1 Supplementary data to the Human Endogenous Retrovirus database (HERVd)

This example shows outputs of the search for repetitive elements in a small region of the human chromosome 12. Part 1.1. brings the numerical description of the repetitive elements found in this region. Part 1.2. is the graphical representation of the elements assembled into one member of the HERV-H family.

1.1 Part of unprocessed RepeatMasker output:

```
   SW  perc perc perc  query     position in query         matching    repeat         position in  repeat
score  div. del. ins.  sequence   begin    end   (left)    repeat      class/family   begin  end (left)  ID

...
 1003  20.0  0.0  1.6  ctg18706  360341 360525 (124660) +  HERVL       LTR/ERVL        3321 3502 (2262)  827
  864  23.4  0.0  1.2  ctg18706  360928 361094 (124091) +  MER41D      LTR/ERV1         393  557    (0)  829
  589  18.2  1.0  0.0  ctg18706  361099 361197 (123988) +  HERVL       LTR/ERVL        4511 4610 (1154)  827
  899  16.2  1.4  0.0  ctg18706  361372 361513 (123672) C  MER4E       LTR/ERV1         (0)  779    636  831
 1835   8.8  0.0  1.7  ctg18706  361514 361752 (123433) C  AluSg1      SINE/Alu        (16)  293     59  832
 3455   6.0  0.4  1.1  ctg18706  361755 362205 (122980) C  LTR7        LTR/ERV1         (2)  448      1  877
 8294   8.4  0.7  1.0  ctg18706  362206 363322 (121863) C  HERVH       LTR/ERV1         (0) 7713   6600  877
  511  12.5  1.2  0.0  ctg18706  363323 363402 (121783) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  521  11.2  1.2  0.0  ctg18706  363403 363482 (121703) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  510  14.8  0.0  0.0  ctg18706  363483 363563 (121622) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  511  16.1  0.0  0.0  ctg18706  363564 363644 (121541) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  512  14.8  0.0  0.0  ctg18706  363645 363725 (121460) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  479  15.0  1.2  0.0  ctg18706  363726 363805 (121380) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  547  12.3  0.0  0.0  ctg18706  363806 363886 (121299) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  527  14.8  0.0  0.0  ctg18706  363887 363967 (121218) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  443  17.5  1.2  0.0  ctg18706  363968 364047 (121138) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  518  13.6  0.0  0.0  ctg18706  364048 364128 (121057) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  503  16.1  0.0  0.0  ctg18706  364129 364209 (120976) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  506  16.1  0.0  0.0  ctg18706  364210 364290 (120895) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  516  13.6  0.0  0.0  ctg18706  364291 364371 (120814) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  434  18.8  1.2  0.0  ctg18706  364372 364451 (120734) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  496  16.1  0.0  0.0  ctg18706  364452 364532 (120653) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  520  14.8  0.0  0.0  ctg18706  364533 364613 (120572) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  517  13.6  0.0  0.0  ctg18706  364614 364694 (120491) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  411  19.2  1.3  0.0  ctg18706  364697 364774 (120411) C  HERVH       LTR/ERV1      (1035) 6678   6600  877
  527  13.6  0.0  0.0  ctg18706  364775 364855 (120330) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  526  13.6  0.0  0.0  ctg18706  364856 364936 (120249) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  526  13.6  0.0  0.0  ctg18706  364937 365017 (120168) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  517  13.6  0.0  0.0  ctg18706  365018 365098 (120087) C  HERVH       LTR/ERV1      (1033) 6680   6600  877
  411  19.2  1.3  0.0  ctg18706  365101 365178 (120007) C  HERVH       LTR/ERV1      (1035) 6678   6600  877
 2895  10.4  0.9  0.2  ctg18706  365179 365619 (119566) C  HERVH       LTR/ERV1      (1033) 6680   6237  877 *
15432  16.7  1.7  2.9  ctg18706  365599 365738 (119447) C  HERVH       LTR/ERV1      (2372) 4157   4017  877
 2421   8.4  0.3  0.0  ctg18706  365739 366049 (119136) +  AluSp       SINE/Alu           2  313    (0)  873
15432  16.7  1.7  2.9  ctg18706  366050 370282 (114903) C  HERVH       LTR/ERV1      (2512) 4017      1  877
 3393   7.6  0.0  2.4  ctg18706  370285 370743 (114442) C  LTR7        LTR/ERV1         (2)  448      1  877
  466   9.7  0.0  0.0  ctg18706  370744 370805 (114380) C  Alu         SINE/Alu       (255)   63      2  878
 2378  16.5  9.6  6.3  ctg18706  370830 370951 (114234) C  MER4D       LTR/ERV1       (247)  625    509  879
 1477  15.1  9.7  2.1  ctg18706  370952 371230 (113955) C  AluJb       SINE/Alu         (2)  300      1  880
 2378  16.5  9.6  6.3  ctg18706  371231 371712 (113473) C  MER4D       LTR/ERV1       (363)  509      1  879
  195  18.6  0.0  0.0  ctg18706  371738 371780 (113405) +  (TA)n       Simple_repeat      1   43    (0)  881
  629  16.7 12.5  0.0  ctg18706  371805 371924 (113261) C  AluJo       SINE/Alu       (175)  137      3  882
...
```

1.2 Graphical representation of one member of the HERV-H family

Corresponding parts of the chromosome sequence and the family consensus sequence are connected. The height of boxes on the chromosome shows Swiss-Waterman score (SW). Blue and orange colors are used for internal and LTR sequences, respectivelly. Amplification of a small region, several deletions and insertions of unrelated elements are well recognized. The unrelated repetitive elements found by RepetMasker are in red.