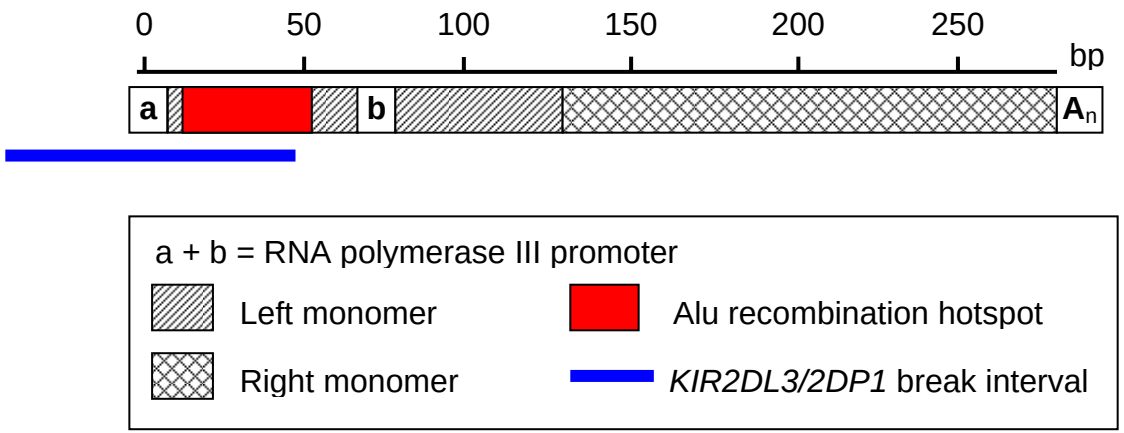
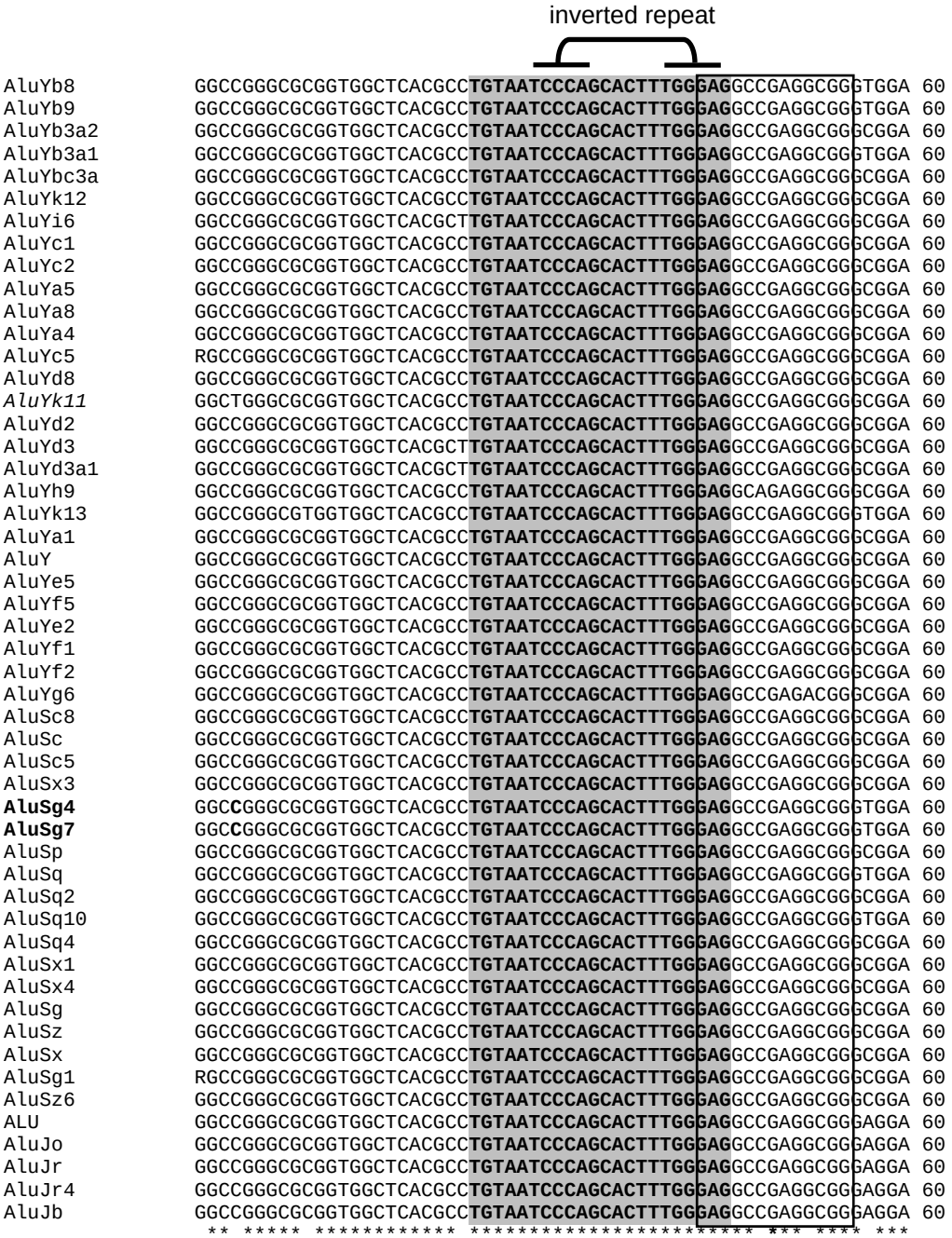
Supplementary figure 1: The structure of an Alu element. The breakpoint interval includes a recombination hotspot intrinsic to the Alu element. The length of the Alu sequence is ~282 bp, excluding the 3' poly(A) tail. The element consists of left (diagonal shaded) and right (diamond shaded) monomers. The left monomer contains an RNA polymerase III promoter (a and b). The Alu recombination hotspot is demarcated by a red box. The KIR2DL3/2DP1 break interval incorporated within the Alu element is represented by a blue line.



Supplementary figure 2: Sequence alignment of Alu element subclasses. The 22-nucleotide Alu recombination hotspot (position 24-45, grey highlighted) is conserved amongst Alu loci(23).The boxed sequence from position 43 to 55 represents a degenerate 13-mer motif that is associated with recombination hotspots and genome instability in humans(28). A potential hairpin-inciting inverted repeat, TCCCA – TGGGA, within the Alu hotspot is indicated (Figure S3).



```
AluYb8          TCAT--GAGGTCAGGAGATCGAGACCATCCTGGCTAACAAGGTGAAACCCCGTCTCTACT 118
AluYb9          TCAT--GAGGTCAGGAGATCGAGACCATCCTGGCTAACAAGGTGAAACCCCGTCTCTACT 118
AluYb3a2        TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYb3a1        TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYbc3a        TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYk12         TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYi6          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYc1          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYc2          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACAAGGTGAAACCCCGTCTCTACT 118
AluYa5          TCAC--GAGGTCAGGAGATCGAGACCATCCCGGCTAAAACGGTGAAACCCCGTCTCTACT 118
AluYa8          TCAC--GAGGTCAGGAGATCGAGACCATCCCGGCTAAAACGGTGAAACCCCGTCTCTACT 118
AluYa4          TCAC--GAGGTCAGGAGATCGAGACCATCCCGGCTAAAACGGTGAAACCCCGTCTCTACT 118
AluYc5          TCAC--GAGGTCAGGAGATCGAGACCA------------CGGTGAAACCCCGTCTCTACT 106
AluYd8          TCAC--GAGGTCAGGAGATCGAGACCA------------CGGTGAAACCCCGTCTCTACT 106
AluYk11         TCAC--AAGGTCAGGAGATCGAGACCATCTTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYd2          TCAC--GAGGTCAGGAGATCGAGACCA------------CGGTGAAACCCCGTCTCTACT 106
AluYd3          TCAC--GAGGTCAGGAGATCGAGACCA------------CGGTGAAACCCCGTCTCTACT 106
AluYd3a1        TCAC--GAGGTCAGGAGATCGAGACCA------------CGGTGAAACCCCGTCTCTACT 106
AluYh9          TCAT--GAGGTCAGGAGATCGAGACCATCCTGGCTAACGCGGTGAAACCCCGCCTCTACT 118
AluYk13         TCAT--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACAGTGAAACCCCGTCTCTACT 118
AluYa1          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluY            TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYe5          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYf5          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYe2          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYf1          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluYf2          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACAGTGAAACCCCGTCTCTACT 118
AluYg6          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluSc8          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACT 118
AluSc           TCAC--GAGGTCAAGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACT 118
AluSc5          TCAC--GAGGTCAGGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACT 118
AluSx3          TCAC--GAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 118
AluSg4          TCAC--GAGGTCAGGAGTTCGAGACCAGCCTGGCCAAGATGGTGAAACCCCGTCTCTACT 118
AluSg7          TCAC--GAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAACCCCGTCTCTACT 118
AluSp           TCACCTGAGGTCGGGAGTTCGAGACCAGCCTGACCAACATGGAGAAACCCCGTCTCTACT 120
AluSq           TCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 120
AluSq2          TCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 120
AluSq10         TCACCTGAGGTCAGGAGTTCTAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 120
AluSq4          TCAC-TGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 119
AluSx1          TCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 120
AluSx4          TCAC-TGAGGTCAGGAGTTCGAGACCAGCCTGGCCAATATGGTGAAACCCCGTCTCTACT 119
AluSg           TCAC--GAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 118
AluSz           TCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 120
AluSx           TCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACT 120
AluSg1          TCAC--GAGGTCAGGAGTTCGAGACCAGCCTGGCCAAGATGGTGAAACCCCGTCTCTACT 118
AluSz6          TCACTTGAGGCCAGGAGTTCGAGACCAGCCTGGCCAACATGGCAAAACCCCGTCTCTACT 120
ALU             TTGCTTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCTACA 120
AluJo           TTGCTTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCTACA 120
AluJr           TCGCTTGAGGCCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCTACA 120
AluJr4          TCGCTTGAGGCCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCTACA 120
AluJb           TCACTTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATGGTGAAACCCCGTCTCTACA 120
                *      **  *  *** ** ******              *  * ****** ******
```

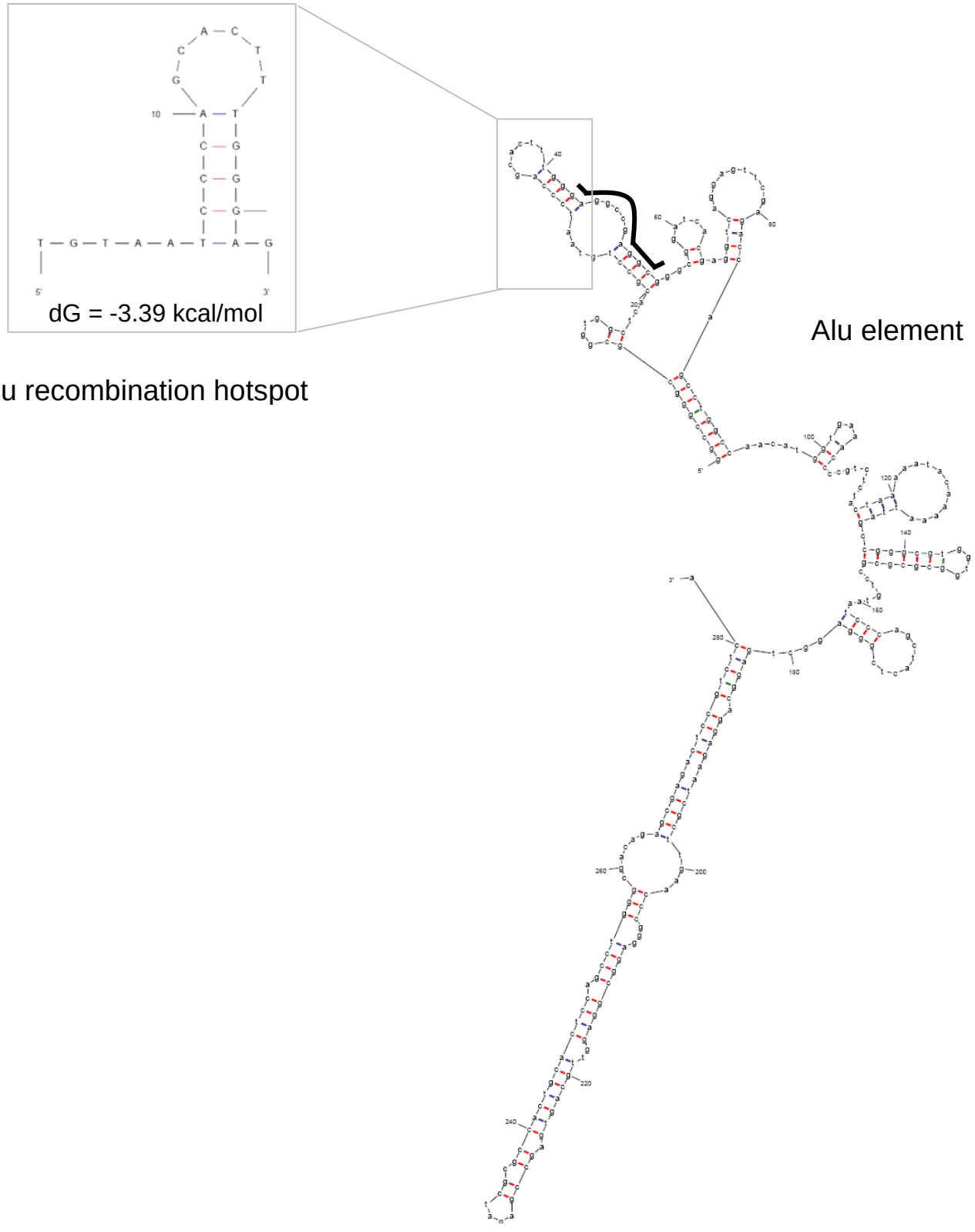
```
AluYb8          AAAAA-TACAAAAAA-TTAGCCGGGCGCGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYb9          AAAAA-TACAAAAAA-TTAGCCGGGCGCGGTGGCGGGCGCCTGTAGTCCCAGCTACTGGG 176
AluYb3a2        AAAAA-TACAAAAAA-TTAGCCGGGCGCGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYb3a1        AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYbc3a        AAAAA-TACAAAAAAATTAGCCGGGCGTGGTAGCGGGCGCCTGTAGTCCCAGCTACTCGG 177
AluYk12         AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTGGTGGGCGCCTGTAATCCCAGCTACTCGG 176
AluYi6          AAAAA-TACAAAAAAATTAGCCGGGCGTGATGGCGGGCGCCTGTAGTCCCAGCTACTCGG 177
AluYc1          AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTAGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYc2          AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTAGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYa5          AAAAA-TACAAAAAA-TTAGCCGGGCGTAGTGGCGGGCGCCTGTAGTCCCAGCTACTTGG 176
AluYa8          AAAAC-TACAAAAAA-T-AGCCGGGCGTAGTGGCGGGCGCCTGTAGTCCTAGCTACTTGG 175
AluYa4          AAAAA-TACAAAAAA-TTAGCCGGGCGTAGTGGCGGGCGCCTGTAGTCCCAGCTACTTGG 176
AluYc5          AAAAA-TACAAAAAA-TTAGCCGGGCGCAGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 164
AluYd8          AAAAA-TACAAAAAA-TTAGCCGGGCGCAGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 164
AluYk11         AAAAA-TACAAAAAA-TTAGCCGGGCGCGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYd2          AAAAA-TACAAAAAA-TTAGCCGGGCGCGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 164
AluYd3          AAAAA-TACAAAAAA-TTAGCCGGGCGCGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 164
AluYd3a1        AAAAA-TACAAAAAA-TTAGCCGGGCGCGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 164
AluYh9          AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTGGCGGGCGCCTGTGGTCCCGGCTACTCGG 176
AluYk13         AAAAA-TACAAAAAA-TTAGCCGGGAGCGGTGGCGGGCTCCTGTAGTCCCAGCTACTTGA 176
AluYa1          AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluY            AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYe5          AAAAA-TACAAAAAA-TTAGCCGGGCGAGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYf5          AAAAA-TACAAAAAA-TTAGCCGGGCGAGGTGGCGGGCGCCTGTAGTCCCAGCTACGCGG 176
AluYe2          AAAAA-TACAAAAAA-TTAGCCGGGCGAGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGG 176
AluYf1          AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACGCGG 176
AluYf2          AAAAAACACAAAAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACGCGG 178
AluYg6          AAAAA-TACAAAAAA-TTAGCCGGGCATGGTGGCGCGCGCCTGTAGTCCCAGCTACACGG 176
AluSc8          AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTGGCACGCGCCTGTAGTCCCAGCTACTCGG 176
AluSc           AAAAA-TACAAAAA--TTAGCTGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGG 175
AluSc5          AAAAA-TACAAAAA--TTAGCTGGGCGTGGTGGCGCGTGCCTGTAATCCCAGCTACTCGG 175
AluSx3          AAAAA-TACAAAAAA-TTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGG 176
AluSg4          AAAAA-TACAAAAA--TTAGCCGGGCGCGGTGGCGGGCGCCTGTAATCCCAGCTACTCGG 175
AluSg7          AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGGGCGCCTGTAATCCCAGCTACTCGG 175
AluSp           AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCATGCCTGTAATCCCAGCTACTCGG 177
AluSq           AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGGGCGCCTGTAATCCCAGCTACTCGG 177
AluSq2          AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGGGCGCCTGTAATCCCAGCTACTCGG 177
AluSq10         AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCAGGCGCCTGTAATCCCAGCTACTCGG 177
AluSq4          AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGGGCGCCTGTAATCCCAGCTACTCGG 176
AluSx1          AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGGGCGCCTGTAATCCCAGCTACTCGG 177
AluSx4          AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGG 176
AluSg           AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGG 175
AluSz           AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGG 177
AluSx           AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGG 177
AluSg1          AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGGGCCTGTAATCCCAGCTACTCGG 175
AluSz6          AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAATCCCAGCTACTCGG 177
ALU             AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGG 177
AluJo           AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGG 177
AluJr           AAAAA-TATAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGG 177
AluJr4          AAAAA-TNTAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCTAGCTACTCGG 177
AluJb           AAAAA-TACAAAAA--TTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGG 177
                ****     *****  * *** ***     * *      *****  ***  *****  *
```

```
AluYb8          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAAGCGGAGCTTGCAGTGAGCCGAGA 234
AluYb9          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAAGCGGAGCTTGCAGTGAGCCGAGA 234
AluYb3a2        -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYb3a1        -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYbc3a        -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCAAGA 235
AluYk12         -GAGGCTGAGGCAGGAGAATGGCATGAACCC-AAGAGGCGGAGCTTGCAGTGAGCCGGGA 234
AluYi6          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 235
AluYc1          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYc2          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYa5          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYa8          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 233
AluYa4          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYc5          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGAAGGCGGAGCTTGCAGTGAGCGGAGA 222
AluYd8          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGAAGGCGGAGCTTGCAGTGAGCGGAGA 222
AluYk11         -GAGGCTGAGGCAGGAGAATGGCGTGAACCT-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYd2          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 222
AluYd3          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 222
AluYd3a1        AGAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 223
AluYh9          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCTGAGG 234
AluYk13         -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-AGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYa1          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluY            -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYe5          -GAGGCTGAGGCAGGAGAATGGCGTGAACCCCGGGGGGCGGAGCCTGCAGTGAGCCGAGA 235
AluYf5          -GAGGCTGAGGCAGGAGAATGGCGTGAACCCCGGGGGGCGGAGCCTGCAGTGAGCCGAGA 235
AluYe2          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYf1          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 234
AluYf2          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGCCGAGA 236
AluYg6          -GAGGCTGAGGCAGGAGAATGGCGTGAACCC-GGGAGGCGGAGCTTGCAGTGAGTCGAGA 234
AluSc8          -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 234
AluSc           -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 233
AluSc5          -GAGGCTGAGGCAGGAGAATCGCTTGAACCA-GGGAGTCGGAGGTTGCAGTGAGCCGAGA 233
AluSx3          -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 234
AluSg4          -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 233
AluSg7          -GAGGCTGAGGCAG-AGAATTGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 232
AluSp           -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCGGTGAGCCGAGA 235
AluSq           -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 235
AluSq2          -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 235
AluSq10         -GGGGCCGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 235
AluSq4          -GAGGCTGAGGCAGGAGAATCGCTTGAAACC-GGAAGGCGGAGGTTGCAGTGAGCCGAGA 234
AluSx1          -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 235
AluSx4          -GAGGCTGAGGCAGAAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 234
AluSg           -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 233
AluSz           -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 235
AluSx           -GAGGCTGAGGCAGGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 235
AluSg1          -GAGGCTGAGGCAG-AGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 232
AluSz6          -GAGGCTGAGGCACGAGAATCGCTTGAACCC-GGGAGGCGGAGGTTGCAGTGAGCCGAGA 235
ALU             -GAGGCTGAGGCAGGAGGATCGCTTGAGCCC-AGGAGTTCGAGGCTGCAGTGAGCTATGA 235
AluJo           -GAGGCTGAGGCAGGAGGATCGCTTGAGCCC-AGGAGTTCGAGGCTGCAGTGAGCTATGA 235
AluJr           -GAGGCTGAGGCGGGAGGATCGCTTGAGCCC-AGGAGTTCGAGGCTGCAGTGAGCTATGA 235
AluJr4          -GAGGCTGAGGCAGGAGGATCGCTTGAGCCC-AGGAGTTCGAGGTTACAGTGAGCTATGA 235
AluJb           -GAGGCTGAGGCAGGAGGATCGCTTGAGCCC-GGGAGGTCGAGGCTGCAGTGAGCCGTGA 235
                 * *** *****   ** ** ** ***  *          ***  * * *****    *
```

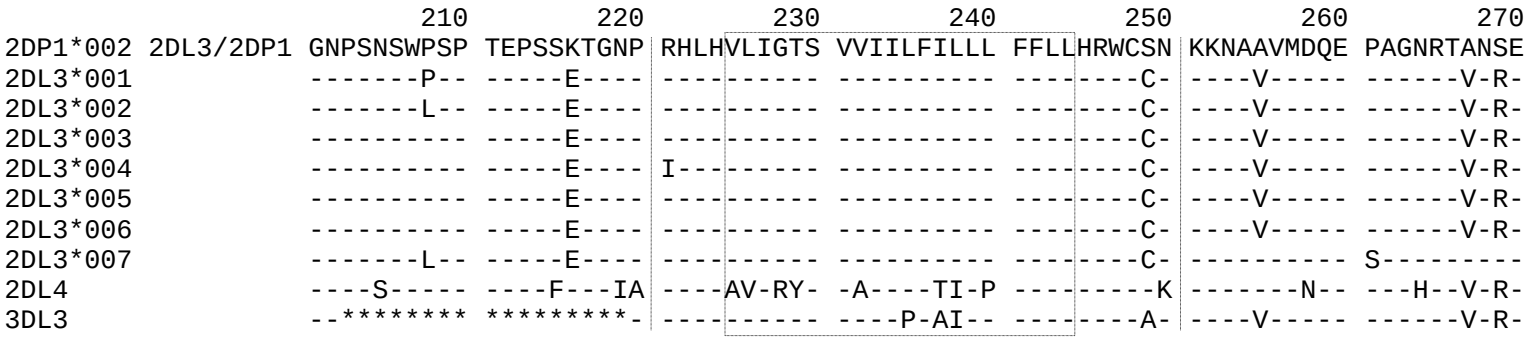
```
AluYb8          TTGCGCCACTGCAGTCCGCAGTCCGGCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 289
AluYb9          TTGCGCCACTGCAGTCCGCAGTCCGGCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 289
AluYb3a2        TAGCGCCACTGCAGTCCG-------GCCTGGGCGAAA-GAGCGAGACTCCGTCTCA---- 282
AluYb3a1        TTGCGCCACTGCAGTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluYbc3a        TAGCGCCACTGCAGTCCA-------GCCTGGGCGAAA-GAGCGAGACTCCGTCTCA---- 283
AluYk12         TAGCGCCACTGCAGTCCA-------GCTTGGGCGAAA-GAGTGAGACTCCGTCTCAAAAA 286
AluYi6          TTGCGCCACTGCACTCCC-------GCCTGGGCCACA-GAGCGAGACTCCGTCTCA---- 283
AluYc1          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluYc2          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluYa5          TCCCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluYa8          TCCCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 281
AluYa4          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluYc5          TCGCGCCACAGCACTCCC-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCAAAAA 274
AluYd8          TCGCGCCACAGCACTCCC-------GCCTGGGCGACA-GAACGAGACTCCGTCTCA---- 270
AluYk11         TTGCGCCACTGCAATCCG-------GCCTGGGCTAAA-GAGCGGGACTCCGTCTCAAAAA 286
AluYd2          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 270
AluYd3          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 270
AluYd3a1        TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 271
AluYh9          TCGCGCCACTGCACCCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluYk13         TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCCAGACGCTGTCTCAAAAA 286
AluYa1          TCCCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluY            TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluYe5          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-G--CGAGACTCCGTCTCA---- 281
AluYf5          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-G--CGAGACTCCGTCTCAAAAA 285
AluYe2          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-G--CGAGACTCCGTCTCA---- 280
AluYf1          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 282
AluYf2          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 284
AluYg6          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAAACTCCGTCTCA---- 282
AluSc8          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCAAAAA 286
AluSc           TCGCGCCACTGCACTCCA-------GCCTGG-CGACA-GAGCGAGACTCCGTCTCA---- 280
AluSc5          TCGCGCCACTGCACTCCA-------GCCTGG-CGACA-GAGCGAGACTCCGTCTCAAAAA 284
AluSx3          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCAAAAA 286
AluSg4          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCAAAAA 285
AluSg7          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCAAAAA 284
AluSp           TCGCGCCATTGCACTCCA-------GCCTGGGCAACAAGAGCGAAACTCCGTCTCA---- 284
AluSq           TCGCGCCACTGCACTCCA-------GCCTGGGCAACAAGAGCGAAACTCCGTCTCA---- 284
AluSq2          TCGCGCCATTGCACTCCA-------GCCTGGGCGACAAGAGCGAAACTCCGTCTCAAAAA 288
AluSq10         TCGCGCCATCGCACTCCA-------GCCTGGGGGACAAGAGCGAGACTTCGTCTCAAAAA 288
AluSq4          TCGCGCCACTGCACTCCA-------GCCTGGGCAACAAGAGCGAAACTCCGTCTCAAAAA 287
AluSx1          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCAAAAA 287
AluSx4          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCAAAAA 286
AluSg           TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 281
AluSz           TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 283
AluSx           TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 283
AluSg1          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCCGTCTCA---- 280
AluSz6          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACTCTGTCTCAAAAA 287
ALU             TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACCCTGTCTCAAAAA 287
AluJo           TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACCCTGTCTCA---- 283
AluJr           TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACCCTGTCTCAAAAA 287
AluJr4          TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACCCTGTCTCAAAAA 287
AluJb           TCGCGCCACTGCACTCCA-------GCCTGGGCGACA-GAGCGAGACCCTGTCTCA---- 283
                *  *****  ***  **        ** ***   * * *      **   ******
```

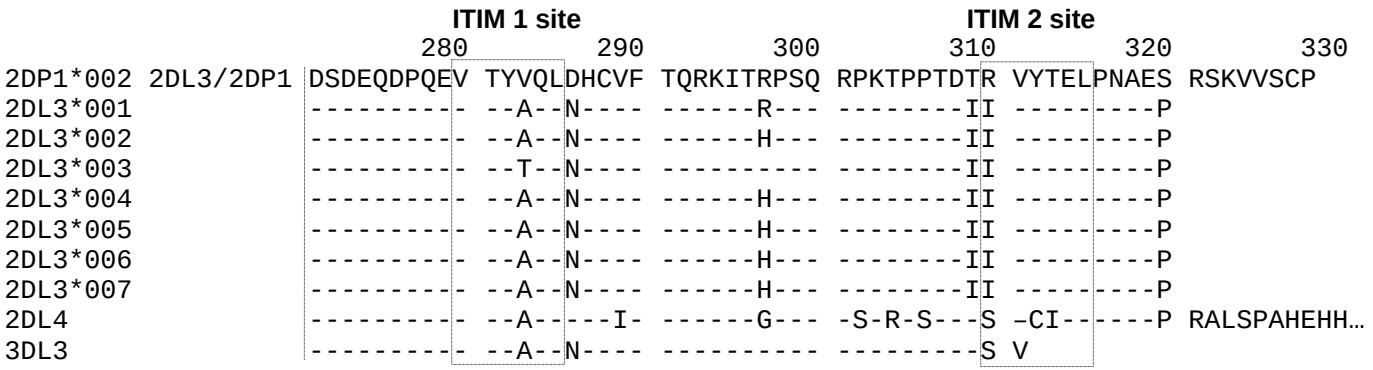
```
AluYb8          ----------------------------
AluYb9          ----------------------------
AluYb3a2        ----------------------------
AluYb3a1        ----------------------------
AluYbc3a        ----------------------------
AluYk12         AAAAAAAAAAAAAAAAAAAAAAAAAA-- 312
AluYi6          ----------------------------
AluYc1          ----------------------------
AluYc2          ----------------------------
AluYa5          ----------------------------
AluYa8          ----------------------------
AluYa4          ----------------------------
AluYc5          AAAAAAAAAAAAAAAAAAAAAAAAA--- 299
AluYd8          ----------------------------
AluYk11         AAAAAAAAAAAAAAAAAAAAAAAAAA-- 312
AluYd2          ----------------------------
AluYd3          ----------------------------
AluYd3a1        ----------------------------
AluYh9          ----------------------------
AluYk13         AAAAAA---------------------- 292
AluYa1          ----------------------------
AluY            ----------------------------
AluYe5          ----------------------------
AluYf5          AAAAAAAAAAAAAAAAAAAAAAAAAA- 312
AluYe2          ----------------------------
AluYf1          ----------------------------
AluYf2          ----------------------------
AluYg6          ----------------------------
AluSc8          AAAAAAAAAAAAAAAAAAAAAAAAAA-- 312
AluSc           ----------------------------
AluSc5          AAAAAAAAAAAAAAAAAAAAAAAAAAAA 312
AluSx3          AAAAAAAAAAAAAAAAAAAAAAAAAA-- 312
AluSg4          AAAAAAAAAAAAAAAAAAAAAAAAAA- 312
AluSg7          AAAAAAAAAAAAAAAAAAAAAAAAAAAA 312
AluSp           ----------------------------
AluSq           ----------------------------
AluSq2          AAAAAAAAAAAAAAAAAAAAAAAA---- 312
AluSq10         AAAAAAAAAAAAAAAAAAAAAAAA---- 312
AluSq4          AAAAAAAAAAAAAAAAAAAAAAAAA--- 312
AluSx1          AAAAAAAAAAAAAAAAAAAAAAAAA--- 312
AluSx4          AAAAAAAAAAAAAAAAAAAAAAAAAA-- 312
AluSg           ----------------------------
AluSz           ----------------------------
AluSx           ----------------------------
AluSg1          ----------------------------
AluSz6          AAAAAAAAAAAAAAAAAAAAAAAAA--- 312
ALU             AAAAAAAAAAAAAAAAAAAAAAAAA--- 312
AluJo           ----------------------------
AluJr           AAAAAAAAAAAAAAAAAAAAAAAAA--- 312
AluJr4          AAAAAAAAAAAAAAAAAAAAAAAAA--- 312
AluJb           ----------------------------
```

Supplementary figure 3: Potential ssDNA secondary structures formed in an Alu sequence. The Alu recombination hotspot is boxed(23). The bold black line depicts a degenerate 13-mer motif that is associated with recombination hot spots and genome instability in humans(28).

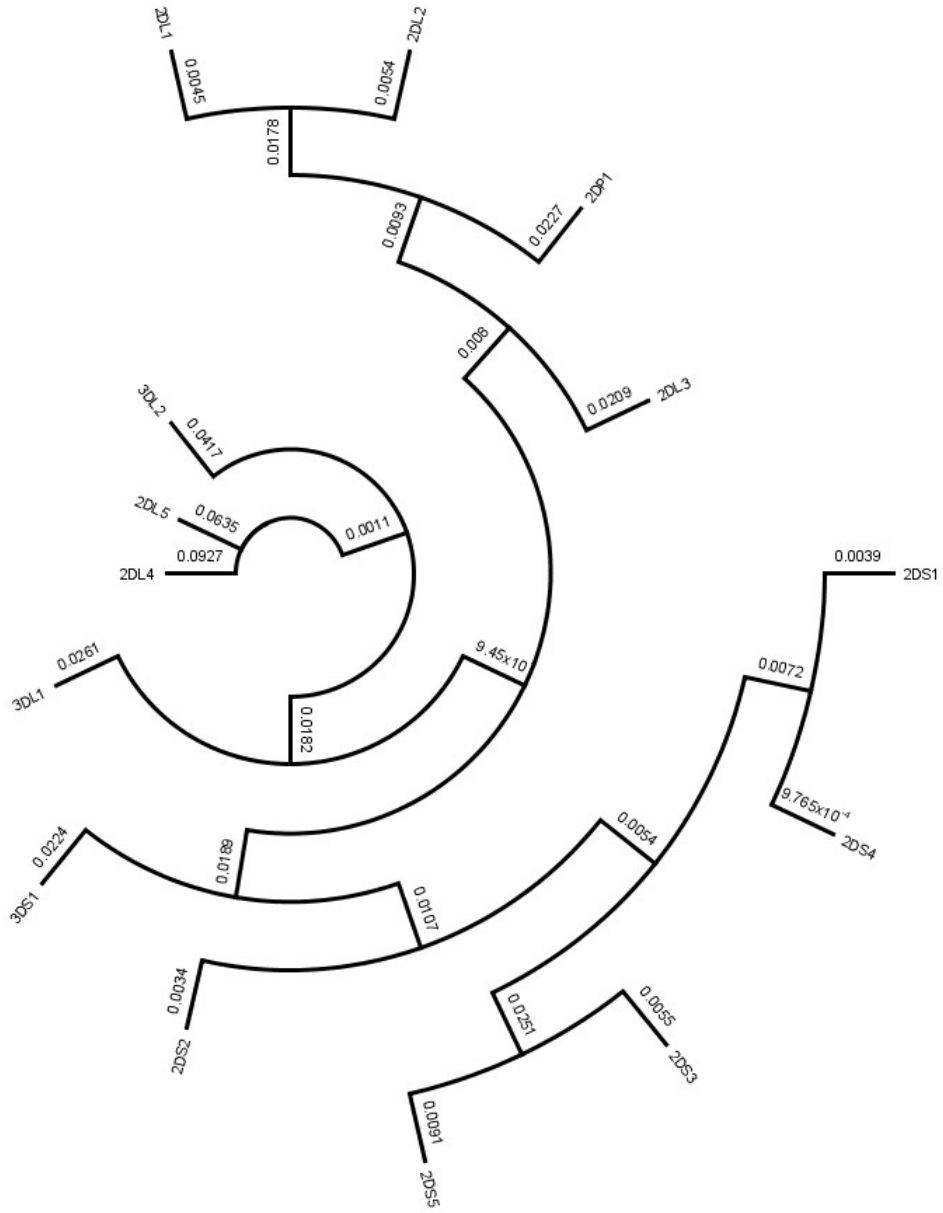


Supplementary figure 4: The KIR2DP1 single ITIM. Sequence alignment of exons 6 - 9 (encoding stem, transmembrane and cytoplasmic domains) of KIR2DL3/2DP1, KIR2DL4, KIR3DL3 and all known alleles of KIR2DL3 is shown. KIR3DL3 also possesses a single ITIM (V/I/L/SxYxxL/V/I/S), whereas KIR2DL4 possesses a single ITIM and carries a charged residue in the transmembrane region.

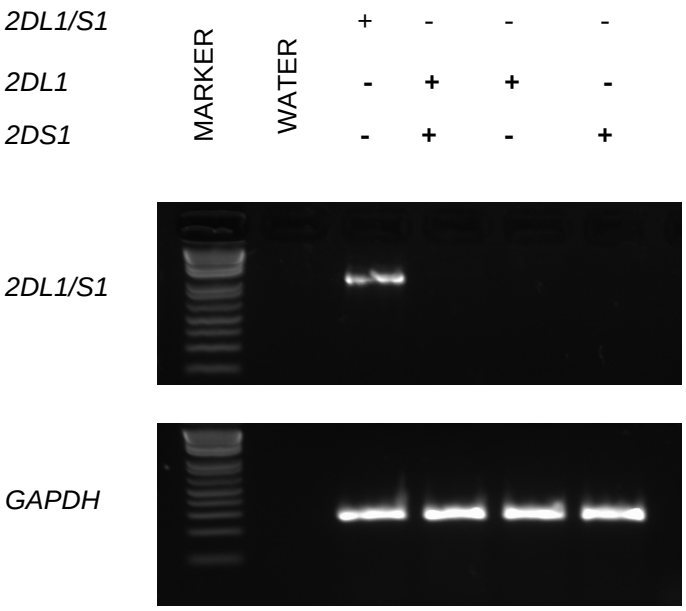




Supplementary figure 5: Phylogenetic relationship of KIR2DP1 with other human KIR groups. Tree distances are shown after each node.



Supplementary figure 6: Transcription of the KIR2DL1/2DS1 hybrid. Using specific primers for RT-PCR, KIR2DL1/S1 transcript could be amplified from NK cDNA.



Supplementary figure 7: KIR2DL1/2DS1 Sp1 and YY1 transcription factor binding sites. Sequence alignments of the region corresponding to the YY1 and SP1-binding sites of the indicated KIR genes, against the consensus binding motifs of the respective transcription factor. Nucleotide numbering corresponds to position relative to the translation initiation codon of the KIR2DL1/S1 gene, where the base A of the ATG codon is denoted nucleotide +1. The Sp1 site spans nucleotides -24 to -33 relative to the translation start site of the KIR transcript. The YY1 binding site is located at nucleotides -172 to -178, downstream of the region where antisense transcription is initiated.

```
                                  -172
                                   |
YY1 consensus                  CCATNTT
2DS1*002                 AG CCGTGTT CC
2DL1/S1                  AG CCGTGTT CC
2DL1                     AG CCGTGTT CC


                                  -24
                                   |
Sp1 consensus                  GRGGCRGGGW
2DL1/2DS1             CT CGGTCGCGGC TG
2DL1                  CT CGGTCGCGGC TG
2DS1*002              CT GGGGCGCGGC CG
3DL1*002              CT GGGGCGCGGC CG
```

Apart from a non-synonymous nucleotide substitution (amino acid change: Leu/Thr) at position +10 of the leader peptide sequence (Figure 11), *KIR2DS1* and *KIR2DL1* are identical across exons 1-3, upstream of the intron 3 *KIR2DL1/2DS1* break site. The biological consequence of the leader peptide amino acid substitution is not known, however other genomic differences in *KIR* proximal promoter sequences are known to regulate functional activity of the bidirectional *KIR* promoters and are primary determinants of receptor expression(PMID: 19008943).

Apart from the non-transcribed *KIR2DL5**002 allele that has a disrupted AML-binding site, the transcription factor binding sites (TFBS) within the core promoter region of *KIR* genes and alleles are conserved. Modulation of *KIR* promoter activity is achieved through polymorphisms in TFBS that flank the core promoter region (outer-core region). To investigate the genetic differences in the regulatory regions of *KIR2DL1/2DS1* hybrid (possessing proximal promoter sequence of *KIR2DL1*) compared to *KIR2DS1*, DNA sequences containing the previously identified bidirectional promoter region

(the 300 bp of upstream sequence from the initiation codon) were examined for sequence variation (Figure 14). A total of ten SNPs and a single 1 bp deletion were identified. None of the sequence variation located to the central area, leaving the core promoter region untouched. All sequence variation was situated in the outer-core regions associated with control of bi-directional promoter activity.

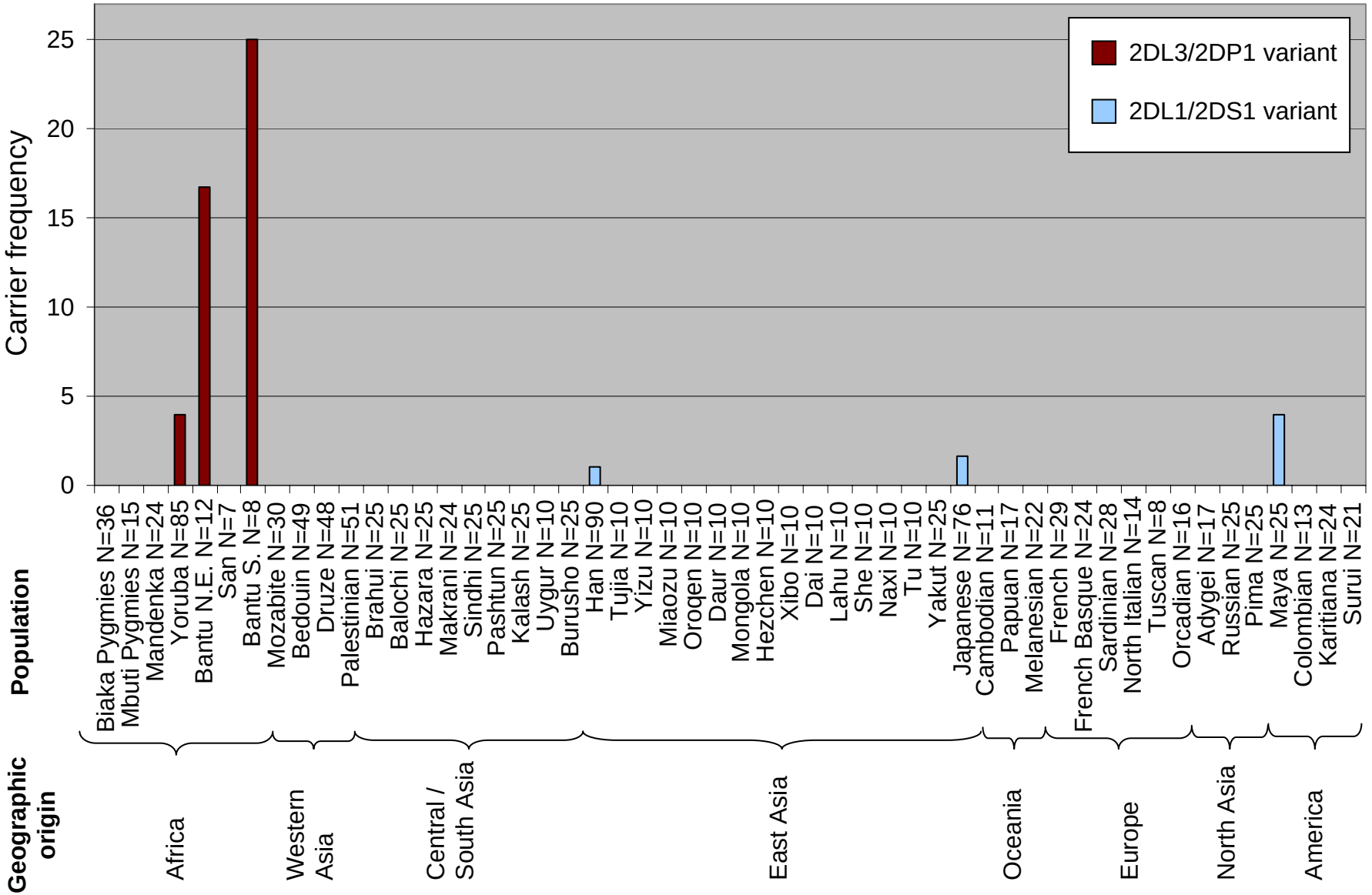
Single nucleotide differences in the Sp1 and YY1 TFBSs in the outer-core promoter region have been shown previously to have a significant effect on *KIR* transcriptional activity(PMID: 19008943). Both the *KIR2DL1/2DS1* hybrid gene (*KIR2DL1**001 promoter) and *KIR2DS1**002 share an identical sequence over the YY1 site and have an A to G nucleotide substitution at the same position with respect to the TFBS consensus sequence (Figure S7). This substitution has previously been shown to abrogate YY1 binding(PMID: 15940669) and is associated with strong antisense promoter activity(PMID: 19008943). However, *KIR2DL1/S1* (*KIR2DL1**001 promoter) and *KIR2DS1**002 differ at the Sp1 site, predicting functional differences in forward promoter activity of the genes. The *KIR2DL1/S1* promoter has three SNPs relative to the canonical Sp1 site (Figure S7). In electrophoretic mobility shift assay (EMSA) analysis, the G to T substitution seen in *KIR2DL1/2DS1* has been shown to reduce Sp1 binding to undetectable levels and is associated with increased forward promoter activity(PMID: 19008943). In contrast, *KIR2DS1**001 has only one nucleotide discrepancy at the Sp1 site with respect to the TFBS consensus sequence. EMSA analysis has shown that the *KIR2DS1* Sp1 site (which is shared by *KIR3DL1**002) maintains binding capacity to Sp1(PMID: 19008943). In the probabilistic model of the bidirectional *KIR* promoter, these sequence differences predict a higher frequency expression of the putatively activating *KIR2DL1/2DS1* relative to the wild type *KIR2DS1**002 gene, which is predicted to have strongly dominant reverse promoter activity and accordingly low gene expression(PMID: 17315044).

The nine additional nucleotide polymorphisms identified in the bidirectional promoter region outside Sp1 and YY1 sites have the potential to differentiate further the expression of the *KIR2DL1/S1* hybrid gene relative to *KIR2DS1* by altering the affinity or specificity of transcription factor binding or by changing the kinetics of transcription (Figure 14). Indeed, seven of the nine polymorphisms disrupted putative TFBS predicted by the MatInspector program. A potential heat-shock transcription factor-binding site, HSF-1, was identified in the *KIR2DS1* promoter, suggesting up-regulation by cellular stress.

Supplementary figure 8: Allele frequencies of KIR3DL2*007 and KIR3DL3*00602 in worldwide populations. Data from www.allelefrequencies.net by kind permission. New Allele Frequency Database: http://www.allelefrequencies.net. Middleton D, Menchaca L, Rood H, Komerofsky R. Tissue Antigens 2003, 61, 403-407. KIR3DL3*00602 is relatively rare in Caucasians suggesting that the ancestral precursor of the *j* and/or *t* haplotypes may originate from a different ethnic population.

Allele	Population	Phenotype Frequency (%)	Allele Frequency	Sample Size
3DL2*007	Brazil Belo Horizonte	21.1		90
3DL2*007	China Zhejiang Han		0.130	104
3DL2*007	Cuban White	30.0		70
3DL2*007	England KIR	33.1		334
3DL2*007	Finland Helsinki	46.2		101
3DL2*007	Hong Kong Chinese	33.0		100
3DL2*007	India North Hindus	57.1		72
3DL2*007	Ireland Northern	32.0	0.180	200
3DL2*007	Ireland Northern	38.3		154
3DL2*007	Japan		0.116	132
3DL2*007	Oman	17.6		99
3DL2*007	South Africa San	5.0		91
3DL2*007	South Africa Xhosa	4.0		50
3DL2*007	South Korea	17.2		154
3DL2*007	USA Caucasian		0.188	75
3DL3*00602	USA Caucasian		0.007	75

Supplementary figure 9: Ethno-geographic distribution of the KIR hybrid genes. Populations, samples sizes and carrier frequencies of recombinant genes. We tested 1,214 unrelated individuals from 52 ethnographically distinct populations.



Supplementary figure 10: KIR3DL1/2v Alu sequence ssDNA secondary structure. Potential ssDNA secondary structure formed from an Alu sequence at the KIR3DL1 breaksite of the 3DL1/2v fusion gene.