIMA41_polputein                   -KRHQRMTGGHVYK-YQALLLN-PDITLKVCQ-TLNPATYLPE-PTGTL-
HERV_H_RGH1_pol                     ------------------------------GL-ITSPA------------
HERVHRGH2_pol                       -HLTHILSAPRLLQ-IYTPFILP--ITIIPGL-DFNPASHIILDTIPDP-
HERV_H48I_polputein                 -SVLSLISAPRLLQ-LYALFIETPHITMLTSS-RLNPATLLPEATTAQDP
HERV_H_RTVLH2_pol                   -HLMHYFLLPGS--------------------------------------
HERVFH21_pol                        -ALSSLP-PSRIQL-LHALFIKNPKFSLAKSA-PLNPASLLPVSSSLPT-
HERVRblike_chr4_109047953           -KGGYWLTAGRLGQ-YQAILLDDPAVKLQTTG-TLNPTTLLPPTGESEEL
ChrY_13790029_693_po_C_Xz12/11      ---GKWMSNARILQ-YQSLLLDQPLLTFSSIR-CLNPAAFLPDPNF----
HERVS71_polputein                   ---GKWMSNAHILQ-YQSLLLDQPWLTFSPTR-CLNPATFLPDPDL----
BaEV_M7_pol                         ---DRWITNARLTH-YQALLLDTDRVQFGPPV-TLNPATLLPVPEN----
HERVR_polputein                     ---DRWITNARLTH-YQALLLDTDRVQFGPPV-TLNPATLLPVPEN----
AKR_MLV_MLOCG_RT                    ---DRWLSNARMTH-YQALLLDTDRVQFGPVV-ALNPATLLPLPE-----
GaLV_pol                            ---DRWMTNARMTH-YQSLLLN-ERVSFAPPA-VLNPATLLPVESE----
Chr3_158587414_1662_po_C_Ez10/      ---HHWLTNARLTK-YQSLLCENPHITIEVCN-TLNPATLLPVZES----
Chr3_79627623_1138_po_C_Ez4/7       ---HHWLTNARLTK-YQSLLCESP-ITIKVCN-TLNPATLLPVSES----
Chr19_20477244_1350_po_C_Ez10/      ---HHWLTNARLTK-YQSLLCENPHINIEVCN-TLNPTTLLPVSES----
Chr5_111890459_1626_po_C_Ez10/      ---HHWLPNASLIK-YQSLLCENPRITTEVCN-TLNPA-LLPVSES----
ChrX_48766518_1600_po_C_Ez6/3       ---HHWLTNARLTK-YQSLLCENPHITTEVCN-TLNPATLLLVSES----
ChrX_152580713_1572_po_C_Ez10/      ---HHWLTNARLTK-YQSLLCENPRITIEVRN-TLNPATLLPVSES----
ChrX_139287533_1154_po_C_Xz8/4      ---HHWLTNARLTK-YQSLVCENPRITIEVCN-TLNPATLLPVSDS----
Chr2_52615786_912_po_C_Xz11/5       ---HHWLTNARLTS-YQSLLFENPYITIEVCN-TLNPATLLPVSEP----
HERV-E_4_1                          ---HHWLTNARLTK-YQSLLCENPHITIEVCN-TLNPATLLPVLEI----
ChrY_21962399_1137_po_C_Xz13/9      ---HHLLMNARLTR-YQISLCENPCITTEVCN-TLNPATLFPVSEP----
ChrY_8634032_950_po_C_Xz14/10       ---HHWLKNALLTK-YQNLLCENLHIDIEICN-TLNPATRLPVSER----
Chr3_126735360_997_po_C_Xz5/10      ---HNWLMNVRLTK-YQTAL-ENPCITIEVCN-TLHPASLLPVSES----
Chr3_126713203_958_po_C_Xz7/8       ---HHWLTDARLIK-YQTLLCENPRITIEVCN-TLHPASLLPVSES----
Chr3_75226118_977_po_C_Xz8/9        ---HHWLTNARLTK-YQTAL-ENPRITIEVCN-TLHPATLLPVSES----
Chr7_6629919_827_po_C_Xz8/7         ---HHWLTNATLTK-YQTLLCENPHITIEVCN-TLHPATLLPVSE-----
Chr4_9166812_811_po_C_Xz8/10        ---HHWLMNARLTK-YQSWLCENPHITIEVCN-SLHPATLLPLSES----
Chr12_8478642_834_po_C_Xz12/7       ---HHWLMNARLTK-DQTLLCENPHITMEVCHPTPPPPTLLQVSES----
Chr10_15187175_787_po_C_Xz11/5      ---HHWLTKARLTK-DETLLCESPHITIEVCN-TLHPATLLQVSES----
Chr11_71819740_854_po_C_Xz11/9      ---HHWLTNATLTD-YQTLLCENPRITIEVCN-TLHPATLLPVSKS----
Chr11_71519974_886_po_C_Xz13/9      ---HHWLT-ARLTK-YQTLLCENPHITTEVCN-SLHPATLLPLSES----
Chr8_7271441_969_po_C_Xz10/6        ---RHWLTNARLTK-YQTLVCENPRITIEVCN-TLHPATLLPVSGS----
Chr8_11770066_826_po_C_Xz9/8        ---HHWLMNARLTK-YQTLLCENPHITSKVCN-TLHPATLLPVSES----
Chr4_4128362_651_po_C_Xz7/11        ---HHWLMNAKLTK-YQTLLCENPRITIEVCN-SLHPATLLLVSES----
Chr14_50203955_850_po_C_Xz10/8      -------------------------------------------IMS----
Chr14_50205291_837_po_C_Xz10/8      -------------------------------------------IMS----
Chr7_6597120_842_po_C_Xz7/7         ---HHWLTNARLTK-YQSLLCENPRITIEVCN-TLHPATLLPLSES----
Chr11_71544658_783_po_C_Xz8/13      ---HHWLTNARLTK-YQTLLCENPRITIEVCN----PATLLPLSES----
Chr11_67999520_760_po_C_Xz8/8       ---HHWLMNARLTK-YQSLLCENPRVTIEVCN----PATLLPLSES----
Chr3_75461975_796_po_C_Xz7/9        ---HHWLTNARLTK-YQTLLCENPRITIEVCN-TLHPATLLPVSES----
Chr12_52228939_692_po_C_Xz3/10      ---HHWLTNARLTK-YQSLLCENLRITIEVWN-TLNPTTLLPVSES----
Chr10_15211234_286_po_C_Xz11/1      ---HDWLTNARLTK-YQS---ENPHITIEVCN-ILHPATLLPLSES----
Chr12_100973302_945_po_C_Xz12/      ---HHWLTNARLTK-YQILLCKNPRITIEVCN-TLNPTTLLPISES----
Chr19_9685420_817_po_C_Xz6/10       ---HHWLTNARLTK-YQSLLCENPHITTEICN-TLSPATLLPVSES----
Chr13_40936820_719_po_C_Xz9/10      ---HHWLTNARLTK-YQTLLCENPCITIEVCN-TLHPATLLPVSES----
Chr7_63858202_1754_po_C_Rz12/9      ---QHWLSNPTMTZ-YQGLLCENPRITLETVN-TLNPATLLPVK------
ERV3_chr7q21_AC073210_polputei      ---QHWLSNPTMTZ-YQGLLCENPRITLETVN-TLNPATLLPVK------
Chr6_110218096_843_po_C_Xz16/8      ---HHWLLNPRMTQ-YQGLLCENPHITLETVN-TLNPANLLPVE------
Chr6_110214535_786_po_C_Xz16/8      ---HHWLLNPRMTQ-YQGLLCENPHITLETVN-TLNPANLLPVE------
ChrY_3018841_881_po_C_Rz9/12        ---HHWLSNLRMTR-YQGLLCENPHITLETVN-TLNPATLFPIE------
ChrX_86876098_639_po_C_Xz7/8        ---HHWLSNLRMTR-YQGLLCENPHITLETVN-TLNPATLFPIE------
Chr14_68899917_965_po_C_Xz6/8       ---HHWLSNPRMTQ-YQGLLRENPHITLETVN-ILNLATLLPIE------
ChrY_23874437_661_po_C_Xz15/9       ---HYZLSNPRVTQ-YQGLICKNPDITLEIIN-TLNLATLVPVK------
ChrY_27518418_649_po_C_Xz15/9       ---HYZLSNPRVTQ-YQGLICKNPDITLEIIN-TLNLATLVPVK------
ChrY_25016892_657_po_C_Xz13/12      ---HYZLSNPRVTQ-YQGLICKNPDITLEIIN-TLNLATLVPVK------
ChrY_19932937_520_po_C_Xz10/14      ---PL-LSSPRMTR-YQVLLCKNSHITFKIIN-TLKLAILLPIE------
ChrY_20104614_518_po_C_Xz10/14      ---PL-LSSPRMTR-YQVLLCKNSHITFKIIN-TLKLAILLPIE------
Chr19_58660155_598_po_C_Xz21/1      ---HHZLSNLRMTQ-YQGLLCEN-LHNFRTVN--PNPAILLPVE------
Chr19_58621153_500_po_C_Xz21/1      ---HHZLSNLRMTW-YQGLLCENPYTTLNTKH--PNLAILLPVE------
Chr7_139686987_710_po_C_Xz11/7      ---HHWLSNLRMTQ-YQGLLCENPYSILETVN-SLNPATLLPVE------
Chr7_137651301_1097_po_C_Pz13/      ---HPZLSNLRMTQ-YQGLLCKN-----PHIN-TLNLATLLPTD------
Chr11_29579197_930_po_C_Pz8/12      ---HHWLSNLRMTQ-YQQLLCENSPITLETVN-TLNPATLLPIE------
Chr4_117329205_488_po_C_Xz3/16      ---HHWLSNPRMTQ-YQGLLCENPCLTLETVN-TLNLATLLPIE------
Chr21_38522185_383_po_C_Xz7/9       ---HHWLSNPGMTQ-YQGLLCKNPHIALETVN-TLNPATLLLIE------
Chr4_70173370_1049_po_C_Xz11/8      ---HHWFSNPRITQ-YQGLLCDNPCVTLETVS-TLNPATLLPIE------
Chr4_70022513_788_po_C_Kz10/12      ---HHWLSNPRITR-YQGLLCENLCIILT-VN-TLNPATLLTIE------
Chr14_105225894_1027_po_C_Xz13      ---HQWLZNPRMIZ-YQELLCENPCITLETVN-TLNLATLP---------
Chr7_119241536_761_po_C_Xz11/1      ---HHWLSNPRMTQ-ZQGLLCENPYITLETVN-TLNPATLLLIE------
ChrY_22564970_567_po_C_Xz9/10       ---HYWLSNPRITDRYQGLLSENLYVIIGTVI-SLNPATLLPIK------
Chr21_14409346_1477_po_C_Pz5/6      ---HHWLSNPKMLR-YQGLLCENTYITLETVN-TLNPATLLPIE---YVE
ChrX_51259700_1396_po_C_Pz7/8       --GHRWLSNPRMLR-YQWLLCENPYITLETVN-TLNLATLLPIE---YAE
```

```
Chr6_64985396_1041_po_C_Xz11/1      ---YRLLSNPRMLR-YQGLFCENPYITLETVN-NLNPATLLPIQ---WVE
ChrY_20355209_659_po_C_Xz7/11       ---HRWLSNP---Q-MEGLLCKNPYITLETVN-TVNPATLLPKE---ZAE
Chr3_34567166_523_po_C_Xz8/6        ---HGWLSNTRMLR-YQGFLCENLYITLETVN-ALNPATLLPIE---WVE
Chr3_122072468_245_po_C_Xz9/9       ---HHWLSNPKMLR-YQGLLCENPYITLETVN-TLNPATLLPTE---WVE
Chr11_58788475_541_po_C_Xz8/22      ---HRWFSNPRIZK-YQGLLCENPYRTLETVN-TLNPAILLPVE---WVE
Chr6_137421042_373_po_C_Xz6/9       ---HCWLSNPRMLR-YQGLLCEIPYLTLETMN-TLNPATLLLIE---WAE
Chr4_115910608_550_po_C_Pz8/12      ---HRWLSNPRMLR-YQGLLCENPYITLEIVN-TLNPATLLPIE---WVE
ChrY_13743185_977_po_C_Xz14/8       ---RRWLSHSRMLK-YQVLLCENLRVHYR-LN-ALNPTTLLPVEEPNWKH
HERV15                              ---RRWLSHSRMLK-YQGLLCENLRVHYR-LN-TLNPATLLPVEEPDWKH
ChrX_117198092_983_po_C_Xz8/8       ---YGWLSNSRMLK-YQGLLCENPQITLENVN-TLNPATLLPVEEPDWKD
Chr19_21166030_1412_po_C_Xz9/1      ---HRCLSNFRMLK-CQSFLCKISRNTTD-CK-FLEPSYLLPVEEPTLKN
Chr19_21213743_913_po_C_Xz14/1      ---HRWLSNSRMLK-YQGLLCANPDNIKT-VN-TLNAATLQPMEEPDWKD
Chr19_15957774_569_po_C_Xz8/10      S------ASRVAGITGPCLACRSPHEEG---M-RVLTQLVATYP------
Chr11_77508681_297_po_C_Xz19/1      PN-WGDEPLSPSPVSLVGVASHSRHLQFPFPN-QLSEPLLLPVL------
Chr13_98250769_738_po_C_Xz7/13      GE-KHKSGFRACPVIVPSLELHYDIDSFAHMF-FADLLLIITLL------
Chr6_39968183_513_po_C_Pz10/9       VS------SNPRMTRYQGPLCZNPHMIIETVN-TLNLATLLPID------
Chr15_66867263_645_po_C_Xz11/1      EE-ILSGDSSAMGSGLASGPLGVRCTPTKQTG-YSPYKVLFRQP------
HERV18_polputein                    -----WTAQTHFSK-VGRLLRTRSTLSTSPLAAELQEVLGPVVLMQ----
HERVIP10FH_polputein                --AKRZLTDSRTLK-YKAILLEKDDLTLTTEN-SLNPAGFLTG-------
MER66_polputein                     -------ITTKMEGPYRVLLSTPTAVKLQGITSWVXLSRIKPVS------


HERVI_polputein                     ---------EEHLCLDLIDYQTKLRPDLGKIP-FKTGRHLFIDGSSQLIE
HERVIP10F_polputein                 -------LRREHTCLDLIDYHTKVRPDLGETP-FWTGRHLFIDGSSRVIE
HERVadp_AC005741                    -------ETSDHNCLNIIEYQTKVRPDFREAP-LHV--VLFVDGSSZVID
HERV-PT47D                          -----------SARLGLPKVLGLHREPQRPAP-PFVN----VDGGVALG-
HERV_W_chr6_141432567_ERV9_lik      ----------EHDCQQIIVQTYATRDDLLEVPLANPDLNLYTDGSSFVEN
ERV9_PH1_RT                         ----------KHDCQQIIVQTYAARDDLLEVPLTNPDLNLYTDGSSFVEN
Chr13_54564403_572_po_C_Yz7/14      ----------EHNCQQVITQTNSTRGDLLEVPLTDPDLNLYTDGSSFVEK
HERV_Y_chr12_51022911_pol           ----------EHNCQQVIAQTCAARGDLLEVPLTDPNLNLYSGGSSFVEK
HERV17_polputein                    ----------EHNCQQVIAQTYAARGDLLEVPLTDPDLNLYTDGSSFVEK
HERV_W_chr7_9105739_syncytin_p      ----------KHNCQQVISQTYATRGDLLEVPLTDPDLNLYTDGSSFVEK
HERV9_polputein                     ----------EHDCQQITVQTYATQDDLLEVPLANPDLNLYTDGSSFAEN
HUERS-P3_polputein                  ----------EHDCEQIVVQTYAAREDLRETPLENPDWTLFMDGSSFVEQ
ERVfrd_AC004022                     ----------TYSCCEILNKIYASWEDLKDQPIDNPDKIWFSDGSSFVRD
PRIMA41_polputein                   ----------DHSCIQVMEQVYSSCPDLKDEPLDNPEVEWFTDGSSFVHQ
HERV_H_RGH1_pol                     --------------------------KAGYAIVSS---TSIIEAT-----
HERVHRGH2_pol                       -----------HDCISLIHLTFTPFPHISFCPVSHPDHTWFIDGS-----
HERV_H48I_polputein                 ----------THFCVNTVQTFLIPFPNLTDQPLPDASFTWFVDGSSFLHQ
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        -----------HSCTDILDHLQPHFPNISSEPLTNPNDQLFIDGSSSGPT
HERVRblike_chr4_109047953           ----------IHDCPEVIDQVFSSPLDLKDTALSCANWTLFANRKSLVIN
ChrY_13790029_693_po_C_Xz12/11      -------TTPVRDCQELLETTETGZTDLQDMPLZ--TYTMFIDG-SSFLE
HERVS71_polputein                   -------TTPVHDCQELLETTZTGRPDLQDVPLKEVDSTLFTDS-SSFLE
BaEV_M7_pol                         -------QPSPHDCRQVLAETHGTREDLKDQELPDADHTWYTDG-SSYLD
HERVR_polputein                     -------QPSPHDCRQVLAETHGTREDLKDQELPDADHTWYTDG-SSYLD
AKR_MLV_MLOCG_RT                    -------EGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDG-SSLLQ
GaLV_pol                            -------ATPVHRCSEILAEETGTRRDLEDQPLPGVP-TWYTDG-SSFIT
Chr3_158587414_1662_po_C_Ez10/      --------PVKHNHVEVLDSAYSSRPDLRDQPWASVDWELYVDWSSFINP
Chr3_79627623_1138_po_C_Ez4/7       --------PVKHNCVEVLDAVYSSRPDLWDQPWAAVDWELYMDRSSFINP
Chr19_20477244_1350_po_C_Ez10/      --------PVEHNCVEVLDSVYSSRPDLRDQPWASVDLELYLDGSSFINP
Chr5_111890459_1626_po_C_Ez10/      --------LVKRNCVEVLDSVYSSRPDFWDQPWASVDWELYVDGSSFINP
ChrX_48766518_1600_po_C_Ez6/3       --------LVEHNCVEVLDSVYSSRPDLWDQPWASVDWELYMDGSSFINP
ChrX_152580713_1572_po_C_Ez10/      --------PVKHNCVEVLDSVYPSRPDLRDQPWASADWELYMDGSNFT-I
ChrX_139287533_1154_po_C_Xz8/4      --------PVEHDCVEALDSVYSSRPDLRDQPRASVDWELYMDGSSFINP
Chr2_52615786_912_po_C_Xz11/5       ---------VEHNCVEALDSVYSSRPDLRDHPWTSVDRELYRDRSSCVNP
HERV-E_4_1                          --------PVEHDCVEVLTQFTLG----------SVDWELYVDGSSFVNP
ChrY_21962399_1137_po_C_Xz13/9      ---------VEHDCVEALDSVYSSRPDLZDHLWTTVDZELCMDGSSFIIP
ChrY_8634032_950_po_C_Xz14/10       --------LIAYNCVEVLDSVNSSRSDLRDKPWASVDWELYMEGTSFINP
Chr3_126735360_997_po_C_Xz5/10      --------PVEPDCVEVLDSVDSSRPDLRDQAWASVDWELYVYGSSFFNP
Chr3_126713203_958_po_C_Xz7/8       --------TVEPDCVEVLDSVDSSRPDLRDQAWASGDWKLYVYGSSFFNP
Chr3_75226118_977_po_C_Xz8/9        --------PVEPDCVEVLDSIDSSRPDLRDQAWASVDWELYVDGSSFFNP
Chr7_6629919_827_po_C_Xz8/7         --------PVEPDYVEMLDSVASSRPDLWDQPWASVDWELYVDGSSFFNP
Chr4_9166812_811_po_C_Xz8/10        --------PVEPDCVEVLDSIDSSRPDLRDQAWASVDWELYVDGSSFFNP
Chr12_8478642_834_po_C_Xz12/7       --------PVEPDCVEVLDSIDSSRPDLQDQPWASVDWELYADGNSFFNP
Chr10_15187175_787_po_C_Xz11/5      --------PVEPDCVEVLDSIDSSRPDFRDQPWASVDWERYVDGSSFFNP
Chr11_71819740_854_po_C_Xz11/9      --------PVKPGCVEVLDSIDSSRPDLWDQPWASVDWELYLDGSSFFNP
Chr11_71519974_886_po_C_Xz13/9      --------PVEPGCVEVLDTIDSSRPDLRGQTWASVDWELYMDGSSFFNP
Chr8_7271441_969_po_C_Xz10/6        --------PVEPDCVEVLDSIDSSRPELRDQTWASVDWEPHVDGSSFFNP
Chr8_11770066_826_po_C_Xz9/8        --------SAEPDCVEMLYSIDCSRPDLRDQTWASVDWEPYVDGSSFFNS
Chr4_4128362_651_po_C_Xz7/11        --------PVEPGCVEVLDTIDSSRPDLRGQPWA-IDWELYVDGSSFFN-
Chr14_50203955_850_po_C_Xz10/8      --------PVEHDCVEVLDSVDSSRPDLRDQPWASVDWELZVDGSSFVNP
Chr14_50205291_837_po_C_Xz10/8      --------PVEHDCVEVLDSVDSSRPDLRDQPWASVDWELZVDGSSFVNP
Chr7_6597120_842_po_C_Xz7/7         --------PVEHDCVEVLDSVDPSRPDIRDQPWASVDWELYVHGSSLVNT
Chr11_71544658_783_po_C_Xz8/13      --------PVEHDCVEVLDSVDSSRPDIRDQLGHQYTGNYTCMGAASSTP
Chr11_67999520_760_po_C_Xz8/8       --------PVEHDCVEVLDSVDSSRPDFWDQPWASVDWELYVDGSSFFNP
Chr3_75461975_796_po_C_Xz7/9        --------PVQADCVEVLDSIDSSRPDLQDQPWASVDWELYVDGSSFFNA
```

```
Chr12_52228939_692_po_C_Xz3/10      --------PVGHDCVEVFDSVDSSRSDLRDQPWASVDWELHVDGSSFINL
Chr10_15211234_286_po_C_Xz11/1      --------PVKHDCVEMLDSVDSSRPDLRVQPWASVDWELZVDGSSFIIP
Chr12_100973302_945_po_C_Xz12/      --------PVEHNCVEVLDSVYSSRPDLQDQSZASVDZDLYVDRSSFINP
Chr19_9685420_817_po_C_Xz6/10       --------PVKRNCVEVLDSVYSSRPDLZDQPWASVDWDLYMDGSSFINP
Chr13_40936820_719_po_C_Xz9/10      --------PVEPDCVEVLDSVDSSRPDLRDQPWASVDWELSVDGSSFFNP
Chr7_63858202_1754_po_C_Rz12/9      ---PG---ITFRDCGN-VDEVFSSRGDLTDQPFKDPDVEYFIDG-SSFVL
ERV3_chr7q21_AC073210_polputei      ---PG---ITFRDCGN-VDEVFSSRGDLTDQPFKDPDVEYFIDG-SSFVL
Chr6_110218096_843_po_C_Xz16/8      ---PG---APLHNCVEIVDEVFLSWGDLMDHPLRVPDVEYFTDG-SSFIL
Chr6_110214535_786_po_C_Xz16/8      ---PG---APLHNCVEIVDEVFLSWGDLMDHPLRVPDVEYFTDG-SSFIL
ChrY_3018841_881_po_C_Rz9/12        ---TG---APLHDCVETVDEVFSSWGDLTDCPLGDPDVEYFTDR-SSFIL
ChrX_86876098_639_po_C_Xz7/8        ---TG---APLHDCVETVDEVFSSWGDLTDCPLGDPDVEYFTDR-SSFIL
Chr14_68899917_965_po_C_Xz6/8       ---G----IPLHDCVETVDEVFSSQGDLTDQPLRDPDVEYFTDG-STFML
ChrY_23874437_661_po_C_Xz15/9       ---PG---VLLHDSVETVDDVFSSQKDLIEQPLWDPDGKYY-------IL
ChrY_27518418_649_po_C_Xz15/9       ---PG---VLLHDSVETVDDVFSSQKDLIEQPLWDPDGKYY-------IL
ChrY_25016892_657_po_C_Xz13/12      ---PG---VLLHDSVETVDDVFSSQKDLIEQPLWDPDGKYY-------IL
ChrY_19932937_520_po_C_Xz10/14      ---PG---VPFHDSMETVDTVLZIQRYLTEQPLWNPDVKCYTYR-SSFTL
ChrY_20104614_518_po_C_Xz10/14      ---PG---VPFHDSMETVDTVLZIQRYLTEQPLWNPDVKCYTYR-SSFTL
Chr19_58660155_598_po_C_Xz21/1      ---QA---APLHNCVKTVDEVFPNEKDLTDRPLRDPDVKYFTGG-SSSVL
Chr19_58621153_500_po_C_Xz21/1      ---QA---ALLHNCVKTVDEVFPNEKDLTDRPLRDPDVKYFTGG-SSSVL
Chr7_139686987_710_po_C_Xz11/7      ---PG---APLHDCVKTVDEVFSSZKDLMDQPLRDPDVEYFTDG-SSFVL
Chr7_137651301_1097_po_C_Pz13/      ---LEP-GGPLHCRVDGVGEVSSSQKDLTHQPLKDPDVEYFTDG-SNFVL
Chr11_29579197_930_po_C_Pz8/12      ---LGP-GAPLHCCVDAVDEVFSSQRDLTDQPLRDPDVEYFTDG-SSFVL
Chr4_117329205_488_po_C_Xz3/16      ---P---GVPLHDCIETVDEIFSSRGDLTDZTLRDPHVEYFTDG-SSFIL
Chr21_38522185_383_po_C_Xz7/9       ---L---GTPLHDCVKMVDEVFLSRGDLTDQPLRDPVVKZFTDG-SSFIL
Chr4_70173370_1049_po_C_Xz11/8      ---SVP-GSPLHCZVDVVDEVFSSQRDLTDRPLGDLDIEYFTDG-SSFIL
Chr4_70022513_788_po_C_Kz10/12      ---SMP-GSPLQCCVDVVGEVFLSQRDLTDQPLGDPDIGYFTDG-SSFIL
Chr14_105225894_1027_po_C_Xz13      --------PSLLC--DVADEVFSSQRDLTDRPLGDPDIEYFTDG-SSFIR
Chr7_119241536_761_po_C_Xz11/1      ---L---GAPLHDCVETVDKVFSSQGDLTDQTLREPGVEYFTDG-SCFIL
ChrY_22564970_567_po_C_Xz9/10       ---RPKDSIPYHYCVDVVEELFCSZKDLTDQPLVGPDVEYFMDG-SSFIS
Chr21_14409346_1477_po_C_Pz5/6      HGKPPLCAPGYHCCVETVDEVFSSQKDLKDQPLKDPDVEYFTDG-SSFIS
ChrX_51259700_1396_po_C_Pz7/8       HGKPPLCAPGYHCCVETVDEVFSSRKDLKDQPLKDPDIEYFTDG-SSFIS
Chr6_64985396_1041_po_C_Xz11/1      HGKPPLCGPGYHCCVETVDEVFSSZKDLKDQPLKYQDVEYFTDE-SSFIS
ChrY_20355209_659_po_C_Xz7/11       HGKPPLCGPEYHCCVETVDEVFSNLKDLKELPLQNPDLN--SDG-SSFIS
Chr3_34567166_523_po_C_Xz8/6        HGKS-RCVAQGITVLETVDEA----KDLKNQPLKDPDVEYFTDG-SSFVS
Chr3_122072468_245_po_C_Xz9/9       HGAP-LCGPGYHCCVETVDEVLSSQKNLTDZPLRDPDVKYFTDG-SSFIS
Chr11_58788475_541_po_C_Xz8/22      HGKPPLZGPGYHCCVETVDEVFSSREDLKDRPLRDPDVEYFTDE-SSFIS
Chr6_137421042_373_po_C_Xz6/9       HGKPPLCGPGYHCCVETVDEVFSSZKDLKDZPLKDPHVKYFADG-SSFIS
Chr4_115910608_550_po_C_Pz8/12      HGKPMLCGPGYDCCVEILDEVFSSRKDIKDQPLKDPDVEYFTDG-SSFIS
ChrY_13743185_977_po_C_Xz14/8       GGLPQCWQDLPHCCINMVDEVFLSWEDLRDTRLESPDVEYFTDG-SSFIT
HERV15                              GGLPQCWQDLPHCCINMVDEVFLSWEDLRDTPLESPDVEYFTDG-SSFIT
ChrX_117198092_983_po_C_Xz8/8       AGLPHSWQDLPHCCINMVDKVFSSQEDLRDSPLESPDVIYFTDG-SSFIT
Chr19_21166030_1412_po_C_Xz9/1      GGLPHCWQDSPLLHCIQWTKCSRAR-ELRDTPLESPDVEYFTDG-SSFIT
Chr19_21213743_913_po_C_Xz14/1      GELPHCZZDLPHCCINTVGKVFSSQEDLRDTPLESPDVEYFTDG-SSFIT
Chr19_15957774_569_po_C_Xz8/10      ---QGFKIWMNPITPIIRLPPQHHLVCHQ----CLRYHLKHL-G-ECRGA
Chr11_77508681_297_po_C_Xz19/1      ---IGVRVSSKLTAALIAPSPSGGLLGIPREVIRLPFRPYRM-G-PALGP
Chr13_98250769_738_po_C_Xz7/13      ---STTFLFAEIMNIIINKNZGNSEAGAG----AGPWCAECR-S-PGPTV
Chr6_39968183_513_po_C_Pz10/9       ---QKTLPZHYCVDIVDEVFSSZRNLKDQP--LGGPHVEYFMDG-SRFVS
Chr15_66867263_645_po_C_Xz11/1      ---PPIISQVKGDLCELGELTLRRQMQALGMAMQKVHGWVWERM-PISLT
HERV18_polputein                    -----------DKAMGPEAPLDPEPSPFKEGHPPIPDGAWYTDG-----S
HERVIP10FH_polputein                -------------------DLNLNYRTKVRPDLGNSLQDRTIDG-SSRVI
MER66_polputein                     -----------YESQAQKEDTMTYICEPLRLPLPIZKNQHSARS-GNVAV


HERVI_polputein                     GKRHNGYPVIDGEILYIRTESRKLPNNWSAQTCELFA-------LSQALK
HERVIP10F_polputein                 GKRHNGYSVIDGETL-VEIESGKLPNSWSAQTCELFA-------LSQALK
HERVadp_AC005741                    GRRHSRYTVID-ENKHSLCEKGRLPNGWLAQTCELHA-------LNQALN
HERV-PT47D                          ----------------------------VRTQEHSD-------LQQPVA
HERV_W_chr6_141432567_ERV9_lik      GIZREH-YAIVS--DV------TVLS----QLAEF-------VALTQALE
ERV9_PH1_RT                         GIRRLV-VTILE--SK------PLPPGTSAQLAEL-------VALTWALE
Chr13_54564403_572_po_C_Yz7/14      GLZKAG-YAVVS--DNGILENNPLTPGTSAQLAEL-------IALTRALE
HERV_Y_chr12_51022911_pol           GLQKAG-YAVVS--DNGILESNPLTPGTSAQLAEL-------IALTRALE
HERV17_polputein                    GLRKAG-YAVVS--DNGILESNPLTPGTSAQLAEL-------IALTRALE
HERV_W_chr7_9105739_syncytin_p      GLRKVG-YAVVS--DNGILESNPLTPGTSAQLAEL-------IALTWALE
HERV9_polputein                     GIZRAG-CDIVS--DVTVPESKPLPPGTSTQLTEL-------VALTZALE
HUERS-P3_polputein                  GIRKAG-YAVVT--LNNVIESAXLSPGTSAQLAEL-------IALTRALE
ERVfrd_AC004022                     GTRYAG-YATES--HHQVVETKALSPGTSVQLAEF-------ITLTRALK
PRIMA41_polputein                   GNGKAG-YAVVS--QHEVIESQALPASTSAQKAEL-------IALIRALQ
HERV_H_RGH1_pol                     --------ALPP--STTS-----------QQAEL-------VTLTQALT
HERVHRGH2_pol                       --------STRP--NRHSPAK--------AGYAEL-------VALIQALT
HERV_H48I_polputein                 GRRHAGYAIVSP--HTHTIEANPLPLGTTSQKAEL-------IALTQALT
HERV_H_RTVLH2_pol                   ----------FS--CTHS-------------------------LLSPTIT
HERVFH21_pol                        GSPKIAGYAVVS--LDQVIEAKPLPPGTSSQKAEL-------IALTRALT
HERVRblike_chr4_109047953           GRRNAA-YAVVS--SEG-IEARTLPTGTSAQKAQS-------IALTRALQ
ChrY_13790029_693_po_C_Xz12/11      QGVQKIGAAFTT--ETDVLWSQVLPAGTSAQKAELIA-------LTQAVZ
HERVS71_polputein                   QGVGKAGAAVTM--ETDVLWAQALPAGTSAQKAELVT-------FTQALZ
BaEV_M7_pol                         SGTRRAGAAVVD--GHNTIWAQSLPPGTSAQKAELIA-------LTKALE
HERVR_polputein                     SGTRRAGAAVVD--GHNTIWAQSLPPGTSAQKAELIA-------LTKALE
```

```
AKR_MLV_MLOCG_RT                    EGQRKAGAAVTT--ETEVIWAKALPAGTSAQRAELIA-------LTQALK
GaLV_pol                            EGKRRAGAPIVD--GKRTVWASSLPEGTSAQKAELVA-------LTQALR
Chr3_158587414_1662_po_C_Ez10/      QGERCAGYAVVT--LDAVIEAKPLPQGTSAQKTELIA-------LTRAQE
Chr3_79627623_1138_po_C_Ez4/7       QGERCAGYAMVT--LDAVTEAKPLPQGTSAQKAELTA-------LTQALE
Chr19_20477244_1350_po_C_Ez10/      QGERCAEYAVVT--LDAVIETKPLPQGTSAQKAELTA-------LTRALE
Chr5_111890459_1626_po_C_Ez10/      ZGERCAGYAVVT--LDAAIEAKLLPQGTSAQKPELIA-------LTRALE
ChrX_48766518_1600_po_C_Ez6/3       QGERCAGYVMVT--LDAVIEAKPLPQGTSAQKAELIA-------LTRALE
ChrX_152580713_1572_po_C_Ez10/      HGERCAGYAVVT--LDAVHZSQTIATGTSAQKAELIA-------LTRALE
ChrX_139287533_1154_po_C_Xz8/4      QGESCAGYVVVT--LDDIIEAKPLPQGTSAQKAELIA-------LTRALE
Chr2_52615786_912_po_C_Xz11/5       PGERCAGYAVVT--LDAVVEAKSLPKGTZAKKAKLIA-------LIRASE
HERV-E_4_1                          QEERCAGYAVVT--LDTVAEARSFPQGTSTQKAELIA-------LIRALE
ChrY_21962399_1137_po_C_Xz13/9      QGERCAGYAVVTRTLKCVIEAKSLPQGTSAQMTELTA-------LIQALE
ChrY_8634032_950_po_C_Xz14/10       QGZRCAZYVVVT--LDTVVEAKPFPQDTSAQKAELIAZ------ITWALE
Chr3_126735360_997_po_C_Xz5/10      QGERGAGYAVIT-LDP-VVEARSLPQTTSAQKAELIA-------FIRALE
Chr3_126713203_958_po_C_Xz7/8       QRERGAGYAVIT-LDT-VFEARSLPQATSAQKAELIA-------FIRALE
Chr3_75226118_977_po_C_Xz8/9        QGERGAGYAVGT-LDT-VVETRSLPQATSAQKAELIA-------FIRALE
Chr7_6629919_827_po_C_Xz8/7         QGERGAGYAVIT-LDT-VVEARSLPQATSAQKAELIA-------LIRALE
Chr4_9166812_811_po_C_Xz8/10        ZGERGAGYAVIT-LDT-VVEARSLPQATSTQKAELIA-------FIRALE
Chr12_8478642_834_po_C_Xz12/7       QGERGAGYAVIT-LGT-VVEAGZLPQATSAQKAELIA-------FIRALE
Chr10_15187175_787_po_C_Xz11/5      QGERGAGYAVIT-LGT-VVEARSLPQATSAQKAELIA-------FIRALE
Chr11_71819740_854_po_C_Xz11/9      Q--RGGGYAGVT-LDT-VVEATSLPQATSAZKAELIA-------FIRALE
Chr11_71519974_886_po_C_Xz13/9      QGERGAGYAVIT-LDT-VVEARSLPQGTSAQKAELIA-------FIRALE
Chr8_7271441_969_po_C_Xz10/6        QGERGAGYAVIT-LDT-VVETRSLPQATSAQKAELNA-------FIWALE
Chr8_11770066_826_po_C_Xz9/8        QGERGSGYAVIT-LDT-VVEARSLPQATSAQKAELIA-------FIGALE
Chr4_4128362_651_po_C_Xz7/11        QGERGAGZAVIT-LDT-VVEARSLPQATSAQKAELIA-------FIRALE
Chr14_50203955_850_po_C_Xz10/8      HGERGAGYSVVT-LDTQQLEARSLPQATSAQZAELIA-------LIRALE
Chr14_50205291_837_po_C_Xz10/8      HGERGAGYSVVT-LDTQQLEARSLPQATSAQZAELIA-------LIRALE
Chr7_6597120_842_po_C_Xz7/7         QGKRGAGYAVVT-LDT-VVEARSLPQATSAQKAELIA-------FIRALE
Chr11_71544658_783_po_C_Xz8/13      K-EREVQGMRDN-LDT-VVEARSMPQGTSAQKAELIA-------FIRALE
Chr11_67999520_760_po_C_Xz8/8       QGERGAGYAVIT-LDT-VVEARSLPQATSAQKAELIA-------FIWALE
Chr3_75461975_796_po_C_Xz7/9        QGKRGAGYAVVT-DTV--VEARSLPQATSAQKAELIA-------LIRALE
Chr12_52228939_692_po_C_Xz3/10      QGERCAGYAVVT-LYT-VVEARSLPQGTSAQKAEPIA-------LIWALE
Chr10_15211234_286_po_C_Xz11/1      QGERGAGYAVVT-LDT-VVEARSMPQATS-QKAELIA-------FIQALE
Chr12_100973302_945_po_C_Xz12/      QGERCAGYAVVT-LNA-VIEAKPLPQGNSAQKAELIA-------LTQVLE
Chr19_9685420_817_po_C_Xz6/10       QGERCAGYAVVT-LDA-VIEAKPLPQGVSAQKAELI--------LTRALE
Chr13_40936820_719_po_C_Xz9/10      QGERGAGYAVVT--LDTVVEARSLPQATSAQKAELIA-------FIQALE
Chr7_63858202_1754_po_C_Rz12/9      EGVCWAGZAVVT--LDSVVEAZPQPAGTSAQKAEL-AA------LTRVLQ
ERV3_chr7q21_AC073210_polputei      EGVCWAGZAVVT--LDSVVEAZPQPAGTSAQKAEL-AA------LTRVLQ
Chr6_110218096_843_po_C_Xz16/8      EGVHWAEYVVVT--LDLVVEAQPLPTRTSAQKAEL-IS------LTRACL
Chr6_110214535_786_po_C_Xz16/8      EGVHWAEYVVVT--LDLVVEAQPLPTRTSAQKAEL-IS------LTRACL
ChrY_3018841_881_po_C_Rz9/12        EGVQWAGYVVVT--LDSVVEAQPLPTEMLTQKAEL-IA------LTRALL
ChrX_86876098_639_po_C_Xz7/8        EGVQWAGYVVVT--LDSVVEAQPLPTEMLTQKAEL-IA------LTRALL
Chr14_68899917_965_po_C_Xz6/8       EGVRQAGYAVVT--LDSVTEAQPLPTETSTQKAEL-ID------LTRALL
ChrY_23874437_661_po_C_Xz15/9       KGVRETGYAVIT--LDSVVEAQ--PTRTSAQKAEL-IA------LMR-AL
ChrY_27518418_649_po_C_Xz15/9       KGVRETGYAVIT--LDSVVEAQ--PTRTSAQKAEL-IA------LMR-AL
ChrY_25016892_657_po_C_Xz13/12      KGVRETGYAVIT--LDSVVEAQ--PTRTSAQKAEL-IA------LMR-AL
ChrY_19932937_520_po_C_Xz10/14      KGVQQAGYAVMI--LDSVVEAQPLPTRMSAQKAEL-IA------LET-IL
ChrY_20104614_518_po_C_Xz10/14      KGVQQAGYAVMI--LDSVVEAQPLPTRMSAQKAEL-IA------LET-IL
Chr19_58660155_598_po_C_Xz21/1      EEFGHAGYAVVT--SDSVVEAQPLPTGTSAQKAEL-IA------LTRALL
Chr19_58621153_500_po_C_Xz21/1      EEFGHARYTVVT--SDSVVEAQPLPTGTSAQKAEL-IA------LTRALL
Chr7_139686987_710_po_C_Xz11/7      EGVQHAGYAVVT--LDSVVEAQSLPTGTSAQKAEL-IA------PTRALL
Chr7_137651301_1097_po_C_Pz13/      EGVRQAGYAVVT--GLSGRGSV-SAHRTSAQKADL-IA------LTRALL
Chr11_29579197_930_po_C_Pz8/12      EGVZQARYAVVT--LDSVIEAS-LSTGTSAHKAEL-IA------LTRALW
Chr4_117329205_488_po_C_Xz3/16      EEVCPAGYAVVT--LDSVVEAQPLPTGTSAQNAEL-IA------LIKALL
Chr21_38522185_383_po_C_Xz7/9       EGVRWAGYAVVT--LDSVIEAQPLPTGASAQKAEL-MA------LTRALL
Chr4_70173370_1049_po_C_Xz11/8      EGVHQTGYAVVT--LDSAVEAQSLSTGTASQKAKL-IG------LIRALW
Chr4_70022513_788_po_C_Kz10/12      EGVRQAGYAVVT--LDSVVEAKSLPTGSSPQKAEL-IA------LTRALQ
Chr14_105225894_1027_po_C_Xz13      G-ARRAGYAVVT--LDSVAEAZSLPTGTSAQKAEL-IA------LTRAHW
Chr7_119241536_761_po_C_Xz11/1      GGVRZAGYMVVT--LDSVVEVQPLPT-----EAEL-IA------LTZALL
ChrY_22564970_567_po_C_Xz9/10       EGIQZPAYTVVT--WDSVVEAQPLPARKSAQKAEL-TA------PTRALL
Chr21_14409346_1477_po_C_Pz5/6      KGVKKARFAIVT--LSSVAKARPLAVGTSAQKA---IA------LTRALL
ChrX_51259700_1396_po_C_Pz7/8       EGIRKAGYAVAT--LNSVAEACPLPVGTSAQKAEL-IA------LIRALL
Chr6_64985396_1041_po_C_Xz11/1      EGVRWAGYAVVT--LNSVAEACPLLVGTSAQRAEL-IA------LTRTLL
ChrY_20355209_659_po_C_Xz7/11       EGVRPAGHAVVT--LNSVAEVCPLTVGSSAQRAEL-IVSLELILLTRALL
Chr3_34567166_523_po_C_Xz8/6        EG-RLAGYAVVT--LNSVAKACPLPVGTLAQRAEL-IA------LNKTLL
Chr3_122072468_245_po_C_Xz9/9       EGVRRAGCAVVT--LNSVAEARPLPVGTLAQRAEL-IA------LTRALL
Chr11_58788475_541_po_C_Xz8/22      EGIRWAGYAVVT--LNSVAAGLPLPVRTLAQRAEL-IA------LTRALL
Chr6_137421042_373_po_C_Xz6/9       KGISWAGYAVVT--LNSAAETCPLPAGTLAQRAEI-IA------VTRVLL
Chr4_115910608_550_po_C_Pz8/12      EGVRWAGYAVVT--LNSVSEAHPLPVGTLAQIAEL-IA------LTRALL
ChrY_13743185_977_po_C_Xz14/8       DGVZYAGYKVVT--QHSVVEAQALTSGTSAQRAEL-IA------LTRALL
HERV15                              DGVZYAGYAVVT--QHSVVEAQALPSGTSAQKAEL-IA------LTRALL
ChrX_117198092_983_po_C_Xz8/8       DGVRYAGYAVVT--QHSVVEAQALPSGTSAKKAEL-IA------LTRALL
Chr19_21166030_1412_po_C_Xz9/1      NGVRYAGYTVVT--QRSVVEAQALTSGTSAQKAELIIA------LSRALL
Chr19_21213743_913_po_C_Xz14/1      HGVZYGGYAVVT--QHLVVEAQALPSGTSTRKAEL-IT------LTRALL
Chr19_15957774_569_po_C_Xz8/10      SEACKPLLPKLL--CGRCS-IQALPSCFT---------------------
Chr11_77508681_297_po_C_Xz19/1      KTLPDSZSALFL--ESSCSSSQDLLRCWVGVLGHG-------ELISPLG
Chr13_98250769_738_po_C_Xz7/13      VSLYFVSVSISF--LSLSSHPTRNTHVWRGR-PPL--------QVSKKKA
```

```
Chr6_39968183_513_po_C_Pz10/9       EGVCZAGYTVLT--WDSVVEAQPLPSGTAQK-------------------
Chr15_66867263_645_po_C_Xz11/1      DPVHPFKPGDFV--LVKKWNPTTLGPICDGPHTVI--------LSTPTAV
HERV18_polputein                    SRGATAAWTAVAVQPSTDTIWFDTGCGQSSQWAEL-----------RAVW
HERVIP10FH_polputein                EGKRHNGYSVSDKETLVEAELGKLPNNWSAQTCELFA-------LSQALK
MER66_polputein                     GGNRSINFSLLPDCNTSFLSLZPTTFSWETPLLSLLG--------VEATK


HERVI_polputein                     HLQNQEGT---IYTDS-----KYAFGVAHTFGK--SWTECGLTNSKG-HD
HERVIP10F_polputein                 YLQNQEGT---IYTDS-----RYAFGVAHTFGK--IWTERGLINSKG-QD
HERVadp_AC005741                    LLEDQEGT---IYTDS-----KYAHGVLHTFGK--IWTEQGLINSRZ-KE
HERV-PT47D                          FLSKVLTQ---SLVDG-----LNASSPSRTAIL--VEESRKLTFGGK-LT
HERV_W_chr6_141432567_ERV9_lik      LGKGKRIN---VYTDR-----KYAYLILYAHVA--IWKEREFLMSGG-TP
ERV9_PH1_RT                         LGKGKRIN---VYTDS-----KYAYLILHAHAA--IWKEREFLTSVG-TP
Chr13_54564403_572_po_C_Yz7/14      LGEGKRVN---IYTDS-----KYAYLVLQAYAA--IWREREFLTSEG-TP
HERV_Y_chr12_51022911_pol           LGEGKRVN---IYTDC-----KYAYLVLHAHAA--IWRERKFLTSEG-TP
HERV17_polputein                    LGEGKRVN---IYTDS-----KYAYLVLHAHAA--IWREREFLTSEG-TP
HERV_W_chr7_9105739_syncytin_p      LGEEKRAN---IYTDS-----KYAYLVLHAHAA--IWKEREFLTSER-TP
HERV9_polputein                     LGKGKRIN---VYTDS-----KYAYLIPHAHAA--IWKEREFLTSGG-TP
HUERS-P3_polputein                  LSKGKVAN---IYTDS-----KYAFLVLHAHAA--IWKERHFLTTNG-SP
ERVfrd_AC004022                     LGERKRIT---IYTDS-----KYAFLVLHAHMA--IWKERRYLTAQD-AP
PRIMA41_polputein                   LGKDLRIN---ISTDS-----KYAFLVLHAHAA--IWKEWGLLTAKG-SP
HERV_H_RGH1_pol                     LAKGLRIN---IYTDS-----KYAFHILQHHAV--IWAERGFLTTQG-SS
HERVHRGH2_pol                       LAKGLRVN---IYTDS-----KYAFHILHHHAV--IWAERGFLTTQV-SS
HERV_H48I_polputein                 LAARQQIN---IYSNS-----HYAFHIVHSHSS--IWKERGFLTAKN-TP
HERV_H_RTVLH2_pol                   IVPGPDFN---PAS-----------HIIP----------------DT-TP
HERVFH21_pol                        LSKGKRVN---IYTDS-----KYAYHILHSHAA--IWQKRGFLTAKG-TP
HERVRblike_chr4_109047953           LSQSKSAN---IYTDS-----KYAFIIVHAQGA--IWKERGLLKADN-TE
ChrY_13790029_693_po_C_Xz12/11      WVKDKCIN---IYTDS-----RYAFATVHVHGA--VYQECGLLTSAG-KI
HERVS71_polputein                   WGKDKRIN---IYTDS-----RYVFPTVHVHKA--IYKSEG--YSPG-KT
BaEV_M7_pol                         LSKGKKAN---IYTDS-----RYAFATAHTHGS--IYERRGLLTSEG-KE
HERVR_polputein                     LSKGKKAN---IYTDS-----RYAFATAHTHGS--IYERRGLLTSEG-KE
AKR_MLV_MLOCG_RT                    MAEGKKLN---VYTDS-----RYAFATAHIHGE--IYRRRGLLTSEG-KE
GaLV_pol                            LAEGKNIN---IYTDS-----RYAFATAHIHGA--IYKQRGLLTSAG-KD
Chr3_158587414_1662_po_C_Ez10/      LSEGKTVN---IYTDS-----QYAFLTLQVHG--ALYKEKGLLNSGG-KD
Chr3_79627623_1138_po_C_Ez4/7       FSE-----------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      LSEGKTVN---IYIDS-----RHAFLTLQVHG--ALYKEKGLLNSGG-KD
Chr5_111890459_1626_po_C_Ez10/      LSEGK-LN---IYTDS-----RYAFLTLQVHG--ALYKEKVLLNSGG-KD
ChrX_48766518_1600_po_C_Ez6/3       LSEGKTVN---IYIDS-----RYAFLTLQVHG--TLYKEKGLLNSGG-KD
ChrX_152580713_1572_po_C_Ez10/      LSEGKTVN---IYTDS-----ZYAFLILQVHG--ALYKZKGLLNSRG-KD
ChrX_139287533_1154_po_C_Xz8/4      LREGKTVN---IYTDS-----RYAFLTLQVHG--ALYKEKALLNSGG-KD
Chr2_52615786_912_po_C_Xz11/5       LSEGKTVN---IYTDS-----RYAFLTLQVG---ALYKEKGLLNSGG-KD
HERV-E_4_1                          LSEGKTVN---IYTDS-----ZYVFLTLQVHG--ALCKEKGLLNSGG-KD
ChrY_21962399_1137_po_C_Xz13/9      LTDGKIVN---IYTDS-----GHAFLTLQVHG--ALYKEKGLLNTGG-KY
ChrY_8634032_950_po_C_Xz14/10       LSEGKTVN---SYTDS-----ZYTFLTLHVHR--ALCKEKP-LNFEG-KN
Chr3_126735360_997_po_C_Xz5/10      LSGCETVN---TYTDS-----RYVFSTLQVHG--AI--EKGLLNSGG-KD
Chr3_126713203_958_po_C_Xz7/8       LSEGETVN---IYTDS-----WYVFSTLQVHG--AZZKEKGLLNSGV-KD
Chr3_75226118_977_po_C_Xz8/9        LSEGETVN---IYTDS-----RYVFLTLQVHG--AIER-KGLLNSGG-KD
Chr7_6629919_827_po_C_Xz8/7         LSEGETVN---IYTDS-----PYAFZTLQVHG--AIER-KGLLNSGG-KD
Chr4_9166812_811_po_C_Xz8/10        LSEGETVN---IYTDS-----RYVFZTFQVHG--AIER-KGLLNSGG-KD
Chr12_8478642_834_po_C_Xz12/7       LSEGETVN---IYTDS-----RYVFLTLQVHG--AZWKEKGLLNSGG-KD
Chr10_15187175_787_po_C_Xz11/5      LSEGETVN---IYTDS-----RYVFLTLQVHV--AZZKEKGLLNSGG-KD
Chr11_71819740_854_po_C_Xz11/9      LSEGETVN---IYTDS-----RYVFSTLQVHG--AZZKEKGLLNSG--RD
Chr11_71519974_886_po_C_Xz13/9      LSEGQTVN---IYTDS-----QYVFSTLQVPG--VZZKEKGLLNSGE-KD
Chr8_7271441_969_po_C_Xz10/6        LSEGETVN---TYTDS-----QYVFLTLQVHG--VZZKEKGLLNSGG-KD
Chr8_11770066_826_po_C_Xz9/8        LSEGETVN---IYTDS-----PYVFLTLQVHG--AIERRR-VLNCGG-KD
Chr4_4128362_651_po_C_Xz7/11        LSEGETVN---IYTDS-----PYVFZTVQVHG--AIER-KGPIELWG-KR
Chr14_50203955_850_po_C_Xz10/8      LSEGKTVN---TYTDS-----WYAFLSLQVHG--VZQKEKGLSNSGG-KG
Chr14_50205291_837_po_C_Xz10/8      LSEGKTVN---TYTDS-----WYAFLSLQVHG--VZQKEKGLSNSGG-KG
Chr7_6597120_842_po_C_Xz7/7         LSEGETVN---IYTDS-----RYAFLTLZVHG--ACZKEKGLLNSGG-KG
Chr11_71544658_783_po_C_Xz8/13      LSEGETVN---IYTDS-----RYVFLTLHVHG--AZZKEKGLLDSGG-KG
Chr11_67999520_760_po_C_Xz8/8       LSEGETVN---IYTDS-----RYVFZTLHMHG--AIER-KGLLNSGV-KD
Chr3_75461975_796_po_C_Xz7/9        LNKGETVN---IYTDS-----PYAFLTLQVHG--AIER-KGLLNSGG-KD
Chr12_52228939_692_po_C_Xz3/10      LHEGKTVN---IDTDS-----RYAFLTLQVHG--ALZKENGLWNSGG-KD
Chr10_15211234_286_po_C_Xz11/1      LSEGETVN---IYTDS-----RYVFLTLQVHG--AIKE-KGLLNSGG-KG
Chr12_100973302_945_po_C_Xz12/      LSEGKTVN---IYTDS-----QYAFLTLQVRG--ALYKEKDLLNSGE-KD
Chr19_9685420_817_po_C_Xz6/10       LSEGKTVN---IYTDS-----RYAFLTLQVHGVWSIYKEKGLLNSGG-KD
Chr13_40936820_719_po_C_Xz9/10      LSEGETVN---IYTDS-----RYVFLTLQVHG--AZZKEKGLLNSGG-KD
Chr7_63858202_1754_po_C_Rz12/9      LAEDKKIS---VYTDS-----KYAFATLHVHGA--IYKERGLLTATG-KE
ERV3_chr7q21_AC073210_polputei      LAEDKKIS---VYTDS-----KYAFATLHVHGA--IYKERGLLTATG-KE
Chr6_110218096_843_po_C_Xz16/8      LARGKRVN---VYTDS-----KYAFSTLHVHGA--IYKERELLTAGG-KE
Chr6_110214535_786_po_C_Xz16/8      LARGKRVN---VYTDS-----KYAFSTLHVHGA--IYKERELLTAGG-KE
ChrY_3018841_881_po_C_Rz9/12        LVKDKRVN---IYTDS-----KYAFATLHVHGA--IYKERGLLTAGG-KE
ChrX_86876098_639_po_C_Xz7/8        LVKDKRVN---IYTDS-----KYAFATLHVHGA--IYKERGLLTAGG-KE
Chr14_68899917_965_po_C_Xz6/8       LAKDKKVN---VYTDS-----KYAFATLRVHGA--IYKERGLITAGE-KZ
ChrY_23874437_661_po_C_Xz15/9       LTKRKMAN---VYTDS-----KYTFATLPDNGA--IYKEKGFLTAVGQKE
ChrY_27518418_649_po_C_Xz15/9       LTKRKMAN---VYTDS-----KYTFATLPDNGA--IYKEKGFLTVVGQKE
ChrY_25016892_657_po_C_Xz13/12      LTKRKMAN---VYTDS-----KYTFATLPDNGA--IYKEKGFLTAVGQKE
```

```
ChrY_19932937_520_po_C_Xz10/14      LAKGEMVN---VYTDS-----KYTFATLHDNGA--IYKEKGLLMAVG-KK
ChrY_20104614_518_po_C_Xz10/14      LAKGEMVN---VYTDS-----KYTFATLHDNGA--IYKEKGLLMAVG-KK
Chr19_58660155_598_po_C_Xz21/1      LAKEK-----SIYTNS-----KYAFTTLHVHGT--IYKERGLLSAR--QE
Chr19_58621153_500_po_C_Xz21/1      LAKEKKV---SIYTNS-----KYAFTTLHVHGT--IYKEKGLLSAR--QE
Chr7_139686987_710_po_C_Xz11/7      LAKDKKVNTVNIYTDS-----KYAFTTLHVHGT--IYKKRGLLTAGG-KZ
Chr7_137651301_1097_po_C_Pz13/      LAKDKKVN---VYTDS-----KYAFATLHVHGA--IYKERGPLTAGG-KE
Chr11_29579197_930_po_C_Pz8/12      S-KDKKVN---VYTDS-----KYAFAMLHVHRA--IYKZRGLTTAEG-KE
Chr4_117329205_488_po_C_Xz3/16      LAKDKKVN---VYTN-----VKYVFATLHVHGA--MYR-RGLLTAEG-KE
Chr21_38522185_383_po_C_Xz7/9       LMRDKKVN---IYTDSKSQRQKCVFATLHVHEA--IYKERGLLTAGK-KE
Chr4_70173370_1049_po_C_Xz11/8      LAKDQKAN---IYTDS-----KYAFATLHVHGA--IYKEKRLLTPGG-KE
Chr4_70022513_788_po_C_Kz10/12      LAKVQKAN---LYRDS-----KYAFATLHVHGA--IYKEKELLTAGG-KE
Chr14_105225894_1027_po_C_Xz13      LAKDQKTN---IHTNS-----KYAFAILNVHGA--IYKERGLLTDGG-KE
Chr7_119241536_761_po_C_Xz11/1      LAKDKKIN---VYTDS-----KYAFAMLHVLGA--IYKERGPLTARD-KE
ChrY_22564970_567_po_C_Xz9/10       LARGKVVN---SYTDF-----KYAFVTLLAHGA--IYKERGLLTAGR-KE
Chr21_14409346_1477_po_C_Pz5/6      LAKRKSAN---IYTDS-----RYAFATLHAHGA--IYKERELLTTEG-KE
ChrX_51259700_1396_po_C_Pz7/8       LAKGKSVN---IYTDS-----KYAFATLHAHGA--IYKEKGLLTTEG-KE
Chr6_64985396_1041_po_C_Xz11/1      LAKGKSVS---IYTDS-----RYVFA-IHAHGA--IYKERGLLTTEG-KE
ChrY_20355209_659_po_C_Xz7/11       LARGKSVN---IYNDL-----RYAFATLHAHGA--IYKERGLLATEG-KE
Chr3_34567166_523_po_C_Xz8/6        LAKGKSVY---IYTDS-----RYAFATLHAHEA--ISKERGLLTTEG-KE
Chr3_122072468_245_po_C_Xz9/9       LAKGKSVN---IYTDS-----RHAFATLHAHRA--KYKERGLLTTEG-KE
Chr11_58788475_541_po_C_Xz8/22      LAKGKSVN---IYTVS-----RYAFATLPVHG---AYKERALLTTEG-KE
Chr6_137421042_373_po_C_Xz6/9       LAK-KSVN---IYTDS-----RYAFATLHAHGA--IYKERGLLTTEG-KE
Chr4_115910608_550_po_C_Pz8/12      LAKGKSVN---IYTDS-----RYAFATLHAHGA--IYKERGLLTTEG-KK
ChrY_13743185_977_po_C_Xz14/8       LAKRKKVS---IYTDS-----RYVFATLHDHGA--IZKERGLLTTEG-ND
HERV15                              LAKRKKVN---IYTDS-----RYAFATLHAHGA--IZKERGLLTTEG-KD
ChrX_117198092_983_po_C_Xz8/8       LAKGKKVN---ICTDS-----KYAFATLRAHGG--IHKERGLLTTEG-KE
Chr19_21166030_1412_po_C_Xz9/1      LAKGKK-N---IYTDS-----NCAFATLHTHGA--IYN-RGLLTTEG-KE
Chr19_21213743_913_po_C_Xz14/1      LAKGKKVN---IYT-S-----KYAFATLHTHGV--TYKERGLZTMRK-KE
Chr19_15957774_569_po_C_Xz8/10      ------LSSSLSFMYSPTP-ISASPPTTAIPPP-----------------
Chr11_77508681_297_po_C_Xz19/1      KFSRRHLGSQVSLGPGLQPPIGKGDSKPVISTP--GWLAEMLQETEEWRD
Chr13_98250769_738_po_C_Xz7/13      QICHEEVTYLGFIVSHEERRLGDERKQAICALP-----------------
Chr6_39968183_513_po_C_Pz10/9       -AELIALTRALLLAKGKVVNIYTDYAFATLHAH-----------------
Chr15_66867263_645_po_C_Xz11/1      KVAGRALDPPQLAETGIPGPVDLPAGRPSDQA--DLZ-------------
HERV18_polputein                    MVITKEVTPMVICTDS-----WAVYRGLLVVNY--LELQKWLVGHRP---
HERVIP10FH_polputein                YLQNQERA---IYTNS-----KLIWTEQGLINS----KEIEIPNLQGFQQ
MER66_polputein                     AQPDTMLS-------------LLILFALPITLIQCGWEQNSIVNISKIIA


HERVI_polputein                     LVHKELIIQVLDNLQLPEEIAIVR-VPGHQKSLSF---------------
HERVIP10F_polputein                 LVHKELITQVLNNLQLPEEIAIVH-VPGHQKSLSF---------------
HERVadp_AC005741                    LVHGELVKQVLESLLLPAEVAIVH-VNDHQIGNTV---------------
HERV-PT47D                          VSKPHQVRTILN--------------------------------------
HERV_W_chr6_141432567_ERV9_lik      IKYHKEIIQLLHAVQKPKEVAVLT-AEAIEKGK-----------------
ERV9_PH1_RT                         IKYHKEIIELLHAMQKHKEVGILH-ZQSHQNRK-----------------
Chr13_54564403_572_po_C_Yz7/14      IKHQEAIRRLLLAVQKPKEVAVLH-CRGHQKGK-----------------
HERV_Y_chr12_51022911_pol           IKHQEAIRRLLLAVQKPKEVAVLH-CRGHQKGK-----------------
HERV17_polputein                    IKHQEAIRRLLLAVQKPKEVAVLH-CRGHQKGK-----------------
HERV_W_chr7_9105739_syncytin_p      IKHQEAIRKLLLAVQKPKEVAVLH-CRGHQKGK-----------------
HERV9_polputein                     IKCHKEIMELLHALQKPKEVAVLH-CQSHQKGE-----------------
HUERS-P3_polputein                  IKYHQEINRLLSSVFLPREVAVMH-CRGHQKGT-----------------
ERVfrd_AC004022                     IKYGSQVLELLEAAHLPQEVAVEH-CEGHHRGF-----------------
PRIMA41_polputein                   IKHHLEIRNLLDAVLLPKEVAIIH-CRGHLKGN-----------------
HERV_H_RGH1_pol                     IINASLIKTLLKTTLLPKEAGVIH-CKGHQKAS-----------------
HERVHRGH2_pol                       IINAS-KKTLLKAALLPKEAGVIH-CKGHQKTS-----------------
HERV_H48I_polputein                 VINGSLISKLLQAARLPQKVAIIH-CRGHQTPD-----------------
HERV_H_RTVLH2_pol                   DPHDCISLIHLTFTPFPHIS--------FFPIP-----------------
HERVFH21_pol                        ITNGSPYLPLLQAAHLPTKAGVIH-CZGHQTGSERPLGNRLISRXXXXXX
HERVRblike_chr4_109047953           IKYAKQVLELPEAIKAPREIAVMR-CPGHQCNN-----------------
ChrY_13790029_693_po_C_Xz12/11      VKNKEEILALLEA------VAVIH-CKGHQKEDLA---------------
HERVS71_polputein                   IKNKEEILALLEAVWLPPQVAVIH-CKCHQTEGMA---------------
BaEV_M7_pol                         IKNKAEIIALLKALFLPQEVAIIH-CPGHQKGQDP---------------
HERVR_polputein                     IKNKAEIIALLKALFLPQEVAIIH-CPGHQKGQDP---------------
AKR_MLV_MLOCG_RT                    IKNKDEILALLKALFLPKRLSIIH-CPGHQKGHSA---------------
GaLV_pol                            IKNKEEILALLEAIHLPRRVAIIH-CPGHQRGSNP---------------
Chr3_158587414_1662_po_C_Ez10/      IK-SQEILQLLEAVZKPQKVAVMH-CRGHQQASTS---------------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      IKYQQEILQLLEAVWKPQKVAVMH-CREYQZGSTS---------------
Chr5_111890459_1626_po_C_Ez10/      IKYQQVILQLLEAVWKPQKAAVMH-CRGHQQASTS---------------
ChrX_48766518_1600_po_C_Ez6/3       IKYQQEILQLLEAVWKPQKVAVMH-CRGHQRASIS---------------
ChrX_152580713_1572_po_C_Ez10/      IKYQQEVLQLLEAVWKPQKVAVMH-CRGHQRASTS---------------
ChrX_139287533_1154_po_C_Xz8/4      IKYQQEILQLLEAVRKPQKVAVMH-CRGHQZASTL---------------
Chr2_52615786_912_po_C_Xz11/5       IKYQZEVLQLLEAVWKPQKVTVMY-CRGHQRASTS---------------
HERV-E_4_1                          IKYQQEILQLLEAVWKPHKVAVIH-CGGHQZASTL---------------
ChrY_21962399_1137_po_C_Xz13/9      IKDQQEILQLLEAVWRPRKVAVMH-CRGHQZASTS---------------
ChrY_8634032_950_po_C_Xz14/10       IKYQQEILKLLEAVWKPQKVAVLH-CSGHQZSSTS---------------
Chr3_126735360_997_po_C_Xz5/10      RKYQQEILQZZEAVLKPHKVAVMH-CRGHQRASIL---------------
Chr3_126713203_958_po_C_Xz7/8       RKYQQEILQZLEAVWKPHKVAVMH-CRGHQRASTL---------------
```

```
Chr3_75226118_977_po_C_Xz8/9        RKYQQEILQZLEAVWKPHKVAVTH-CRGHQRASTL---------------
Chr7_6629919_827_po_C_Xz8/7         RKYQQEILQZLEAVWKPHKVAVMH-CRGHQRASTL---------------
Chr4_9166812_811_po_C_Xz8/10        RKCQQEILQZLEAVWKPHKVAVTH-CRGHQRASTL---------------
Chr12_8478642_834_po_C_Xz12/7       RKYQQEILQZLEAVWRPHKVAVRH-CRGHQRASTL---------------
Chr10_15187175_787_po_C_Xz11/5      RKYQQEIFQZLEAVWRPHKVAVKH-CRGHQRASTL---------------
Chr11_71819740_854_po_C_Xz11/9      RKYQQEILQZLEAVWKPHKVAVMH-CRGHQRASTW---------------
Chr11_71519974_886_po_C_Xz13/9      RKYQQEILQZLEAVWKLHKVAVMH-CRGHQZASTL---------------
Chr8_7271441_969_po_C_Xz10/6        RKYPQEILQZLEAVWKPHKVAVRH-CRGHQRASTL---------------
Chr8_11770066_826_po_C_Xz9/8        RKYQQEVFQZLEAVLKPHKVEVRH-CRGHQRASIL---------------
Chr4_4128362_651_po_C_Xz7/11        QKISTRNLAMIRNSMETPQGGNYA-LQRTPASFHL---------------
Chr14_50203955_850_po_C_Xz10/8      RKYQPEIFHLLEAVWKPHKVAVMH-CRGHQRASTL---------------
Chr14_50205291_837_po_C_Xz10/8      RKYQPEIFHLLEAVWKPHKVAVMH-CRGHQRASTL---------------
Chr7_6597120_842_po_C_Xz7/7         RKYQPEIFHVLEAVWKPHKVAVMH-CRGHQRASTL---------------
Chr11_71544658_783_po_C_Xz8/13      TKYQPEIFHVLEAVWKPHKVAVMH-CRGHQRASTL---------------
Chr11_67999520_760_po_C_Xz8/8       RKYQPEIFYVLEVVWKPHKVAVMH-CRGHQRASTL---------------
Chr3_75461975_796_po_C_Xz7/9        RKYQQVILQZLEAVWKPHKVAVMH-CRGHQQASTL---------------
Chr12_52228939_692_po_C_Xz3/10      IKYQQEILQLLEAVWKPHKVAGMH-CRGHQRASTL---------------
Chr10_15211234_286_po_C_Xz11/1      RKZQPEILHLLEAVWKPHKVAVMH-CRGHQRASTL---------------
Chr12_100973302_945_po_C_Xz12/      IKYQQETLQLLEAMWKPQKVAVMH-A-GHQZASTS---------------
Chr19_9685420_817_po_C_Xz6/10       IKYQQDILQLLEAVWKPQKVAVMH-CRGHQZASTS---------------
Chr13_40936820_719_po_C_Xz9/10      KISTRNLAIIRNSMETPNVTGGVHPSILFVILWRD---------------
Chr7_63858202_1754_po_C_Rz12/9      IKWKEZILQLLDAVWIPEKVAAMH-CRGHQRAGTS---------------
ERV3_chr7q21_AC073210_polputei      IKWKEZILQLLDAVWIPEKVAAMH-CRGHQRAGTS---------------
Chr6_110218096_843_po_C_Xz16/8      IKYKEEILQLLDAVWAPKKVAIIH-CRGNQRGNTGG--------------
Chr6_110214535_786_po_C_Xz16/8      IKYKEEILQLLDAVWAPKKVAIIH-CRGNQRGNTGG--------------
ChrY_3018841_881_po_C_Rz9/12        IKYKEEILQLLEAVZALKKVAIMH-CRG-QRSRTL---------------
ChrX_86876098_639_po_C_Xz7/8        IKYKEEILQLLEAVZALKKVAIMH-CRGHQRSRTL---------------
Chr14_68899917_965_po_C_Xz6/8       IKYKEEVLQLLDAVWP-KEVAVMH-CMGHQRSRTL---------------
ChrY_23874437_661_po_C_Xz15/9       KKVKEKILQLLDAVWTPKKTAAMH-CKGHQKPGSL---------------
ChrY_27518418_649_po_C_Xz15/9       KKVKEKILQLLDAVWTPKKTAAMH-CKGHQKPGSL---------------
ChrY_25016892_657_po_C_Xz13/12      KKVKEKILQLLDAVWTPKKTAAMH-CKGHQKPGSL---------------
ChrY_19932937_520_po_C_Xz10/14      IKYKEKILQLLDAVCASKKMAVMH-CRGHQKAGTL---------------
ChrY_20104614_518_po_C_Xz10/14      IKYKEKILQLLDAVCASKKMAVMH-CRGHQKAGTL---------------
Chr19_58660155_598_po_C_Xz21/1      IKYKEEILQLLEAVZAPEKVAIMH-CRGHQMARMP---------------
Chr19_58621153_500_po_C_Xz21/1      IKYKEEILQHLEAVZLPEKVAIMH-CRGHQKARMP---------------
Chr7_139686987_710_po_C_Xz11/7      IKYKKKILQFLEAAWAPEKVVVMH-CRGHQKAGMP---------------
Chr7_137651301_1097_po_C_Pz13/      IKYEEELLQLLDAVZAPKKVAVMH-CR-HQKAGTL---------------
Chr11_29579197_930_po_C_Pz8/12      IKYKEKILQLLNVVLAPKKVAVIL-CKGHQRAGTL---------------
Chr4_117329205_488_po_C_Xz3/16      IKYKEKILQLLDTVWAPKKVAVMH-CRRHEKAGTL---------------
Chr21_38522185_383_po_C_Xz7/9       IKYKEEILQLLGTVZAPKKVAVMH-CRGHQKARTL---------------
Chr4_70173370_1049_po_C_Xz11/8      IKCNEEILRLLDVVWAPKQVAVMH-CR--QKAGTL---------------
Chr4_70022513_788_po_C_Kz10/12      IKYKEEILTLRHCIS-PKQVAVMQ-CR--QKAGTL---------------
Chr14_105225894_1027_po_C_Xz13      RKYKEEILQPLDAVWAPKQVAVMH-CRGHLRAGTS---------------
Chr7_119241536_761_po_C_Xz11/1      IKHKEEILQLLDAVWAPEKVAIMH-CRSHQKAGKL---------------
ChrY_22564970_567_po_C_Xz9/10       IKNKEKILQLLDAAWAPEKVAVIH-CKGHKASNFE---------------
Chr21_14409346_1477_po_C_Pz5/6      IKNKKEIGQVLEAVWAPKEVAVFH-CKGHQTGGSD---------------
ChrX_51259700_1396_po_C_Pz7/8       IKNKKEIEQLLEAVWTPKKVALIH-CNGHQTGGSN---------------
Chr6_64985396_1041_po_C_Xz11/1      IKNEKEIEQLLEAIWAPKEVAVIY-CKGHQTGGRD---------------
ChrY_20355209_659_po_C_Xz7/11       IKNKKKVQQLSEAIWAPKEVAVIH-RKGHQTGGGD---------------
Chr3_34567166_523_po_C_Xz8/6        VKNKKEIEQLLEAVWAPKEVAVIH-CKRHRTGGGD---------------
Chr3_122072468_245_po_C_Xz9/9       IKNKKEIEKLLEAVWAPKEVAVIH-CKEQQTGGGD---------------
Chr11_58788475_541_po_C_Xz8/22      IKNKKKIEQFLEAIWAPKEVAVIH-CKMHKTVGGD---------------
Chr6_137421042_373_po_C_Xz6/9       IKNKKEIEQLLEAVWAPKEVAVIH-CKGHQTGGHD---------------
Chr4_115910608_550_po_C_Pz8/12      IKNZKEIEQLLEAVWVPKEVTVIY-CKGHQTEGGD---------------
ChrY_13743185_977_po_C_Xz14/8       IKNKEEILQLIEDIWAPEKVAVIH-CKGHZIRKGY---------------
HERV15                              IKNKEEILQLLEDIWAPEKVAVIH-CKGHQIRKSY---------------
ChrX_117198092_983_po_C_Xz8/8       IKNKEEILQLLEAVWAPEKVAVIH-CKGHQIRKSY---------------
Chr19_21166030_1412_po_C_Xz9/1      IQNKEEILQLLEAIWAPEKVALIH-CKGHQIRKSY---------------
Chr19_21213743_913_po_C_Xz14/1      IKNKKEILQLLEALWAPEKVAAIH-CRGHQTGKSY---------------
Chr19_15957774_569_po_C_Xz8/10      ----------------------------PRRSVHYW--------------
Chr11_77508681_297_po_C_Xz19/1      QYESTGRYLFLGTHWLSGFISKNL-SIKQRQSRVFI--------------
Chr13_98250769_738_po_C_Xz7/13      -------------TPTTRHQIREF-LGAAGFCHIWI--------------
Chr6_39968183_513_po_C_Pz10/9       ------------------------------RAIYFL--------------
Chr15_66867263_645_po_C_Xz11/1      --------------------QDQV-ATENNNSPALV--------------
HERV18_polputein                    -IWGQAMWQDLWEMGHQKDVTIYH-VSGHMPLAAP---------------
HERVIP10FH_polputein                KZSLLKVNSVTCIILTSNLVALDSLVHRHEGSSLWKR-------------
MER66_polputein                     SGNHLHGCWICHQHPQDREFHLLAYPENLTAIS-----------------


HERVI_polputein                     --------------ESRGNNLTNQIAKQAAVSSETP--------------
HERVIP10F_polputein                 --------------ESRGNNLADQVAKQAAVSSEMR--------------
HERVadp_AC005741                    --------------EAIGNRLADEAAKQASLEEEIR--------------
HERV-PT47D                          ------------------QRAGRWLTDSSAISFETK--------------
HERV_W_chr6_141432567_ERV9_lik      ------------ER---GEQQHKQLAEAAERKE-----------------
ERV9_PH1_RT                         ------------ER---GEQQHKQLAEAAERKEE----------------
Chr13_54564403_572_po_C_Yz7/14      ------------EREIEGNCQADIEAKRAARKPLGDYY-----------W
HERV_Y_chr12_51022911_pol           ------------EREIEGNCQADIEAKRAARQDPP---------------
```

```
HERV17_polputein                    ------------EREIEGNRQADIEAKRAARQDPP---------------
HERV_W_chr7_9105739_syncytin_p      ------------EREIEENCQADIEAKRAARQDPP---------------
HERV9_polputein                     ------------GEKAEGNRQADAEAKIAARWILP---------------
HUERS-P3_polputein                  ------------DEIAKGNKLADQAAKSAARKPQG---------------
ERVfrd_AC004022                     ------------YZTTWGNRLADKNAEEAAMSNDT---------------
PRIMA41_polputein                   ------------SSVTKGNSFADAAAKATALKDPVG--------------
HERV_H_RGH1_pol                     ------------DPIALGNAYADKVARQAASSPTS---------------
HERVHRGH2_pol                       ------------DPIALGNAYADKVARQAASSPTF---------------
HERV_H48I_polputein                 ------------NPISAGNALADQVAKQVALQP-----------------
HERV_H_RTVLH2_pol                   ------------HPITLD--------------------------------
HERVFH21_pol                        KAAXNRISGGIRNEISRGNRKADEAAKEASLSS-----------------
HERVRblike_chr4_109047953           ------------SEVARGNAFLDCTARHLASSNVELRVPLIPQIDLAASN
ChrY_13790029_693_po_C_Xz12/11      --------------IAHGNQRADSVTRKAAWLPVMPLTLL----------
HERVS71_polputein                   --------------IACGNQKAGSAAREAAWLPVTPLTLL----------
BaEV_M7_pol                         --------------VAVGNRQADRVARQAAMAEVLTLATE----------
HERVR_polputein                     --------------VAVGNRQADRVARQAAMAEVLTLATE----------
AKR_MLV_MLOCG_RT                    --------------EARGNRMADQAARKAAITETPDTSTL----------
GaLV_pol                            --------------VATGNRRADEAAKQAALSTRVLAGTT----------
Chr3_158587414_1662_po_C_Ez10/      --------------VALGNSQVDSEAZKATSTPYWASVA-----------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      --------------VALGNSRADSEARKAASTPYZASVA-----------
Chr5_111890459_1626_po_C_Ez10/      --------------VALGNSZADSEAQNAASTPYWASVA-----------
ChrX_48766518_1600_po_C_Ez6/3       --------------VALGSSRADSEAZKAASTPYQALVA-----------
ChrX_152580713_1572_po_C_Ez10/      --------------VALGNSRADSEAZKAASTPYWASVA-----------
ChrX_139287533_1154_po_C_Xz8/4      --------------VSQGNSRADTEARKAASTPYRASVT-----------
Chr2_52615786_912_po_C_Xz11/5       --------------IALGTSZADSEAQKAAFTQYWASVT-----------
HERV-E_4_1                          --------------VGLGNSCTDLEAQKAALP-FRASVT-----------
ChrY_21962399_1137_po_C_Xz13/9      --------------ISLGNSQADSEAZKSVSTPYWASVT-----------
ChrY_8634032_950_po_C_Xz14/10       --------------VALGYPZADSEAQKAASTSYWASVA-----------
Chr3_126735360_997_po_C_Xz5/10      --------------LGLGNSHSDSEVRKAASAPFWASV------------
Chr3_126713203_958_po_C_Xz7/8       --------------LGLGNSRADSEAQKAASAPFRASV------------
Chr3_75226118_977_po_C_Xz8/9        --------------LGLGNPRADSEAQKAASAPFWASV------------
Chr7_6629919_827_po_C_Xz8/7         --------------LGLGNSCADSEAQKAASAPFQASV------------
Chr4_9166812_811_po_C_Xz8/10        --------------LGLGNSRADSEARKAASAPFRASV------------
Chr12_8478642_834_po_C_Xz12/7       --------------LGLGNPRADSEAZKAASAPFWASV------------
Chr10_15187175_787_po_C_Xz11/5      --------------LGLGNSGADSEARKAASAPFWASV------------
Chr11_71819740_854_po_C_Xz11/9      --------------LGLRNSRADSEARKAASSPFWASG------------
Chr11_71519974_886_po_C_Xz13/9      --------------LGLGNSRADSDARKAASAPFWASV------------
Chr8_7271441_969_po_C_Xz10/6        --------------LCLGNSRAESEARKAATAPFWASV------------
Chr8_11770066_826_po_C_Xz9/8        --------------LDLGNSHADSEARKAASVPFWASV------------
Chr4_4128362_651_po_C_Xz7/11        --------------AGFGEFPRDSEAGKAASAPFGASV------------
Chr14_50203955_850_po_C_Xz10/8      --------------VGLGNSRADSEAGKAASAPFZASVT-----------
Chr14_50205291_837_po_C_Xz10/8      --------------VGLGNSRADSEAGKAASAPFZASVT-----------
Chr7_6597120_842_po_C_Xz7/7         --------------VGLGNSRADSEARKAASAPFQASVT-----------
Chr11_71544658_783_po_C_Xz8/13      --------------VGLGNSRADSEAQKQHLP-LSRSVT-----------
Chr11_67999520_760_po_C_Xz8/8       --------------VGLGNSRADSZAGKAASARFQASVT-----------
Chr3_75461975_796_po_C_Xz7/9        --------------LGLGNSCADSEAPKAASAPFWTSV------------
Chr12_52228939_692_po_C_Xz3/10      --------------LGLGNSRADSEARKAASAPFRQSVT-----------
Chr10_15211234_286_po_C_Xz11/1      --------------VGLGNSRADSEAAKAASVPFRASVI-----------
Chr12_100973302_945_po_C_Xz12/      --------------VALGNSRANSEAPKAASTPYWASVA-----------
Chr19_9685420_817_po_C_Xz6/10       --------------VALGNSRADSETRKAASTPYRASVA-----------
Chr13_40936820_719_po_C_Xz9/10      --------------IALNTTVGVHTWGTSTGILFVISSK-----------
Chr7_63858202_1754_po_C_Rz12/9      --------------EAKGNRKADREAKWAAMVTPHFKKE-----------
ERV3_chr7q21_AC073210_polputei      --------------EAKGNRKADREAKWAAMVTPHFKKE-----------
Chr6_110218096_843_po_C_Xz16/8      --------------QPKGNKKADREAKWAAMTTPHLREE-----------
Chr6_110214535_786_po_C_Xz16/8      --------------QPKGNKKADREAKWAAMTTPHLREE-----------
ChrY_3018841_881_po_C_Rz9/12        --------------EAKRNGKTDREARQAATTTAHSRKK-----------
ChrX_86876098_639_po_C_Xz7/8        --------------EAKRNGKTDREARQAATTTAHSRKE-----------
Chr14_68899917_965_po_C_Xz6/8       --------------EAKGNRKADREAKQAAVTTLHPGEE-----------
ChrY_23874437_661_po_C_Xz15/9       --------------EAKRNKKVDRYAKGAAMTTHR-QTT-----------
ChrY_27518418_649_po_C_Xz15/9       --------------EAKRNKKVDRYAKGAAMTTHR-QTT-----------
ChrY_25016892_657_po_C_Xz13/12      --------------EAKRNKKVDRYAKGAAMTTHR-QTT-----------
ChrY_19932937_520_po_C_Xz10/14      --------------EAKKNKKIDREARGAAMTTLQFKKK-----------
ChrY_20104614_518_po_C_Xz10/14      --------------EAKKNKKIDREARGAAMTTLQFKKK-----------
Chr19_58660155_598_po_C_Xz21/1      --------------EATGNRKA--EAKQAAMITLHFKEE-----------
Chr19_58621153_500_po_C_Xz21/1      --------------EAKGNRKADREAKQAAMITLHFKEE-----------
Chr7_139686987_710_po_C_Xz11/7      --------------EAKGNRKTGRGAKRAAMIMPHFKEK-----------
Chr7_137651301_1097_po_C_Pz13/      --------------EAKGNRKADMEAKQAGMTTPHCKKE-----------
Chr11_29579197_930_po_C_Pz8/12      --------------EAKGNRKVDREVAMA---TLHCEEE-----------
Chr4_117329205_488_po_C_Xz3/16      --------------EAKGNRKADKEAKRAAMTTPPFKKE-----------
Chr21_38522185_383_po_C_Xz7/9       --------------EAKGNRKADREAKRTAMITPHFEKE-----------
Chr4_70173370_1049_po_C_Xz11/8      --------------EAKGNRKADKEAKQAAMATTFYSKEE----------
Chr4_70022513_788_po_C_Kz10/12      --------------EAKGSRKTDKEAKQAAMATPP-SKQE----------
Chr14_105225894_1027_po_C_Xz13      --------------EVKGNRKAD-EAKQTAVVTLP-SKEE----------
Chr7_119241536_761_po_C_Xz11/1      --------------ETMGNRKADREAKWATMTTLHFKKE-----------
ChrY_22564970_567_po_C_Xz9/10       ---------------AKANRKADKKAQWASMATFPPKEG-----------
Chr21_14409346_1477_po_C_Pz5/6      --------------ETTGNRKADKEVKRAAMTEITKKE------------
```

```
ChrX_51259700_1396_po_C_Pz7/8       --------------EATGNRKVDREAKRAAITKMTEKR------------
Chr6_64985396_1041_po_C_Xz11/1      --------------EARGNRK-ATEAKRAAMTEVTKKEE-----------
ChrY_20355209_659_po_C_Xz7/11       --------------EARGNRKGDREAKGAAMTEVTKMEE-----------
Chr3_34567166_523_po_C_Xz8/6        --------------EARGKRKVDREAKRAAMTDVTKKEE-----------
Chr3_122072468_245_po_C_Xz9/9       --------------KARGNRKADREAKRAVMTEVTKKEK-----------
Chr11_58788475_541_po_C_Xz8/22      --------------ETRGNGKADREAKRTAMAEVTKKEE-----------
Chr6_137421042_373_po_C_Xz6/9       --------------KARENRKAEREAKRAAMTAITKKEE-----------
Chr4_115910608_550_po_C_Pz8/12      --------------EARGNRKVDREAKRAAMTEVTKKEK-----------
ChrY_13743185_977_po_C_Xz14/8       --------------EAQGNRZADQEAQQAAMSKVLPEER-----------
HERV15                              --------------EAQGNRKADQEAQQAAMSKVLPEER-----------
ChrX_117198092_983_po_C_Xz8/8       --------------EVQGNRKADQEARQAAMSKASPEER-----------
Chr19_21166030_1412_po_C_Xz9/1      --------------EAQGNRRADQEVRWAAMSKVLPEEI-----------
Chr19_21213743_913_po_C_Xz14/1      --------------KVQGNRKVDPEARQAAMNKALPVKR-----------
Chr19_15957774_569_po_C_Xz8/10      --------------SPMRPQPKALS-------------------------
Chr11_77508681_297_po_C_Xz19/1      --------------SGLTKVKKAVNHIVHNLWPCIAGGLG----------
Chr13_98250769_738_po_C_Xz7/13      --------------PNFSLMPKPLYEATKG-----GEKEP----------
Chr6_39968183_513_po_C_Pz10/9       --------------KSZGKHNHKKG-------------------------
Chr15_66867263_645_po_C_Xz11/1      --------------TPKADQSTHGGNLRKQ---------Q----------
HERV18_polputein                    -----------------SNDEADALAKVQWLESASTZ-------------
HERVIP10FH_polputein                -------------MVIIFEKKKGGGGRIYVKGMLYGKFLSZ---------
MER66_polputein                     --------------PDLLTNHSNPEVPKPLLVKWNSPPLN----------


HERVI_polputein                     ----MFHLTPCLPSPT--AISFFSSIEKEEKIEPKEKTRMKTAVTRPKGN
HERVIP10F_polputein                 ----IFHLTPYLPPPTN-PHFLFTEKEKLIKIGAKENSEGKWILPDQREM
HERVadp_AC005741                    ----LFSLIPDISKVVLRPQFTKEEKEELDKIGVTQTEDGKWVLPAGR-E
HERV-PT47D                          ----TLFIT-----------------------------------------
HERV_W_chr6_141432567_ERV9_lik      -----RKRQEVREREKE--------KETERQSKKEREEETKMESERKRET
ERV9_PH1_RT                         -TGSQRKRQRGRDKKLERRGTDTEVKERVKKRGRDKEEVKERDGSSKEKT
Chr13_54564403_572_po_C_Yz7/14      LYRNLKRAVLHQEAIRRLLLAVQKPKEVAVLHCRGHQKGKEREIEGNCQA
HERV_Y_chr12_51022911_pol           -LEMLIEGPLVWGNPLQETKPQYSEEEIEWGTSZGHSFLLRMASHQRRKN
HERV17_polputein                    -LEMLIEGPLVWGNPLRETKPQYSAEEIEWGTSRGHSFLPSGWLATEEGK
HERV_W_chr7_9105739_syncytin_p      -LEMLIKQPLVZGNPLREPSPS-TQQEKQNGEPHEDSFLPSGRLATEEGK
HERV9_polputein                     -LEIPMEGPLEWNKPLQEIKPQYSPNETEWGLSWGHSFLPSGWLTTKEGK
HUERS-P3_polputein                  --INTLQAPLIWEGSIREIKPQYSPTEIEWATSRGYTFQPSEWLQQRMAN
ERVfrd_AC004022                     -----FMGDLFLSLPSELPFLQYTKEEIDWSSQHGYKEETSGWYRLEEFS
PRIMA41_polputein                   ------LVGMLVPSATVITEPRYTKEEQEWAKGQGLIQDPSGWLIND-NK
HERV_H_RGH1_pol                     ---------VPHSQFFSFASVTPTYSPTETSTY--QSLPTQGKWFLDQGK
HERVHRGH2_pol                       ---------VPHGQFFSFTSVTPTYSTAETSTY--QALPPQGKWFLDQGK
HERV_H48I_polputein                 ----------VQGQFLSLSLFSPLYSSEEKEDFRAQNLQKQGPWYVKEGC
HERV_H_RTVLH2_pol                   ---------------LLMASTRPNCHTPAKADY----------------A
HERVFH21_pol                        ---------APASLLLIPLQSNPXTLPLRKGFTTTTRNLLPRGLDSQKSK
HERVRblike_chr4_109047953           LGTVLRTKRLPRTKGLFRMKNDWKLNDKGLIWVSEIKFLLSYSFCLHLVL
ChrY_13790029_693_po_C_Xz12/11      ----S--AVSFPQPNL----AAPEKKRNKLQIFRPINQ-EGWWILPDSRI
HERVS71_polputein                   ----P--TVSFPQPDLPDHPQTPQRKKNKLQIFGPVNIRKVWWILPDSRI
BaEV_M7_pol                         -------PDNTSHITIEHT-YTSEDQE-EARAIGATENKDTRNWEKEGKI
HERVR_polputein                     -------PDNTSHITIEHT-YTSEDQE-EARAIGATENKDTRNWEKEGKI
AKR_MLV_MLOCG_RT                    ----L--IENSSPYTSEHFHYTVTDIK-DLTKLGAIYDKTKKYWVYQGKP
GaLV_pol                            ----K--PQEP----------IEPAQE-KTRPRELTPDR--------GK-
Chr3_158587414_1662_po_C_Ez10/      ------A-PLLPQTPDLVPTYSKEEKDFFH-AEGGQAIKGGWIRLPDERV
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      ------A-PLLPQIPDLVPTZ--ERKRLLP-HRRGQVIKGGWIRLPDGRV
Chr5_111890459_1626_po_C_Ez10/      ------A-PLLPQTPDLVPTYSKEEKDFFH-AEGGQVIKGGWIRLPDGRV
ChrX_48766518_1600_po_C_Ez6/3       ------A-PLLPQTSDLVPTYSKEEKDFFH-AEGGZVIKGGWIRLPDGRV
ChrX_152580713_1572_po_C_Ez10/      ------T-PLLPQTPDLVPTYSKEERLLPL-RKGASKKKEDGSDC-QGRV
ChrX_139287533_1154_po_C_Xz8/4      ------A-PLLPQTPDLVPTYSKEEKDFLQ-AEGGQTIKERCIKLPDGRI
Chr2_52615786_912_po_C_Xz11/5       ------AHPLLPQARDLVPTYSEEEKDFLQ-AEGGQVIEEGZFQLLDRRI
HERV-E_4_1                          ------A-PLLPQAPDLVPTYSKEEKDFLQ-AEGGQVMEEGWIWLPDGRV
ChrY_21962399_1137_po_C_Xz13/9      ------A-PLLPQAPDIVPTYSKEEKDFLQ-AEGGQVIEEGGIWLLGRRI
ChrY_8634032_950_po_C_Xz14/10       ------I-PLLPQIPDLVSAYSKVEKDFFH-AEVGQVIKGGD-QTARFEG
Chr3_126735360_997_po_C_Xz5/10      ----------FPQAPDLGPTSSKEEKDFLQ-VEGRTSDGGRMDSVTRWET
Chr3_126713203_958_po_C_Xz7/8       ----------LPQAPDLGPTSSKEEKDFLQ-VEGRTSDGGRMDSVTRWET
Chr3_75226118_977_po_C_Xz8/9        ----------LPQAPDLGPASSKEEKDFLQ-VEGRTSDEGRMDSVTRWEV
Chr7_6629919_827_po_C_Xz8/7         ----------LPQAPDLGPTYSKEEKDFVQ-VEGRTSDGGRMNSVTRWES
Chr4_9166812_811_po_C_Xz8/10        ----------LPQAPDLEPTYSKEEKNFLQ-VEGRTSDGGRMDSVTDGRV
Chr12_8478642_834_po_C_Xz12/7       ----------LPQAPNLGPASSKEEKDFLQ-VEGRTSDEEGWTRLPEGRE
Chr10_15187175_787_po_C_Xz11/5      ----------LPQAPDLGPASSKEEKDCLQ-VEGRTSDGGGWIRLPHGRE
Chr11_71819740_854_po_C_Xz11/9      ----------LPQAPDLGPTSSNEEKDFLQ-VEGRTSDEEGWIRLADGRV
Chr11_71519974_886_po_C_Xz13/9      ----------LPQAPDLGPT-SKEEKDFLQ-VEGRTSDGGRMDSVSDGRV
Chr8_7271441_969_po_C_Xz10/6        ----------LPQAPDLGPASSKEERTFSRZREGQVMEEGWIQLPDR-RV
Chr8_11770066_826_po_C_Xz9/8        ----------LLQAPDLGPASSKEEKDFLQ-VEGRTSDGGRMDLVTRWEV
Chr4_4128362_651_po_C_Xz7/11        ----------LPQAPDLGPTSSKEEKDFLQ-VEGRTSDGGRMDSVTRWES
Chr14_50203955_850_po_C_Xz10/8      ------A-PLLPQAPDLGPAYSKEEKDFLQ-VEGRTSDGGRMDSVTRGRV
Chr14_50205291_837_po_C_Xz10/8      ------A-PLLPQAPDLGPAYSKEEKDFLQ-VEGRTSDGGRMDSVTRGRV
Chr7_6597120_842_po_C_Xz7/7         ------A-PLLPQAPDLGPAYSKGEKDFVQ-VEGRTSDGGRMDLVTGWES
Chr11_71544658_783_po_C_Xz8/13      ------A-PLLPQAPDLGPAYSKEEKDFLQ-VEGRTSDGGGWIRLPDGRV
Chr11_67999520_760_po_C_Xz8/8       ------A-PLLPQAPDLGPAYSKEEKDFVQ-VEGRTVMEEGWR------V
```

```
Chr3_75461975_796_po_C_Xz7/9        ----------FPQAPDLGCTYSKEEKDFLQ-VEGRTSDGGRWIRLPDGRV
Chr12_52228939_692_po_C_Xz3/10      ------A-TLLPQAPDLVPTYSKEEKDFLQ-VEGRTSDEEGWIRLPDGRV
Chr10_15211234_286_po_C_Xz11/1      ------A-PLLPQAPDLGPAYSKEZKDFLQ-VEGRTSDGGRMESVTRWER
Chr12_100973302_945_po_C_Xz12/      ------A-PLLPQTPELVPTYSREER--LLPCRRGQVIKGGWIRLPDGKV
Chr19_9685420_817_po_C_Xz6/10       ------A-PLLPQTPDLVLPILRKKK--TSSAEGGQVIKGGWNRLPDGRV
Chr13_40936820_719_po_C_Xz9/10      ------D--TAPSITVGVPHVCTPVILFVISMVNITSSNPQRVHPVIFFI
Chr7_63858202_1754_po_C_Rz12/9      ----ALAMPLLPEPPLQEVSSYSPNEKAVLPKKLENILKEDGGNF-DGRS
ERV3_chr7q21_AC073210_polputei      ----ALAMPLLPEPPLQEVSSYSPNEKAVLPKKLENILKEDGGNF-DGRS
Chr6_110218096_843_po_C_Xz16/8      ----ALAMALLPEPPLPVVLSYSPNAKAWFSRESGKYIERGGGDS-PVGD
Chr6_110214535_786_po_C_Xz16/8      ----ALAMALLPEPPLPVVLSYSPNAKAWFSRESGKYIERGGGDS-PVGD
ChrY_3018841_881_po_C_Rz9/12        ----ALVIPLLLKPPLGRFQVTLQVRRPG-AQESGKYIEGGWZKFSDRRL
ChrX_86876098_639_po_C_Xz7/8        ----ALVIPLLPEPPLGRFQVTLQVRRPG-AQESGKYIEGGWZKFSDGRL
Chr14_68899917_965_po_C_Xz6/8       ----ALAMPPLPEPPLLEVSSYSPNEKAWFAQESGKYIEVGWWKFSDGRL
ChrY_23874437_661_po_C_Xz15/9       ----AIADSPPKYAPLRGFKLP-PNEKLCFSQGSGKYIKRD-GEFSNRRL
ChrY_27518418_649_po_C_Xz15/9       ----AIADSPPKYAPLRGFKLP-PNEKLCFSQGSGKYIKRD-GEFSNRRL
ChrY_25016892_657_po_C_Xz13/12      ----AIADSPPKYAPLRGFKLP-PNEKLCFSQGSGKYIKRD-GEFSNRRL
ChrY_19932937_520_po_C_Xz10/14      ----AIAMPLLPKPLLSEVSSY-LQEKPCFAQESGKYIKCY-GENSPRRL
ChrY_20104614_518_po_C_Xz10/14      ----AIAMPLLPKPLLSEVSSY-LQEKPCFAQESGKYIKCY-GENSPRRL
Chr19_58660155_598_po_C_Xz21/1      ----ALAMPLLPEPPLQDNPSYTPNQRSWFAQEDKKYIKRRZZKFFDRRL
Chr19_58621153_500_po_C_Xz21/1      ----ALAMPLLPTSSPR-RPKLYSKSKILVCQEDIKCMKRKZKKFFYRRL
Chr7_139686987_710_po_C_Xz11/7      ----ALAMPLLPEPPLQEDPTYTPNERAWFAQEAGKYIKGGWZKLFDGKS
Chr7_137651301_1097_po_C_Pz13/      ----ALALPLLP-----EVPNYSPNERAWCAKQTGHYIEERWWKFSDERL
Chr11_29579197_930_po_C_Pz8/12      ----AIALPLLPEPPLLEVPNYSPNERAWFAKDTGSYIKGGWWKFSNRRL
Chr4_117329205_488_po_C_Xz3/16      ----ALTMPLLPEPPLSEIPSYSPNKKAWFAREIGKYTESMVEILT--RL
Chr21_38522185_383_po_C_Xz7/9       ----SLAMLLLPEPSLPEIPSYSPNEKAWFAQETGKCTEGGWWKFSDGRL
Chr4_70173370_1049_po_C_Xz11/8      ----ALAMPLLLEIPLPGTPSYTP-IKGLDLPRKMGITEGRWWKFSNGRL
Chr4_70022513_788_po_C_Kz10/12      ----ALAMTLLPEIPLPETPSYTQ-MKGLGLQENGNYVEGGZWKFSNGKL
Chr14_105225894_1027_po_C_Xz13      ----ALAMPLLPEIPLPETPSYTPNERAFFAQENGSYVEGGWWKFSNGRL
Chr7_119241536_761_po_C_Xz11/1      ----TLDMPLLPDPSLPEVSSILQMRRPGLPENLESTLKE-----NGGNS
ChrY_22564970_567_po_C_Xz9/10       ----ALAMPLFPEISLPEIPSYTLNERSCFAQENGNYTEGGZWKFSDGRL
Chr21_14409346_1477_po_C_Pz5/6      ----ETYVMPSLEPPLAETPNYSSSEKVWFTQENGCYQKGGWGSFQMGRR
ChrX_51259700_1396_po_C_Pz7/8       ----ETYVMPLLEPPLAETPNYTSSEKAWFAQENGDSKKGGWWKFSDGRL
Chr6_64985396_1041_po_C_Xz11/1      ----ETLTMPLLELPLTEPTNYSSNEKA-FEQDSEVTRKEVGGSSR-WEA
ChrY_20355209_659_po_C_Xz7/11       ----ETLTMPLLELSLMEPPKYSFNEKASFEQESVSYQKGGCQKFR-WEV
Chr3_34567166_523_po_C_Xz8/6        ----ETLTTPLLELPLTEPPNYSSNEKAWFEQESRSYQKGSLWKFSDGRL
Chr3_122072468_245_po_C_Xz9/9       ----ETLTMPLLELPLTELPNYSSNEKASFEQESGSYQKGCZWKFSDGRL
Chr11_58788475_541_po_C_Xz8/22      ----KTLTMPLLELPLTEPPNYSLMKRLSLRGEWKLPER-RLVEVLRGRL
Chr6_137421042_373_po_C_Xz6/9       ----ETLTMPLLELPLTEPPNYSSNEKAWFEQESGSYQKGSWWKFSNGRL
Chr4_115910608_550_po_C_Pz8/12      ----ETLTMPLLELPLTEPPNYSANEKAWFEQKNGIYQKGGWWKFSDGRL
ChrY_13743185_977_po_C_Xz14/8       ----TPAMPLLIEPPYLRYPITLZVKKLGFITKQETYIKDSWWLFSDGRL
HERV15                              ----TPAMPLLIEPPLLEYPITLQVKKLVFIRKQETYIKDGWWLFSDGRL
ChrX_117198092_983_po_C_Xz8/8       ----TLAMPLLIEPPLLEVPPYSSSEKACFVLETGKYIKGGWWLFSDGRL
Chr19_21166030_1412_po_C_Xz9/1      ----TLAMPLFVEPPLPEVPNYSSIEKAWFGQETGKYIKGGWWLFSDGRL
Chr19_21213743_913_po_C_Xz14/1      ----TLAMPLLIEPPLPEVPNYSLSEKAWFRQETGKYIKSGWWLFSDGRL
Chr19_15957774_569_po_C_Xz8/10      ------WKDLPSLRRWIWA-------------------------------
Chr11_77508681_297_po_C_Xz19/1      ----AVSKENQELAKYRHLZNIIMLKKPGKEVTVKEFVFLFFFFLQPCSG
Chr13_98250769_738_po_C_Xz7/13      ----LLWKTNQEIAFKZIKEALTQAPALGLPDITKPFFLZ----------
Chr6_39968183_513_po_C_Pz10/9       -------HTHKKGYRYRRLKK-----------------------------
Chr15_66867263_645_po_C_Xz11/1      ----TLLZSPQKLTSLHTAEAZ----------------------------
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                --------NKLTGCLKKGMFATSQKVEACQRIVCKSREKKVIKGFMQEML
MER66_polputein                     ----------GFHLNSXTTITWTWEVGYLYLQCTNDSIFCTHCCVNDTKA


HERVI_polputein                     VIQARYV-------------------------------------------
HERVIP10F_polputein                 LSKPLMR-------------------------------------------
HERVadp_AC005741                    MIKPLMR-------------------------------------------
HERV-PT47D                          --CTRVN-------------------------------------------
HERV_W_chr6_141432567_ERV9_lik      EKEGKRERG-----------------------------------------
ERV9_PH1_RT                         VTYSFKSQG-----------------------------------------
Chr13_54564403_572_po_C_Yz7/14      DIEAKRAAR-----------------------------------------
HERV_Y_chr12_51022911_pol           TFLPTANQW-----------------------------------------
HERV17_polputein                    ILLPAANQW-----------------------------------------
HERV_W_chr7_9105739_syncytin_p      ILLPATIQW-----------------------------------------
HERV9_polputein                     VLIPEASQW-----------------------------------------
HUERS-P3_polputein                  STCQ-PPQG-----------------------------------------
ERVfrd_AC004022                     IYLPMESLQGKAFHWEVDGEIGGEMKKLAGRWRKVWGILVDAELVVGVCL
PRIMA41_polputein                   LLTPGANQW-----------------------------------------
HERV_H_RGH1_pol                     YLLPASQAH-----------------------------------------
HERVHRGH2_pol                       YLLPASQAH-----------------------------------------
HERV_H48I_polputein                 FVLPHSQTI-----------------------------------------
HERV_H_RTVLH2_pol                   IVQATSPP------------------------------------------
HERVFH21_pol                        ARPPPEQTK-----------------------------------------
HERVRblike_chr4_109047953           FLHPLLTSFYNGHRTMQTASSMIPVRSVACYPFLAPPWWVSPMQGKDWRH
ChrY_13790029_693_po_C_Xz12/11      F-IPRALKE-----------------------------------------
HERVS71_polputein                   F-IAPSPLE-----------------------------------------
BaEV_M7_pol                         V-LPQKEAL-----------------------------------------
```

```
HERVR_polputein                     V-LPQKEAL-----------------------------------------
AKR_MLV_MLOCG_RT                    V-MPDQFTF-----------------------------------------
GaLV_pol                            --------------------------------------------------
Chr3_158587414_1662_po_C_Ez10/      A-VPQLLGA-----------------------------------------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      A-VLQLLGA-----------------------------------------
Chr5_111890459_1626_po_C_Ez10/      A-MAQLLGA-----------------------------------------
ChrX_48766518_1600_po_C_Ez6/3       A-VPQLLGA-----------------------------------------
ChrX_152580713_1572_po_C_Ez10/      A-VPQLLGA-----------------------------------------
ChrX_139287533_1154_po_C_Xz8/4      A-VPQLLGA-----------------------------------------
Chr2_52615786_912_po_C_Xz11/5       V-VPQLLGV-----------------------------------------
HERV-E_4_1                          A-VPQLLGA-----------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      A-MPQMLGA-----------------------------------------
ChrY_8634032_950_po_C_Xz14/10       S-SATVAGS-----------------------------------------
Chr3_126735360_997_po_C_Xz5/10      S-CATAARS-----------------------------------------
Chr3_126713203_958_po_C_Xz7/8       S-CATAARS-----------------------------------------
Chr3_75226118_977_po_C_Xz8/9        A-VPQLLGA-----------------------------------------
Chr7_6629919_827_po_C_Xz8/7         S-CATAARS-----------------------------------------
Chr4_9166812_811_po_C_Xz8/10        A-APQLLGA-----------------------------------------
Chr12_8478642_834_po_C_Xz12/7       A-VPQLLGA-----------------------------------------
Chr10_15187175_787_po_C_Xz11/5      A-VPQLLGA-----------------------------------------
Chr11_71819740_854_po_C_Xz11/9      A-VPQLLGA-----------------------------------------
Chr11_71519974_886_po_C_Xz13/9      A-VPQLLGA-----------------------------------------
Chr8_7271441_969_po_C_Xz10/6        A-VPQLLGA-----------------------------------------
Chr8_11770066_826_po_C_Xz9/8        A-VPQLLGA-----------------------------------------
Chr4_4128362_651_po_C_Xz7/11        A-APQLLGA-----------------------------------------
Chr14_50203955_850_po_C_Xz10/8      A-VPQLRGA-----------------------------------------
Chr14_50205291_837_po_C_Xz10/8      A-VPQLRGA-----------------------------------------
Chr7_6597120_842_po_C_Xz7/7         A-VPQLPGA-----------------------------------------
Chr11_71544658_783_po_C_Xz8/13      A-VPQLLGA-----------------------------------------
Chr11_67999520_760_po_C_Xz8/8       A-APQLLGA-----------------------------------------
Chr3_75461975_796_po_C_Xz7/9        A-VPQLPGA-----------------------------------------
Chr12_52228939_692_po_C_Xz3/10      A-VPQLLGA-----------------------------------------
Chr10_15211234_286_po_C_Xz11/1      S-CATAVRN-----------------------------------------
Chr12_100973302_945_po_C_Xz12/      A-VPQLLGA-----------------------------------------
Chr19_9685420_817_po_C_Xz6/10       A-VLQLMGA-----------------------------------------
Chr13_40936820_719_po_C_Xz9/10      I-SZGDIAA-----------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      A-IPKMVAP-----------------------------------------
ERV3_chr7q21_AC073210_polputei      A-IPKMVAP-----------------------------------------
Chr6_110218096_843_po_C_Xz16/8      A-IPEMVAP-----------------------------------------
Chr6_110214535_786_po_C_Xz16/8      A-IPEMVAP-----------------------------------------
ChrY_3018841_881_po_C_Rz9/12        A-IPEMLAP-----------------------------------------
ChrX_86876098_639_po_C_Xz7/8        A-IPEMLAP-----------------------------------------
Chr14_68899917_965_po_C_Xz6/8       A-IPEMVAS-----------------------------------------
ChrY_23874437_661_po_C_Xz15/9       A-NHEMLAP-----------------------------------------
ChrY_27518418_649_po_C_Xz15/9       A-NHEMLAP-----------------------------------------
ChrY_25016892_657_po_C_Xz13/12      A-NHEMLAP-----------------------------------------
ChrY_19932937_520_po_C_Xz10/14      A-NHETVAL-----------------------------------------
ChrY_20104614_518_po_C_Xz10/14      A-NHETVAL-----------------------------------------
Chr19_58660155_598_po_C_Xz21/1      A-IPKMLAP-----------------------------------------
Chr19_58621153_500_po_C_Xz21/1      A-IPKMLAP-----------------------------------------
Chr7_139686987_710_po_C_Xz11/7      A-IPEMLAP-----------------------------------------
Chr7_137651301_1097_po_C_Pz13/      A-IPEIVTP-----------------------------------------
Chr11_29579197_930_po_C_Pz8/12      A-IPETVAP-----------------------------------------
Chr4_117329205_488_po_C_Xz3/16      A-ISEMVAP-----------------------------------------
Chr21_38522185_383_po_C_Xz7/9       A-IPEMVAP-----------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      A-IPEILAP-----------------------------------------
Chr4_70022513_788_po_C_Kz10/12      A-ILEMLAP-----------------------------------------
Chr14_105225894_1027_po_C_Xz13      A-ITEMVAP-----------------------------------------
Chr7_119241536_761_po_C_Xz11/1      P-MGDZLSL-----------------------------------------
ChrY_22564970_567_po_C_Xz9/10       A-IPEMMTP-----------------------------------------
Chr21_14409346_1477_po_C_Pz5/6      A-IPEAIAP-----------------------------------------
ChrX_51259700_1396_po_C_Pz7/8       A-IPEAVAP-----------------------------------------
Chr6_64985396_1041_po_C_Xz11/1      R-LPIAPPP-----------------------------------------
ChrY_20355209_659_po_C_Xz7/11       Y-HTRSNTP-----------------------------------------
Chr3_34567166_523_po_C_Xz8/6        A-IPETIAP-----------------------------------------
Chr3_122072468_245_po_C_Xz9/9       A-IAEAIAP-----------------------------------------
Chr11_58788475_541_po_C_Xz8/22      A-IPEAIAL-----------------------------------------
Chr6_137421042_373_po_C_Xz6/9       ASYPRSNSP-----------------------------------------
Chr4_115910608_550_po_C_Pz8/12      A-IPEAIAPPP---------------------------------------
ChrY_13743185_977_po_C_Xz14/8       A-IPETIAP-----------------------------------------
HERV15                              A-IPETIAP-----------------------------------------
ChrX_117198092_983_po_C_Xz8/8       D-VPETIAS-----------------------------------------
Chr19_21166030_1412_po_C_Xz9/1      A-VPETIAP-----------------------------------------
Chr19_21213743_913_po_C_Xz14/1      A-IPETIAP-----------------------------------------
Chr19_15957774_569_po_C_Xz8/10      --------------------------------------------------
Chr11_77508681_297_po_C_Xz19/1      GEGCLQPIPLP---------------------------------------
```

```
Chr13_98250769_738_po_C_Xz7/13      --------------------------------------------------
Chr6_39968183_513_po_C_Pz10/9       --------------------------------------------------
Chr15_66867263_645_po_C_Xz11/1      --------------------------------------------------
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                YNLKVLGLLN----------------------------------------
MER66_polputein                     EVSLRCSDP-----------------------------------------


HERVI_polputein                     --------------------------------------------------
HERVIP10F_polputein                 --------------------------------------------------
HERVadp_AC005741                    --------------------------------------------------
HERV-PT47D                          --------------------------------------------------
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      --------------------------------------------------
HERV_Y_chr12_51022911_pol           --------------------------------------------------
HERV17_polputein                    --------------------------------------------------
HERV_W_chr7_9105739_syncytin_p      --------------------------------------------------
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  --------------------------------------------------
ERVfrd_AC004022                     VSLLVRVYSESTLQTVLGRAPISIPFFFSAIMQIGAIPIKVILEKMGPKM
PRIMA41_polputein                   --------------------------------------------------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 --------------------------------------------------
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           LQVFLDLKHWTGSQMMG--------------VTKANVQNGCTTFNDPGHG
ChrY_13790029_693_po_C_Xz12/11      --------------------------------------------------
HERVS71_polputein                   --------------------------------------------------
BaEV_M7_pol                         --------------------------------------------------
HERVR_polputein                     --------------------------------------------------
AKR_MLV_MLOCG_RT                    --------------------------------------------------
GaLV_pol                            --------------------------------------------------
Chr3_158587414_1662_po_C_Ez10/      --------------------------------------------------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      --------------------------------------------------
Chr5_111890459_1626_po_C_Ez10/      --------------------------------------------------
ChrX_48766518_1600_po_C_Ez6/3       --------------------------------------------------
ChrX_152580713_1572_po_C_Ez10/      --------------------------------------------------
ChrX_139287533_1154_po_C_Xz8/4      --------------------------------------------------
Chr2_52615786_912_po_C_Xz11/5       --------------------------------------------------
HERV-E_4_1                          --------------------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      --------------------------------------------------
ChrY_8634032_950_po_C_Xz14/10       --------------------------------------------------
Chr3_126735360_997_po_C_Xz5/10      --------------------------------------------------
Chr3_126713203_958_po_C_Xz7/8       --------------------------------------------------
Chr3_75226118_977_po_C_Xz8/9        --------------------------------------------------
Chr7_6629919_827_po_C_Xz8/7         --------------------------------------------------
Chr4_9166812_811_po_C_Xz8/10        --------------------------------------------------
Chr12_8478642_834_po_C_Xz12/7       --------------------------------------------------
Chr10_15187175_787_po_C_Xz11/5      --------------------------------------------------
Chr11_71819740_854_po_C_Xz11/9      --------------------------------------------------
Chr11_71519974_886_po_C_Xz13/9      --------------------------------------------------
Chr8_7271441_969_po_C_Xz10/6        --------------------------------------------------
Chr8_11770066_826_po_C_Xz9/8        --------------------------------------------------
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      --------------------------------------------------
Chr14_50205291_837_po_C_Xz10/8      --------------------------------------------------
Chr7_6597120_842_po_C_Xz7/7         --------------------------------------------------
Chr11_71544658_783_po_C_Xz8/13      --------------------------------------------------
Chr11_67999520_760_po_C_Xz8/8       --------------------------------------------------
Chr3_75461975_796_po_C_Xz7/9        --------------------------------------------------
Chr12_52228939_692_po_C_Xz3/10      --------------------------------------------------
Chr10_15211234_286_po_C_Xz11/1      --------------------------------------------------
Chr12_100973302_945_po_C_Xz12/      --------------------------------------------------
Chr19_9685420_817_po_C_Xz6/10       --------------------------------------------------
Chr13_40936820_719_po_C_Xz9/10      --------------------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      --------------------------------------------------
ERV3_chr7q21_AC073210_polputei      --------------------------------------------------
Chr6_110218096_843_po_C_Xz16/8      --------------------------------------------------
Chr6_110214535_786_po_C_Xz16/8      --------------------------------------------------
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        --------------------------------------------------
Chr14_68899917_965_po_C_Xz6/8       --------------------------------------------------
ChrY_23874437_661_po_C_Xz15/9       --------------------------------------------------
ChrY_27518418_649_po_C_Xz15/9       --------------------------------------------------
```

```
ChrY_25016892_657_po_C_Xz13/12      --------------------------------------------------
ChrY_19932937_520_po_C_Xz10/14      --------------------------------------------------
ChrY_20104614_518_po_C_Xz10/14      --------------------------------------------------
Chr19_58660155_598_po_C_Xz21/1      --------------------------------------------------
Chr19_58621153_500_po_C_Xz21/1      --------------------------------------------------
Chr7_139686987_710_po_C_Xz11/7      --------------------------------------------------
Chr7_137651301_1097_po_C_Pz13/      --------------------------------------------------
Chr11_29579197_930_po_C_Pz8/12      --------------------------------------------------
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      --------------------------------------------------
Chr4_70022513_788_po_C_Kz10/12      --------------------------------------------------
Chr14_105225894_1027_po_C_Xz13      --------------------------------------------------
Chr7_119241536_761_po_C_Xz11/1      --------------------------------------------------
ChrY_22564970_567_po_C_Xz9/10       --------------------------------------------------
Chr21_14409346_1477_po_C_Pz5/6      --------------------------------------------------
ChrX_51259700_1396_po_C_Pz7/8       --------------------------------------------------
Chr6_64985396_1041_po_C_Xz11/1      --------------------------------------------------
ChrY_20355209_659_po_C_Xz7/11       --------------------------------------------------
Chr3_34567166_523_po_C_Xz8/6        --------------------------------------------------
Chr3_122072468_245_po_C_Xz9/9       --------------------------------------------------
Chr11_58788475_541_po_C_Xz8/22      --------------------------------------------------
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------------------
ChrY_13743185_977_po_C_Xz14/8       --------------------------------------------------
HERV15                              --------------------------------------------------
ChrX_117198092_983_po_C_Xz8/8       --------------------------------------------------
Chr19_21166030_1412_po_C_Xz9/1      --------------------------------------------------
Chr19_21213743_913_po_C_Xz14/1      --------------------------------------------------
Chr19_15957774_569_po_C_Xz8/10      --------------------------------------------------
Chr11_77508681_297_po_C_Xz19/1      --------------------------------------------------
Chr13_98250769_738_po_C_Xz7/13      --------------------------------------------------
Chr6_39968183_513_po_C_Pz10/9       --------------------------------------------------
Chr15_66867263_645_po_C_Xz11/1      --------------------------------------------------
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                --------------------------------------------------
MER66_polputein                     --------------------------------------------------


HERVI_polputein                     ----EDLVSTYLRTHWGSQAMCDAVLR--------VYGCIRIYTLAKQVT
HERVIP10F_polputein                 ----EVLSQLHQGTHWGPQAMCDAVLR--------VYGCIGIYTLAKQVT
HERVadp_AC005741                    ----ELMSILHKGSHWGPQALCDAILR--------NYVCIGIZTLTKQVC
HERV-PT47D                          ----DIFPSYLQPRKQNSHTRNAHTIR--------IYIZ-----------
HERV_W_chr6_141432567_ERV9_lik      ----RDRQREVRKRETKK---KSKRKSG--------SKEKTVY-------
ERV9_PH1_RT                         ----KFLSTHARQFSIGS---QPISASQ--------TVEEIMKSILTLQS
Chr13_54564403_572_po_C_Yz7/14      ----QDLPLEMLIEDPZYGVIPSRKPSP--------STQKKKZNEEPHED
HERV_Y_chr12_51022911_pol           ----ILLKILHQTFHLGIDSTIKMAKSL--------FTGAGLVKTIKQIV
HERV17_polputein                    ----KLLKTLHQTFHLGIDSTHQMAKSL--------FTGPGLFKTIKQIV
HERV_W_chr7_9105739_syncytin_p      ----KLLKTLHQTFHLGIDSTHQMAKSL--------FTGPGLFKTIKQIV
HERV9_polputein                     ----KILKTLHQTFHMGIENTHQMATSL--------FTGPNLLWAFQQVV
HUERS-P3_polputein                  ----K-LKILHK-LHLGKDKTYQCAQRL--------FSGENLLKTVKQVV
ERVfrd_AC004022                     EPPNGVIKSFYNFFHLAKDSLGQICKZV--------FSGKGLNKTIQQVC
PRIMA41_polputein                   ----KIVKHLHDSTHLGRDSLFQLMSQL--------FIGKGLLKTVKQVT
HERV_H_RGH1_pol                     ----SILSSFHNLFHVGYKPLDRLLEPL--------ISFPSWKSILKEIT
HERVHRGH2_pol                       ----SILSSFHNLFHVGYKPLACLLGPL--------ISFPSWKSILKEIT
HERV_H48I_polputein                 ----PHLQSLHNSFHVGYKPLLQLLRPI--------LTCPHLSSRVREIT
HERV_H_RTVLH2_pol                   ------------------------LRTL--------ISFPSWKSIPKEIT
HERVFH21_pol                        ----EILTSLHQSFHIGARPLYLLLCPY--------FSXPXLFTSLKDIT
HERVRblike_chr4_109047953           SFLHPILKYIHDGTTFGQDASLTSYRSI--------RKG--MRAHLESII
ChrY_13790029_693_po_C_Xz12/11      ----TLISRLHSTTHLR-GVKLAQLLRSHFR------SLRLQNSANQAAL
HERVS71_polputein                   ----TLINHLHSATHLG-GIKLAQLLRSHFN------IPHLQDLTNQAAL
BaEV_M7_pol                         ----AMIQQMHAWTHLG-NRKLKLLIEKTD-----FLIPRASTLIEQVTS
HERVR_polputein                     ----AMIQQMHAWTHLG-NRKLKLLIEKTD-----FLIPRASTLIEQVTS
AKR_MLV_MLOCG_RT                    ----ELLDFLHQLTHLS-FSKMKALLERSHSP---YYMLNRDRTLKNITE
GaLV_pol                            ----EFIKRLHQLTHLG-PEKLLQLVNRTS-----LLIPNLQSAVREVTS
Chr3_158587414_1662_po_C_Ez10/      ----TLIGQARNHASR--QESLEKLLG------RLLSLTLASPCQSSSTT
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      ----TIVLAMHETTHLG-QESFEKLLA------GTSTSHASQPLPKQYQT
Chr5_111890459_1626_po_C_Ez10/      ----TIILAMHETTHLG-QESHEKLLS------CYFYISYLPALAKAVAQ
ChrX_48766518_1600_po_C_Ez6/3       ----TIVLAMHETTHPG-QESLEKLLG------RYFYISQLPALAKAVTQ
ChrX_152580713_1572_po_C_Ez10/      ----TIVLAMHETTHLA-QESLEKLLG------QYFYVSHLPALAKAVAQ
ChrX_139287533_1154_po_C_Xz8/4      ----AVVLAVHETTHVG-QESLEKLZS------RYFYISHLPALAKTVAQ
Chr2_52615786_912_po_C_Xz11/5       ----AVVLAVREITHLG-QE----LLG------QCFYVWHLSALAKTVVR
HERV-E_4_1                          ----AVVLAVHKTTHLG-QESLEKLLG------WYFYISHLSALAKTVTQ
ChrY_21962399_1137_po_C_Xz13/9      ----AVILAVHETIHLG-QMSFEKLLG------WYFYILHLSALAKTVVQ
ChrY_8634032_950_po_C_Xz14/10       ----HSRIEHAQNTHLG-QESLEELLG------QYFYISHLPALAKAVAR
Chr3_126735360_997_po_C_Xz5/10      ----CSCTGCARN-HPL-RSGVTKFVR------LVFLH--LASVSPCHNG
```

```
Chr3_126713203_958_po_C_Xz7/8       ----CSCTGCVRTTHRG-QESLEKLLG------WYFFISPLSALAKTVRQ
Chr3_75226118_977_po_C_Xz8/9        ----AVVLAVQETTHRG-QQSLEKLLG------RYFYISSFSALAKTVRQ
Chr7_6629919_827_po_C_Xz8/7         ----CSVLAVQETSHRG-QESLEKLLG------WYFSISPLSALAKTVRQ
Chr4_9166812_811_po_C_Xz8/10        ----AVVLAVQETTHRG-QESLEKLLG------RYFYISPLSALAKTVRQ
Chr12_8478642_834_po_C_Xz12/7       ----AVVLAVQETTHRG-QESLEKLLG------RYFYITPWSALAKTVKQ
Chr10_15187175_787_po_C_Xz11/5      ----AVVLAVQETTHRG-QESLEKLLG------RYFYISPLSALAKTVRQ
Chr11_71819740_854_po_C_Xz11/9      ----AVVLAVQETTHRG-QESLEKLLG------RYFYISPLSALAKTVRQ
Chr11_71519974_886_po_C_Xz13/9      ----AVVLAVQETTHRG-QESLEKLLG------WYFYISPLSALAKTVRQ
Chr8_7271441_969_po_C_Xz10/6        ----AVVLAVQENTHRG-QESLEKLLG------RYFYISPLSTLAKTVRQ
Chr8_11770066_826_po_C_Xz9/8        ----AVVLAVQETTHRG-QESLEKLLG------WYFYISPLSALAKTVRQ
Chr4_4128362_651_po_C_Xz7/11        ----AVVLAVLHLTFVS-P-------------------------AKTVRQ
Chr14_50203955_850_po_C_Xz10/8      ----AVVLAVQETTHLG-QESLKKLLG------QYFYISPFSALAKTVRQ
Chr14_50205291_837_po_C_Xz10/8      ----AVVLAVQETTHLG-QESLKKLLG------QYFYISPFSALAKTVRQ
Chr7_6597120_842_po_C_Xz7/7         ----AVVLAVQETTHLG-QESLKKLLG------WYFSISPLSALAKTVRQ
Chr11_71544658_783_po_C_Xz8/13      ----AVVLAVQETTHLG-QESLQKVVR------PVVLHSPFLALAKTVRQ
Chr11_67999520_760_po_C_Xz8/8       ----AVVLAVQETTYRG-QESLEKLLG------RYFYISRFSALAKTVRQ
Chr3_75461975_796_po_C_Xz7/9        ----AVVLAVQETTHVG-QESLEKLLA------RYFYILRFPALAKTVRQ
Chr12_52228939_692_po_C_Xz3/10      ----AVVLAVHETTHLS-ES-LEKLLG------WYFYISHLSALAKTVTQ
Chr10_15211234_286_po_C_Xz11/1      ----CSCTG----THLC-QESLEKLLG------WYFYISHLSALAKTVRQ
Chr12_100973302_945_po_C_Xz12/      ----TIVLAMHETTHLG-QESLEKLLG------RYFYMSHLPALQTSS-I
Chr19_9685420_817_po_C_Xz6/10       ----TIVLAMHETTHLG-QESLEKLLG------RYFYISHLPALKAVA-Q
Chr13_40936820_719_po_C_Xz9/10      ----NNTVGVHHVCTLC-D-MIAYILG------RYSFZYHSECTPCDIIH
Chr7_63858202_1754_po_C_Rz12/9      ----KFVKZLHH-ELVG-KTALETLLG------RHFYVPRLTAITQAVCE
ERV3_chr7q21_AC073210_polputei      ----KFVKZLHH-ELVG-KTALETLLG------RHFYVPRLTAITQAVCE
Chr6_110218096_843_po_C_Xz16/8      ----KFVZQFHQGNHMG-ETALETLLG------HRFYGPRLTAITZAICE
Chr6_110214535_786_po_C_Xz16/8      ----KFVZQFHQGNHMG-ETALETLLG------HRFYGPRLTAITZAICE
ChrY_3018841_881_po_C_Rz9/12        ----KFIRQFHZGTHMG-KMALETLLG------HHFYVPRLTAITQAICK
ChrX_86876098_639_po_C_Xz7/8        ----KFVRQFHZGTHMG-KMALETLLG------RHFYVPRLTAITQAICK
Chr14_68899917_965_po_C_Xz6/8       ----KFVRQFHQGTHME-KTALEMLLG------HHFYVLRLTAITQAVCE
ChrY_23874437_661_po_C_Xz15/9       ----KFVKQFHQGTHMG-KMSLKTLLG------HHFYVPQLTVITZSVCK
ChrY_27518418_649_po_C_Xz15/9       ----KFVKQFHQGTHMG-KMSLKTLLG------HHFYVPQLTVITZSVCK
ChrY_25016892_657_po_C_Xz13/12      ----KFVKQFHQGTHMG-KMSLKTLLG------HHFYVPQLCHHL-IVCK
ChrY_19932937_520_po_C_Xz10/14      ----KFVRQFYQGTHMG-KTSLKTLLG------HHVYVLRLPAITZSICK
ChrY_20104614_518_po_C_Xz10/14      ----KFVRQFYQGTHMG-KTSLKTLLG------HHVYVLRLPAITZSICK
Chr19_58660155_598_po_C_Xz21/1      ----KFVKQTHQETHIR-KTALNVMKVPFLT-KVPFYVLWLSAVIZAICK
Chr19_58621153_500_po_C_Xz21/1      ----KFVKQIHQENHIR-KTALATLCS--AN-ESAFYVPWLSAVTQDVCE
Chr7_139686987_710_po_C_Xz11/7      ----KFVKQIHQGTHMK-KRALKMLLK------CHLYVPRVSAITZAVCK
Chr7_137651301_1097_po_C_Pz13/      ----KFVRQFHQGTHIIRKTALERLLG------HHFYVLWLTAITRAVCK
Chr11_29579197_930_po_C_Pz8/12      ----RFVKQFHQGTHIG-KTALETLLE------HHFYMPRLTTIPQAVCE
Chr4_117329205_488_po_C_Xz3/16      ----KFVKEFHQGTHMG-KMALEMLLR------S--FLSFLCAMTHCHHP
Chr21_38522185_383_po_C_Xz7/9       ----KFVKQFHQETHMG-KTALQTLLG------CHLYVPRLNALTQAVCR
Chr4_70173370_1049_po_C_Xz11/8      ----RFAKQF-QGTHMR-KTVLETZLG------CHFYVQWLTAITZAICK
Chr4_70022513_788_po_C_Kz10/12      ----RFVKQFHQGTHMG-KMALETLLG------HHFYVPQLTAITQAVCK
Chr14_105225894_1027_po_C_Xz13      ----RFVKQFHQGTHMG-KTALETLLG------CHFYVSQLTASLEQFVN
Chr7_119241536_761_po_C_Xz11/1      ----KWEFQVCKTIPLR--NSYGENGT------RNYYVPWLTAITQDICK
ChrY_22564970_567_po_C_Xz9/10       ----RFVNQFHQGTHMG--RQLETLLG------WHFYVQRLSPITZAICE
Chr21_14409346_1477_po_C_Pz5/6      ----QFIMQFHQGTHMG-KTALEILVG------QYFYVPCLTAITRAICE
ChrX_51259700_1396_po_C_Pz7/8       ----RLIKQLHQGTQIR-KTALETLVG------ZYFFVPHLTAVTGAICE
Chr6_64985396_1041_po_C_Xz11/1      ----AQFIKFHQGMHMG-KTALETLVG------QHFYVLHLTAIAQAICE
ChrY_20355209_659_po_C_Xz7/11       ----RFIKQFHHGMHVR-KTALETLAG------QPFYVPRLTAITRAICE
Chr3_34567166_523_po_C_Xz8/6        ----QFIKQFHQETHMG-KTALETRVG------WHFYVPCLTAITRALCE
Chr3_122072468_245_po_C_Xz9/9       ----RFIKQFHQGKHMG-KTALETLIG------RHFYVLLPTAITQAVCE
Chr11_58788475_541_po_C_Xz8/22      ----WFIKQFYQGTHMG-KIALETLVG------WHFYVPHLTAITGAVSE
Chr6_137421042_373_po_C_Xz6/9       ----WFIKQFHQGTHMG-KTALESCR-------MAFLLLSLTAITRAIYE
Chr4_115910608_550_po_C_Pz8/12      ----QFIKQFHQGTHME-KTALETLVG------WHFYVPRLIAITRTICE
ChrY_13743185_977_po_C_Xz14/8       ----RFVKQIHQGTHIG-RTALEILIG------QHFYVPWLSAITHAVCE
HERV15                              ----RFVKQIHQGTHIG-RTALETLIG------QHFYVPWLSAITHAVCE
ChrX_117198092_983_po_C_Xz8/8       ----RFVKQIHQRTHTG----LETLIG------QHFYVPRLSAITHAVCE
Chr19_21166030_1412_po_C_Xz9/1      ----RFVKQIHQGTHAG-STALETLIG------QHFYVPRLSAITRAVCE
Chr19_21213743_913_po_C_Xz14/1      ----RFVKQIHQGTHIG-RMALETDRD------TSMCHSSLQPPVLFMNN
Chr19_15957774_569_po_C_Xz8/10      ----------------ETSPKSWG--------------------------
Chr11_77508681_297_po_C_Xz19/1      ----WLLKQHYLITVLEVSLLGRGKLVLFFL-TLALPVTFLGVNECIFIF
Chr13_98250769_738_po_C_Xz7/13      --------------VHERKGMALEVLTQVIR-LWHCPVAYLSKQLDSVAL
Chr6_39968183_513_po_C_Pz10/9       ----------------EWPP------------------------------
Chr15_66867263_645_po_C_Xz11/1      -----------GVIIKZVNVAGN---------------------------
HERV18_polputein                    ----DVALWLHRKLGHAGGKLMQQVNK--------CWGLSLPTQDIWEAC
HERVIP10FH_polputein                ------VKLLRNSLCARCIRKVKYTFG-----------------KRIIRR
MER66_polputein                     ----TLVAKFPRPQQGKWDSTCKRGDH--------IWCLLLDQNIQGKNN


HERVI_polputein                     DSCLICKKTSKQILRKPPLGK---RFR-AKTISKCS----INYT-----E
HERVIP10F_polputein                 DSCLVCKKTNRHTIKRLPLGG---RNPGLRPFQSIQ----VDYT-----E
HERVadp_AC005741                    GSCVICQSLNKMLARKQAMGG---RPPGLRLFQSIQ----VDFR-----E
HERV-PT47D                          --CTACPIINVZLGKPCRRGG---RWGNPLSTYQAL----ILDN-----H
HERV_W_chr6_141432567_ERV9_lik      ----------PIPLKARATPS---KSG-----------------------
ERV9_PH1_RT                         QIRLFGSSDSPKPLGLDLLTA---EKGGLCTFLGEE----CCFY-----T
Chr13_54564403_572_po_C_Yz7/14      IVSSPQDGQPPKKEKYLCLQL---TSRNYLKPFTKP----FTLA-----L
```

```
HERV_Y_chr12_51022911_pol           RACEVCQRNNPLPYCZAPSGE---QRTGHYPGEGWQ----LDFT-----H
HERV17_polputein                    RACEVCQRNNPLPYRQAPSGE---QRTGHYPGEDWQ----LDFT-----H
HERV_W_chr7_9105739_syncytin_p      RACEVCQRNNPLPYRQAPSGE---QRTGHYPGEDWQ-----DFT-----H
HERV9_polputein                     KTCEVCQRNNPLVHCKAHLGE---QRIGHYPGEDWQ----LDFT-----I
HUERS-P3_polputein                  NACEVCLKNNPLNRRLLPPQT---QRMESYLGEDWQ----IDFT-----H
ERVfrd_AC004022                     QAYTLCTINNSQRGKPPPLISP-IQRRSTYPGEDWQ----MDFT-----H
PRIMA41_polputein                   QACELCARNNPNNQSLPPPLVRPVQHRGTYPGEDWQ----IDYT-----Q
HERV_H_RGH1_pol                     SQCSIA-STTPQGLFRPPPFP-THQARGFAPAQDWQ----IDFT-----H
HERVHRGH2_pol                       SQCSICYSATPQGLFRPPPFP-THKARGFAPAQDWQ----IDFT-----H
HERV_H48I_polputein                 QSCSICHSVSPQGSLRPPPFP-THQARGQVPGQDWQ----VDFT-----H
HERV_H_RTVLH2_pol                   SQCSICYSTNPQGLFRPPPFP-THQAZGFAPTQDWQ----ISFTZ----H
HERVFH21_pol                        SNCHICSVTSSQGALXSPSIP-THQLRGTLPGEDWQ----IDFT-----H
HERVRblike_chr4_109047953           LRCHFCTKNEPNNHNTGQPGQ---QGRGKQPLENWQ----IDFTQ----M
ChrY_13790029_693_po_C_Xz12/11      W-CMACAQVNAKQGPKPISG---HCLRGNSPGKRZE----IDFT--EIKP
HERVS71_polputein                   W-CMDCAQVNTKQVLSPAQ----STLQGGSPRERWE----VDFT--EIKP
BaEV_M7_pol                         A-CKVCQQVNAGATRVPAG----KRTRGNRPGVYWE----IDFT--EVKP
HERVR_polputein                     A-CKVCQQVNAGATRVPAG----KRTRGNRPGVYWE----IDFT--EVKP
AKR_MLV_MLOCG_RT                    T-CKACAQVNASKSAVKQG----TRVRGHRPGTHWE----IDFT--EIKP
GaLV_pol                            Q-CQACAMTNAVTTYRETG----KRQRGDRPGVYWE----VDFT--EIKP
Chr3_158587414_1662_po_C_Ez10/      VCYLLTAQCEAR--PHFPPS---IQVYGAASFEDLHRNAFEDFT--EMLK
Chr3_79627623_1138_po_C_Ez4/7       ------AQCEAR--PHCSAR---HTSLWSGSFEDLQ----VDFT--KMLK
Chr19_20477244_1350_po_C_Ez10/      VRYLPTAQYEAM--SHCSAW---HTSYVAAPFQDPQ----VDFT--EMC-
Chr5_111890459_1626_po_C_Ez10/      Q-CIICQQHNARZGPTVPPG---IQAYGAAPFEDFE----VDFT--EMPK
ChrX_48766518_1600_po_C_Ez6/3       QQCVTCQQHNARQGPTVPPS---IQAYGAAPFEDLQ----VDFT--EMPK
ChrX_152580713_1572_po_C_Ez10/      W-CVTCPQHIDRQGPTVPPG---IQAYGVAPFEDLQ----VEFT--EMPK
ChrX_139287533_1154_po_C_Xz8/4      R-CVTCRQHNASKAPLFLPA---YKAYGAAPFEDLQ----VDFT--EMPK
Chr2_52615786_912_po_C_Xz11/5       Q-RVTCQQHNARQGPTVTPG---MQTYGAAPFEDLQ----VDFT--EMPK
HERV-E_4_1                          R-CVTCRQHNARQGPAVPPG---IQAYGAAPFEDLQ----VDFT--EMPK
ChrY_21962399_1137_po_C_Xz13/9      Q-CVTCWQHNARQGPTAPPG---IQSYGAAPFEDLQ----VDFT--EMPK
ChrY_8634032_950_po_C_Xz14/10       W--CVTCQANARQGPTVPPF---IQAYGAAPFEDIQG----DFT--EMLQ
Chr3_126735360_997_po_C_Xz5/10      E-AGVLPADSMMRGKVQPFRP--AYELRAALFEGLQ----VDFT--EVPM
Chr3_126713203_958_po_C_Xz7/8       R-CVTCRQHDARQGPAVPHG---IRAYGAAPFEGLQ----VDFR--EMPK
Chr3_75226118_977_po_C_Xz8/9        R-CVTCRQHDVRQGPAVPPG---IRAYGAALFEGLQ----VDFT--EMPK
Chr7_6629919_827_po_C_Xz8/7         W-CVTCQQHDARQRPAVPPG---IRAYGAAPFEGLQ----LDFT--EMPK
Chr4_9166812_811_po_C_Xz8/10        R-CVTCZQHDARQGPAIPPG---IRAYGAAPFEGLK----VDFT--EMPK
Chr12_8478642_834_po_C_Xz12/7       R-CAADSMMRG--KVAVPPG---IRAYGAAP-EGLQ----VDFT--ETPK
Chr10_15187175_787_po_C_Xz11/5      R-CVTCRQHDAGQGAAVPPG---IRAYGAAPFEGLQ----VDFT--ETPK
Chr11_71819740_854_po_C_Xz11/9      P-CVTCRQHDMRQGLAVLPG---IRASGAAPFEGLQ----VDFT--EMSK
Chr11_71519974_886_po_C_Xz13/9      R-CVTCRQHDARQCPAVPHG---IZAYGAALFEGLQ----VDFT--EMPK
Chr8_7271441_969_po_C_Xz10/6        R-CVTCZQQDGSQGPAVPPG---IRACGAAPFEGLQ----VDFTFTEMPK
Chr8_11770066_826_po_C_Xz9/8        V-CYLR-QHDARQGPAVPPG---IRAYGAAPFEGLQ----VDFT--EMSK
Chr4_4128362_651_po_C_Xz7/11        R-CVTCRLHDARQGPAVPPG------------------------------
Chr14_50203955_850_po_C_Xz10/8      R-CVSCRQHHARQGPAVPPG---IQAYGAAAFEDLQ----VDFT--EMPE
Chr14_50205291_837_po_C_Xz10/8      R-CVSCRQHHARQGPAVPPG---IQAYGAAAFEDLQ----VDFT--EMPE
Chr7_6597120_842_po_C_Xz7/7         R-CVSCRQHDARPGPAVPPG---IQAYGAAPFEDLQ----VDFT--EMPK
Chr11_71544658_783_po_C_Xz8/13      R-CVSCRQHHARQGPAVPPG---IQAYGAAPFEDLQ----VDFT--EMPK
Chr11_67999520_760_po_C_Xz8/8       R-CVTCRQHKARQGPAVPPG---IQAYGAAPLEDLQ----VDFT--EMPK
Chr3_75461975_796_po_C_Xz7/9        P-CVTCRQHDARQGPAVPPG---IRAYGEAPFEDLQ----VDFT--EMPK
Chr12_52228939_692_po_C_Xz3/10      R-CVTCZQHNARQGPAVPPG---IKAYGAAPFEDLQ----VDFT--EMPK
Chr10_15211234_286_po_C_Xz11/1      W-CVTCLQPNARQGPAVQPS---VQAYGAVPFEDLQ----VDFT--EMPK
Chr12_100973302_945_po_C_Xz12/      V-VCDLLTAQCEAGPTVLPG---IQAYGVAPFEDLQ----VNFT--EMPK
Chr19_9685420_817_po_C_Xz6/10       W-CITCRQHNARQGLTVPPG---IQAY-VRLLLRIF----VDFT--EMPK
Chr13_40936820_719_po_C_Xz9/10      NILERCAANITEGVPPVTLFV--ISZRDVTPNVTGG-----VHS--VILF
Chr7_63858202_1754_po_C_Rz12/9      Q-YLTCAQNNPZQWPTRPPG---IQEIGTMPCENLL----MD-T--ELPQ
ERV3_chr7q21_AC073210_polputei      Q-YLTCAQNNPZQWPTRPPG---IQEIGTMPCENLL----MD-T--ELPQ
Chr6_110218096_843_po_C_Xz16/8      Q-CLICAQNNPZQGLTRPPG---IQEMGATPCENLL----MDFT--ELPS
Chr6_110214535_786_po_C_Xz16/8      Q-CLICAQNNPZQGLTRPPG---IQEMGATPCENLL----MDFT--ELPS
ChrY_3018841_881_po_C_Rz9/12        Q-YLTCARNNPRQTPTRPPG---FRKYGPHPVN-LL----MDFT--ELPQ
ChrX_86876098_639_po_C_Xz7/8        Q-YLTCAR-QPTTEAHQAPG---IQEIGAAPCKNLL----MDFT--ELPQ
Chr14_68899917_965_po_C_Xz6/8       Q-CLTCAQNNPZQGPTRPPG---IQEMGATLCENLL----MDFT--ELPR
ChrY_23874437_661_po_C_Xz15/9       Q-CLTCAQNIPZQGPTWPLG---IQEMGATPCENLL----RNFT--EMPQ
ChrY_27518418_649_po_C_Xz15/9       Q-CLTCAQNIPZQGPTWPLG---IQEMGATPCENLL----RNFT--EMPQ
ChrY_25016892_657_po_C_Xz13/12      Q-CLTCAQNIPZQGPTWPLG---IQEMGATPCENLL----RNFT--EMPQ
ChrY_19932937_520_po_C_Xz10/14      Q-CLTCAQNKLZQGPTRPLG---IQETGATPCEKLL----INFT--KLPQ
ChrY_20104614_518_po_C_Xz10/14      Q-CLTCAQNKLZQGPTRPLG---IQETGATPCEKLL----INFT--KLPQ
Chr19_58660155_598_po_C_Xz21/1      Q-CLTCAQNNRYGGPLGP-Q---IQEAGAMPCENLV----INFT--ELLR
Chr19_58621153_500_po_C_Xz21/1      Q-ZLTCAENNPRQGPTQP-Q---IQEVGAVPCENLL----INFS--ELPR
Chr7_139686987_710_po_C_Xz11/7      Q-YLTCAQNNLPQGPTQPPG---IQEIGAMPCK-TY----LQTT--ELPR
Chr7_137651301_1097_po_C_Pz13/      Q-YLICAQQ-PTARAHSAPG---IQZIGATPCKNLL----MDFT--ELPR
Chr11_29579197_930_po_C_Pz8/12      Q-CLICVQNNPWQGPTQPPG---IQEMGATPCENLV----MDFT--ELPR
Chr4_117329205_488_po_C_Xz3/16      R-CLQTMSN-------LCPK---QPMTRAHSA--PR----IQKM--GASP
Chr21_38522185_383_po_C_Xz7/9       Q-CLTCAQNPIITY--SAPG---NLGNWSHTLZKPA----YGLH--KAAL
Chr4_70173370_1049_po_C_Xz11/8      Q-RLTCAQNNPZKGPTQPPG---IQEKGAMPCENLL----MDFS--ELPQ
Chr4_70022513_788_po_C_Kz10/12      Q-CLTCTQSNSTRAYLAPGE---FRKQWPCPVITCF----WTSP--NCPE
Chr14_105225894_1027_po_C_Xz13      G-VZLVLITTSTRAHLTPQE---FRKQEPSLCKRAY----GLHQ--TVLS
Chr7_119241536_761_po_C_Xz11/1      Q-CLTCAQNNPRQGPTWPPG---NRKWGPHPVK-TY----LQIS--KMPQ
ChrY_22564970_567_po_C_Xz9/10       Q-CLICPQNNPWQGPTWPPA---VQKR-ATPCENLL----VDFT--KLPQ
```

```
Chr21_14409346_1477_po_C_Pz5/6      Q-CVTCIQNNLRKGPTQTPG---IQETGAVPCENLL----VNFT--KLPQ
ChrX_51259700_1396_po_C_Pz7/8       Q-CVICAQNTPDKGPLGPQE---FKKQEQLPCENLL----VEFT--QLPW
Chr6_64985396_1041_po_C_Xz11/1      Q-CLTCAQNNPWQGPTWPPG---IQETGATPCENLL----VDFT--KLPQ
ChrY_20355209_659_po_C_Xz7/11       Q-CLTCAQNNPSR-PTRPPG---IQETGATPCENLF----VYFT--KLPQ
Chr3_34567166_523_po_C_Xz8/6        Q-CLTCAQNNPQQGPTWPPG---IQETGATPCENLL----VDFT--ZLPR
Chr3_122072468_245_po_C_Xz9/9       Q-CLACAQNNPWQRPTRPPG---IQETGVIPCENLL----VDFM--ELPQ
Chr11_58788475_541_po_C_Xz8/22      Q-FLAGAQNNPWQGPTRPPR---IQETGATPCENLL----VDFT--KLSQ
Chr6_137421042_373_po_C_Xz6/9       Z-CLTCAPNNPQQGPTWPPG---IQETGATPCENLH----VGFT--ELPR
Chr4_115910608_550_po_C_Pz8/12      ----------PLQGPTQPPG---IQKTEATPCENLL----VDFT--KLPR
ChrY_13743185_977_po_C_Xz14/8       Q-CLSCAZNNPRQEPTZSPG---IQEVGAVPCENLL----IDFT--KLPQ
HERV15                              Q-CLSCVRNNPKQGPTRPPG---IQEMGAVPCENLL----VDFT--ELPR
ChrX_117198092_983_po_C_Xz8/8       Z-CLSCVWNKPKQGPIRPPG---IQEMGAVPCENLL----VDFT--ELPZ
Chr19_21166030_1412_po_C_Xz9/1      Q-CLSCEQSNPKQGPTRPPE---IQEMGAVPCENLL----VDLT--ELPZ
Chr19_21213743_913_po_C_Xz14/1      V-CLSCALNNPKZEPTQPPG---IQEMGAVPFENLL----VDFT--ELPQ
Chr19_15957774_569_po_C_Xz8/10      -------QEWSRRRWAGRAQ---RRT---------------------WSL
Chr11_77508681_297_po_C_Xz19/1      K-FLPQSQHFGRLRWVDHEV---RRLDQPGQHGETL----SLLK--IHKL
Chr13_98250769_738_po_C_Xz7/13      G-WPPCLKALAATILLTQEA---NRFTLG--------------Q--QLTI
Chr6_39968183_513_po_C_Pz10/9       ----ERIKYYGPATWAEDET---WGYCTP--------------------I
Chr15_66867263_645_po_C_Xz11/1      ----LRPIIFPILLIVLLLF---CCFSQP-----------------PPGK
HERV18_polputein                    QKCPACAQAYPRRRQLPSVTQ--QVTVGRMPLTRWQ----IDYIG-PLPK
HERVIP10FH_polputein                HKNVDFYLHZKVKKFFVLKVZ---ASFETLIVKEIL--------------
MER66_polputein                     CLIWEWEYRPLLENKDVRPPPLEQGKYLYDPTWQQR----IDITP-----


HERVI_polputein                     MPPIGHLKYLLVIIDHFTYWVETIPHSNATTSNVVKALI-----------
HERVIP10F_polputein                 MPPIGRLKYLLVIVDHLTHWVEAIPFSNATANNVVKALI-----------
HERVadp_AC005741                    MPKIVRLKYLLVIVDHLSGWMEGFPLPTATTRN-VKIIL-----------
HERV-PT47D                          ZPVRKYSEKEILIICNIZ--------KINZTSNVHEEEK-----------
HERV_W_chr6_141432567_ERV9_lik      ----------QKVEQALLKRRHLCSQKT----------------------
ERV9_PH1_RT                         NQSGIAZDATQRLQEKASEIRRLSNSYTNLWSWATWLLPFLGPVAAILLL
Chr13_54564403_572_po_C_Yz7/14      IAP-IRWPNYYLLDRSFKTIKRIVRACEVCQR-----------SNPALQA
HERV_Y_chr12_51022911_pol           MPKSQGFQYLLVWVDTFTGWTEAFPCKTEKFQ------------EVIKAL
HERV17_polputein                    MPKSQGFQYLLVWVDTFTGWAEAFPCRTEKAQ------------EVIKAL
HERV_W_chr7_9105739_syncytin_p      KPKPQGFQYLLVWVDTFTGWAEAFPCRTEKAQ------------EVIKAL
HERV9_polputein                     WPKSKGFQYLLVCVDTFTNWIEAFPCKTEKTQ------------EVIEVL
HUERS-P3_polputein                  MPKTKGIQYLLXWVXTFTNWVEAFPCRTEKAS------------EVVKVL
ERVfrd_AC004022                     LPAGFRZKYLLVCVDTFTKWVEACSTKTEKAQ------------EVAKFL
PRIMA41_polputein                   MPPCKVFKYLLVFVDTFSGWIEAFPTWSEKAI------------EVSKLL
HERV_H_RGH1_pol                     MPRVRKLKY-LVWVDTFTGWVEAFPTGS-EAT------------AVISSL
HERVHRGH2_pol                       MPRVRKLKYLLVWVDTFTGWVEAFPIESEKAT------------AVISSL
HERV_H48I_polputein                 MPPDKRLRYLLVFVCTFSGWVEAFPTTSEGAN------------VVTQTL
HERV_H_RTVLH2_pol                   APSQITXKYLLVZVDTFTGZVDAFPTGSEKAT------------VVISSL
HERVFH21_pol                        MPPVKRTKFLLTLIDTFSGWVEAFPTSSEKAA------------EVSQIL
HERVRblike_chr4_109047953           LPAPGGYKYLLVLVDTFSGWVEAYPCCSEWAM------------EVVKVL
ChrY_13790029_693_po_C_Xz12/11      HW--TKZKYLLVLADTFSGWTEAFAPKNETATTVARFL------------
HERVS71_polputein                   HW--AGYKYLLVLIDTFSGWTKAFTTGNETATMVVRLL------------
BaEV_M7_pol                         HY--AGYKYLLVFVDTFSGWVEAFPTRQETAHIVAKKI------------
HERVR_polputein                     HY--AGYKYLLVFVDTFSGWVEAFPTRQETAHIVAKKI------------
AKR_MLV_MLOCG_RT                    GL--YGYKYLLVFIDTFSGWIEAFPTKKETAKVVTKKL------------
GaLV_pol                            GR--YGNKYLLVFIDTFSGWVEAFPTKTETALIVCKKI------------
Chr3_158587414_1662_po_C_Ez10/      CG---GNTYLLVLVCTYSGWVEAYLTRTEKAYEVTRVL------------
Chr3_79627623_1138_po_C_Ez4/7       CG---GNKYLLVLVCTYSGWVEAYPTZTEKAYEVTRVL------------
Chr19_20477244_1350_po_C_Ez10/      -----GNKYLLVPVCTYSGWVEAYPKRTEKAPKVTCVL------------
Chr5_111890459_1626_po_C_Ez10/      CG---GNKYLLVLVCTYSEWVEAYPTQTEKAYEVTHVF------------
ChrX_48766518_1600_po_C_Ez6/3       CG---GNRYLLVRVCTYSGWVEAYPTRTEKAYEVTRVL------------
ChrX_152580713_1572_po_C_Ez10/      CG---GNKYLLVLVCTYSGWVEAYPTRTEKAYEVTRVL------------
ChrX_139287533_1154_po_C_Xz8/4      CK---GNKYLPVLVCTYSRZVEAYPSRTEKAHEVTRVL------------
Chr2_52615786_912_po_C_Xz11/5       CG---GNKYLLVLVSTYSGWMEAYLTRTEKAREITDVL------------
HERV-E_4_1                          CG---GNKYLLVLVCTYSGQVEAYPTRTEKAHEVTRVL------------
ChrY_21962399_1137_po_C_Xz13/9      HG---GNKYLLVLVCTYSGWVEAYPTRTEKACEVTRML------------
ChrY_8634032_950_po_C_Xz14/10       CR---GNKYLRVLLCTYTGWVEAYPTRTEEAYEVKHLL------------
Chr3_126735360_997_po_C_Xz5/10      FG---GNKYVLVLGRTYSGWVEAYPTRTEKTREVTTVL------------
Chr3_126713203_958_po_C_Xz7/8       CG---GNKYVLVLGRTYSGWVEAYPTRTEKAREVTPVL------------
Chr3_75226118_977_po_C_Xz8/9        CG---GNKYVLVLGRTYSGWVEAYPERTEKAREVTSVL------------
Chr7_6629919_827_po_C_Xz8/7         CG---GNKYSLVLGRTYSGWVZAYPTLTEKAREVTPVL------------
Chr4_9166812_811_po_C_Xz8/10        CG---GNKYVLVLGRNYSGWVEAYPTRTEKAREVTPVL------------
Chr12_8478642_834_po_C_Xz12/7       CG---GNKYILVLGRTYSGZVEAYPTRTEKAREVTPVL------------
Chr10_15187175_787_po_C_Xz11/5      CG---GNKYVLVLGRTCSGWVEAYPTZTEKAHEVTPVL------------
Chr11_71819740_854_po_C_Xz11/9      CG---GNKYVVVLGHTYSGWEEAYQTRTEKSREVTPVL------------
Chr11_71519974_886_po_C_Xz13/9      CG---GNKYLLVVMCTYSGWVEAYPTLTEKAREVTPVL------------
Chr8_7271441_969_po_C_Xz10/6        CG---GNKHVLVLGRTYSGWVEAYPTRTEKAGEVIPVL------------
Chr8_11770066_826_po_C_Xz9/8        CG---GNKYVLVLGRTYSGWVEAYPTRTEKAHEVTPVL------------
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      CG---GNKYLPVLGRTYSGWVETYPTRAEKAREVTRVL------------
Chr14_50205291_837_po_C_Xz10/8      CG---GNKYLPVLGRTYSGWVETYPTRAEKAREVTRVL------------
Chr7_6597120_842_po_C_Xz7/7         CG---GNKYSLVLGRTYSGWVEAYLTRTEKAREVTHVL------------
Chr11_71544658_783_po_C_Xz8/13      CG---GNKYLLVLGRTZSVWVEAYLTZTEKAREVSRVL------------
```

```
Chr11_67999520_760_po_C_Xz8/8       CG---GNKYLLVLGRTYSGWVEAYPT-TEKAREVTRVL------------
Chr3_75461975_796_po_C_Xz7/9        CG---GNKYLLVLGRTYSGWVEAYPTZTEKASEVTPVL------------
Chr12_52228939_692_po_C_Xz3/10      CG---GQKYLLVLGCTCSGWVEAYPTRTKKAREVTRVF------------
Chr10_15211234_286_po_C_Xz11/1      CG---GNKYLLVLGRTYSGWVEAYPTRTGKPREETRVL------------
Chr12_100973302_945_po_C_Xz12/      CG---DNKYLLVLVCTYSGYVEAYPTRTEKAYEVTCVL------------
Chr19_9685420_817_po_C_Xz6/10       CG---GNNYLLVLVCTYSGWVEAYPTSTEKVYEVPVCF------------
Chr13_40936820_719_po_C_Xz9/10      II---SRDVLVLKLZVPARWIAGKPGGGGNKAE-NPGP------------
Chr7_63858202_1754_po_C_Rz12/9      AG---SYRYMLVLVYTFSGWAKAFPTRSKNSZEVTKI------------L
ERV3_chr7q21_AC073210_polputei      AG---SYRYMLVLVYTFSGWAKAFPTRSKNSZEVTKI------------L
Chr6_110218096_843_po_C_Xz16/8      GE---GCQYMLVFICTFLGWVEAFPTQTEKAREVTKI------------L
Chr6_110214535_786_po_C_Xz16/8      GE---GCQYMLVFICTFLGWVEAFPTQTEKAREVTKI------------L
ChrY_3018841_881_po_C_Rz9/12        VG---DYQYLLVLIFTFSGWVKALPTRTEKARQVTKV------------L
ChrX_86876098_639_po_C_Xz7/8        VG---DYQYLLVLIFTFSGWVKALPTRTEKAREVTKV------------L
Chr14_68899917_965_po_C_Xz6/8       AG---DYRYMLVLICTFSGWIEAFPTRIEKTZEVTKV------------L
ChrY_23874437_661_po_C_Xz15/9       SV---GCKYILVIFCTFSGZVEAFPTLTEKSZEVTEV------------L
ChrY_27518418_649_po_C_Xz15/9       SV---GCKYILVIFCTFSGZVEAFPTLTEKSZEVTEV------------L
ChrY_25016892_657_po_C_Xz13/12      SV---GCKYILVIFCTFSGZVEAFPTLTEKSZEVTEV------------L
ChrY_19932937_520_po_C_Xz10/14      SG---DYQYLLGLVCTFLGWFEAFHTPTEKSGEVTEV------------R
ChrY_20104614_518_po_C_Xz10/14      SG---DYQYLLGLVCTFLGWFEAFHTPTEKSGEVTEV------------R
Chr19_58660155_598_po_C_Xz21/1      ASDR-GCRYMLVLICTFFKIGZSFPHQDRKS-MRSDZ-------------
Chr19_58621153_500_po_C_Xz21/1      A--R-GYGYMLLLICTFSRZVKAFSTRGKVQ-EVTKVL------------
Chr7_139686987_710_po_C_Xz11/7      AG---GYRYMLVLICTFPGWVEAFPTRTKKAREVTKARZ----------K
Chr7_137651301_1097_po_C_Pz13/      ARG---YPYMLVLVCTFSGWVETFSTRTEKAZEVTKVL------------
Chr11_29579197_930_po_C_Pz8/12      AGG---YWYMLVLIFTISGCVEAFPTRTEKAREVTKVL------------
Chr4_117329205_488_po_C_Xz3/16      CEN---LLMNFTKLLRAGG---YQYMLVLSAWEVTKVL------------
Chr21_38522185_383_po_C_Xz7/9       SRG---LLYMLVLFCTLSGZVKAFPTQTEKAREVTKVM------------
Chr4_70173370_1049_po_C_Xz11/8      AGV---YRYILVLVCTFSGWVMAFSVRTEKAZELMKVL------------
Chr4_70022513_788_po_C_Kz10/12      GG----YQYMLVLVCTFSGWVKVFPTZTZKAQEVTKG-------------
Chr14_105225894_1027_po_C_Xz13      RG----YQYMLVLVCTFSGWVEAFSTQTKKAZEATKVL------------
Chr7_119241536_761_po_C_Xz11/1      MRG---YQYMLVLVCNFSGWVEAFPTRTERAZEVTKVL------------
ChrY_22564970_567_po_C_Xz9/10       AG---GHWYLLVLVCTFSLWVEVFPNRTEKAWEVTRVQ------------
Chr21_14409346_1477_po_C_Pz5/6      AG---GCQYILVFTFSG--WIKAFPTRTEKAZEITRIL------------
ChrX_51259700_1396_po_C_Pz7/8       AG---DYQYMLVLVCTFZRWVEAFPTRTEKAREVTRIL------------
Chr6_64985396_1041_po_C_Xz11/1      AR---GYWYMLVFAFTFSGWVEAFPTRTEKAQEVTKIL------------
ChrY_20355209_659_po_C_Xz7/11       AR---GYRYMLVFVCSFSGWVETFLTRTGKTREVTKIL------------
Chr3_34567166_523_po_C_Xz8/6        AR---GYRYILMFVCTFSGWAEAFRTRTEKVREVTRIL------------
Chr3_122072468_245_po_C_Xz9/9       AR---GYWYMLVFVCTFSGWVKAFPTRKEKAQEVTRIL------------
Chr11_58788475_541_po_C_Xz8/22      AR---GYRYMLVFVCTFL--------------------------------
Chr6_137421042_373_po_C_Xz6/9       TG---GYRCMLVFVCTFSGWVEAFPTRTKKAREVTRIL------------
Chr4_115910608_550_po_C_Pz8/12      VG---GYRYR-RFVCTFSGWVESFPTVTEKAREVTRIL------------
ChrY_13743185_977_po_C_Xz14/8       AG---GYWYMLVFVCTFLGCVEAFPTRNEKAREVTKVI------------
HERV15                              AG---GYRYMLVFVCTFSGWVEAFPTRTEKAREVTKVL------------
ChrX_117198092_983_po_C_Xz8/8       AG---GYRYMLVFLCTFSGWVEAFPTRTEKAREVTKVL------------
Chr19_21166030_1412_po_C_Xz9/1      AG---GYWYMLVFVT-FSGZVEAFPIRTEKAQEVTKVL------------
Chr19_21213743_913_po_C_Xz14/1      AE---CYWYMLVSVCTFSRWVEAFPNRTEKAZEVTKLL------------
Chr19_15957774_569_po_C_Xz8/10      LGQAGRLPGREVAGAGFLTMSECFKVGDEAMQWATLEFK----------T
Chr11_77508681_297_po_C_Xz19/1      ARRGGTLLZSQLLGG--LRQGNCLNLGGGGCSEPRLCHY----------T
Chr13_98250769_738_po_C_Xz7/13      WEPHSVITLTDQRGHHWLSNPKMTQYQGFLCDNTYIILV----------N
Chr6_39968183_513_po_C_Pz10/9       YVINFIIWLQAVLEVITSETSRALNLLAMQATQMRNALY----------E
Chr15_66867263_645_po_C_Xz11/1      TSLLPMLGVRMLLFALLLMVPISMLEGEGPIEGCPHCMH----------T
HERV18_polputein                    SQ----GYTHALTAVDMATSVFAYPCRVADQQHTIQAL------------
HERVIP10FH_polputein                ------CVNILAKVKGVSSSFSELDIKIKAQQVFLKALTCS---------
MER66_polputein                     ---------RRASVCAPTGLIFVCGHEWEEVTX-----------------


HERVI_polputein                     -ENIVPRFGLIQSTDSDNGTHFTAYVIKKLAQVLDIKWKNHIPWHL---S
HERVIP10F_polputein                 -ENIVPRFGLIENIDSDNGTHFTTHIIKKLSQTLDIRWEHHTPWHP---P
HERVadp_AC005741                    -EZIVPRFGLAENIDSDNGSHFTSRVLRGIMEGLQIRWDYRTPWHP---P
HERV-PT47D                          -KTILPRN----ITDDSNKRKGIIWILDRILQIAKISFFQINNIFF---A
HERV_W_chr6_141432567_ERV9_lik      -----------------------HEKILIITGHQRNANQNHK------EI
ERV9_PH1_RT                         LAFGPRIFNLLVKFVWSRIEAIKLQMVLQIEPQMSSTNNFYRGPL---DZ
Chr13_54564403_572_po_C_Yz7/14      IHFNPCIFNLLVKFDSSRIKTVKLQIVLQMEPQMQSMTKIHCGPL---DR
HERV_Y_chr12_51022911_pol           VHEIIPRFGLPZGLRSDNGPAFKAAVTQVVSQALGIRYHLHCAWR---PQ
HERV17_polputein                    VHEIIPRFGLPRGLQSDNGPAFKAAVTQGVSQALGIQYHLHCAWR---PQ
HERV_W_chr7_9105739_syncytin_p      VHEIIPRFGLPRGLQSDNSPAFQATVTQGVSQALGIRYHLHCAZR---PQ
HERV9_polputein                     IDEIIPRFGLPZGLQSDNGP-FKATITQGMSRLLGIQYHLHCARR---PQ
HUERS-P3_polputein                  VNEITPRFGLPKYLQSDNGPSFKAAVTQGVSKALGYS-ISPLAWR---PQ
ERVfrd_AC004022                     LKECISWFELPRSLQSDNGPSFISZVTQQVSRAVSKHKCTFTLPG---DQ
PRIMA41_polputein                   LKEIIPRFGLPKSLQSDNGPSFTATITQNISSALGIQYRLHSAWR---PQ
HERV_H_RGH1_pol                     LSDIIPRFGLPTSIQSDNGPAFTSQITQAVSQALGIQWKLHIPYH---PQ
HERVHRGH2_pol                       LSDIIPWFGLPFSIQSDNGPAFTSK-SQAVSQALGIQWNLHIPZH---PQ
HERV_H48I_polputein                 IMHIIPRFGLPTSIQSDNGPAFISQITQGVSTSLXIKWVLHTPYR---PQ
HERV_H_RTVLH2_pol                   LSDIIPRFGLPTSIQSHSRPAFIGEISQAFFQALSIQZNLYIPYS---PQ
HERVFH21_pol                        VTEIIPRFGLPGSIQSDNGPSFISQITQQVSQSLGIQWRLHIPCW---PQ
HERVRblike_chr4_109047953           LKEIIPRYGFPDIIQSDNGPSFTZEIAQQVRKVLGIKWKLHTAWR---PQ
ChrY_13790029_693_po_C_Xz12/11      LDEIILZHRLPAAIRSDNRPAFTLSIPQSVSKNS--------------EL
HERVS71_polputein                   LIKIISQHGLPVAIGSDNGPAFTSSMAQSVSKALNIKWKLHCTYZ---PQ
```

```
BaEV_M7_pol                         LEEIFPRFGLPKVIGSDNGPAFVSQVSQGLARILGINWKLHCAYR---PQ
HERVR_polputein                     LEEIFPRFGLPKVIGSDNGPAFVSQVSQGLARILGINWKLHCAYR---PQ
AKR_MLV_MLOCG_RT                    LEEIFPRFGMPQVLGTDNGPAFVSKVSQTVADLLGIDWKLHCAYR---PQ
GaLV_pol                            LEEILPRFGIPKVLGSDNGPAFVAQVSQGLATQLGINWKLHCAYR---PQ
Chr3_158587414_1662_po_C_Ez10/      LZDLIPRFGLPLZISSDNRPAFVAVLVQKTAKALGITWKLCATYR---PQ
Chr3_79627623_1138_po_C_Ez4/7       LRDLIPRFELPLZIVSDNGPAFVPDLVQKTAKPLGISWNLHAAYR---LQ
Chr19_20477244_1350_po_C_Ez10/      LZDLIPRFGLPLZIGSDNGPVFVADLVQKTAKALGITWKLHATYR---PQ
Chr5_111890459_1626_po_C_Ez10/      LZDFIPRFGLPLZIGSDNGRAFVADLVQKTAKALGITWKLCATYZ---PQ
ChrX_48766518_1600_po_C_Ez6/3       LZDLIPRFGLPLZIGSDNGPAFVADLVQKTAKALEITWKLHAAYR---PQ
ChrX_152580713_1572_po_C_Ez10/      LZDLIPRFGLPLZIGSDNGPAFVADLVQKTAKALGITWKLHAAYZ---PQ
ChrX_139287533_1154_po_C_Xz8/4      LRDLIPRFGLPQWIGSDNGRAFVADLVQRTAKVLRITWKLRAAYR---PQ
Chr2_52615786_912_po_C_Xz11/5       LRDLIPRFGLPLZIGSDNRPAFVADLVQKTAKVLGITWKLYTAYZ---PQ
HERV-E_4_1                          L-EILPRFGLPLRIGSDNGLVFVADLVQKTAKVLGITWKLHAAYQ---PQ
ChrY_21962399_1137_po_C_Xz13/9      PRDLITRFGLPLZISSDNGPVFVADLVZKTAKPLGITWKL--TYR---PQ
ChrY_8634032_950_po_C_Xz14/10       LZNLIPRFGLPLZISSNNWLAFLPDLVQKTAKALGLTWKLHAVYG---PQ
Chr3_126735360_997_po_C_Xz5/10      LZDVIPRFRPPLRIGSDNGPAFLAALVQKTAKILGITRKLHAASR---PQ
Chr3_126713203_958_po_C_Xz7/8       LRDLIPRFRPPLWIGSDNGPAFLAALVQKTAKVLGITRKLHATSR---PQ
Chr3_75226118_977_po_C_Xz8/9        LRDLIPRFRRP-RIGSDNGPAFLAALVQKTAKVLGITGKLHAASR---PQ
Chr7_6629919_827_po_C_Xz8/7         LRDLIPRFRPPLRIGSDKGPAYLAAFVQKTAKVLGITRKLHAASR---PQ
Chr4_9166812_811_po_C_Xz8/10        LZDLIPRFRPPLGIGSDKGPAFLAALVQKTAKVLGITRKLHAASQ---PQ
Chr12_8478642_834_po_C_Xz12/7       LRDVIPRFRPPLRIGSDNRPAFLAALVQKTAKVLGITRKLHAASR---PQ
Chr10_15187175_787_po_C_Xz11/5      LRDVIPRFRPPLRIGSDNGPVFLAALVQKTAKVLGITWKLHAAYR---PQ
Chr11_71819740_854_po_C_Xz11/9      LRDLIPRFRRPLGIGSDNGPAFLVALVQKTAKVLGITRKLHAASR---PQ
Chr11_71519974_886_po_C_Xz13/9      LZDLIPRFRPPLWIGSDYGPAFLAALVQKTAKVLGITRKLHAASR---PQ
Chr8_7271441_969_po_C_Xz10/6        LRDVIPRFRPPLWIGSDNGPAFLAALVQKTAKVLGIARKLHAASR---PQ
Chr8_11770066_826_po_C_Xz9/8        LHDWIPRFRRALGIGSDSGPVFLAALVQKMAKVLGITRKLHAASR---PQ
Chr4_4128362_651_po_C_Xz7/11        ----------------------IELMEQPHLKVS---RWLHRDVK---VW
Chr14_50203955_850_po_C_Xz10/8      LRDLIPRLELPFRIGSDNGPAFVADLLQKTATVLGITRKLHAASR---PQ
Chr14_50205291_837_po_C_Xz10/8      LRDLIPRLELPFRIGSDNGPAFVADLLQKTATVLGITRKLHAASR---PQ
Chr7_6597120_842_po_C_Xz7/7         LRDLIRRLGQPFRIGSDNGPAVVADLLQKTAKVLGITRKVHAASR---PQ
Chr11_71544658_783_po_C_Xz8/13      LRDLIRRLGLPFRIGSDNGPAFVADLLQKTAKVLGITRKLHAASR---SQ
Chr11_67999520_760_po_C_Xz8/8       LRDLIRRLGLPFRIGSDNGPAFVADLLQKTAKELGITRKLPAASR---PQ
Chr3_75461975_796_po_C_Xz7/9        LZGMIPRFZLPLRMGSDNGPAFLAAWVQKTAKVLGITWKLHAASR---PQ
Chr12_52228939_692_po_C_Xz3/10      LRDLIPRFGLPLRISSDNGPAFVADLVQKTAKVSGITWKRHAAYQ---PQ
Chr10_15211234_286_po_C_Xz11/1      LRGLIPRLGLPFRIGSDNGPAFVADLLQKTAKVLGITRKLHAASQ---PQ
Chr12_100973302_945_po_C_Xz12/      LZDLIPRFRLPLZISSDNGPAFVADLVQKTAKASGITWKTHATYZ---PQ
Chr19_9685420_817_po_C_Xz6/10       SEILFLGLDCPY------------------ESAQNMAWRLWLTWY---RL
Chr13_40936820_719_po_C_Xz9/10      GSNTASHLQLSFCLFLPLLYRLLHSFFLSLHPSFSSTCRWHPLVM---PZ
Chr7_63858202_1754_po_C_Rz12/9      LRNIIPRFGLSVTLGSDNRLAFLAEIIQELTRVLKIKWKLHTAYR---LQ
ERV3_chr7q21_AC073210_polputei      LRNIIPRFGLSVTLGSDNRLAFLAEIIQELTRVLKIKWKLHTAYR---LQ
Chr6_110218096_843_po_C_Xz16/8      LRDIIPRFGLLLTLDSDNRPAFVAEISGTNMAVSSS-------------Q
Chr6_110214535_786_po_C_Xz16/8      LRDIIPRFGLLLTLDSDNRPAFVAEISGTNMAVSSS-------------Q
ChrY_3018841_881_po_C_Rz9/12        LTDIILRFGLPLTLGSDNGLRFVAKISGTNA-LLKIKWKLHTAID---RR
ChrX_86876098_639_po_C_Xz7/8        LRDIILRFGLPLTLGSDKGLRFVAEIVQERMRLLKIKWKLHTAYR---SQ
Chr14_68899917_965_po_C_Xz6/8       LR-HYPRFGLPLTLGSDNGPAFLATIVQKLTQLLKVRZKLHTAYQ---PQ
ChrY_23874437_661_po_C_Xz15/9       LRDDIPRLRMPLTLGSENGLKFVTETVQGL--------------------
ChrY_27518418_649_po_C_Xz15/9       LRDDIPRLRMPLTLGSENGLKFVTETVQGL--------------------
ChrY_25016892_657_po_C_Xz13/12      LRDDIPRLRMPLTLGSENGLKFVTETVQGL--------------------
ChrY_19932937_520_po_C_Xz10/14      LRDIIPRFRLPLTLRSENGLEFTAEIIQGLTWLLKIKWKLHS-LP---VT
ChrY_20104614_518_po_C_Xz10/14      LRDIIPRFRLPLTLRSENGLEFTAEIIQGLTWLLKIKWKLHS-LP---VT
Chr19_58660155_598_po_C_Xz21/1      --RLLPRFGLPLSLKSDNGPAFAAEIVQDLTRLLKIKZKLHIAYW---PE
Chr19_58621153_500_po_C_Xz21/1      IKDIIPRFGLALPLKSDNGPAFVAEIVZDFTSLLKIKZKLHTDYQ---SQ
Chr7_139686987_710_po_C_Xz11/7      TLSIIPRFGLPLTLGSDNGPAFVAEIVQDLTRLFFFFSFLYIYMY-----
Chr7_137651301_1097_po_C_Pz13/      LRDIITRFGLPLTPGSDNGPAFVAEIAQDLARLLKIKWKLHIAYK---PQ
Chr11_29579197_930_po_C_Pz8/12      IRDIISRFGLPLILGSDNGPVFVAEIVQELIZLLKIKWKLHTAYQ---PQ
Chr4_117329205_488_po_C_Xz3/16      LRDIIPRFZLPLTLKSGNGPAFVAEIVHELAQLLKIRWKLHTAYQ---PQ
Chr21_38522185_383_po_C_Xz7/9       LRDIIPRFGLP-QLGSDNGPALVAEIVQELVWLLKIKWKLHTAYQ---LQ
Chr4_70173370_1049_po_C_Xz11/8      LRDIIPRFGLPLTLGSDNGLAFVAEIVQDLTRLLKIKWKLHPVYQ---SV
Chr4_70022513_788_po_C_Kz10/12      --GIIPRFGLLLNLGSDNGLAFVAEIVQDLTZLLKIKWKLHTAYQ---PQ
Chr14_105225894_1027_po_C_Xz13      LRDVIPRFGLPLTLGLDNGLAFVAEVVQDLTRLLGIKWKLYTACS---LQ
Chr7_119241536_761_po_C_Xz11/1      LRYIIPRFGLPLTLGSDNGLAFVAEIVQELTRLLKIKZKLH-AYR---LQ
ChrY_22564970_567_po_C_Xz9/10       LKDCIPRFGLPVTIGSENGQAFVAEIVQQLTQLLKIKLKWHTAYH---LQ
Chr21_14409346_1477_po_C_Pz5/6      LKDIILRFRMPLPLGSDNRPAFVAEVVQQLTQILZIKZKLHTAYQ---PQ
ChrX_51259700_1396_po_C_Pz7/8       LKDIIPRFRLPLTLGSDNRPAFVAEVVQQLTQVK---RKLRTAYH---PQ
Chr6_64985396_1041_po_C_Xz11/1      LKDIISRFZLTLTLGSDSGPVFVVEIVQQLTQMLKIKWKLHTAYH---LQ
ChrY_20355209_659_po_C_Xz7/11       LKDIIPRFRLPLISGSDSGPALVAEIVQZLTQMLKTKWNLHIAYH---PQ
Chr3_34567166_523_po_C_Xz8/6        LKDIIPRFZLPLTLGPDNGPAFVAKIVQQLTQMLKIKWKLPTAYLS--PT
Chr3_122072468_245_po_C_Xz9/9       LKDIIPKFELPLTLGTDNGPAFVTEIVQQLMQMLKIKWKLHTAYR---PQ
Chr11_58788475_541_po_C_Xz8/22      ---------------SDNGSAFVAEIVQQLTQMLKIKWKLHTAYC---QQ
Chr6_137421042_373_po_C_Xz6/9       LKDIIPRFRLPLTLGSDNGPAFVAEIVQQLTQVKN-QMKTAYSLS---PI
Chr4_115910608_550_po_C_Pz8/12      LKDIIPRFGLPLTLESDNGPAFMAKIVQQLMQMLKIKWKLHTAYH---PQ
ChrY_13743185_977_po_C_Xz14/8       LKDIIPRFRLTLILGSDN-PAFVAELVQQLTZLLKIKWKQHTASS---PQ
HERV15                              LKDIIPRFGLPLTLGSDNGPAFVAEVVQQLTQLLKIKWKLHTAYZ---PQ
ChrX_117198092_983_po_C_Xz8/8       LKDIIPRFGLPLTLGSDNGPAFVAEVVLQLTQLLKIKWKLHTAYZ---PQ
Chr19_21166030_1412_po_C_Xz9/1      LKDIISRFGLPLTLGSDNGPAFAAEVVQQLTQLLKIKWKLHTAYZ---PK
Chr19_21213743_913_po_C_Xz14/1      LR-HIARFGZPLTIGSDNGPAFVAEVVQQLTQHFFFVFIDHLGCF---SE
Chr19_15957774_569_po_C_Xz8/10      SCQAWHHLACPQAQNNEED----MQLVEDWTAIFMTFT------------
```

```
Chr11_77508681_297_po_C_Xz19/1      --PAWQQRKTPPQKKKKEKKMVKMEIVMFILPQFFFFFZRNELSSKAKTW
Chr13_98250769_738_po_C_Xz7/13      -----TFNPATLLPVERELPSMTVETVDEVFSSWKDLTDG----------
Chr6_39968183_513_po_C_Pz10/9       NSLALNYLLASEGRVCGKFNLNCCLEIDDNGQAVMEITAR----------
Chr15_66867263_645_po_C_Xz11/1      TWIGNTVVKTLLYHRSLSGQNNHTQLMALTKYKPLPKRED----------
HERV18_polputein                    -QHLCALYGRPLAIESDRGTHFTGQQVQQWAQQMDIKWGFHVPYN---SQ
HERVIP10FH_polputein                LTKIIKGZKESIKILPYGQTLKLDKYVYKVLLKLSLTLITTNIKVKFSLS
MER66_polputein                     --RPLPTPQGPPVLLGVAFPCISKTWNRGECTLATLAPPGVTVYN---PI


HERVI_polputein                     SSRRVERMNQTLKSHLTKLVLET-ZLPWTKCLPIALLRIRTAPQRDTGLS
HERVIP10F_polputein                 SSGRVERMNQT-KEPLTKLVLET-RLPWTKCLPIALLRIRTAPRKDIGLS
HERVadp_AC005741                    SSGKVERMNQTFKKHITKLILET-KMPCTTCLQIALLSIRTAPRKD-GIA
HERV-PT47D                          IPN----------EIAKRFFLEI-RQNYFKIZ------VEKQMHKD----
HERV_W_chr6_141432567_ERV9_lik      PSHTS---------------------------------------------
ERV9_PH1_RT                         PAGTSPGLESS---PLKDTTTAKPLLR-----------------------
Chr13_54564403_572_po_C_Yz7/14      PASPCSDVNDSKAPLLRKSQLHNPYYAPIQQETEQSLANFPNSTWVFLLR
HERV_Y_chr12_51022911_pol           SSGKVKKMDETLKZYLKKLTQETHLAWPALLPIALLRIQNS------PQK
HERV17_polputein                    SSGKVEKMNETLKRHLKKLTQETHLAWPALLPIALLRIRNS------PQK
HERV_W_chr7_9105739_syncytin_p      SSGKVEKMNETLKGHLKKQTQETHLTWPALLPIALKRICNF------PQK
HERV9_polputein                     SSGKVKKANETLE-ALRKLTQGTS-PWPTLLPMALLRIRNS------PHK
HUERS-P3_polputein                  SSGKVEKTNDIIK-TLRKLSQETHLPWVTLLPMALLQVRNX------PSK
ERVfrd_AC004022                     PSGKVERTKETLEHILNKLCQKTAQPWVDLLLLAPLWIHIAP-----QTP
PRIMA41_polputein                   SSGKVERANQTLKRTLAKLCXETSETWLSLLPVALLRVZMAP-----KGN
HERV_H_RGH1_pol                     SSGKVERTNGLLKPHLTKLSLQLKKDWTVLLPLALLRIRACP-----RDA
HERVHRGH2_pol                       SSGKVKPTNGLLKTHLTKLSLQLKKDWTVLLPLALLRIRTCP-----RDA
HERV_H48I_polputein                 SSGKVEKINSVLKAQLTKLALETHQSWTKNLPFALMRLRATP-----KAP
HERV_H_RTVLH2_pol                   SSGKVEQTNSLLKTHLTSSATNLKRTRQYFYHFPFSEFRPVL-----RML
HERVFH21_pol                        TSGKVERANGILKAQLTKLTLEVQKPWTSLLPIALESIRASP-----KAP
HERVRblike_chr4_109047953           SSRQTERINHTLKTTLAKLCQETRLKWLQLLGIPLLQVRIAP-----RSG
ChrY_13790029_693_po_C_Xz12/11      CTGRMHKPHP--KNTFTKLILKTG-EDZVRLL-LVLLKVRZTSYQAR-FS
HERVS71_polputein                   SSGQVERMNHTIKSTLTKLILETS-ENWVKLLPLALLRVRYTTYWAG-FS
BaEV_M7_pol                         SSGQVERMNRTIKETLTKLTLETGLKDWRRLLSLALLRARNTPNRFG-LT
HERVR_polputein                     SSGQVERMNRTIKETLTKLTLETGLKDWRRLLSLALLRARNTPNRFG-LT
AKR_MLV_MLOCG_RT                    SSGQVERMNRTIKETLTKLTLATGSRDWVLLLPLALYRARNTPGPHG-LT
GaLV_pol                            SSGQVERMNRTIKETLTKLALETGGKDWVTLLPLALLRARNTPGRFG-LT
Chr3_158587414_1662_po_C_Ez10/      SSGKVERMNRTIKNSLGKVCQET-GLEWIQALPMVLFKIRCTPSKKKGYS
Chr3_79627623_1138_po_C_Ez4/7       RSRKVEQMNQTIKNSLGKVYQET-GIKWY-ALPMVWFKIRCNPPRK-GCS
Chr19_20477244_1350_po_C_Ez10/      SSGKVEPMIRTIKNSLGKVCQEA-GLKWIQALPMVLFKIRCTPSKKTGYS
Chr5_111890459_1626_po_C_Ez10/      SSGKDGVMNRTIKNSLGKVCQEI-GLKQIQSLPMILFKIRCTPSKKTGHS
ChrX_48766518_1600_po_C_Ez6/3       SSGKVERMNRTIKNSLGKVCQET-GLKWIQALPMVLFKIRCTPSKKTGYS
ChrX_152580713_1572_po_C_Ez10/      SSGKVERMNYTIK--IGKVCQET-GLKCIQALPMVLFKIRCTPSKKTGYS
ChrX_139287533_1154_po_C_Xz8/4      SSGKVERMNRTIKNSLRKVCQET-GLKWIQALPMVLFKIRCTPSKKTGYS
Chr2_52615786_912_po_C_Xz11/5       SSRKVEQMYRTTKNSLGKVCQEI-GLKWVZTLPMVLFKIRCTPSKRIGYS
HERV-E_4_1                          SSGKVERMNRTIKNSLGKVCQET-GLKWIQALPMVLFKIRCTPSKRTGYS
ChrY_21962399_1137_po_C_Xz13/9      SSGKVERMNRTIKNSLGKACQET-RLKWVQALPMVLFTIRCTPSKRIGYS
ChrY_8634032_950_po_C_Xz14/10       SSGKVERMNQTIKNCLGKVSQET-GFKWIQALPMVLFKIRCTSSKKTGZC
Chr3_126735360_997_po_C_Xz5/10      SSGKVERMNLTIKN------------------STVVFPAGYLKQHHKGRQ
Chr3_126713203_958_po_C_Xz7/8       SSGKVEWMNRTIKN------------------STIVFPAGYLKNTTRGVK
Chr3_75226118_977_po_C_Xz8/9        SSGKVERMNQTIKN------------------STLSSPLDIENNTTRGVK
Chr7_6629919_827_po_C_Xz8/7         HSGKVERMNWTIKN------------------STIVFPAVYLKNTARGVK
Chr4_9166812_811_po_C_Xz8/10        RSGKVERMNRTIKN------------------STIVFPAGYWKQQHKVRQ
Chr12_8478642_834_po_C_Xz12/7       SSGKVERMNRTIKN------------------STIVFPAGYLKNTARGVK
Chr10_15187175_787_po_C_Xz11/5      SSGKVERMNRTFKN------------------STIVFPAGYVKNTTRGVK
Chr11_71819740_854_po_C_Xz11/9      SSGKVEWMNRTIKN------------------STIVFPSGYLKNTTRGIK
Chr11_71519974_886_po_C_Xz13/9      HSGKVERMNRTIKN------------------STIVFPTGCKTTPQAASN
Chr8_7271441_969_po_C_Xz10/6        SSGKVERMNRTIKN------------------STTVFPAGYLKQQHKGRQ
Chr8_11770066_826_po_C_Xz9/8        SSGKVERMNRTIKN------------------SSIVFPAGYLKQHHKGRQ
Chr4_4128362_651_po_C_Xz7/11        R-------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      SSGKVERMNRTIKN------------------SIMVFPAGYVKNTTRGVK
Chr14_50205291_837_po_C_Xz10/8      SSGKVERMNRTIKN------------------SIMVFPAGYVKNTTRGVK
Chr7_6597120_842_po_C_Xz7/7         SSRKVERMNRTIKN------------------NIIVFPTGCKTTPRGASN
Chr11_71544658_783_po_C_Xz8/13      SSGKVEWMNRTIKN------------------NIIVFPAGYVKNTTRGVK
Chr11_67999520_760_po_C_Xz8/8       SSGKVEGMNRTIKI------------------VLLSSAMDMZNNTTRGVK
Chr3_75461975_796_po_C_Xz7/9        SSGKVERMNRTIKN------------------STIVFTCGYLKNTTRGIK
Chr12_52228939_692_po_C_Xz3/10      SSGKVERMNRTIKN------------------SIIIFPTGYKTIPQGASN
Chr10_15211234_286_po_C_Xz11/1      SSGKVERMNWTIKN------------------STIVFPSGYVKHHHEGRQ
Chr12_100973302_945_po_C_Xz12/      SSGKVEZMNQTIKS------------------SLGKMLGNRIKVDRGGWS
Chr19_9685420_817_po_C_Xz6/10       AG----CISSTCIP------------------SYLRLRHEDLEPGRWRFQ
Chr13_40936820_719_po_C_Xz9/10      ANRPAANLGAEAMR----------------LPSNAALPLARAGVQGSRYP
Chr7_63858202_1754_po_C_Rz12/9      SLGKVECMNWTLKQLLKKYCQET-PLRWDQVLPMVLLQVRCTPTKQTGYS
ERV3_chr7q21_AC073210_polputei      SLGKVECMNWTLKQLLKKYCQET-PLRWDQVLPMVLLQVRCTPTKQTGYS
Chr6_110218096_843_po_C_Xz16/8      SSGKVEHMNQTLKQLLKKYCZET-HLRWDQVLPMVLLZVRCIPTKQTGYL
Chr6_110214535_786_po_C_Xz16/8      SSGKVEHMNQTLKQLLKKYCZET-HLRWDQVLPMVLLZVRCIPTKQTGYL
ChrY_3018841_881_po_C_Rz9/12        AQEKWRAZTRHS-NATKEILPRN-PSEMGLGPAHGPPPGQVHPHNKLGIH
ChrX_86876098_639_po_C_Xz7/8        SSGKVESMNQTQ-IATKEILPRN-PSEMGLVLPMVLLQVRCTPTKQTGYS
Chr14_68899917_965_po_C_Xz6/8       SSGKVEHMNQTLKQPLKKYCQET-HLRWDQVLPMVLLRVR-HPYQTTGYS
ChrY_23874437_661_po_C_Xz15/9       ----------------------------------------FSTGHPQSZL
```

```
ChrY_27518418_649_po_C_Xz15/9       ----------------------------------------FSTGHPQSZL
ChrY_25016892_657_po_C_Xz13/12      ----------------------------------------FSTGHPQSZL
ChrY_19932937_520_po_C_Xz10/14      KLRKSRIHELDTQTASKKYCQET-HLEWNQVLPMVLLZVRYTPTKQTGYL
ChrY_20104614_518_po_C_Xz10/14      KLRKSRIHELDTQTASKKYCQET-HLEWNQVLPMVLLZVRYTPTKQTGYL
Chr19_58660155_598_po_C_Xz21/1      SSRKVKYMNQTLSSYZRNTAKKL-TZ--DKIRSFIWLSSESGAPPPNKLG
Chr19_58621153_500_po_C_Xz21/1      SSRKVKRMNWTLKQLLKKYCQET-HLRZDQLFRMVLLQVRCTPTKQTRYS
Chr7_139686987_710_po_C_Xz11/7      -----VYIYYTLSSRVHVHNLQV-CY------ICIHVPCWCAAPIN----
Chr7_137651301_1097_po_C_Pz13/      SSGKVEHMNWTLIQLLKKLAKNS-SE-MGSVLPMVLLZVRCTPTKQTGYL
Chr11_29579197_930_po_C_Pz8/12      SSRKVERMNQTLQKLLKKYCQET-HLRWDQVLPMVLLRVRCTPSQAL---
Chr4_117329205_488_po_C_Xz3/16      NSSKVERMNRTLKQLLKKYCQET-HLRWDQVLPMVLLQVRCTPTKQTAYL
Chr21_38522185_383_po_C_Xz7/9       SSKNIECMNWTLKR-LKKILSET-HLRWDQVLPMVLLZVRCTPTKQTRYS
Chr4_70173370_1049_po_C_Xz11/8      LSKDGAHEPDTQAAAEEILSGNSSEVGPGLAHGLSPSQVHPHKTNHTGYL
Chr4_70022513_788_po_C_Kz10/12      SSEKMEHMNRTLKAAEEIWSGISSEVGSGLIHGLPPSQVHPHKKKQTGYL
Chr14_105225894_1027_po_C_Xz13      SSGKVEHMNWTLKQLLKKFCEET-HLRWDQVLPVVLLQVGCTRTKQTGCS
Chr7_119241536_761_po_C_Xz11/1      SSGKVKHMNWTLKQLLKKYCQET-HLRWDQILPMVPLSL-MHPQQTNWVF
ChrY_22564970_567_po_C_Xz9/10       SSGKVEWMPNTQATTEEILSRNSSQVGSGFARS---PPLSQVPHYQTNWV
Chr21_14409346_1477_po_C_Pz5/6      SSGKVERMNKTPKQLLKKFCQET-HLRWDQVLPMVLLZVRCTPTTLNGYS
ChrX_51259700_1396_po_C_Pz7/8       SSRKVERMNQTLKQLWKKFCQET-HLRWDQVLPMVLLZVRCTPTKLTGYS
Chr6_64985396_1041_po_C_Xz11/1      GSGRVERMNWTLKQLLKKFCQET-HLRZDQVLPVVLLHVRCTPTKLTGYS
ChrY_20355209_659_po_C_Xz7/11       NSGKGERINWTLKQLLKKFCQET-HLRWDQVLPMFLLQVRRTPTKLTGYS
Chr3_34567166_523_po_C_Xz8/6        EFSKVERMNQTLKQLLKKFCQET-HLRWDQVLPMVLLRIRCTPTKLTGYS
Chr3_122072468_245_po_C_Xz9/9       SSGKVKRMNQTLKQLLK-FCQET-HLRWDQVLPMVLLZARZTPTKVTGYS
Chr11_58788475_541_po_C_Xz8/22      SYGKVERMNWTLKQLLKKFCKET-HLKWVQVLPMVLLZVRCTPIKLTGYS
Chr6_137421042_373_po_C_Xz6/9       EFWKVERINRTLKQPLKKFCHET-HLRWDQVLPMVLLZVRCTPTKLTGYS
Chr4_115910608_550_po_C_Pz8/12      SSEKAEKMNQTFKQLLKKFCQET-HLRWDQVLPMVLLRVRCTPTKLT-YS
ChrY_13743185_977_po_C_Xz14/8       RLGKVKQMKQTLKQLLKKFCQET-HLQWDQVLPMVLLQVRCTPTKQTRYL
HERV15                              SSGKAXRMNQTLKQ-LKKVLPET-HLQWDQVLPMVLLQVRCTPTKQTGYL
ChrX_117198092_983_po_C_Xz8/8       SSGKAEWMNQTLKQLLKMFCQET-HLRWDQVLPMVLLQVRCTLTEQIGYS
Chr19_21166030_1412_po_C_Xz9/1      SSEKVERMNQTLKQLLKNFASKL-TYDG-QVLPMVLLQVRCTPTKQTGYL
Chr19_21213743_913_po_C_Xz14/1      R-GMWQGHRIIVERRSADTHEQR-SLVFLGRGPCGLPQCLCPWVLEIR--
Chr19_15957774_569_po_C_Xz8/10      ------SVGR------FLLPHPVVLPLYAH--ALLLQALGISSRRRHP--
Chr11_77508681_297_po_C_Xz19/1      INVECISVSQRSKSDGFLLYKDSTYIHYERZNYRDEKQVGRNSVRRIEZV
Chr13_98250769_738_po_C_Xz7/13      -PLKDPAVEYFTDGSSFVLEEVRTCQVHSSNIDSVVEAQPLPTRTPAQKA
Chr6_39968183_513_po_C_Pz10/9       ----MQKLAHVSVQTWKGLSPDSLFRGWFSSFDGFKTLIGGFMLILGRRL
Chr15_66867263_645_po_C_Xz11/1      -------LPCHEGLNNSDAFYNSIYRKHQRGEMRQEIQEKZNZKEKNKLS
HERV18_polputein                    AAGMIEZYNGLLKNGLHLHVTPPSLRGWSSRLDLVLQILNEWPRKGGPAP
HERVIP10FH_polputein                GIKIIQEALSNIKWCLAFFGLKTIKIGAKGNFSVRRHQTIKSTADVPTFK
MER66_polputein                     RPRNTRSKRARINSGGNRGSNRT---------SSTLGWLCLPRVNPKNLT


HERVI_polputein                     PYEMLYGLPY-----LYSTANIPTFKIK---------------DQFLKNY
HERVIP10F_polputein                 PYEMLYGLPY-----LHSTADIPTFETK---------------DQFLKNY
HERVadp_AC005741                    PYKLLYGLPY-----LGRATDLPTMETK---------------DQFLRNY
HERV-PT47D                          ------NQEN-----FEKTRNLP--------------------RHILKNI
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      GGTERQDZ-------LDFLGQLRIPKPSGK------------VTASTFKH
HERV_Y_chr12_51022911_pol           AGLSPYEM-------LYGRP-LTNNLVLDR------------EMANLVAD
HERV17_polputein                    AGLSPYEM-------LYGRPFLTNDLVLDR------------ETANLVAD
HERV_W_chr7_9105739_syncytin_p      AGLSPYEM-------LYGRPFITNDLVLP-------------KTANLVAD
HERV9_polputein                     MGLSPYEM-------LYGQHFLTNDLLLDE------------EKTNLFKD
HUERS-P3_polputein                  LGXSPFEM-------LYGRXFLTNDFLLDQ------------ETSELVKH
ERVfrd_AC004022                     LQLIPFEA-------LYGKPFLYSDLMLDE------------KTAKITYC
PRIMA41_polputein                   LQLSPFEI-------MYGRPFLTTDFLIDI------------DTFKLQNY
HERV_H_RGH1_pol                     TGYSPFEL-------LYGRTFLLGPSLIPD------------TSPLGDYL
HERVHRGH2_pol                       QG-TPFEL-------LYGRTFLLGPNLIPD------------TSPLGDYL
HERV_H48I_polputein                 SFYSPFEI-------MYGRTFVLGPPPLPD------------SEPLGNYL
HERV_H_RTVLH2_pol                   QG-TAHLS-------SYRYSFLLGPSLIPD------------TRPI---W
HERVFH21_pol                        SFLSPFEL-------IYGRPFLLQ------------------NRP-----
HERVRblike_chr4_109047953           IKZSPCEI-------IFRRPFVANLSQVAR------------MSPDKELA
ChrY_13790029_693_po_C_Xz12/11      PFKIIYK------RVPPILPKLKDTHLAE------------ISQANLLHY
HERVS71_polputein                   PFEIMYG------GSS-ILPKLRDTNLAE------------ISQANLFS-
BaEV_M7_pol                         PYEILYG------GPPPLSTLLN-SFSPS------------NSKTDLQAR
HERVR_polputein                     PYEILYG------GPPPLSTLLN-SFSPS------------NSKTDLQAR
AKR_MLV_MLOCG_RT                    PYEILYG------APPPLVNFPDPDMTRV------------TNSPSLQAH
GaLV_pol                            PYEILYG------GPPPILESGETLGPDD------------RFLPVLFTH
Chr3_158587414_1662_po_C_Ez10/      PYEILYHR------PPPIPRELPGTPQE-------------LGEIELQRQ
Chr3_79627623_1138_po_C_Ez4/7       PYEKLYHR------PPSILWELPGTPRE-------------LGEIELQZQ
Chr19_20477244_1350_po_C_Ez10/      PYEILYHR------PPPILWGLPGIPZE-------------LGEIELQZQ
Chr5_111890459_1626_po_C_Ez10/      PZEIPYHR------PPPILLGLSGTPQV-------------SGEIELQRQ
ChrX_48766518_1600_po_C_Ez6/3       PYEILYHR------PPPILRGGFQALPE-------------LGEIELQRQ
ChrX_152580713_1572_po_C_Ez10/      PYDILYHR------PPPILWELPGTPZE-------------LGGIELQRQ
ChrX_139287533_1154_po_C_Xz8/4      SYEILYRR------PPPILRGLPGTPQE-------------LGEIELQRQ
Chr2_52615786_912_po_C_Xz11/5       PYVILYHIYIYIYMPPPILQGLPDTPZE-------------LGEIELQQQ
HERV-E_4_1                          PYEILYHR------PPPILRGLPGTPRE-------------LGEIELQRZ
ChrY_21962399_1137_po_C_Xz13/9      PYEILYHS------LLPILWGLPGTPRE-------------LGEIELQZQ
ChrY_8634032_950_po_C_Xz14/10       PYEILYHR------PPPILWELPGTPRE-------------LGEIELQRQ
```

```
Chr3_126735360_997_po_C_Xz5/10      TTAKFEG--------MLSSPHLP-RPRI-------------LETITQGZC
Chr3_126713203_958_po_C_Xz7/8       PPAKFEG--------MLSSPHLP-QPRI-------------LETITQGZC
Chr3_75226118_977_po_C_Xz8/9        PPAKFEG--------MLSSPHLP-RPRI-------------LETITQGZC
Chr7_6629919_827_po_C_Xz8/7         PPAKFEG--------MLSSPPLP-RPRI-------------LETITKRZC
Chr4_9166812_811_po_C_Xz8/10        TPATFQG--------MLSSPPIP-RPRI-------------LETITQGZC
Chr12_8478642_834_po_C_Xz12/7       PPAKFER--------ILSSPSLP-RPRI-------------LETITQGZC
Chr10_15187175_787_po_C_Xz11/5      PPAKFEG--------ILSSAPLP-RPRT-------------LETITQGZC
Chr11_71819740_854_po_C_Xz11/9      PPAKFEG--------MLSSPLLP-RPRI-------------LETITQGZC
Chr11_71519974_886_po_C_Xz13/9      DLLNLRE--------CYPLPHSP-GPDI-------------RD-NNKGVM
Chr8_7271441_969_po_C_Xz10/6        TTCZIGGN-------VILSPPPP-APDI-------------RDNNTEVMY
Chr8_11770066_826_po_C_Xz9/8        TTCZIGGN-------VILSPPPP-APDI-------------RDNNTGVMY
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      PPAKFEG--------MLSSPRLP-LPRI-------------LETVTQGZC
Chr14_50205291_837_po_C_Xz10/8      PPAKFEG--------MLSSPRLP-LPRI-------------LETVTQGZC
Chr7_6597120_842_po_C_Xz7/7         HLLHLRE--------CYPLPHPP-VPDI-------------RGNNTGVMY
Chr11_71544658_783_po_C_Xz8/13      PPATFKG--------MLSSPPLH-PSRI-------------LEAKTQGZC
Chr11_67999520_760_po_C_Xz8/8       PPATFEG--------MLSSPPLP-RPRI-------------LEAITQGZC
Chr3_75461975_796_po_C_Xz7/9        PPAKLEG--------MLSSPPRPQRPRI-------------LETVTQ-VC
Chr12_52228939_692_po_C_Xz3/10      HLPNLREC-------YSVLPPPP---DI-------------RDNNIGVMN
Chr10_15211234_286_po_C_Xz11/1      TPAKFGN--------VILFPAPP-APDI-------------RDNNTRVMY
Chr12_100973302_945_po_C_Xz12/      QDGRIGTA-------PVYSSQHE-RHRR-------------QVISAFPTE
Chr19_9685420_817_po_C_Xz6/10       LAQ---IV-------PLHSSLED-GTR-----------------------
Chr13_40936820_719_po_C_Xz9/10      TQLLHDGS-------GSLSAKQSVELIF-------------RGAYKTVPS
Chr7_63858202_1754_po_C_Rz12/9      PYEILFG------KPSPIISQIKGNLR-------------ELGELTLRRQ
ERV3_chr7q21_AC073210_polputei      PYEILFG------KPSPIISQIKGNLR-------------ELGELTLRRQ
Chr6_110218096_843_po_C_Xz16/8      PHEILFG------QPPSIIGQIKGDVH-------------ELGELTLRRQ
Chr6_110214535_786_po_C_Xz16/8      PHEILFG------QPPSIIGQIKGDVH-------------ELGELTLRRQ
ChrY_3018841_881_po_C_Rz9/12        PMRSCSA------GHPPSZ-------------------------VRLRVT
ChrX_86876098_639_po_C_Xz7/8        PYEILFG------WPPPIIGQIKGDLR-------------ELGELTLRRQ
Chr14_68899917_965_po_C_Xz6/8       PYEILFS------QPPPIIGQIKGDLR-------------ELRDLTLRKQ
ChrY_23874437_661_po_C_Xz15/9       KLGVTL-------QTR---------------------------ELTLRRQ
ChrY_27518418_649_po_C_Xz15/9       KLGVTL-------QTR---------------------------ELTLRRQ
ChrY_25016892_657_po_C_Xz13/12      KLGVTL-------QTR---------------------------ELTLRRQ
ChrY_19932937_520_po_C_Xz10/14      SYGIFC-------RPPPIIAQIRGDIC-------------ELGKLTLRRQ
ChrY_20104614_518_po_C_Xz10/14      SYGIFC-------RPPPIIAQIRGDIC-------------ELGKLTLRRQ
Chr19_58660155_598_po_C_Xz21/1      IHLMEFVQ-----SASPIISQIKSDLQ-------------ELEELTLRKQ
Chr19_58621153_500_po_C_Xz21/1      PYZSLFSR------PALIISQIKGDLQ-------------ELEELTLSKQ
Chr7_139686987_710_po_C_Xz11/7      ---------------SSFILGIS---------------------------
Chr7_137651301_1097_po_C_Pz13/      PYEILFG------WPSPIIGQIKGDLH-------------ELGELTLRRQ
Chr11_29579197_930_po_C_Pz8/12      -YEILFG------QPPPIIGQINCDLR-------------ELGKLSLRRQ
Chr4_117329205_488_po_C_Xz3/16      PYEILFS------RPPGIRGQLKGDLR-------------ELGELTLRKQ
Chr21_38522185_383_po_C_Xz7/9       PYEILFG------ATP-IIGQVKGDLH-------------ELGELTLRRQ
Chr4_70173370_1049_po_C_Xz11/8      PNEILFS------QPPPIISQIKGNLR-------------ELQKLTLRRQ
Chr4_70022513_788_po_C_Kz10/12      PCEILFG------WPPPIIRQIKGNLR-------------ELQKLTLRRQ
Chr14_105225894_1027_po_C_Xz13      PYEILFG------RPPPITGQIKGDLC-------------ELR--TLRKQ
Chr7_119241536_761_po_C_Xz11/1      YLETLFS------QPPPIIGQIKGDLC-------------ELGKLTLRRQ
ChrY_22564970_567_po_C_Xz9/10       FTLGDFVQ-----QTTPSITQVKGDLH-------------DLGELNLTRQ
Chr21_14409346_1477_po_C_Pz5/6      PYEIVYS------RPTALISQVKGNLK-------------EIGELTLRRQ
ChrX_51259700_1396_po_C_Pz7/8       PYEIVYG------RPPQLIAQIRGDLK-------------EIGELTLRRQ
Chr6_64985396_1041_po_C_Xz11/1      PCEIVFG------ZPPSIIIQIKGDLK-------------ZIGELTLRRQ
ChrY_20355209_659_po_C_Xz7/11       PYEIVFG------QPPPIITQLKGDLK-------------EIGKLTLRRZ
Chr3_34567166_523_po_C_Xz8/6        LYEIFS-------RPPPIITQIKGDLK-------------EIGGLTLRRQ
Chr3_122072468_245_po_C_Xz9/9       PYEIVFG------QAPLMITQIKGDLR-------------ESGELTLRRQ
Chr11_58788475_541_po_C_Xz8/22      PYEIVFA-------DPNHNSDKRGFZ--------------KIGELTLRRQ
Chr6_137421042_373_po_C_Xz6/9       PYKIVFG------RHPDHNSDKRGF---------------KIGELTLRRQ
Chr4_115910608_550_po_C_Pz8/12      RYEIVFS------ZPPAINSDKRGLEQDTVSKKKKKKRKKEIGGLTLRRQ
ChrY_13743185_977_po_C_Xz14/8       SYDILLR------RPSPII-QIRGDLK-------------EVGELTLKRQ
HERV15                              SYEILFG------RPPPIINQIRGDLK-------------ELGELTLKRQ
ChrX_117198092_983_po_C_Xz8/8       PYEILFG------RPPPIINQIRGZLK-------------ELGELTLRRQ
Chr19_21166030_1412_po_C_Xz9/1      PYEILFR------RPPAIINQIRGHLK-------------ELGELTLRTQ
Chr19_21213743_913_po_C_Xz14/1      --EVMTL------NEHAAFKHLFNKVH-------------LAPPLIFNPE
Chr19_15957774_569_po_C_Xz8/10      ---------------------------------------------SLEPY
Chr11_77508681_297_po_C_Xz19/1      NHRTFZDG-----DIILCNTIMSDMIHDT-----------LHLSEPLEPK
Chr13_98250769_738_po_C_Xz7/13      ELIALTRA-----PLLAKDKKVNIYTDSK-----------YAFTTLHIHG
Chr6_39968183_513_po_C_Pz10/9       ILPCLLP----------------------------------LLIRSIQST
Chr15_66867263_645_po_C_Xz11/1      ILGZHPRG-----------------------------------SSRQSPD
HERV18_polputein                    VEALLHR------------------------------------------A
HERVIP10FH_polputein                TKDQFLRN----------------------------------YILGLSST
MER66_polputein                     QPESLATN----------------------------------TGQALKRI


HERVI_polputein                     ILGPYSTFSSLKTK---GLLAQAPPLEFPAHZRQPGDHVLIKGW-----K
HERVIP10F_polputein                 ILGLSSTFSSLKTK---GLLAQAPPLEFPVHQHQPGDHVLIKSW-----K
HERVadp_AC005741                    LLAISSTLSSLRLK---GLLIQVP-LEFMVT-FQLGDLVLVKTW-----K
HERV-PT47D                          IKLZDESIMLPQPE---KQIAQQVKKKKRVHKGTEINREICYRR-----K
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
```

```
Chr13_54564403_572_po_C_Yz7/14      GACNLTHTQPIRELTK-------MLTRQKQEVKKZPIIYCLKAQREGQGS
HERV_Y_chr12_51022911_pol           VTCLAKYQQVLKTLQG-------ACPREEGKELFHPGDMLLAKSLPFNSP
HERV17_polputein                    ITSLAKYQQVLKTLQG-------TCPREEGKELFHPGDMVLVKSLPSNSP
HERV_W_chr7_9105739_syncytin_p      ITSLAKYQQVLKTLQG-------TYPZEEGKELFHPCDMVLVKSLPSNSP
HERV9_polputein                     ITSSAKYQQNPKNLPE-------GCHREKGTELFQPGDLVLVKSLPSTSP
HUERS-P3_polputein                  VTSLAHFQQELTQLAE-------AQPQEIGPPLFNPGXLVLVKXTXSLSF
ERVfrd_AC004022                     VSSLAGFQQALQEFWLQR---KPKSEGEKYQPLYPPGSLVLIKAWRDGTL
PRIMA41_polputein                   VINLGQVQNALLEYGNQR---LPSPTKEENLVTTQPGDWVLLKTWKEGSP
HERV_H_RGH1_pol                     PVLQQARQEIHQASNL-----LLPTPDRQ---PYEDNLAGRSVVVKNLTP
HERVHRGH2_pol                       PVLQQARQESHQAANL-----LLPTPDPQ---PYEDTLAGRSVLVKNLTP
HERV_H48I_polputein                 PSLIQTRSFIREAANE-----AMPLPVDTSLSSQHNCLAGTDVFICQPDP
HERV_H_RTVLH2_pol                   TVLQKTCH------------------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           IKTYATHLEQTLNPLHKFSSNRNVIDSVEAHHPFQSGDQVLQKEWKEAGP
ChrY_13790029_693_po_C_Xz12/11      LQSLQKVQDIIQPLGLRAHPNPVPDQMGPCHSFQPGDLVFVKKS-----R
HERVS71_polputein                   TCSLSNRYETIQPLVQGAHSNPVPDQTGP-HFFQPGDLVZVKKF-----Q
BaEV_M7_pol                         LKGLQAVQAQIWAPLAELYR-PGHR---TSHPFQVGDSVYVRRH-----R
HERVR_polputein                     LKGLQAVQAQIWAPLAELYR-PGHR---TSHPFQVGDSVYVRRH-----R
AKR_MLV_MLOCG_RT                    LQALYLVQHEVWRPLAAAYQ-EQLDRPVVPHPYRVGDTVWVRRH-----Q
GaLV_pol                            LKALEIVRTQIWDQIKEVYK-PGTVT--IPHPFQVGDQVLVRRH-----R
Chr3_158587414_1662_po_C_Ez10/      LQALGKITQTISTW---VNERCPISLFSPVHPFSP-GDHVWIKD----WN
Chr3_79627623_1138_po_C_Ez4/7       LQALGKITQTISTW---VNERCPINLFSPVHPFSP-GDRMWIKD----WK
Chr19_20477244_1350_po_C_Ez10/      LQALGKITQTISTW---VNERCPISLFSPVHLFSP-GDCIWIKD----CN
Chr5_111890459_1626_po_C_Ez10/      LQALGKITQTISTW---VNERYPISLLSPVHPFSP-GDRMWIKD----WN
ChrX_48766518_1600_po_C_Ez6/3       LQALGKITQTISTW---VNEKCPISLFSPVHPFSP-GDHVWMKD----WN
ChrX_152580713_1572_po_C_Ez10/      LQALGKITQTISTW---VNERCPISLFAPIHTFSP-GGLVWIKD----WN
ChrX_139287533_1154_po_C_Xz8/4      LQALGKITQTISTW---VNERCPRSAYSPVHPFSP-GDCVWIKD----WN
Chr2_52615786_912_po_C_Xz11/5       LQALGKITQKISAW---VNERHPISLFSPVHPFSP-GDQVWIKD----WN
HERV-E_4_1                          LQASGKITQTISAW---VNERCPVNLFSPVHPFSP-GDLVWIKD----ZN
ChrY_21962399_1137_po_C_Xz13/9      LQALGKITHNFSLG---KWE-VPHQLISPVHPFSP-GDQVWIKD----WN
ChrY_8634032_950_po_C_Xz14/10       LQALGKITQTISTW---VNERCPISLFCPVHPFSP-GDCLWTKD----WN
Chr3_126735360_997_po_C_Xz5/10      TPTALLGVISSSAF---LDIRNNVTLCVYACR---------EIQ----WN
Chr3_126713203_958_po_C_Xz7/8       TPTALLGVISSLPS---WILG---TISHCACTP--------VAK----FN
Chr3_75226118_977_po_C_Xz8/9        TPTALLGVISSSPF---LDIRNNITLCVYACR---------EIQ----WN
Chr7_6629919_827_po_C_Xz8/7         MPTALLGVISSSPF---LDIRNNITLCVYACR---------EIQ----WN
Chr4_9166812_811_po_C_Xz8/10        TPTALLGVISS-PF---LDIRNNITLCVYACR---------EIQ----WN
Chr12_8478642_834_po_C_Xz12/7       TTTALLGVISSSPF---LDIRNNITLCVYACR---------EIQ----WN
Chr10_15187175_787_po_C_Xz11/5      TPTALLGVISSSPF---LDIRNNITLCVYACR---------EIQ----WN
Chr11_71819740_854_po_C_Xz11/9      TPTALLGVISSSPF---LDIRNNITLCVYACR---------EIQ----WN
Chr11_71519974_886_po_C_Xz13/9      YTHALLGVISSSPF---LDIRNNITLCVYACR---------EIQ----WN
Chr8_7271441_969_po_C_Xz10/6        THCFIVSNIPLPSW---IVGTVSHCACRPVAK----------FN----GM
Chr8_11770066_826_po_C_Xz9/8        THCFIGSNIILSFL---GYZQQYHTVRVGVSR----------NS----VE
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      TPTALLGVMSSSPF---LDIRNNITLCVYACR---------EIQ----WN
Chr14_50205291_837_po_C_Xz10/8      TPTALLGVMSSSPF---LDIRNNITLCVYACR---------EIQ----WN
Chr7_6597120_842_po_C_Xz7/7         T-HCFIGINVISAF---LGIRNNITLCVYACH---------EIQ----WN
Chr11_71544658_783_po_C_Xz8/13      TPTPLLAVMSSSAF---LDIRNNITVCVEACR---------EIQ----WN
Chr11_67999520_760_po_C_Xz8/8       TPTALLGVMPSSAF---LDIRNNITVCVYACR---------EIQ----WY
Chr3_75461975_796_po_C_Xz7/9        TPTALLGVIHPLPS---WILG-TVSHCACTPV---------AKF----NG
Chr12_52228939_692_po_C_Xz3/10      PLLYWEKYHPLPSW---ILGTISQCACTPVTK-----------Q----WN
Chr10_15211234_286_po_C_Xz11/1      NHCFIGSNIISSLL---GYZEQYHTVRVRLSRN--------SME----CH
Chr12_100973302_945_po_C_Xz12/      VLGSSHZGVSDSGC---RTVGAVHNVCQSRARHHLIREAQGVRE----YP
Chr19_9685420_817_po_C_Xz6/10       -------------------------LCLKKN-------------------
Chr13_40936820_719_po_C_Xz9/10      LTTAIIYQTNEEAL---YPGEIGWLLSAAPH-------------------
Chr7_63858202_1754_po_C_Rz12/9      MQALGIAMZSVHGW---VQKNAHKPDR-PIHPFKPRDSLWVKKZ-----N
ERV3_chr7q21_AC073210_polputei      MQALGIAMZSVHGW---VQKNAHKPDR-PIHPFKPRDSLWVKKZ-----N
Chr6_110218096_843_po_C_Xz16/8      MQALGIGMQGVWLG---TGENTHKPDR-LSMPFKPGDSVWVEKW-----N
Chr6_110214535_786_po_C_Xz16/8      MQALGIGMQGVWLG---TGENTHKPDR-LSMPFKPGDSVWVEKW-----N
ChrY_3018841_881_po_C_Rz9/12        SEN-----------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        MQTLGIAMQGVHGW---VWERMPISLRDPVHPFKPGDSVWVMKW-----N
Chr14_68899917_965_po_C_Xz6/8       MQGLGIAMQDVYGW---VRERMPISLADPAHPFKPGDSVWVKKL-----N
ChrY_23874437_661_po_C_Xz15/9       RQALKRAMQEVHGW---VWERLSISPTNTAHHFRPGDSVZVRN------K
ChrY_27518418_649_po_C_Xz15/9       RQALKRAMQEVHGW---VWERLSISPTNTAHHFRPGDSVZVRN------K
ChrY_25016892_657_po_C_Xz13/12      RQALKRAMQEVHGW---VWERLSISPTNTHTTLDLGTLCRSEI------K
ChrY_19932937_520_po_C_Xz10/14      MQALKTAMQVVHGZ---VWEKLPISLAKPAYSFRPGASVHIRN------K
ChrY_20104614_518_po_C_Xz10/14      MQALKTAMQVVHGZ---VWEKLPISLAKPAYSFRPGASVHIRN------K
Chr19_58660155_598_po_C_Xz21/1      MQALIIAMQEAHNN---MRKNADKS-DRSSTPLKPGDSVZVKKZ-----N
Chr19_58621153_500_po_C_Xz21/1      MQALRTAMQEVHNZ---IQERMPISLTDQVHPFQHGDSVZVKKZ-----N
Chr7_139686987_710_po_C_Xz11/7      -----------PNA---IPPPFPHPQQAPVCDIP----LPVSKC-----S
Chr7_137651301_1097_po_C_Pz13/      MQALGISHAKNPWL---GTGKNAHKPNRPIHPFRPGDSVZVKK-----WN
Chr11_29579197_930_po_C_Pz8/12      MKALGIAMQEIHGW---VWERMPISLTDPTHPFKPRDSVWVKK-----WN
Chr4_117329205_488_po_C_Xz3/16      MQSLRIAMQDVHGZ---AWKRMPISLTDPHTPLTLATLFRSKKSIQPLWD
Chr21_38522185_383_po_C_Xz7/9       MQALGIAMQCVYGW---VWERMPISLTDLAHPFKPGDAIWVKK-----WN
Chr4_70173370_1049_po_C_Xz11/8      MQALDMAMQKVYGZ---AQERMPISLADPVHPFKPEDFVZVKKZ-----T
Chr4_70022513_788_po_C_Kz10/12      MQALGMAMQKVNGZ---TQERMPISLTDPVHPFKPEDSFZVRKZ-----N
Chr14_105225894_1027_po_C_Xz13      MQALGLAMQKVHGW---VQERMPIGLTDPVHPFKPGDLVGVKKW-----N
Chr7_119241536_761_po_C_Xz11/1      MQALGIVMQEIHGZ---VRERMPISLTEPANPFKPGD-VWVKQW-----N
```

```
ChrY_22564970_567_po_C_Xz9/10       IEALGLAMQKVHGW---VREMMPVSLTDPICPLKSGDLVZVMKQ-----N
Chr21_14409346_1477_po_C_Pz5/6      MQALGEVIQEVQGW---VRETIPVSLTDAIHPFKPADSVWVKR-----WN
ChrX_51259700_1396_po_C_Pz7/8       MQALGEVMQEVQGW---VRQRIPVSLTDAVQPFQPRDSVWVKR-----WN
Chr6_64985396_1041_po_C_Xz11/1      MQALSETMZZIQGW---VRKRIPVSLTDAVYPFQPGDSVWVKZ-----WN
ChrY_20355209_659_po_C_Xz7/11       MQALGEDMQETEGZ---VRERIPLSLRDAIHPFQP--TLLHQM-----ME
Chr3_34567166_523_po_C_Xz8/6        MQALGEAVQEIQGW---VRERIPVSLTDAVHHFQPGDSVZIKZ-----WN
Chr3_122072468_245_po_C_Xz9/9       MQALGEAIQEIZGZ---VRERIPVSLTDAVHPFQPGDSGSNN-------R
Chr11_58788475_541_po_C_Xz8/22      MQALGEAMQEIQGW---VRERIPVSLTDAVHPFQPGDSVWIKR-----WN
Chr6_137421042_373_po_C_Xz6/9       MQALGEVSQEMQGW---VRERIPVSLTDAVQPFZPGDSVWVKQ-----WN
Chr4_115910608_550_po_C_Pz8/12      MQALGEAMQEIQGW---VRKRIPVSLTDAVHPFQPEDSVRSND-----E-
ChrY_13743185_977_po_C_Xz14/8       MQALGVAMKEVQSZ---VRERIPISLTDLVHPHKPGDSVWVKR-----WN
HERV15                              MQALGVAMQEVQSX---VRERIPISLTDPVHPHKPGDSVWVKR-----WN
ChrX_117198092_983_po_C_Xz8/8       MQALGVAMQEVQGW---VSERIPLSLTDPVHPHKPGILSGLKCGIQ--WN
Chr19_21166030_1412_po_C_Xz9/1      MQAZGVAMQEVKGW---VRERILISLRDVVHPHKPGDSVWVKR-----WN
Chr19_21213743_913_po_C_Xz14/1      LTQH--MFQKAQGWG--ZGHRLTASQGRRIFLSTEQNGVSYVY-------
Chr19_15957774_569_po_C_Xz8/10      PCPNVRRIPNMAQTS--PAEEKDYLK------------------------
Chr11_77508681_297_po_C_Xz19/1      VCKFKKIFRSSDRZS--QDEIKNVTKZSNCIMNKTQZRKWEKIADLSNLE
Chr13_98250769_738_po_C_Xz7/13      AINKERGLLTAGGKE--IKYKEEILQLLQAAEGPEK--------------
Chr6_39968183_513_po_C_Pz10/9       IEAIVARHTTTQIMA--LTKYQPLPKEK----------------------
Chr15_66867263_645_po_C_Xz11/1      PGSLDNIFKEARAQR--DVLWRLSQHSLN---------------------
HERV18_polputein                    AAPIQLQIHTKDDL------------------------------------
HERVIP10FH_polputein                FLSLKTKSLLAQVP----------PLEFLVNQHQPEDHVLIKG-----WK
MER66_polputein                     QESLDSLANVVLDN--------RLALDYLLAE-----------------Q


HERVI_polputein                     EGKLKPAWEGPYLVLLNTKTAVRTANEDG---------------------
HERVIP10F_polputein                 EEKLEPAWEGPYLVLLTTETTVRTAERGW---------------------
HERVadp_AC005741                    EDKLHPSRDGPYQVLLTTETAMRTAEWGWTHYTGQMTGKRDPRRERERPV
HERV-PT47D                          NFKSD----GHVINGIGTISSIPEK-------------------------
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      G--YEPRHSSRQQQPLWVPSLPMGALFSLYFTLLNLATALFSSMFVTARA
HERV_Y_chr12_51022911_pol           S--LDTSWDGPYPVILSTPTSVKVAGVESWIPHTRAKSWILPKTEKSRRQ
HERV17_polputein                    S--LDTSWEGPYPVILSTPTAVKVAGVESWIHHTRVKPWILQRNPKIQET
HERV_W_chr7_9105739_syncytin_p      S--LDTSWEGPYPVILSTPTAVKVAGVESWIHTLESNPGYCQRNLKIRRQ
HERV9_polputein                     S--MDSLWEGSYSIILSTHTAVKVVGVKSWIHH-----------------
HUERS-P3_polputein                  S--KPX-QGGPHTVLLPTPPAVKVTGIDSWIHHTQVKAWKAEGATPDSPE
ERVfrd_AC004022                     N--SQLTPVRRG---LFTLRAIKVSEIVSWIHHLSEAMERLQNTRSGSNA
PRIMA41_polputein                   ADQLSPKWKGSYQVLLSTPTAVKLLGINSWVHLSELNLSLMKSHRPVEHK
HERV_H_RGH1_pol                     -QTLQPRWTGPYLVIIVP-----------------------DCRPPAGSS
HERVHRGH2_pol                       -QTLQPRWTGPYLVIYSTLTAVRLQDPPHWVHHSRIKLCPLDSQPNPSSS
HERV_H48I_polputein                 HKKLQPKWTGPYTVILSMPTAVRVQGLPHWVHRTR---VKLTPKATPSSK
HERV_H_RTVLH2_pol                   ----------PYYLLSS---------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           AQQLQEKZKEPYDGLLTTDSALKLAGIKPWVHHMWVKIPVAQELHHRNTC
ChrY_13790029_693_po_C_Xz12/11      KK--DSCLEKPHTVILTTPMALKVEGMLTFMTPPSTKHV-----------
HERVS71_polputein                   KEGFTPAZKGPHTVILTTPMALKVDGIPSWIHLSHQKGEQSPARNMDPQA
BaEV_M7_pol                         SQGLEPRWKGPYIVLLTTPTAIKVDGIATWIHASHAKAAPGTPGPTSSGT
HERVR_polputein                     SQGLEPRWKGPYIVLLTTPTAIKVDGIATWIHASHAKAAPGTPGPTSSGT
AKR_MLV_MLOCG_RT                    TKNLEPRWKGPYTVLLTTPTALKVDGIAAWIHAAHVKAADPGGGPSSRLT
GaLV_pol                            PSSLEPRWKGPYLVLLTTPTAVKVDGIAAWVHASHLKPAPPSAPDES---
Chr3_158587414_1662_po_C_Ez10/      RAPLRPWWKGPQTVIWTTPTAVKVEGIPAWIHHSHVKP-----AAAETWE
Chr3_79627623_1138_po_C_Ez4/7       VAPLWPWWKGPQTVILTAPMAVNVEGIPAZIHHSHVKP-----AAAETWE
Chr19_20477244_1350_po_C_Ez10/      IAPLRPRWKGPQTVILTTPTAVKVEGIPAWIHHSCVKP-----AAAETWE
Chr5_111890459_1626_po_C_Ez10/      IAPLWPWWKGPQTIILTTPKAVKVEGIPSWVHHSHVNP-----AAAETWE
ChrX_48766518_1600_po_C_Ez6/3       VAPLRPQWKGPQTVILTTPKAVKVEGIPAWIHHSQZNLGGE-TKPGQPLQ
ChrX_152580713_1572_po_C_Ez10/      VAPLRPRWKGPQTIFLTTSMAVKVEGIPAWIHHSRVKP-----AATETWE
ChrX_139287533_1154_po_C_Xz8/4      LAPLRPZWKGPQTLNLTTPTAVKVEGIPAWIHHGRIKP-----AATGTWE
Chr2_52615786_912_po_C_Xz11/5       VVPLRPWWKGPQTIVLTTPTAIKIEGIPAWIYHSRIKP-----AAPETWE
HERV-E_4_1                          VACLCPRWKGPQTVILSTPTAVKVEGIPTWIHHSRVKP-----AVPETWE
ChrY_21962399_1137_po_C_Xz13/9      IAPLWPQWKGPQTNILTTPTAI------AWIHHSCIKP-----AAPVTWK
ChrY_8634032_950_po_C_Xz14/10       VAPLGPRWKGPQIIILTTPMAVNVEGIPAWIHHSCVKP-----AASETWE
Chr3_126735360_997_po_C_Xz5/10      VILRLPG-------------------------------------------
Chr3_126713203_958_po_C_Xz7/8       EMSSAPPWI-----------------------------------------
Chr3_75226118_977_po_C_Xz8/9        VILCLPGYDEQDHGGCTTSEILGVISSSP-------------LWKLGTIS
Chr7_6629919_827_po_C_Xz8/7         VILRLPGYDEQYHRGCTTSEILGVTSSSS-------------LWKLGTIS
Chr4_9166812_811_po_C_Xz8/10        VILRLPGYDEQYHRVCTTSEILGVISSSP-------------LWKLGTIS
Chr12_8478642_834_po_C_Xz12/7       VILRLPGYDEQYHGGCTTSEILGVIPPSP-------------LWKLGSIS
Chr10_15187175_787_po_C_Xz11/5      VILRLPGYDEQYHGGCTTSEILGVIPPSP-------------LWKLGTIS
Chr11_71819740_854_po_C_Xz11/9      VILRLPGYDEQYHGGCITSEILGVISSSPFELRDNISGIVYTFWDIGISS
Chr11_71519974_886_po_C_Xz13/9      VILRLPGYDEQYHGGCTTSEILGVISSSS-------------LWKLGTIS
Chr8_7271441_969_po_C_Xz10/6        SSCASLDMTNNITG-DTTSEILGVMSSSP-------------LWKLGTIS
Chr8_11770066_826_po_C_Xz9/8        CHPAPSWIDEQYHGGCTTSEILRMISSSP-------------LWKLGTIS
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      VILRLPGYDEQYHGGCTTSEILGVISSSP-------------LWKLGTIP
Chr14_50205291_837_po_C_Xz10/8      VILRLPGYDEQYHGGCTTSEILGVISSSP-------------LWKLGTIP
Chr7_6597120_842_po_C_Xz7/7         VILCLPGYDEQYHGGCTTSEILGVISSSP-------------LWKLGTIS
```

```
Chr11_71544658_783_po_C_Xz8/13      VILCLPGYDEQYHGACKTSEILEZSHP------------------LPSRS
Chr11_67999520_760_po_C_Xz8/8       VILCLPGYDEQYHRGCTTSEISGVISSSS-------------LWKLGTIS
Chr3_75461975_796_po_C_Xz7/9        MSSCASLYDEQYHGGCTTSEILGVISSSA-------------LWKLGTIS
Chr12_52228939_692_po_C_Xz3/10      VILCLPGYDEQCHGGCTTSEILGVIPSSP-------------LWKLGTIS
Chr10_15211234_286_po_C_Xz11/1      PVLPWIMN-SKYHGGCTTSEILGVTSSSP-------------IWNLGTVS
Chr12_100973302_945_po_C_Xz12/      FLVKERGDRRHLENLVTPTLMLHFFKGLS---------------KRHTRR
Chr19_9685420_817_po_C_Xz6/10       -LIIEKZNNNCHQNNVS---VFKWVKSLN---------------MLENV-
Chr13_40936820_719_po_C_Xz9/10      --------------------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      PTTLGPIWDGLHTVILSIPTVVKVAGIVPWIHPSSQLKPAAQDKWTSQQD
ERV3_chr7q21_AC073210_polputei      PTTLGPIWDGLHTVILSIPTVVKVAGIVPWIHPSSQLKPAAQDKWTSQQD
Chr6_110218096_843_po_C_Xz16/8      PTTLGPIWECPHTVMLATPTAIKVTGIVPWIHHS-RLKPAIQDKWTSPQD
Chr6_110214535_786_po_C_Xz16/8      PTTLGPIWECPHTVMLATPTAIKVTGIVPWIHHS-RLKPAIQDKWTSPQD
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        PTTLGPIWVGLHTVILSTPSIVKVAGIVPWIHHS-WLKPEIQDKWTSQQN
Chr14_68899917_965_po_C_Xz6/8       PTTLGPLWDGPHTVILSTPTAVRVAGIVPWIHPS-WQKPAVQDKWTSQQD
ChrY_23874437_661_po_C_Xz15/9       PTTLEPIWDGPHPATLSTSTAVNIAEIVRWIHHS-QLKLAAQDKWTSKHD
ChrY_27518418_649_po_C_Xz15/9       PTTLEPIWDGPHPATLSTSTAVNIAEIVRWIHHS-QLKLAAQDKWTSKHD
ChrY_25016892_657_po_C_Xz13/12      PTTLEPIWDGPHPATLSTSTAVNIAEIVRWIHHS-QLKLAAQDKWTSKHD
ChrY_19932937_520_po_C_Xz10/14      PTTLGPTWDGPYTLTLSTSTAVKVAETVRWILHS-QLKLAAND-------
ChrY_20104614_518_po_C_Xz10/14      PTTLGPTWDGPYTLTLSTSTAVKVAETVRWILHS-QLKLAAND-------
Chr19_58660155_598_po_C_Xz21/1      PTSLRPIZNGPYIVILSTPTAIKFAGIMPZIHHS-QLKLAAQGKWTSQQD
Chr19_58621153_500_po_C_Xz21/1      PTSLGLIZDGPCTVILSTPTAVKFAGIMPZVHHS-QLKPAVQTSGTASK-
Chr7_139686987_710_po_C_Xz11/7      HCSIPTYEZ-----------EHAVFGFCP--CNS-LLRMMVSASSMSLQR
Chr7_137651301_1097_po_C_Pz13/      PTTLGPIZDGPHTVILSTPTAVKVAGTVPWIHHS-QLKPAAQDKWTSQQD
Chr11_29579197_930_po_C_Pz8/12      PTTLGPIZDGPHTVILSTPTAVKVGGILPWIHHS-QLKPAAQTSGPANRN
Chr4_117329205_488_po_C_Xz3/16      PWVHTVILSTPTA-------------------------------------
Chr21_38522185_383_po_C_Xz7/9       PTTLRPIWDRPHT-------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      PTILSPIWDGPHIVILSTPTVVKIAGIVPWIHHS-QLKLSAQDQWTNQQD
Chr4_70022513_788_po_C_Kz10/12      PPTLGPIWDGPHTVVLSTPTAV--AGIIPWIHHS-WLKPAAQDVDQPVGP
Chr14_105225894_1027_po_C_Xz13      -HNLGPICDGPHVVILSTPTAVK-VEIVPWIHHD-WLKPAAQDQRTSZQD
Chr7_119241536_761_po_C_Xz11/1      PTTLGPIWDGLHTIILSTPTAVKVAGIVPWIYHS-ZLKPEARDKWTSHWD
ChrY_22564970_567_po_C_Xz9/10       PTTLGPLWDGPHAVILSTSTAVKVAGIILWVHHS-WPKLAALDKWTSQED
Chr21_14409346_1477_po_C_Pz5/6      PTTLGPLWDGPHIVIMSTPTTVKVASITPWIHT--ADKPAASV--QDQWT
ChrX_51259700_1396_po_C_Pz7/8       PITLRPLWDGPHVVIMYRGKQPDIRRNSPPIFH--VVSFLFSL--VSAGL
Chr6_64985396_1041_po_C_Xz11/1      PTTLGPLSDGPHIVILST-TAVEVAGITPWVHHSQLKPAVTTT--QDQWT
ChrY_20355209_659_po_C_Xz7/11       PSHLEPLWDGPHTVTLSTPTAVKVAGITPRVHHSQLKPAVTAT--QDQWP
Chr3_34567166_523_po_C_Xz8/6        PSTLRPLZDGPHIVILSTPTAVKVSGITPWVHHSWLKPAVATS--QDQWT
Chr3_122072468_245_po_C_Xz9/9       TQPLGPLWDGPHIVILSIRTAVKVAGIIPWVHHSWLKPAAAAI--QDQWT
Chr11_58788475_541_po_C_Xz8/22      PTTFGPLWDGPHIVILPTPAAVKVAGIIPWVLHSQLKPTAATA--QDQWT
Chr6_137421042_373_po_C_Xz6/9       PTTLGPLWDSPHIVILSTPTAVKVAGIIPWVHHSRLKPEAATT--QDQWT
Chr4_115910608_550_po_C_Pz8/12      TQPLGPLWDGPHIVILSTTTAVKVAGITPWVHHSQLKPAAATA--QNQWT
ChrY_13743185_977_po_C_Xz14/8       PTTWGPLRDGPHIVIMSTPTALKVAGVTPWIHH-LAETSSSSDSDETSGL
HERV15                              PTTLGPLWDGPHIVIMSTPTALKVAGVTPWIHHSWLKPAAAVTPDDDQWI
ChrX_117198092_983_po_C_Xz8/8       PTTLGPLZDGLHIVIMSTPTAVKVAGVTPWIHHSWLKPAAAVTPHDDQWT
Chr19_21166030_1412_po_C_Xz9/1      PTTLGPLWDGPHIVIMSTPTAVKVAGVTPSIHHSQPKPVAAVTPDNDQEL
Chr19_21213743_913_po_C_Xz14/1      ---FFLHRHSNNLISLSFPHTSPFSSTKPPLSS-WPILDGCCLFGAVGYT
Chr19_15957774_569_po_C_Xz8/10      -------QNRRQTP-------------------HPQILLATAEP------
Chr11_77508681_297_po_C_Xz19/1      MGGSCKTQCKRSTDZGAIWPSRGLQKLSSLQEHQIEQLFTQKTPSLQPKI
Chr13_98250769_738_po_C_Xz7/13      ---VAVMHYRRHQKAR-----------------MPEAKGNRKADREAKRA
Chr6_39968183_513_po_C_Pz10/9       -----YVPTKEETNDS-----------------VPSIKIHLKRHQRGE--
Chr15_66867263_645_po_C_Xz11/1      ------IRIRKSEVPP--------------QIPLPSVZLYLALQAL----
HERV18_polputein                    ---LRPGMGTNGNLLLPAPTPLKRGTE-----------------------
HERVIP10FH_polputein                EGKLEPAWEGPYLVLLTTETAVYSGKGMDSSHPSQESAAPSRVVGHSPRG
MER66_polputein                     GGVVQLLIKPAAHILTTLDVEVNIQKIYEQATWLHRITRA----------


HERVI_polputein                     -----------LITPASKRRCHLQ---------------------------
HERVIP10F_polputein                 -----------LTTPESRKRHPLQSRG------------------------
HERVadp_AC005741                    KSARVTQEPLKLTQKNLKRKYGLASSGS----------------------
HERV-PT47D                          ----------------DKRNLDLHIT------------------------
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      ELFARRPPLLLPLSQTRP--------------------------------
HERV_Y_chr12_51022911_pol           R-------------------------------------------------
HERV17_polputein                    TLA-----------------------------------------------
HERV_W_chr7_9105739_syncytin_p      R-------------------------------------------------
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  ERPEYQCRNRRS--------------------------------------
ERVfrd_AC004022                     LSSRTHLRGT----------------------------------------
PRIMA41_polputein                   RLIPFIPVSQSVTSDSCSE-------------------------------
HERV_H_RGH1_pol                     PLGS----------------------------------------------
HERVHRGH2_pol                       SWKSQLLSPTS---------------------------------------
HERV_H48I_polputein                 TLTAQHPRS-----------------------------------------
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           SHSMGAEPLEDLKFLFRKR-------------------------------
ChrY_13790029_693_po_C_Xz12/11      --------------------------------------------------
```

```
HERVS71_polputein                   WAMPLKTAHKSSEA------------------------------------
BaEV_M7_pol                         WRLRRSEDPLKIRLSRT---------------------------------
HERVR_polputein                     WRLRRSEDPLKIRLSRT---------------------------------
AKR_MLV_MLOCG_RT                    WRVQRSQNPLKIRLTRE---------------------------------
GaLV_pol                            WELEKTDHPLKLRIRRRRDESA----------------------------
Chr3_158587414_1662_po_C_Ez10/      AKPS----------LDNPCKVT----------------------------
Chr3_79627623_1138_po_C_Ez4/7       AKLS----------LDNPCKVTLRRTTSPAPVTPGS--------------
Chr19_20477244_1350_po_C_Ez10/      AKPS----------PDTPYKVTLRRTTSPAPVTPRS--------------
Chr5_111890459_1626_po_C_Ez10/      AKQS----------LDNPCKVTLRR-------------------------
ChrX_48766518_1600_po_C_Ez6/3       SDSE----------DDKPCSSHTRKLTGLRTAEA----------------
ChrX_152580713_1572_po_C_Ez10/      AKPS----------PDKPCKVTLRRTTSPVPVTPGS--------------
ChrX_139287533_1154_po_C_Xz8/4      AKPS----------LDNPCKVT----------------------------
Chr2_52615786_912_po_C_Xz11/5       VKPS----------VDNPCKVTLR--------------------------
HERV-E_4_1                          ARPSRKPL------QSDPKKTTSPA-------------------------
ChrY_21962399_1137_po_C_Xz13/9      VRPSPDNPQSDSEEDDKPCSSHMWKLTGPCMAEA----------------
ChrY_8634032_950_po_C_Xz14/10       AKPS----------PDNPAKRL----------------------------
Chr3_126735360_997_po_C_Xz5/10      --------------------------------------------------
Chr3_126713203_958_po_C_Xz7/8       --------------------------------------------------
Chr3_75226118_977_po_C_Xz8/9        QG------------------------------------------------
Chr7_6629919_827_po_C_Xz8/7         QG------------------------------------------------
Chr4_9166812_811_po_C_Xz8/10        QV------------------------------------------------
Chr12_8478642_834_po_C_Xz12/7       QG------------------------------------------------
Chr10_15187175_787_po_C_Xz11/5      QG------------------------------------------------
Chr11_71819740_854_po_C_Xz11/9      SRPLD---------------------------------------------
Chr11_71519974_886_po_C_Xz13/9      QGZCTPSDVRTNIILPPTGY------------------------------
Chr8_7271441_969_po_C_Xz10/6        QG------------------------------------------------
Chr8_11770066_826_po_C_Xz9/8        QG------------------------------------------------
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      QG------------------------------------------------
Chr14_50205291_837_po_C_Xz10/8      QG------------------------------------------------
Chr7_6597120_842_po_C_Xz7/7         QGZCTPSVTLGLISSSRPLD------------------------------
Chr11_71544658_783_po_C_Xz8/13      --------------------------------------------------
Chr11_67999520_760_po_C_Xz8/8       QG------------------------------------------------
Chr3_75461975_796_po_C_Xz7/9        QG------------------------------------------------
Chr12_52228939_692_po_C_Xz3/10      QGZCTPS-------------------------------------------
Chr10_15211234_286_po_C_Xz11/1      QG------------------------------------------------
Chr12_100973302_945_po_C_Xz12/      LYPVPGLEGPTPTEPHSLLAQQSEIKLQGSSEAGGGAPAIAEASVGKQSS
Chr19_9685420_817_po_C_Xz6/10       --------------------------------------------------
Chr13_40936820_719_po_C_Xz9/10      --------------------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      L-----DHATQLILRGT---------------------------------
ERV3_chr7q21_AC073210_polputei      L-----DHATS---------------------------------------
Chr6_110218096_843_po_C_Xz16/8      P-----DHPSWLILRRDQVA------------------------------
Chr6_110214535_786_po_C_Xz16/8      P-----DHPSWLILRRDQVA------------------------------
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        P-----DHPSRLILRREPVA------------------------------
Chr14_68899917_965_po_C_Xz6/8       P-----DHPTRLIL------------------------------------
ChrY_23874437_661_po_C_Xz15/9       P-----DHLSWLIL------------------------------------
ChrY_27518418_649_po_C_Xz15/9       P-----DHLSWLIL------------------------------------
ChrY_25016892_657_po_C_Xz13/12      P-----DHLSWLIL------------------------------------
ChrY_19932937_520_po_C_Xz10/14      P-----DHWSQLIL------------------------------------
ChrY_20104614_518_po_C_Xz10/14      P-----DHWSQLIL------------------------------------
Chr19_58660155_598_po_C_Xz21/1      P-----DHPTQLILRRPSCC------------------------------
Chr19_58621153_500_po_C_Xz21/1      ---------TQIIQLSRS--------------------------------
Chr7_139686987_710_po_C_Xz11/7      ---------HELIISYGAZ-------------------------------
Chr7_137651301_1097_po_C_Pz13/      P-----DHLTQLILRWDQVA------------------------------
Chr11_29579197_930_po_C_Pz8/12      ---------QTIPLG-----------------------------------
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      -----IDHLTRMILQ-----------------------------------
Chr4_70022513_788_po_C_Kz10/12      RPSDLADHAMRPRCQWR---------------------------------
Chr14_105225894_1027_po_C_Xz13      -----PDHPTQLTLRRDWVPSGEDNSPALLTSE-----------------
Chr7_119241536_761_po_C_Xz11/1      P------PSPRLILC-----------------------------------
ChrY_22564970_567_po_C_Xz9/10       P-----DHPTGLILRGTKSC------------------------------
Chr21_14409346_1477_po_C_Pz5/6      SQQ-DPDHPSTDLAE--EPRHSRKR-------------------------
ChrX_51259700_1396_po_C_Pz7/8       RNK-GTEYKREKZRDRVQKREILKLGVRGTS-------------------
Chr6_64985396_1041_po_C_Xz11/1      SIKPRPPDTADPVVRPSHY-------------------------------
ChrY_20355209_659_po_C_Xz7/11       SQQ-NPDHLTWMILR-----------------------------------
Chr3_34567166_523_po_C_Xz8/6        SQQ-NPDHPTRLILQ-----------------------------------
Chr3_122072468_245_po_C_Xz9/9       SQQ-NQDHPTWLILQETKSLLTRTTSLL----------------------
Chr11_58788475_541_po_C_Xz8/22      SQQ-YPDHPTRLSYSEIKPLLTRT--------------------------
Chr6_137421042_373_po_C_Xz6/9       SQQ-NPDHSTWLILW-----------------------------------
Chr4_115910608_550_po_C_Pz8/12      SQQ-NPDHPTLETKPRLTRTTALL--------------------------
ChrY_13743185_977_po_C_Xz14/8       ANK-TQIAPPE---------------------------------------
HERV15                              SQK-TQIA------------------------------------------
ChrX_117198092_983_po_C_Xz8/8       SQQ-DPDCPTRTVLWQNPLPVRRTTALL----------------------
Chr19_21166030_1412_po_C_Xz9/1      ANK-TQITPPE---------------------------------------
Chr19_21213743_913_po_C_Xz14/1      CRK-AVASHLED--------------------------------------
```

```
Chr19_15957774_569_po_C_Xz8/10      --------------------------------------------------
Chr11_77508681_297_po_C_Xz19/1      R-------------------------------------------------
Chr13_98250769_738_po_C_Xz7/13      A-------------------------------------------------
Chr6_39968183_513_po_C_Pz10/9       --------------------------------------------------
Chr15_66867263_645_po_C_Xz11/1      --------------------------------------------------
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                KPYQTKAKKNLTLFHLFYYSSS----------------------------
MER66_polputein                     --------------------------------------------------


HERVI_polputein                     --------------------------------------------------
HERVIP10F_polputein                 --------------------------------------------------
HERVadp_AC005741                    --------------------------------------------------
HERV-PT47D                          --------------------------------------------------
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------------------
ERV9_PH1_RT                         --------------------------------------------------
Chr13_54564403_572_po_C_Yz7/14      --------------------------------------------------
HERV_Y_chr12_51022911_pol           --------------------------------------------------
HERV17_polputein                    --------------------------------------------------
HERV_W_chr7_9105739_syncytin_p      --------------------------------------------------
HERV9_polputein                     --------------------------------------------------
HUERS-P3_polputein                  --------------------------------------------------
ERVfrd_AC004022                     --------------------------------------------------
PRIMA41_polputein                   --------------------------------------------------
HERV_H_RGH1_pol                     --------------------------------------------------
HERVHRGH2_pol                       --------------------------------------------------
HERV_H48I_polputein                 --------------------------------------------------
HERV_H_RTVLH2_pol                   --------------------------------------------------
HERVFH21_pol                        --------------------------------------------------
HERVRblike_chr4_109047953           --------------------------------------------------
ChrY_13790029_693_po_C_Xz12/11      --------------------------------------------------
HERVS71_polputein                   --------------------------------------------------
BaEV_M7_pol                         --------------------------------------------------
HERVR_polputein                     --------------------------------------------------
AKR_MLV_MLOCG_RT                    --------------------------------------------------
GaLV_pol                            --------------------------------------------------
Chr3_158587414_1662_po_C_Ez10/      --------------------------------------------------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------------------
Chr19_20477244_1350_po_C_Ez10/      --------------------------------------------------
Chr5_111890459_1626_po_C_Ez10/      --------------------------------------------------
ChrX_48766518_1600_po_C_Ez6/3       --------------------------------------------------
ChrX_152580713_1572_po_C_Ez10/      --------------------------------------------------
ChrX_139287533_1154_po_C_Xz8/4      --------------------------------------------------
Chr2_52615786_912_po_C_Xz11/5       --------------------------------------------------
HERV-E_4_1                          --------------------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      --------------------------------------------------
ChrY_8634032_950_po_C_Xz14/10       --------------------------------------------------
Chr3_126735360_997_po_C_Xz5/10      --------------------------------------------------
Chr3_126713203_958_po_C_Xz7/8       --------------------------------------------------
Chr3_75226118_977_po_C_Xz8/9        --------------------------------------------------
Chr7_6629919_827_po_C_Xz8/7         --------------------------------------------------
Chr4_9166812_811_po_C_Xz8/10        --------------------------------------------------
Chr12_8478642_834_po_C_Xz12/7       --------------------------------------------------
Chr10_15187175_787_po_C_Xz11/5      --------------------------------------------------
Chr11_71819740_854_po_C_Xz11/9      --------------------------------------------------
Chr11_71519974_886_po_C_Xz13/9      --------------------------------------------------
Chr8_7271441_969_po_C_Xz10/6        --------------------------------------------------
Chr8_11770066_826_po_C_Xz9/8        --------------------------------------------------
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------------------
Chr14_50203955_850_po_C_Xz10/8      --------------------------------------------------
Chr14_50205291_837_po_C_Xz10/8      --------------------------------------------------
Chr7_6597120_842_po_C_Xz7/7         --------------------------------------------------
Chr11_71544658_783_po_C_Xz8/13      --------------------------------------------------
Chr11_67999520_760_po_C_Xz8/8       --------------------------------------------------
Chr3_75461975_796_po_C_Xz7/9        --------------------------------------------------
Chr12_52228939_692_po_C_Xz3/10      --------------------------------------------------
Chr10_15211234_286_po_C_Xz11/1      --------------------------------------------------
Chr12_100973302_945_po_C_Xz12/      WEARTGWSPPQLKEACLPLZAPPLGAGHSQTKGSRNLCGLKCPCLTALKR
Chr19_9685420_817_po_C_Xz6/10       --------------------------------------------------
Chr13_40936820_719_po_C_Xz9/10      --------------------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      --------------------------------------------------
ERV3_chr7q21_AC073210_polputei      --------------------------------------------------
Chr6_110218096_843_po_C_Xz16/8      --------------------------------------------------
Chr6_110214535_786_po_C_Xz16/8      --------------------------------------------------
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------------------
ChrX_86876098_639_po_C_Xz7/8        --------------------------------------------------
Chr14_68899917_965_po_C_Xz6/8       --------------------------------------------------
```

```
ChrY_23874437_661_po_C_Xz15/9       --------------------------------------------------
ChrY_27518418_649_po_C_Xz15/9       --------------------------------------------------
ChrY_25016892_657_po_C_Xz13/12      --------------------------------------------------
ChrY_19932937_520_po_C_Xz10/14      --------------------------------------------------
ChrY_20104614_518_po_C_Xz10/14      --------------------------------------------------
Chr19_58660155_598_po_C_Xz21/1      --------------------------------------------------
Chr19_58621153_500_po_C_Xz21/1      --------------------------------------------------
Chr7_139686987_710_po_C_Xz11/7      --------------------------------------------------
Chr7_137651301_1097_po_C_Pz13/      --------------------------------------------------
Chr11_29579197_930_po_C_Pz8/12      --------------------------------------------------
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      --------------------------------------------------
Chr4_70022513_788_po_C_Kz10/12      --------------------------------------------------
Chr14_105225894_1027_po_C_Xz13      --------------------------------------------------
Chr7_119241536_761_po_C_Xz11/1      --------------------------------------------------
ChrY_22564970_567_po_C_Xz9/10       --------------------------------------------------
Chr21_14409346_1477_po_C_Pz5/6      --------------------------------------------------
ChrX_51259700_1396_po_C_Pz7/8       --------------------------------------------------
Chr6_64985396_1041_po_C_Xz11/1      --------------------------------------------------
ChrY_20355209_659_po_C_Xz7/11       --------------------------------------------------
Chr3_34567166_523_po_C_Xz8/6        --------------------------------------------------
Chr3_122072468_245_po_C_Xz9/9       --------------------------------------------------
Chr11_58788475_541_po_C_Xz8/22      --------------------------------------------------
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------------------
ChrY_13743185_977_po_C_Xz14/8       --------------------------------------------------
HERV15                              --------------------------------------------------
ChrX_117198092_983_po_C_Xz8/8       --------------------------------------------------
Chr19_21166030_1412_po_C_Xz9/1      --------------------------------------------------
Chr19_21213743_913_po_C_Xz14/1      --------------------------------------------------
Chr19_15957774_569_po_C_Xz8/10      --------------------------------------------------
Chr11_77508681_297_po_C_Xz19/1      --------------------------------------------------
Chr13_98250769_738_po_C_Xz7/13      --------------------------------------------------
Chr6_39968183_513_po_C_Pz10/9       --------------------------------------------------
Chr15_66867263_645_po_C_Xz11/1      --------------------------------------------------
HERV18_polputein                    --------------------------------------------------
HERVIP10FH_polputein                --------------------------------------------------
MER66_polputein                     --------------------------------------------------


HERVI_polputein                     --------------------------------------
HERVIP10F_polputein                 --------------------------------------
HERVadp_AC005741                    --------------------------------------
HERV-PT47D                          --------------------------------------
HERV_W_chr6_141432567_ERV9_lik      --------------------------------------
ERV9_PH1_RT                         --------------------------------------
Chr13_54564403_572_po_C_Yz7/14      --------------------------------------
HERV_Y_chr12_51022911_pol           --------------------------------------
HERV17_polputein                    --------------------------------------
HERV_W_chr7_9105739_syncytin_p      --------------------------------------
HERV9_polputein                     --------------------------------------
HUERS-P3_polputein                  --------------------------------------
ERVfrd_AC004022                     --------------------------------------
PRIMA41_polputein                   --------------------------------------
HERV_H_RGH1_pol                     --------------------------------------
HERVHRGH2_pol                       --------------------------------------
HERV_H48I_polputein                 --------------------------------------
HERV_H_RTVLH2_pol                   --------------------------------------
HERVFH21_pol                        --------------------------------------
HERVRblike_chr4_109047953           --------------------------------------
ChrY_13790029_693_po_C_Xz12/11      --------------------------------------
HERVS71_polputein                   --------------------------------------
BaEV_M7_pol                         --------------------------------------
HERVR_polputein                     --------------------------------------
AKR_MLV_MLOCG_RT                    --------------------------------------
GaLV_pol                            --------------------------------------
Chr3_158587414_1662_po_C_Ez10/      --------------------------------------
Chr3_79627623_1138_po_C_Ez4/7       --------------------------------------
Chr19_20477244_1350_po_C_Ez10/      --------------------------------------
Chr5_111890459_1626_po_C_Ez10/      --------------------------------------
ChrX_48766518_1600_po_C_Ez6/3       --------------------------------------
ChrX_152580713_1572_po_C_Ez10/      --------------------------------------
ChrX_139287533_1154_po_C_Xz8/4      --------------------------------------
Chr2_52615786_912_po_C_Xz11/5       --------------------------------------
HERV-E_4_1                          --------------------------------------
ChrY_21962399_1137_po_C_Xz13/9      --------------------------------------
```

```
ChrY_8634032_950_po_C_Xz14/10       --------------------------------------
Chr3_126735360_997_po_C_Xz5/10      --------------------------------------
Chr3_126713203_958_po_C_Xz7/8       --------------------------------------
Chr3_75226118_977_po_C_Xz8/9        --------------------------------------
Chr7_6629919_827_po_C_Xz8/7         --------------------------------------
Chr4_9166812_811_po_C_Xz8/10        --------------------------------------
Chr12_8478642_834_po_C_Xz12/7       --------------------------------------
Chr10_15187175_787_po_C_Xz11/5      --------------------------------------
Chr11_71819740_854_po_C_Xz11/9      --------------------------------------
Chr11_71519974_886_po_C_Xz13/9      --------------------------------------
Chr8_7271441_969_po_C_Xz10/6        --------------------------------------
Chr8_11770066_826_po_C_Xz9/8        --------------------------------------
Chr4_4128362_651_po_C_Xz7/11        --------------------------------------
Chr14_50203955_850_po_C_Xz10/8      --------------------------------------
Chr14_50205291_837_po_C_Xz10/8      --------------------------------------
Chr7_6597120_842_po_C_Xz7/7         --------------------------------------
Chr11_71544658_783_po_C_Xz8/13      --------------------------------------
Chr11_67999520_760_po_C_Xz8/8       --------------------------------------
Chr3_75461975_796_po_C_Xz7/9        --------------------------------------
Chr12_52228939_692_po_C_Xz3/10      --------------------------------------
Chr10_15211234_286_po_C_Xz11/1      --------------------------------------
Chr12_100973302_945_po_C_Xz12/      VVVLPAHSLRSENGQNASSSGSLTPEZPNWEATPSRGT
Chr19_9685420_817_po_C_Xz6/10       --------------------------------------
Chr13_40936820_719_po_C_Xz9/10      --------------------------------------
Chr7_63858202_1754_po_C_Rz12/9      --------------------------------------
ERV3_chr7q21_AC073210_polputei      --------------------------------------
Chr6_110218096_843_po_C_Xz16/8      --------------------------------------
Chr6_110214535_786_po_C_Xz16/8      --------------------------------------
ChrY_3018841_881_po_C_Rz9/12        --------------------------------------
ChrX_86876098_639_po_C_Xz7/8        --------------------------------------
Chr14_68899917_965_po_C_Xz6/8       --------------------------------------
ChrY_23874437_661_po_C_Xz15/9       --------------------------------------
ChrY_27518418_649_po_C_Xz15/9       --------------------------------------
ChrY_25016892_657_po_C_Xz13/12      --------------------------------------
ChrY_19932937_520_po_C_Xz10/14      --------------------------------------
ChrY_20104614_518_po_C_Xz10/14      --------------------------------------
Chr19_58660155_598_po_C_Xz21/1      --------------------------------------
Chr19_58621153_500_po_C_Xz21/1      --------------------------------------
Chr7_139686987_710_po_C_Xz11/7      --------------------------------------
Chr7_137651301_1097_po_C_Pz13/      --------------------------------------
Chr11_29579197_930_po_C_Pz8/12      --------------------------------------
Chr4_117329205_488_po_C_Xz3/16      --------------------------------------
Chr21_38522185_383_po_C_Xz7/9       --------------------------------------
Chr4_70173370_1049_po_C_Xz11/8      --------------------------------------
Chr4_70022513_788_po_C_Kz10/12      --------------------------------------
Chr14_105225894_1027_po_C_Xz13      --------------------------------------
Chr7_119241536_761_po_C_Xz11/1      --------------------------------------
ChrY_22564970_567_po_C_Xz9/10       --------------------------------------
Chr21_14409346_1477_po_C_Pz5/6      --------------------------------------
ChrX_51259700_1396_po_C_Pz7/8       --------------------------------------
Chr6_64985396_1041_po_C_Xz11/1      --------------------------------------
ChrY_20355209_659_po_C_Xz7/11       --------------------------------------
Chr3_34567166_523_po_C_Xz8/6        --------------------------------------
Chr3_122072468_245_po_C_Xz9/9       --------------------------------------
Chr11_58788475_541_po_C_Xz8/22      --------------------------------------
Chr6_137421042_373_po_C_Xz6/9       --------------------------------------
Chr4_115910608_550_po_C_Pz8/12      --------------------------------------
ChrY_13743185_977_po_C_Xz14/8       --------------------------------------
HERV15                              --------------------------------------
ChrX_117198092_983_po_C_Xz8/8       --------------------------------------
Chr19_21166030_1412_po_C_Xz9/1      --------------------------------------
Chr19_21213743_913_po_C_Xz14/1      --------------------------------------
Chr19_15957774_569_po_C_Xz8/10      --------------------------------------
Chr11_77508681_297_po_C_Xz19/1      --------------------------------------
Chr13_98250769_738_po_C_Xz7/13      --------------------------------------
Chr6_39968183_513_po_C_Pz10/9       --------------------------------------
Chr15_66867263_645_po_C_Xz11/1      --------------------------------------
HERV18_polputein                    --------------------------------------
HERVIP10FH_polputein                --------------------------------------
MER66_polputein                     --------------------------------------
```

Alignment made 2004/01/03 11:55:49 with FASTA file C:\RETROVIRALA SEKVENSER FÖR DATABAS\HERV COLLECTIONS\ERV3 BROADLY 040102_FASTA POLPUTEIN2 W MANY GAMMARETROVIRAL POLS.TXT).

NJ tree made from this alignment, .DND format:

```
(
(
(
(
(
(
(
Chr7_63858202_1754_po_C_Rz12/9:0.00666,
ERV3_chr7q21_AC073210_polputei:0.00561)
:0.14292,
(
Chr6_110218096_843_po_C_Xz16/8:0.00000,
Chr6_110214535_786_po_C_Xz16/8:0.00000)
:0.17464)
:0.00724,
(
Chr14_68899917_965_po_C_Xz6/8:0.12957,
(
ChrY_3018841_881_po_C_Rz9/12:0.06596,
ChrX_86876098_639_po_C_Xz7/8:0.04072)
:0.14084)
:0.00644)
:0.00734,
(
(
(
ChrY_23874437_661_po_C_Xz15/9:0.00190,
ChrY_27518418_649_po_C_Xz15/9:0.00015)
:0.00654,
ChrY_25016892_657_po_C_Xz13/12:0.01241)
:0.17669,
(
ChrY_19932937_520_po_C_Xz10/14:0.00000,
ChrY_20104614_518_po_C_Xz10/14:0.00000)
:0.17152)
:0.09162)
:0.00317,
(
Chr7_139686987_710_po_C_Xz11/7:0.22562,
(
Chr19_58660155_598_po_C_Xz21/1:0.16945,
Chr19_58621153_500_po_C_Xz21/1:0.14334)
:0.06711)
:0.03189)
:0.00324,
(
(
(
(
(
(
(
(
(
(
(
(
Chr3_158587414_1662_po_C_Ez10/:0.08493,
Chr3_79627623_1138_po_C_Ez4/7:0.10404)
:0.00925,
Chr19_20477244_1350_po_C_Ez10/:0.11461)
:0.00943,
(
(
(
Chr5_111890459_1626_po_C_Ez10/:0.08868,
ChrX_48766518_1600_po_C_Ez6/3:0.07348)
:0.00767,
ChrX_152580713_1572_po_C_Ez10/:0.08879)
:0.00536,
ChrX_139287533_1154_po_C_Xz8/4:0.09472)
:0.00323)
:0.01291,
(
Chr2_52615786_912_po_C_Xz11/5:0.12221,
HERV-E_4_1:0.10343)
```

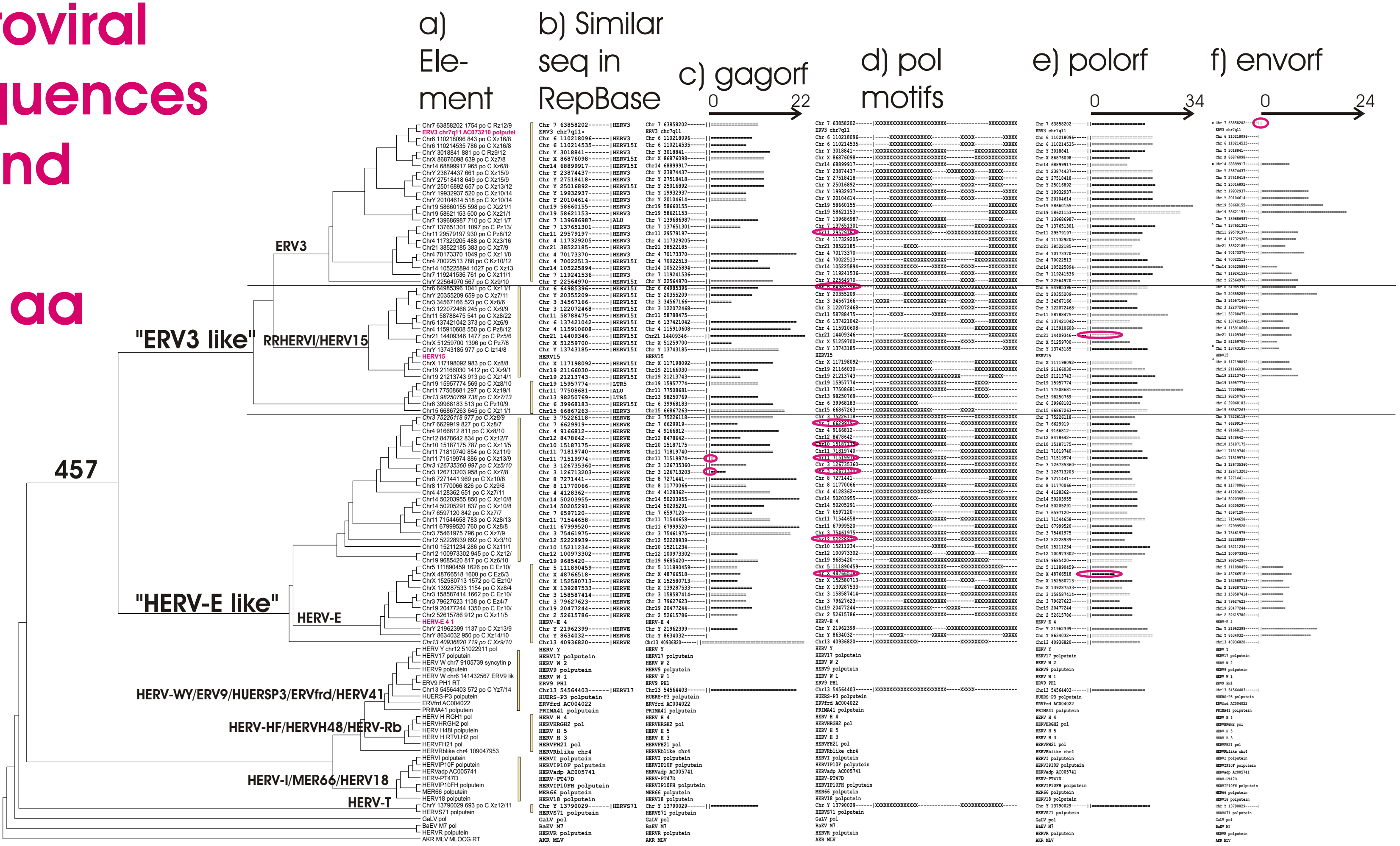
```
:0.00204)
:0.00870,
ChrY_21962399_1137_po_C_Xz13/9:0.15482)
:0.01170,
ChrY_8634032_950_po_C_Xz14/10:0.17634)
:0.02880,
(
(
(
(
(
(
(
(
Chr3_126735360_997_po_C_Xz5/10:0.10540,
Chr3_126713203_958_po_C_Xz7/8:0.07281)
:0.01287,
(
(
(
(
Chr3_75226118_977_po_C_Xz8/9:0.06647,
Chr7_6629919_827_po_C_Xz8/7:0.06885)
:0.00317,
Chr4_9166812_811_po_C_Xz8/10:0.08986)
:0.00452,
(
Chr11_71819740_854_po_C_Xz11/9:0.07856,
(
Chr12_8478642_834_po_C_Xz12/7:0.06071,
Chr10_15187175_787_po_C_Xz11/5:0.05140)
:0.02049)
:0.01045)
:0.00255,
Chr11_71519974_886_po_C_Xz13/9:0.08787)
:0.01410)
:0.01370,
(
Chr8_7271441_969_po_C_Xz10/6:0.10287,
Chr8_11770066_826_po_C_Xz9/8:0.10687)
:0.01295)
:0.01001,
Chr4_4128362_651_po_C_Xz7/11:0.13419)
:0.00890,
(
(
(
(
Chr14_50203955_850_po_C_Xz10/8:0.00000,
Chr14_50205291_837_po_C_Xz10/8:0.00000)
:0.09597,
Chr7_6597120_842_po_C_Xz7/7:0.09870)
:0.00300,
(
Chr11_71544658_783_po_C_Xz8/13:0.10530,
Chr11_67999520_760_po_C_Xz8/8:0.09169)
:0.00986)
:0.02494,
Chr3_75461975_796_po_C_Xz7/9:0.10903)
:0.00932)
:0.03343,
Chr12_52228939_692_po_C_Xz3/10:0.14306)
:0.00445,
Chr10_15211234_286_po_C_Xz11/1:0.22867)
:0.04192,
(
Chr12_100973302_945_po_C_Xz12/:0.16674,
Chr19_9685420_817_po_C_Xz6/10:0.14795)
:0.04551)
:0.01232)
:0.01851,
Chr13_40936820_719_po_C_Xz9/10:0.30616)
:0.06958,
(
(
(
ChrY_13790029_693_po_C_Xz12/11:0.17446,
```

```
HERVS71_polputein:0.15622)
:0.11387,
(
(
(
BaEV_M7_pol:0.00000,
HERVR_polputein:0.00000)
:0.16884,
AKR_MLV_MLOCG_RT:0.16791)
:0.01110,
GaLV_pol:0.16976)
:0.07735)
:0.03553,
(
(
(
(
(
(
(
(
Chr13_54564403_572_po_C_Yz7/14:0.23695,
(
HERV_W_chr6_141432567_ERV9_lik:0.14606,
ERV9_PH1_RT:0.17126)
:0.05120)
:0.05484,
(
(
HERV_W_chr7_9105739_syncytin_p:0.09486,
(
HERV_Y_chr12_51022911_pol:0.07995,
HERV17_polputein:0.04518)
:0.01539)
:0.07161,
HERV9_polputein:0.18484)
:0.04678)
:0.03319,
HUERS-P3_polputein:0.22375)
:0.03135,
ERVfrd_AC004022:0.32134)
:0.00720,
PRIMA41_polputein:0.29585)
:0.01330,
(
(
HERV_H_RTVLH2_pol:0.32657,
(
(
HERV_H_RGH1_pol:0.08547,
HERVHRGH2_pol:0.09064)
:0.17881,
HERV_H48I_polputein:0.22414)
:0.01565)
:0.01498,
HERVFH21_pol:0.26355)
:0.04933)
:0.00824,
HERVRblike_chr4_109047953:0.31393)
:0.01482,
(
(
(
(
(
HERVI_polputein:0.14200,
HERVIP10F_polputein:0.11800)
:0.08763,
HERVadp_AC005741:0.22944)
:0.06045,
HERV-PT47D:0.38866)
:0.03964,
(
HERVIP10FH_polputein:0.43920,
MER66_polputein:0.48413)
:0.05872)
:0.01054,
```

```
HERV18_polputein:0.44678)
:0.03561)
:0.01536)
:0.01675)
:0.03943,
(
(
(
Chr13_98250769_738_po_C_Xz7/13:0.28043,
(
Chr19_15957774_569_po_C_Xz8/10:0.29503,
Chr11_77508681_297_po_C_Xz19/1:0.34949)
:0.05529)
:0.02257,
Chr6_39968183_513_po_C_Pz10/9:0.27981)
:0.02052,
Chr15_66867263_645_po_C_Xz11/1:0.36463)
:0.00937)
:0.02833,
(
(
(
Chr21_14409346_1477_po_C_Pz5/6:0.12700,
ChrX_51259700_1396_po_C_Pz7/8:0.13264)
:0.03197,
(
(
(
Chr6_64985396_1041_po_C_Xz11/1:0.14250,
ChrY_20355209_659_po_C_Xz7/11:0.16093)
:0.00774,
(
Chr3_34567166_523_po_C_Xz8/6:0.12174,
Chr3_122072468_245_po_C_Xz9/9:0.11826)
:0.01147)
:0.02584,
(
Chr4_115910608_550_po_C_Pz8/12:0.16409,
(
Chr11_58788475_541_po_C_Xz8/22:0.17937,
Chr6_137421042_373_po_C_Xz6/9:0.22958)
:0.04106)
:0.05695)
:0.01290)
:0.04554,
(
(
Chr19_21166030_1412_po_C_Xz9/1:0.14707,
(
ChrX_117198092_983_po_C_Xz8/8:0.10323,
(
ChrY_13743185_977_po_C_Xz14/8:0.06351,
HERV15:0.04265)
:0.06903)
:0.00783)
:0.03603,
Chr19_21213743_913_po_C_Xz14/1:0.23241)
:0.05142)
:0.01223)
:0.00066,
ChrY_22564970_567_po_C_Xz9/10:0.23633)
:0.01573)
:0.00798,
(
(
(
Chr7_137651301_1097_po_C_Pz13/:0.16191,
Chr11_29579197_930_po_C_Pz8/12:0.15730)
:0.01091,
(
Chr4_117329205_488_po_C_Xz3/16:0.18217,
Chr21_38522185_383_po_C_Xz7/9:0.16852)
:0.13443)
:0.00504,
(
(
Chr4_70173370_1049_po_C_Xz11/8:0.13287,
```

```
Chr4_70022513_788_po_C_Kz10/12:0.14736)
:0.03301,
Chr14_105225894_1027_po_C_Xz13:0.18127)
:0.02182)
:0.00332,
Chr7_119241536_761_po_C_Xz11/1:0.21687);
```

Segregation of 76 ERV3-related retroviral sequences found via Pol aa



The cladogram shows

a) RetroTector defined elements from the hg15 version of the human genome, written as chromosome, position, RetroTector chain score, RetroTector genus assignment, "po" for "pol", PBS (X=not determined) "z",number of stops in *pol*,"/",number of frameshifts in *pol*.

b) Similar elements detected in RepBase, based on nucleotide identity,

c) Sum of stops and frameshifts in *gag*, shown graphically as columns of X:es. The elements most intact in *gag* are encircled in red.

d) Presence of motifs in the Pol putein is shown as "XX" for presence and as "--" for absence. The most intact elements in *pol* are encircled in red.

e) Sum of stops and frameshifts in *pol*, shown graphically as columns of X:es. The elements most intact in *pol* are encircled in red.

f) Sum of stops and frameshifts in *env*, shown graphically as columns of X:es. The elements most intact in *env* are encircled in red.

Reference elements are shown without the information in a-e

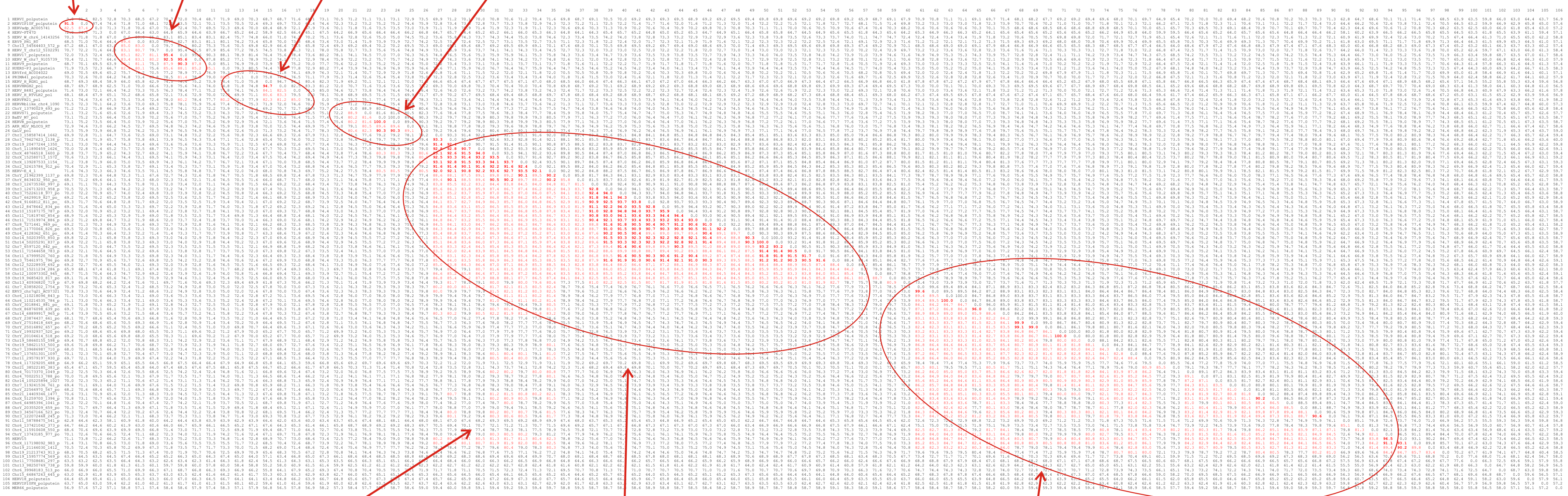
Similarity of ERV3-like and HERVE-like sequences in Pol to each other and to other gammaretroviruses

HERV-W/-Y/ERV9

HERVI-like

HERVH-like

MuLV-like



HERVE-like

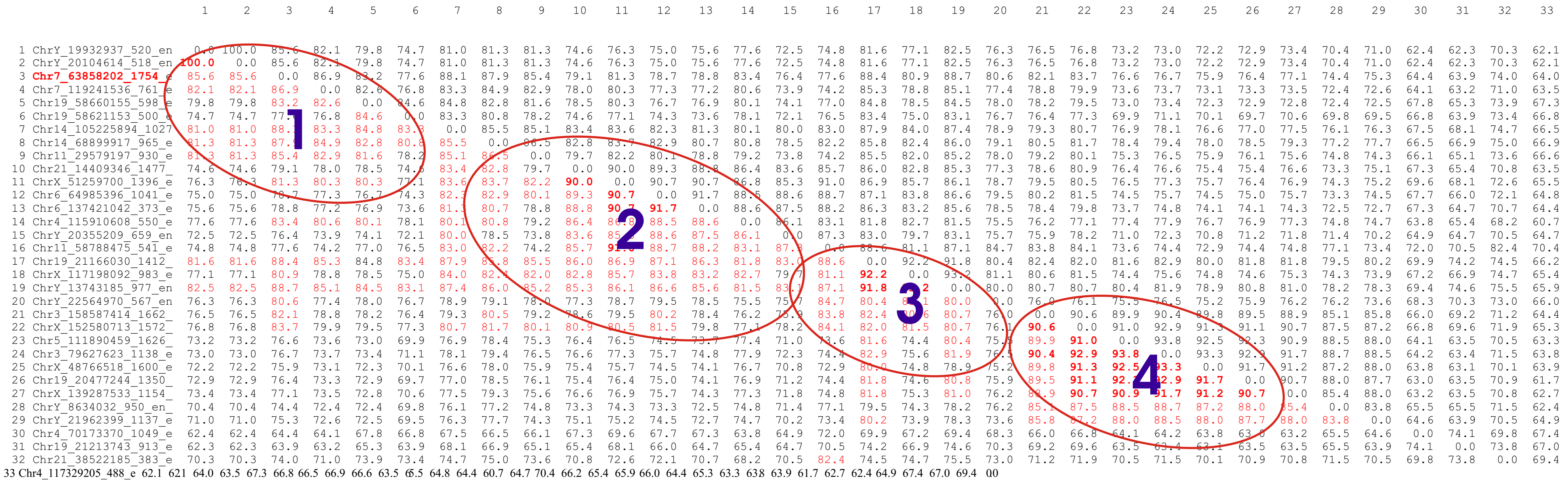
ERV3-like

HERVE-like and ERV3-like elements do not segregate completely.

Similarities >80% are shown in red.

Distance matrix of env puteins from 33 ERV3-related sequences.

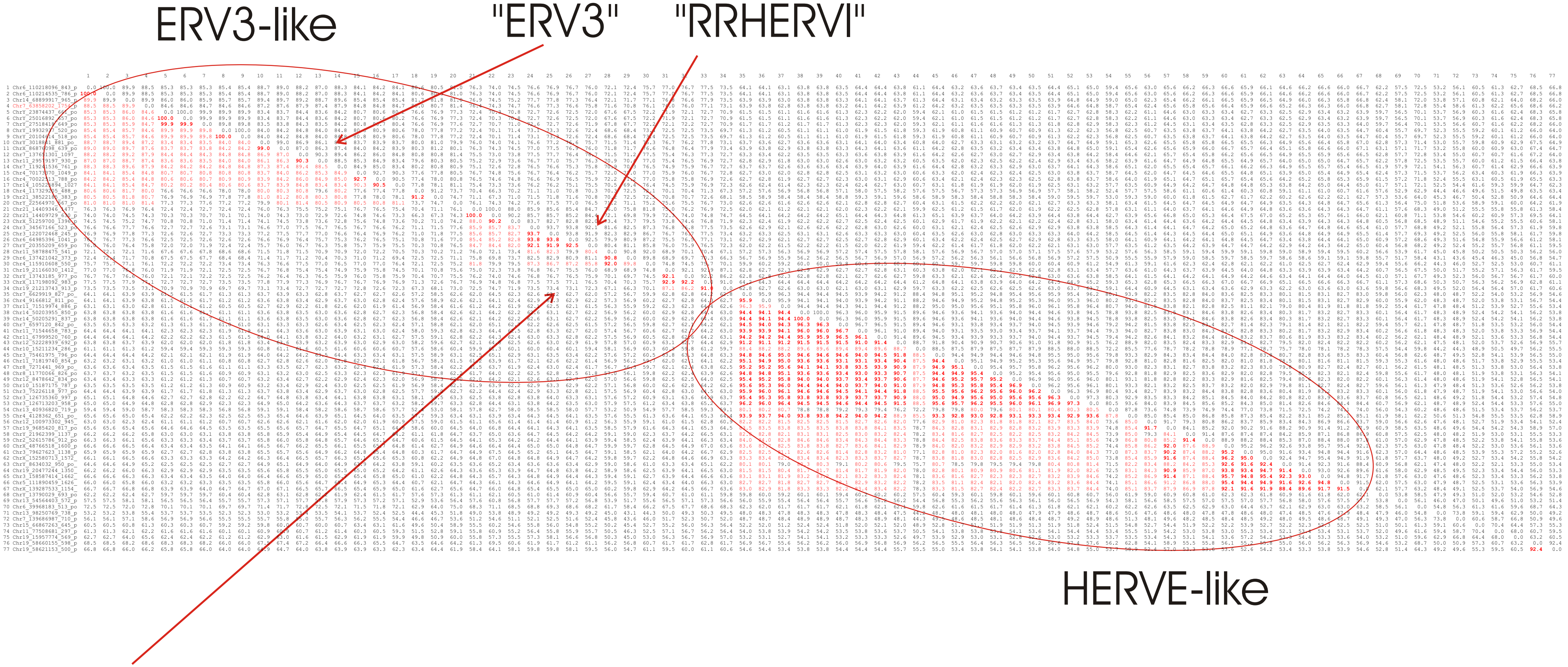
Distances (% of self similarity) were calculated using the PAM250 matrix and a Clustal W alignment.



		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1	ChrY_19932937_520_en	0.0	100.0	85.6	82.1	79.8	74.7	81.0	81.3	81.3	74.6	76.3	75.0	75.6	77.6	72.5	74.8	81.6	77.1	82.5	76.3	76.5	76.8	73.2	73.0	72.2	72.9	73.4	70.4	71.0	62.4	62.3	70.3	62.1
2	ChrY_20104614_518_en	**100.0**	0.0	85.6	82.1	79.8	74.7	81.0	81.3	81.3	74.6	76.3	75.0	75.6	77.6	72.5	74.8	81.6	77.1	82.5	76.3	76.5	76.8	73.2	73.0	72.2	72.9	73.4	70.4	71.0	62.4	62.3	70.3	62.1
3	**Chr7_63858202_1754_e**	85.6	85.6	0.0	86.9	83.2	77.6	88.1	87.9	85.4	79.1	81.3	78.7	78.8	83.4	76.4	77.6	88.4	80.9	88.7	80.6	82.1	83.7	76.6	76.7	75.9	76.4	77.1	74.4	75.3	64.4	63.9	74.0	64.0
4	Chr7_119241536_761_e	82.1	82.1	86.9	0.0	82.6	76.8	83.3	84.9	82.9	78.0	80.3	77.3	77.2	80.6	73.9	74.2	85.3	78.8	85.1	77.4	78.8	79.9	73.6	73.7	73.1	73.3	73.5	72.4	72.6	64.1	63.2	71.0	63.5
5	Chr19_58660155_598_e	79.8	79.8	83.2	82.6	0.0	84.6	84.8	82.8	81.6	78.5	80.3	76.7	76.9	80.1	74.1	77.0	84.8	78.5	84.5	78.0	78.2	79.5	73.0	73.4	72.3	72.9	72.8	72.4	72.5	67.8	65.3	73.9	67.3
6	Chr19_58621153_500_e	74.7	74.7	77.6	76.8	84.6	0.0	83.3	80.8	78.2	74.6	77.1	74.3	73.6	78.1	72.1	76.5	83.4	75.0	83.1	76.7	76.4	77.3	69.9	71.1	70.1	69.7	70.6	69.8	69.5	66.8	63.9	73.4	66.8
7	Chr14_105225894_1027	81.0	81.0	88.1	83.3	84.8	83.3	0.0	85.5	85.1	83.4	83.6	82.3	81.3	80.1	80.0	83.0	87.9	84.0	87.4	78.9	79.3	80.7	76.9	78.1	76.6	77.0	77.5	76.1	76.3	67.5	68.1	74.7	66.5
8	Chr14_68899917_965_e	81.3	81.3	87.9	84.9	82.8	80.8	85.5	0.0	86.5	82.8	83.7	82.9	80.7	80.8	78.5	82.2	85.8	82.4	86.0	79.1	80.5	81.7	78.4	79.4	78.0	78.5	79.3	77.2	77.7	66.5	66.9	75.0	66.9
9	Chr11_29579197_930_e	81.3	81.3	85.4	82.9	81.6	78.2	85.1	86.5	0.0	79.7	82.2	80.1	78.8	79.2	73.8	74.2	85.5	82.0	85.2	78.0	79.2	80.1	75.3	76.5	75.9	76.1	75.6	74.8	74.3	66.1	65.1	73.6	66.6
10	Chr21_14409346_1477_	74.6	74.6	79.1	78.0	78.5	74.6	83.4	82.8	79.7	0.0	90.0	89.3	88.8	86.4	83.6	85.7	86.0	82.8	85.3	77.3	78.6	80.9	76.4	76.6	75.4	75.4	76.6	73.3	75.1	67.3	65.4	70.8	63.5
11	ChrX_51259700_1396_e	76.3	76.3	81.3	80.3	80.3	77.1	83.6	83.7	82.2	**90.0**	0.0	90.7	90.7	86.8	85.3	91.0	86.9	85.7	86.1	78.7	79.5	80.5	76.5	77.3	75.7	76.4	76.9	74.3	75.2	69.6	68.1	72.6	65.5
12	Chr6_64985396_1041_e	75.0	75.0	78.7	77.3	76.7	74.3	82.3	82.9	80.1	89.3	**90.7**	0.0	91.7	88.5	88.6	88.7	87.1	83.8	86.6	79.5	80.2	81.5	74.5	75.7	74.5	75.0	75.7	73.3	74.5	67.7	66.0	72.1	64.8
13	Chr6_137421042_373_e	75.6	75.6	78.8	77.2	76.9	73.6	81.3	80.7	78.8	88.8	**90.7**	**91.7**	0.0	88.6	87.5	88.2	86.3	83.2	85.6	78.5	78.4	79.8	73.7	74.8	74.1	74.1	74.3	72.5	72.7	67.3	64.7	70.7	64.4
14	Chr4_115910608_550_e	77.6	77.6	83.4	80.6	80.1	78.1	80.1	80.8	79.2	86.4	86.8	88.5	88.6	0.0	86.1	83.1	81.8	82.7	81.5	75.5	76.2	77.1	77.4	77.9	76.7	76.9	77.3	74.8	74.7	63.8	65.4	68.2	60.7
15	ChrY_20355209_659_en	72.5	72.5	76.4	73.9	74.1	72.1	80.0	78.5	73.8	83.6	85.3	88.6	87.5	86.1	0.0	87.3	83.0	79.7	83.1	75.7	75.9	78.2	71.0	72.3	70.8	71.2	71.8	71.4	70.2	64.9	64.7	70.5	64.7
16	Chr11_58788475_541_e	74.8	74.8	77.6	74.2	77.0	76.5	83.0	82.2	74.2	85.7	**91.0**	88.7	88.2	83.1	87.3	0.0	88.6	81.1	87.1	84.7	83.8	84.1	73.6	74.4	72.9	74.4	74.8	77.1	73.4	72.0	70.5	82.4	70.4
17	Chr19_21166030_1412_	81.6	81.6	88.4	85.3	84.8	83.4	87.9	85.8	85.5	86.0	86.9	87.1	86.3	81.8	83.0	88.6	0.0	92.2	91.8	80.4	82.4	82.0	81.6	82.9	80.0	81.8	81.8	79.5	80.2	69.9	74.2	74.5	66.2
18	ChrX_117198092_983_e	77.1	77.1	80.9	78.8	78.5	75.0	84.0	82.4	82.0	82.8	85.7	83.8	83.2	82.7	79.7	81.1	**92.2**	0.0	93.2	81.1	80.6	81.5	74.4	75.6	74.8	74.6	75.3	74.3	73.9	67.2	66.9	74.7	65.4
19	ChrY_13743185_977_en	82.5	82.5	88.7	85.1	84.5	83.1	87.4	86.0	85.2	85.3	86.1	86.6	85.6	81.5	83.1	87.1	**91.8**	**93.2**	0.0	80.0	80.7	80.7	80.4	81.9	78.9	80.8	81.0	78.2	78.3	69.4	74.6	75.5	65.9
20	ChrY_22564970_567_en	76.3	76.3	80.6	77.4	78.0	76.7	78.9	79.1	78.0	77.3	78.7	79.5	78.5	75.5	75.7	84.7	80.4	81.1	80.0	0.0	76.0	76.1	75.5	76.5	75.2	75.9	76.2	76.2	73.6	68.3	70.3	73.0	66.0
21	Chr3_158587414_1662_	76.5	76.5	82.1	78.8	78.2	76.4	79.3	80.5	79.2	78.6	79.5	80.2	78.4	76.2	75.9	83.8	82.4	80.6	80.7	76.0	0.0	90.6	89.9	90.4	89.8	89.5	88.9	85.1	85.8	66.0	69.2	71.2	64.4
22	ChrX_152580713_1572_	76.8	76.8	83.7	79.9	79.5	77.3	80.7	81.7	80.1	80.9	80.5	81.5	79.8	77.1	78.2	84.1	82.0	81.5	80.7	76.1	**90.6**	0.0	91.0	92.9	91.3	91.1	90.7	87.5	87.2	66.8	69.6	71.9	65.3
23	Chr5_111890459_1626_	73.2	73.2	76.6	73.6	73.0	69.9	76.9	78.4	75.3	76.4	76.5	74.5	73.7	77.4	71.0	73.6	81.6	74.4	80.4	75.5	89.9	**91.0**	0.0	93.8	92.5	92.3	90.9	88.5	88.0	64.1	63.5	70.5	63.3
24	Chr3_79627623_1138_e	73.0	73.0	76.7	73.7	73.4	71.1	78.1	79.4	76.5	76.6	77.3	75.7	74.8	77.9	72.3	74.4	82.9	75.6	81.9	76.5	**90.4**	**92.9**	**93.8**	0.0	93.3	92.9	91.7	88.7	88.5	64.2	63.4	71.5	63.8
25	ChrX_48766518_1600_e	72.2	72.2	75.9	73.1	72.3	70.1	76.6	78.0	75.9	75.4	75.7	74.5	74.1	76.7	70.8	72.9	80.0	74.8	78.9	75.2	89.8	**91.3**	**92.5**	**93.3**	0.0	91.7	91.2	87.2	88.0	63.8	63.1	70.1	63.9
26	Chr19_20477244_1350_	72.9	72.9	76.4	73.3	72.9	69.7	77.0	78.5	76.1	75.4	76.4	75.0	74.1	76.9	71.2	74.4	81.8	74.6	80.8	75.9	89.5	**91.1**	**92.3**	**92.9**	**91.7**	0.0	90.7	88.0	87.7	63.0	63.5	70.9	61.7
27	ChrX_139287533_1154_	73.4	73.4	77.1	73.5	72.8	70.6	77.5	79.3	75.6	76.6	76.9	75.7	74.3	77.3	71.8	74.8	81.8	75.3	81.0	76.2	88.9	**90.7**	**90.9**	**91.7**	**91.2**	**90.7**	0.0	85.4	88.0	63.2	63.5	70.8	62.7
28	ChrY_8634032_950_en_	70.4	70.4	74.4	72.4	72.4	69.8	76.1	77.2	74.8	73.3	74.3	73.3	72.5	74.8	71.4	77.1	79.5	74.3	78.2	76.2	85.1	87.5	88.5	88.7	87.2	88.0	85.4	0.0	83.8	65.5	65.5	71.5	62.4
29	ChrY_21962399_1137_e	71.0	71.0	75.3	72.6	72.5	69.5	76.3	77.7	74.3	75.1	75.2	74.5	72.7	74.7	70.2	73.4	80.2	73.9	78.3	73.6	85.8	87.2	88.0	88.5	88.0	87.7	88.0	83.8	0.0	64.6	63.9	70.5	64.9
30	Chr4_70173370_1049_e	62.4	62.4	64.4	64.1	67.8	66.8	67.5	66.5	66.1	67.3	69.6	67.7	67.3	63.8	64.9	72.0	69.9	67.2	69.4	68.3	66.0	66.8	64.1	64.2	63.8	63.0	63.2	65.5	64.6	0.0	74.1	69.8	67.4
31	Chr19_21213743_913_e	62.3	62.3	63.9	63.2	65.3	63.9	68.1	66.9	65.1	65.4	68.1	66.0	64.7	65.4	64.7	70.5	74.2	66.9	74.6	70.3	69.2	69.6	63.5	63.4	63.1	63.5	63.5	65.5	63.9	74.1	0.0	73.8	67.0
32	Chr21_38522185_383_e	70.3	70.3	74.0	71.0	73.9	73.4	74.7	75.0	73.6	70.8	72.6	72.1	70.7	68.2	70.5	82.4	74.5	74.7	75.5	73.0	71.2	71.9	70.5	71.5	70.1	70.9	70.8	71.5	70.5	69.8	73.8	0.0	69.4
33	Chr4_117329205_488_e	62.1	621	64.0	63.5	67.3	66.8	66.5	66.9	66.6	63.5	65.5	64.8	64.4	60.7	64.7	70.4	66.2	65.4	65.9	66.0	64.4	65.3	63.3	63.8	63.9	61.7	62.7	62.4	64.9	67.4	67.0	69.4	00

The four major clusters correspond to
1-3. "ERV3-like" sequences, consisting of ERV3 (1-2) and RRHERVI (3) elements
and
4. "HERV-E-like"

Similarity of ERV3-like and HERVE-like sequences in *pol*



Some ERV3-like sequences do not group homogeneously according to nucleotide similarity

Similarities >80% are shown in red.

```
CLUSTAL W (1.81) multiple sequence alignment

                50:1                                              100:4
5ltr            TGAGGCAGGAAATATAAAAGGAAAAACAAGTAAAGGGAAAACAAGTCCTTCCCTGATCAG
3ltr            TGAAGCAGGAAATATAAAAGGAAAAACAAGTAAAGGGAAAACAAGTCCTTTCCTGACCAG
                *** ********************************************** ***** ***

                                                        150:5
5ltr            TCTGACTCACTCCAAAGTCCTGCTGGAGCTATGATAACATTATCTGCAAGGCCAGGCAGG
3ltr            TCTGACTCACTCCAAAGTCCTGCTGGAGCTATGATA--ATTATCTGCAAGGCCAGGCAGG
                ************************************  **********************

                                              200:5
5ltr            GACCCCAAAAGAATGGGCTCCAGGAGCAGAGATGAGAAAAACAAGTTCTCCTTATCAGTT
3ltr            GGCTCCGAAGGAGTGGGCTCCAGGAGCAGGGATGAGAAAAACAAGTTCTCCTTATCAGTT
                * * ** ** ** **************** ******************************

                                    250:1
5ltr            TCCGC-CTTGAAATTCTTTCCCCATACCATTATTCTTTGTTCTGCTCTCACAACTATTTT
3ltr            TCCCTGTTTGAAATTCTCTCCCCATAACATTATTCTTTGTTCTGCTCTCACAACTATTTT
                ***    ********** ******** *********************************

                          300:5
5ltr            TGTAACTATTTCTGCAAGTTTGCAAAGATTTCATAAGTTCCTGTTTTTCTTTCTGTAGCA
3ltr            TGTAACTATTTCTGCAAGTCTGTAAAGATTTTGTAAGTTCTTGTTTTTCTTTCTGTAGCA
                ******************* ** ********  ******* *******************

                350:4                                             400:1
5ltr            CGGCAAGGTCACAAGACATGCTTAAGTAAGATAGGGTCATGTTGCAAATCCTGTTGTAAA
3ltr            TGGCAAGGTCACAAGACATGTTTAAGTAAGGTAGGCTCATGTTGCAAATCCTGTTGTAAA
                 ******************* ********* **** ************************

                                                        450:10
5ltr            ACCTGTCACGGTATGATTAACTGTCTTTGTTCTGCTTCTGTAAGACTGCTTTCCTGTCTC
3ltr            ACCTGTCACGGTATGATTAACTGCCTTTGTTCTGCTTCTGTAAGACTGCTTTCTCACCTC
                *********************** *****************************    ***

                                              500:4
5ltr            ACAGGTTTCATGCCAAAAACCTGACCCGCCCCTGTTGGTTGCATGTATAAAAGTCAAGCC
3ltr            GCAGGTTTTGCGCCAAAAACCCGACTTGCCCCTGCCTGATGCATGTATAAAAGTCAAGCC
                 *******   ********** ***  *******   * *********************

                                    550:7
5ltr            CTGTCATTGTTCAGGGCTCAGCCTTTGGATGTTCATCGGCTGGGCTGGTGGTCACCTAAA
3ltr            CTGTCTTTGTTCAGGGCTCAGCCTTTGGATGTTAATCCGCTGGGCCAGTGGCCACCTAAA
                ***** *************************** *** *******  **** ********

                          591:4
5ltr            TAAAATCCTCCTGTTCCACCAAGTGGTCTCTCCAGCCTCCTGATTCCCACAACA
3ltr            TAAAACCTTCCTGTTGCACCCAGTGATCTCTCCGGCCTCCTGATTCCCACAACA
                ***** * ******* **** **** ******* ********************


Identities :  540 of Total  591
```

JB 040810

ConSite analysis of ERV3 LTRs

5´LTR

Predicted sites unique to the 5´LTR are shown in red.

c-FOS	1	R	TGAGGCAG
HFH-2	20	R	GGAAAAACAAGT
c-REL	20	R	GGAAAAACAA
HFH-2	24	R	AAACAAGTAAAG
HNF-3beta	24	R	AAACAAGTAAAG
Sox-5	24	N	AAACAAG
FREAC-4	37	N	GAAAACAA
Sox-5	39	N	AAACAAG
c-FOS	62	N	CTGACTCA
c-FOS	63	R	TGACTCAC
Evi-1	86	N	GAGCTATGATAACA
Sox-5	95	N	TAACATT
SOX17	96	N	AACATTATC
Thing1-E47	101	N	TATCTGCAAG
AML-1	121	R	GACCCCAAA
Sox-5	128	N	AAAGAAT
c-FOS	152	N	ATGAGAAA
HFH-2	155	R	AGAAAAACAAGT
Sox-5	159	N	AAACAAG
Sox-5	192	R	ATTCTTT
Sox-5	208	R	ATTATTC
HFH-1	210	N	TATTCTTTGTT
HFH-2	210	N	TATTCTTTGTTC
HNF-3beta	210	N	TATTCTTTGTTC
Sox-5	211	R	ATTCTTT
SOX17	212	N	TTCTTTGTT
Sox-5	230	N	CAACTAT
HFH-2	231	N	AACTATTTTTGT
HNF-3beta	231	N	AACTATTTTTGT
MEF2	233	N	CTATTTTTGT
Sox-5	235	R	ATTTTTG
E4BP4	237	N	TTTTGTAACTA
Sox-5	242	N	TAACTAT
E4BP4	265	R	AGATTTCATAA
HLF	266	N	GATTTCATAAGT
RORalfa-1	301	N	GGCAAGGTCA
E4BP4	318	R	TGCTTAAGTAA
E4BP4	321	N	TTAAGTAAGAT
RORalfa-1	329	N	GATAGGGTCA
HLF	351	R	TGTTGTAAAACC
c-FOS	372	N	ATGATTAA
SOX17	376	N	TTAACTGTC
SOX17	382	N	GTCTTTGTT
E4BP4	395	N	TTCTGTAAGAC
c-FOS	414	R	TGTCTCAC
Thing1-E47	429	R	ATGCCAAAAA
Thing1-E47	436	R	AAACCTGACC
Myf	448	R	CCCCTGTTGGTT
RORalfa-1	466	N	ATAAAAGTCA
Irf-1	468	N	AAAAGTCAAGCC
SOX17	482	N	GTCATTGTT
HFH-1	484	N	CATTGTTCAGG
Sox-5	485	R	ATTGTTC

AML-1	524	N	CTGGTGGTC
RORalfa-2	530	R	GTCACCTAAATAAA
RORalfa-1	531	R	TCACCTAAAT
HFH-3	533	R	ACCTAAATAAAA
Max	556	N	CACCAAGTGG
Myc-Max	556	N	CACCAAGTGGT
Myc-Max	557	R	ACCAAGTGGTC
Tal1beta-E47S	557	R	ACCAAGTGGTCT
Max	558	R	CCAAGTGGTC
USF	558	R	CCAAGTG
AML-1	559	N	CAAGTGGTC
ARNT	559	R	CAAGTG
ARNT	559	N	CAAGTG
USF	559	N	CAAGTGG
n-MYC	559	R	CAAGTG
n-MYC	559	N	CAAGTG

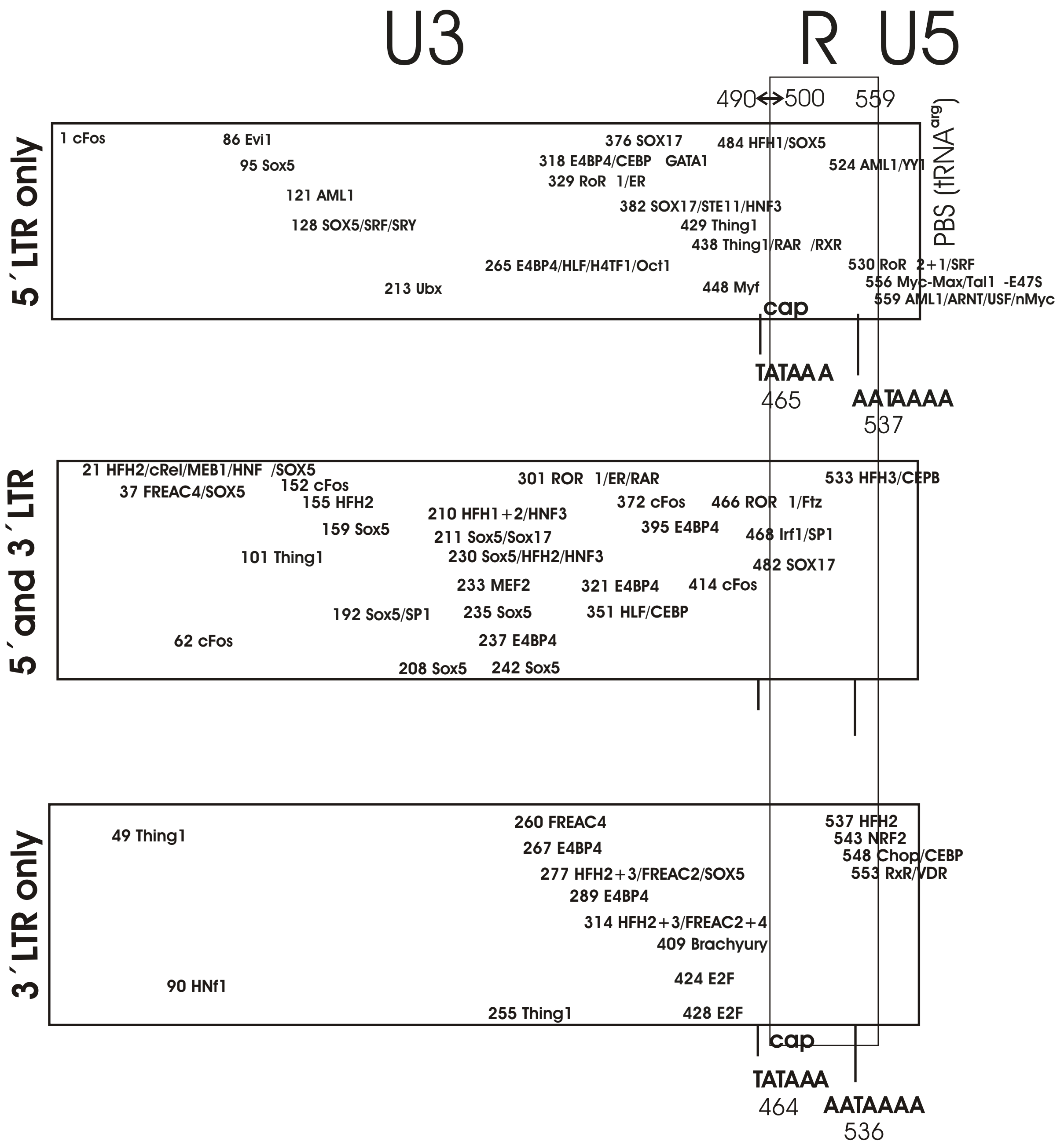
3´LTR

HFH-2	20	R	GGAAAAACAAGT
c-REL	20	R	GGAAAAACAA
HFH-2	24	R	AAACAAGTAAAG
HNF-3beta	24	R	AAACAAGTAAAG
Sox-5	24	N	AAACAAG
FREAC-4	37	N	GAAAACAA
Sox-5	39	N	AAACAAG
Thing1-E47	49	R	TTTCCTGACC
c-FOS	62	N	CTGACTCA
c-FOS	63	R	TGACTCAC
HNF-1	90	R	TATGATAATTATCT
Thing1-E47	99	N	TATCTGCAAG
c-FOS	150	N	ATGAGAAA
HFH-2	153	R	AGAAAAACAAGT
Sox-5	157	N	AAACAAG
Sox-5	203	N	TAACATT
Sox-5	207	R	ATTATTC
HFH-1	209	N	TATTCTTTGTT
HFH-2	209	N	TATTCTTTGTTC
HNF-3beta	209	N	TATTCTTTGTTC
Sox-5	210	R	ATTCTTT
SOX17	211	N	TTCTTTGTT
Sox-5	229	N	CAACTAT
HFH-2	230	N	AACTATTTTTGT
HNF-3beta	230	N	AACTATTTTTGT
MEF2	232	N	CTATTTTTGT
Sox-5	234	R	ATTTTTG
E4BP4	236	N	TTTTGTAACTA
Sox-5	241	N	TAACTAT
Thing1-E47	255	N	AGTCTGTAAA
FREAC-4	260	N	GTAAAGAT
E4BP4	267	N	TTTTGTAAGTT
HFH-2	277	N	TCTTGTTTTTCT
HFH-3	277	N	TCTTGTTTTTCT
FREAC-2	278	R	CTTGTTTTTCTTTC
Sox-5	278	R	CTTGTTT
E4BP4	289	N	TTCTGTAGCAT
RORalfa-1	300	N	GGCAAGGTCA
HFH-2	314	N	ACATGTTTAAGT
HFH-3	314	N	ACATGTTTAAGT
FREAC-2	315	R	CATGTTTAAGTAAG
FREAC-4	316	R	ATGTTTAA

E4BP4	320	N	TTAAGTAAGGT
HLF	350	R	TGTTGTAAAACC
c-FOS	371	N	ATGATTAA
E4BP4	394	N	TTCTGTAAGAC
c-FOS	408	R	TTTCTCAC
Brachyury	409	R	TTCTCACCTCG
E2F	424	N	TTTTGCGC
E2F	428	R	GCGCCAAA
RORalfa-1	465	N	ATAAAAGTCA
Irf-1	467	N	AAAAGTCAAGCC
SOX17	481	N	GTCTTTGTT
HFH-3	532	R	ACCTAAATAAAA
HFH-2	537	R	AATAAAACCTTC
NRF-2	543	R	ACCTTCCTGT
Chop-cEBP	548	R	CCTGTTGCACCC
RXR-VDR	553	R	TGCACCCAGTGATCT

Search settings are described in Materials and Methods.

Transcription factor binding sites predicted by ConSite



U3

R U5

490↔500 559

PBS (tRNA^{arg})

5´LTR only

1 cFos
86 Evi1
95 Sox5
121 AML1
128 SOX5/SRF/SRY
213 Ubx
265 E4BP4/HLF/H4TF1/Oct1
318 E4BP4/CEBP GATA1
329 RoR 1/ER
376 SOX17
382 SOX17/STE11/HNF3
429 Thing1
438 Thing1/RAR /RXR
448 Myf
484 HFH1/SOX5
524 AML1/YY1
530 RoR 2+1/SRF
556 Myc-Max/Tal1 -E47S
559 AML1/ARNT/USF/nMyc

cap

TATAA A
465

AATAAAA
537

5´and 3´LTR

21 HFH2/cRel/MEB1/HNF /SOX5
37 FREAC4/SOX5
62 cFos
101 Thing1
152 cFos
155 HFH2
159 Sox5
192 Sox5/SP1
208 Sox5
210 HFH1+2/HNF3
211 Sox5/Sox17
230 Sox5/HFH2/HNF3
233 MEF2
235 Sox5
237 E4BP4
242 Sox5
301 ROR 1/ER/RAR
321 E4BP4
351 HLF/CEBP
372 cFos
395 E4BP4
414 cFos
466 ROR 1/Ftz
468 Irf1/SP1
482 SOX17
533 HFH3/CEPB

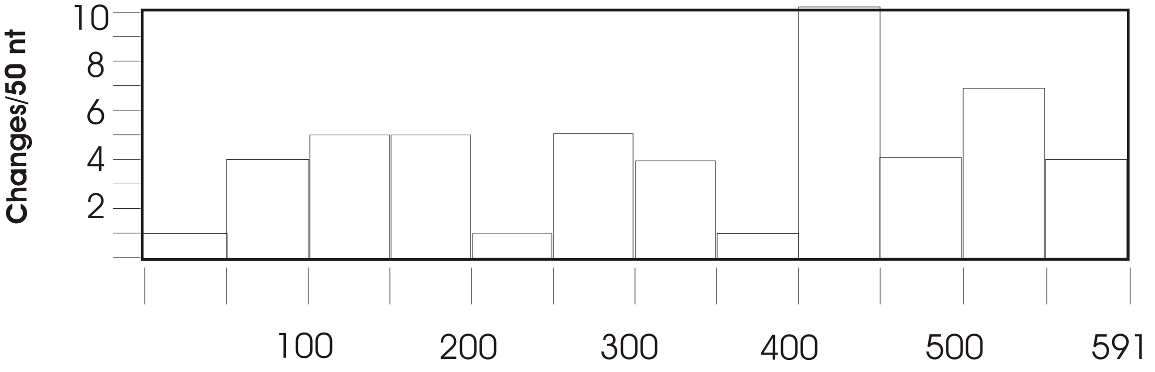
3´LTR only

49 Thing1
90 HNf1
255 Thing1
260 FREAC4
267 E4BP4
277 HFH2+3/FREAC2/SOX5
289 E4BP4
314 HFH2+3/FREAC2+4
409 Brachyury
424 E2F
428 E2F
537 HFH2
543 NRF2
548 Chop/CEBP
553 RxR/VDR

cap

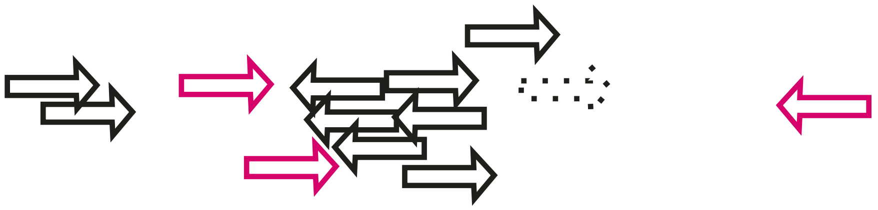
TATAAA
464

AATAAAA
536



Distribution of 5´/3´ differences

SOX5 motifs



Red arrows depict predicted SOX5 sites unique to the 5´LTR, black ones those common between 5´and 3´LTR whereas the dotted LTR is unique to the 3´LTR.

```
!§ CHR:  4
!§ CHAINSTART:  70022513
!§ STRAND: S
!§ CHAIN:
SubGene PBS, type B, score=1 , hotspot 70022513
   PBS:B (tRNALys12-HSRV,MPMV,SRV1,VILV,): Score=1 at 70022513 frame 2
[70022513-70022496]
      TGGcgcccaacgtggggc scored against
      tggggcccaaggcaggac (18 bases)

!§ CHR:  4
!§ CHAINSTART: 115910608
!§ STRAND: S
!§ CHAIN:
SubGene PBS, type CD, score=153 , hotspot 115910193
   PBS:CD (tRNAPro-MLV,HTLV,HuersP): Score=153 at 115910193 frame 2
[115910193-115910176]
      TGGgggctcgtccgggat scored against
      tgggggctcatccgggat (18 bases)

!§ CHR:  6
!§ CHAINSTART:  39968183
!§ STRAND: S
!§ CHAIN:
SubGene PBS, type CD, score=96 , hotspot 39968183
   PBS:CD (tRNAPro-MLV,HTLV,HuersP): Score=96 at 39968183 frame 3
[39968183-39968166]
      TGGgggctcgtccgggat scored against
      tgggggctcatccaggat (18 bases)

!§ CHR:  7
!§ CHAINSTART:  63858202
!§ STRAND: S
!§ CHAIN:
SubGene PBS, type C, score=141 , hotspot 63857606
   PBS:C (tRNAArg-ERV3, ver020308, Snakehead RV): Score=141 at 63857606
frame 2 [63857606-63857589]
      TGGtgagccagccaggag scored against
      tggcgagccagccaggag (18 bases)

!§ CHR:  7
!§ CHAINSTART: 137651301
!§ STRAND: S
!§ CHAIN:
SubGene PBS, type CD, score=152 , hotspot 137651125
   PBS:CD (tRNAPro-MLV,HTLV,HuersP): Score=152 at 137651125 frame 3
[137651125-137651108]
      TGGgggctcgtccgggat scored against
      tgggggctcgcccgggat (18 bases)

!§ CHR:  X
!§ CHAINSTART:  51259700
!§ STRAND: S
!§ CHAIN:
SubGene PBS, type CD, score=151 , hotspot 51259175
   PBS:CD (tRNAPro-MLV,HTLV,HuersP): Score=151 at 51259175 frame 2
[51259175-51259158]
      TGGgggctcgtccgggat scored against
      tgggggctcatccgggat (18 bases)

!§ CHR:  Y
```

```
!§ CHAINSTART:   3018841
!§ STRAND: S
!§ CHAIN:
SubGene PBS, type C, score=200 , hotspot 3018841
   PBS:C (tRNAArg-ERV3, ver020308, Snakehead RV): Score=200 at 3018841
frame 3 [3018841-3018824]
      TGGtgagccagccaggag scored against
      tggtgagccagccaggag (18 bases)

!§ CHR: 11
!§ CHAINSTART:  29579197
!§ STRAND: P
!§ CHAIN:
SubGene PBS, type CD, score=64 , hotspot 29579839
   PBS:CD (tRNAPro-MLV,HTLV,HuersP): Score=64 at 29579839 frame 2
[29579839-29579856]
      TGGgggctcgtccgggat scored against
      tgggggctcctgtgggat (18 bases)

!§ CHR: 21
!§ CHAINSTART:  14409346
!§ STRAND: P
!§ CHAIN:
SubGene PBS, type CD, score=118 , hotspot 14409729
   PBS:CD (tRNAPro-MLV,HTLV,HuersP): Score=118 at 14409729 frame 2
[14409729-14409746]
      TGGgggctcgtccgggat scored against
      tgggggctcatctgggat (18 bases)
```

manually added item:

The RRHERVI/HERV15 element at chrY_13743185 (HERV15Yq1)
Has a PBS highly similar to tRNAIle-RTVLI, RRHERV-I and HML5
TGGcccatatggggattg
tggcccagatgggggttg