








Additional File 7. DRJ analysis

A. DNA motifs found in the direct repeated sequences flanking the IV segments inserted in wasp genomes. Analysis was performed using the DNAMINDA2 webserver (<http://bmbi.sdstate.edu/DMINDA2/annotate.php>); the input dataset was composed of 99 DRJ sequences (right junctions of HdIV and CsIV segments). A total of 89 motifs were obtained; only those whose occurrence exceed 70% of the DRJs are reported.

Motif	Length (nt)	Consensus logo	Consensus sequence	Nb of occurrence in DRJs (n=99 DRJ sequences)	Nb of DRJs containing at least one motif (/99)
Motif-62	6		TGAYGA	1231	97
Motif-56	6		CAAGAA	638	97
Motif-58	6		TGCAAT	380	90
Motif-68	7		AATGTTG	281	82
Motif-88	9		GTTGTCATY	205	76
Motif-10	8		CAATMATG	197	75
Motif-59	6		AATGCA	195	73

B. Result of genome search using motifs predicted with DMINDA 2.0 webserver

Occurrence rate of the motif in DRJ and whole genome sequences. Each motif was search among the 6 bp kmers present in the whole genome (201,969,604) and in the DRJs (33,930). The significance was evaluated using a Chi2 (taking into account the ratio of these motifs / all the other motifs in the DRJS and in the genome).

Motif	DRJ	Whole genome	P-value
TC[G,A]T CA	61	291,908	0.1015
CAAGAA	31	169,153	0.6959

C. Manual analysis of the regions containing an excision site

A CLUSTAL O(1.2.4) multiple sequence alignment was performed on the DRJs. The sequences used for alignment have 3 different origins: (i) DRJs from *H. didymator* genome; (ii) DRJs from *H. didymator* BAC clones; (iii) HdIV segments, e.g. PCR products (cloned in pGEM and Sanger sequenced) from encapsidated HdIV DNA template.

Color code:

DRJ1R (or right junction) underlined in dark grey; DRJ1L (or left junction) underlined in light grey
Nucleotides that differ between DRJ1R and DRJ1L are indicated in red (in blue, differences between genome and BAC sequences)

Regions in the segment sequence where potentially occurred a switch between right/left junctions are underlined.

Segment Hd30

Sequences used for alignment:

(i) DRJs from genome: Hd30_DRJ1R, Hd30_DRJ1L

(ii) DRJs from BAC clone # AB-06P08: right_AB, left_AB

(iii) HdIV segments: Contig_AB-20, Contig_AB-17, Contig_AB-16; plus sequence SH_AB previously sequenced

Alignment

left_AB	AGCCGACGTGATATATGCTGAAACAGGCTCGACTCGCTACGAGGCTAGTTGGTTGGTCC
Hd30_DRJ1L	AGCCGACGTGATATATGCTGAAACAGGCTCGACTCGCTACGAGGCTAGTTGGTTGGTCC
Contig_AB-17	AGTCAACATGGTGTATGTTTCAGACCAAGTCTCCCTCGCTACGAGGCTAGTTGGTTGGTCC
Contig_AB-20	AGTCAACATGGTGTATGTTTCAGACCAAGTCTCCCTCGCTACGAGGCTAGTTGGTTGGTCC
SH_AB	AGTCAACATGGTGTATGTTTCAGACCAAGTCTCCCTCGCTACGAGGCTAGTTGGTTGGTCC
Contig_AB-16	AGTCAACATGGTGTATGTTTCAGACCAAGTCTCCCTCGCTACGAGGCTAGTTGGTTGGTCC
right_AB	AGTCAACATGGTGTATGTTTCAGACCAAGTCTCCCTCGCTACGAGGCTAGTTGGTTGGTCC
Hd30_DRJ1R	AGTCAACATGGTGTATGTTTCAGACCAAGTCTCCCTCGCTACGAGGCTAGTTGGTTGGTCC ** * ** *

left_AB	ATCAGTGCATTGCACGCTCGCTACGGGTTGGCACCAGCAGCATCGTGCACACTGCGCAACTC
Hd30_DRJ1L	ATCAGTGCATTGCACGCTCGCTACGGGTTGGCACCAGCAGCATCGTGCACACTGCGCAACTC
Contig_AB-17	ATCAGTGCATTGCACGCTCGCTACGGGTTGGCACCAGCAGCATCGTGCACACTGCGCAACTC
Contig_AB-20	ATCCGTGCATTGCACGCTCGCTGCGGGTTGGCACCAGCAGCATCGTGCACACTACTGCGACTC
SH_AB	ATCCGTGCATTGCACGCTCGCTGCGGGTTGGCACCAGCAGCATCGTGCACACTACTGCGACTC
Contig_AB-16	ATCCGTGCATTGCACGCTCGCTGCGGGTTGGCACCAGCAGCATCGTGCACACTACTGCGACTC
right_AB	ATCCGTGCATTGCACGCTCGCTGCGGGTTGGCACCAGCAGCATCGTGCACACTACTGCGACTC
Hd30_DRJ1R	ATCCGTGCATTGCACGCTCGCTGCGGGTTGGCACCAGCAGCATCGTGCACACTACTGCGACTC *** ***** *

left_AB	CAGCACCTTACAACCTGCAGCGATAATACACTTCGTGCTGTGGAGCACGGTGTCAACGGTGTTC
Hd30_DRJ1L	CAGCACCTTACAACCTGCAGCGATAATACACTTCGTGCTGTGGAGCACGGTGTCAACGGTGTTC
Contig_AB-17	CAACACCTTACAACCTGCAGCGATAATACACTTCGTGCTGTGGAGCACGGTGTCAACGGTGTTC
Contig_AB-20	CGGCACCTTACGAGCTGCAGCGATAAGACACTTCGTGCTGTGGAGCACGGTGTCAACGGTGTTC
SH_AB	CGGCACCTTACGAGCTGCAGCGATAAGACACTTCGTGCTGTGGAGCACGGTGTCAACGGTGTTC
Contig_AB-16	CGGCACCTTACAACCTGCAGCGATAATACACTTCGTGCTGTGGAGCACGGTGTCAACGGTGTTC
right_AB	CGGCACCTTACGAGCTGCAGCGATAAGACACTTCGTGCTGTGGAGCACGGTGTCAACGGTGTTC
Hd30_DRJ1R	CGGCACCTTACGAGCTGCAGCGATAAGACACTTCGTGCTGTGGAGCACGGTGTCAACGGTGTTC * ***** *

Segment Hd12

Sequences used for alignment:

(i) DRJs from genome: Hd12_DRJ1R, Hd12_DRJ1L

(ii) DRJs from BAC clone # BG-42L09+BL-56G14): left_jct_BGBL1, right_jct_BGBL1

(iii) HdIV segments: Contig_BGBL1-11, Contig_BGBL1-16, Contig_BGBL1-19

Alignment

```

left_jct_BGBL1      --AGTACGGTAATGCGGCCAGTCTTTGGCGTTATACTGTTTGGCGCTTATCTGGAGTTG
Hd12_DRJ1L         ATAGTACGATAATGCGGCCAGTCTTTGGCGTACTACTGTTTACGGCTTATCTGCAGTTG
Contig_BGBL1-19    --AGAACGCTTATCACCACCAATCTTTGGCGTCATACTGTTTGGCGCTTATCTGCAGTTG
Contig_BGBL1-11    --AGAACGCTTATCACCACCAATCTTTGGCGTCATACTGTTTGGCGCTTATCTGCAGTTG
Contig_BGBL1-16    --AGAACGCTTATCACCACCAATCTTTGGCGTCATACTGTTTGGCGCTTATCTGCAGTTG
right_jct_BGBL1    --AGAACGCTTATCACCACCAATCTTTGGCGTCATACTGTTTGGCGCTTATCTGCAGTTG
Hd12_DRJ1R         ATAGAACGCTTATCACCACCAATCTTTGGCGTCATACTGTTTGGCGCTTATCTGCAGTTG
                    ** ** * * * * * ** * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
  
```

```

left_jct_BGBL1      TCACTCATGGTCAAGTTAAGAAAACCTCATTTAGATGTCTCATATTCCTGTATGGTCT
Hd12_DRJ1L         TCACTCATGAGTCAAGTTAAGAAAACCTCATTTAGATGTCTCATATTCCTGTATGGTCT
Contig_BGBL1-19    GTACTTATGGATCAAGTTGAGGAAAACCTTATCCAGGCGTCTCATACCCTTGTATGGTCT
Contig_BGBL1-11    GTACTTATGGATCAAGTTGAGGAAAACCTTATCCAGGCGTCTCATACCCTTGTATGGTCT
Contig_BGBL1-16    GTACTTATGGATCAAGTTGAGGAAAACCTTATCCAGGCGTCTCATACCCTTGTATGGTCT
right_jct_BGBL1    GTACTTATGGATCAAGTTGAGGAAAACCTTATCCAGGCGTCTCATACCCTTGTATGGTCT
Hd12_DRJ1R         GTACTTATGGATCAAGTTGAGGAAAACCTTATCCAGGCGTCTCATACCCTTGTATGGTCT
                    ** ** * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
  
```

```

left_jct_BGBL1      GCTCCGTACCTAAACGTCGCGGGTACACGTACTCAGAAAGGAATGCACACCCAGGCCTTG
Hd12_DRJ1L         GCTCCGTACCTAAACGTCGCGGGTACACGTACTCAGAAAGGAATGCACACCCAGGCCTTG
Contig_BGBL1-19    GCTCCGTACCTAAACGTCGCGGGTACACGTACTCAGAAAGGAATGCACACCCAGGCCTTG
Contig_BGBL1-11    GCTCCGTACCTAAACGTCGCGGGTACACGTACTCAGAAAGGAATGCACACCCAGGCCTTG
Contig_BGBL1-16    GCTCCGTACCTAAACGTCGCGGGTACACGTACTCAGAAAGGAATGCACACCCAGGCCTTG
right_jct_BGBL1    GCTCCGTAACCTAAACGTCGCGGGTACACGTACTCAGAAAGGAATGCACACCCAGGCCTTA
Hd12_DRJ1R         GCTCCGTAACTAAACGTCGCGGGTACACGTACTCAGAAAGGAATGCACACCCAGGCCTTA
                    ***** ** * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
  
```

```

left_jct_BGBL1      CTCAAGACATATGATGAGCATATTCGTAAGCATAGTTTGCAGATATACAACGGAGCTTTGC
Hd12_DRJ1L         CTCAAGACATATGATGAGCATATTCGTAAGCATAGTTTGCAGATATACAACGGAGCTTTGC
Contig_BGBL1-19    CTCAAGACATATGATGAGCATATTCGTAAGCATAGTTTGCAGATATACAACGGAGCTTTGC
Contig_BGBL1-11    CTCAAGACATATGATGAGCATATTCGTAAGCATAGTTTGCAGATATACAACGGAGCTTTGC
Contig_BGBL1-16    CTCAAGACATATGATGAGCATATTCGTAAGCATAGTTTGCAGATATACAACGGAGCTTTGC
right_jct_BGBL1    CTCAAGACATATGATGAGCATATTCGTAAGCATAGTTTGCAGATATACAACGGAGCTTTGC
Hd12_DRJ1R         CTCAAGACATATGATGAGCATATTCGTAAGCATAGTTTGCAGATATACAACGGAGCTTTGC
                    **** * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
  
```

```

left_jct_BGBL1      ATCCCATGTGCACATAACCATCAGAGATGAACGGGGAGATCTCCAATGCTATAAGGCC
Hd12_DRJ1L         ATCCCATGTGCACATAACCATCAGAGATGAACGGGGAGATCTCCAATGCTATAAGGCC
Contig_BGBL1-19    ATCCCATGTGCACATAACCATCAGAGATGAACGGGGAGATCTCCAATGCTATAAGGCC
Contig_BGBL1-11    ATCCCATGTGCACATAACCATCAGAGATGAACGGGGAGATCTCCAATGCTATAAGGCC
Contig_BGBL1-16    ATTTTCATGTGCACATAACCATCAGAGATGAACGGGGAGATCTCCAATGCTATAAGGCC
right_jct_BGBL1    ATCCCATGTGCACATAACCATCAGAGATGAACGGGGAGATCTCCAATGCTATAAGGCC
Hd12_DRJ1R         ATCCCATGTGCACATAACCATCAGAGATGAACGGGGAGATCTCCAATGCTATAAGGCC
                    ** * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
  
```

```

left_jct_BGBL1      CTAGTGGTTCCTACAAAGCGACGGTTTTACGTATCGGTCAACTGCTCCAGCTGTCATAGG
Hd12_DRJ1L         CTAGTGGTTCCTACAAAGCGACGGTTTTACGTATCGGTCAACTGCTCCAGCTGTCATAGG
Contig_BGBL1-19    CTAGTGGTTCCTACAAAGCGACGGTTTTACGTATCGGTCAACTGCTCCAGCTGTCATAGG
Contig_BGBL1-11    CTAGTGGTTCCTACAAAGCGACGGTTTTACGTATCGGTCAACTGCTCCAGCTGTCATAGG
Contig_BGBL1-16    CTAGTGGTTCCTACAAAGCGACGGTTTTACGTATCGGTCAACTGCTCCAGCTGTCATAGG
right_jct_BGBL1    CGAGTGGTTCCTACAGAGCGCGCTTTTCTTATCGGTCAACTAATAAAGCTGCTTAGA
Hd12_DRJ1R         CGAGTGGTTCCTACAGAGCGCGCTTTTCTTATCGGTCAACTAATAAAGCTGCTTAGA
                    * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
  
```

```

left_jct_BGBL1      TGTGTTGTTGTTGCTCGGTGCTGCTTGCAAGTAATCCATCCTGCTCGGCAGAGA
Hd12_DRJ1L         TGTGTTGTTGTTGCTCGGTGCTGCTTGCAAGTAATCCATCCTGCTCGGCAGAGA
Contig_BGBL1-19    TGTGTTGTTGTTGCTCGGTGCTGCTTGCAAGTAATCCATCCTGCTCGGCAGAGA
Contig_BGBL1-11    TGTGTTGTTGTTGCTCGGTGCTGCTTGCAAGTAATCCATCCTGCTCGGCAGAGA
Contig_BGBL1-16    TGTGTTGTTGTTGCTCGGTGCTGCTTGCAAGTAATCCATCCTGCTCGGCAGAGA
right_jct_BGBL1    TGTGTTT--TTGCTCGGTGCTGCTCGCACAAATCCATCCTGCTCGGCAGAGA
Hd12_DRJ1R         TGTGTTT--TTGCTCGGTGCTGCTCGCACAAATCCATCCTGCTCGGCAGAGA
                    ***** * * * * * * * * * * * * * * * * * * * * * * *
  
```

Segment Hd16

Sequences used for alignment:

- (i) DRJs from genome: Hd16_DRJ1R, Hd16_DRJ1L
- (ii) DRJs from BAC clone # BG-42L09+BL-56G14: left_jct_BGBL2, right_jct_BGBL2
- (iii) HdIV segments: Contig_BGBL2-12, Contig_BGBL2-16, Contig_BGBL2-19, SH_BGBL2

Alignment

```
left_jct_BGBL2 -----
Hd16_DRJ1L -----
Contig_BGBL2-19 ATAGAGTTGTCATGAGCGTAACTACATTTAAACTTGAAGCACTGTAAGAAGGAACGGTCT
Contig_BGBL2-12 ATAGAGTTGTCATGAGCGTAACTACATTTAAACTTGAAGCACTGTAAGAAGGAACGGTCT
Contig_BGBL2-16 ATAGAGTTGTCATGAGCGTAACTACATTTAAACTTGAAGCACTGTAAGAAGGAACGGTCT
SH_BGBL2 ATAGAGTTGTCATGAGCGTAACTACATTTAAACTTGAAGCACTGTAAGAAGGAACGGTCT
right_jct_BGBL2 ATAGAGTTGTCATGAGCGTAACTACATTTAAACTTGAAGCACTGTAAGAAGGAACGGTCT
Hd16_DRJ1R -----TTGTCATGAGCGTAACTACATTTAAACTTGAAGCACTGTAAGAAGGAACGGTCT
```

```
left_jct_BGBL2 -----ATAAAATTGTCATGAGAG-----GAACTAA
Hd16_DRJ1L -----TTGTCATGAGAG-----GAACTAA
Contig_BGBL2-19 CACAGTACCTTACCTACACAAGTCTCCATCATCTTACACTAATAAGGATAAGAAAAATAA
Contig_BGBL2-12 CACAGTACCTTACCTACACAAGTCTCCATCATCTTACACTAATAAGGATAAGAAAAATAA
Contig_BGBL2-16 CACAGTACCTTACCTACACAAGTCTCCATCATCTTACACTAATAAGGATAAGAAAAATAA
SH_BGBL2 CACAGTACCTTACCTACACAAGTCTCCATCATCTTACACTAATAAGGATAAGAAAAATAA
right_jct_BGBL2 CACAGTACCTTACCTACACAAGTCTCCATCATCTTACACTAATAAGGATAAGAAAAATAA
Hd16_DRJ1R CACAGTACCTTACCTACACAAGTCTCCATCATCTTACACTAATAAGGATAAGAAAAATAA
* * * * * * * * * *
```

```
left_jct_BGBL2 ATTTAAACTTGTACCAACGTAAGAGG-----AAACGCTCTCACAGTACC-TTTAGCTGC
Hd16_DRJ1L ATTTAAACTTGTACCAACGTAAGAGG-----AAACGCTCTCACAGTACC-TTTAGCTGC
Contig_BGBL2-19 AAGTAAACACATTCACACAAAACCTTTTATTTCAATCTAAGGAACAGTATATGCTGCACTA
Contig_BGBL2-12 AAGTAAACACATTCACACAAAACCTTTTATTTCAATCTAAGGAACAGTATATGCTGCACTA
Contig_BGBL2-16 AAGTAAACACATTCACACAAAACCTTTTATTTCAATCTAAGGAACAGTATATGCTGCACTA
SH_BGBL2 AAGTAAACACATTCACACAAAACCTTTTATTTCAATCTAAGGAACAGTATATGCTGCACTA
right_jct_BGBL2 AAGTAAACACATTCACACAAAACCTTTTATTTCAATCTAAGGAACAGTATATGCTGCACTA
Hd16_DRJ1R AAGTAAACACATTCACACAAAACCTTTTATTTCAATCTAAGGAACAGTATATGCTGCACTA
* * * * * * * * * *
```

```
left_jct_BGBL2 AAGCTTTAGATTAGATTCAACTGCACCATGGCCGGAACGTGTTGCGAGCCGAGTCAAGCAA
Hd16_DRJ1L AAGCTTTAGATTAGATTCAACTGCACCATGGCCGGAACGTGTTGCGAGCCGAGTCAAGCAA
Contig_BGBL2-19 TTAACACTACTCAGATTAAGCTGCACCATGGCCGGAACGTGTTGCGAGCCGAGTCAAGCAA
Contig_BGBL2-12 TTAACACTACTCAGATTAAGCTGCACCATGGCCGGAACGTGTTGCGAGCCGAGTCAAGCAA
Contig_BGBL2-16 TTAACACTACTCAGATTAAGCTGCACCATGGCCGGAACGTGTTGCGAGCCGAGTCAAGCAA
SH_BGBL2 TTAACACTACTCAGATTAAGCTGCACCATGGCCGGAACGTGTTGCGAGCCGAGTCAAGCAA
right_jct_BGBL2 TTAACACTACTCAGATTAAGCTGCACCATGGCCGGAACGTGTTGCGAGCCGAGTCAAGCAA
Hd16_DRJ1R TTAACACTACTCAGATTAAGCTGCACCATGGCCGGAACGTGTTGCGAGCCGAGTCAAGCAA
* * * * * * * * * *
```

```
left_jct_BGBL2 CACTTCTGTGTCCGTACGTTACGCCATAGAATAATATTCCAAGACGTCTGGATTGTTTTG
Hd16_DRJ1L CACTTCTGTGTCCGTACGTTACGCCATAGAATAATATTCCAAGACGTCTGGATTGTTTTG
Contig_BGBL2-19 CACTTCTGTGTCCGTACGTTACGCCATAGAATAATATTCCAAGACGTCTGGATTGTTTTG
Contig_BGBL2-12 CACTTCTGTGTCCGTACGTTACGCCATAGAATAATATTCCAAGACGTCTGGATTGTTTTG
Contig_BGBL2-16 CACTTCTGTGTCCGTACGTTACGCCATAGAATAATATTCCAAGACGTCTGGATTGTTTTG
SH_BGBL2 CACTTCTGTGTCCGTACGTTACGCCATAGAATAATATTCCAAGACGTCTGGATTGTTTTG
right_jct_BGBL2 CACTTCTGTGTCCGTACGTTACGCCATAGAATAATATTCCAAGACGTCTGGATTGTTTTG
Hd16_DRJ1R CACTTCTGTGTCCGTACGTTACGCCATAGAATAATATTCCAAGACGTCTGGATTGTTTTG
*****
```

```
left_jct_BGBL2 ATGCACATTTGCAATGATCTCAGTGAACATCTCTTTAGATTCTTTAAGAAGAATCGCGGG
Hd16_DRJ1L ATGCACATTTGCAATGATCTCAGTGAACATCTCTTTAGATTCTTTAAGAAGAATCGCGGG
Contig_BGBL2-19 ATGCACATTTGCAATGATCTCAGTGAACATCTCTTTAGATTCTTTAAGAAGAATCGCGGG
Contig_BGBL2-12 ATGCACATTTGCAATGATCTCAGTGAACATCTCTTTAGATTCTTTAAGAAGAATCGCGGG
Contig_BGBL2-16 ATGCACATTTGCAATGATCTCAGTGAACATCTCTTTAGATTCTTTAAGAAGAATCGCGGG
SH_BGBL2 ATGCACATTTGCAATGATCTCAGTGAACATCTCTTTAGATTCTTTAAGAAGAATCGCGGG
right_jct_BGBL2 ATGCACATTTGCAATGATCTCAGTGAACATCTCTTTAGATTCTTTAAGAAGAATCGCGGG
Hd16_DRJ1R ATGCACATTTGCAATGATCTCAGTGAACATCTCTTTAGATTCTTTAAGAAGAATCGCGGG
*****
```

```
left_jct_BGBL2 GATCATATATTTAGTCAACCATGCGGAGACGTGCGCCGTACAGAAATGATGCCAGTGGAG
Hd16_DRJ1L GATCATATATTTAGTCAACCATGCGGAGACGTGCGCCGTACAGAAATGATGCCAGTGGAG
Contig_BGBL2-19 GATCATATATTTAGTCAACCATGCGGAGACGTGCGCCGTACAGAAATGATGCCAGTGGAG
Contig_BGBL2-12 GATCATATATTTAGTCAACCATGCGGAGACGTGCGCCGTACAGAAATGATGCCAGTGGAG
Contig_BGBL2-16 GATCATATATTTAGTCAACCATGCGGAGACGTGCGCCGTACAGAAATGATGCCAGTGGAG
SH_BGBL2 GATCATATATTTAGTCAACCATGCGGAGACGTGCGCCGTACAGAAATGATGCCAGTGGAG
right_jct_BGBL2 GATCATATATTTAGTCAACCATGCGGAGACGTGCGCCGTACAGAAATGATGCCAGTGGAG
Hd16_DRJ1R GATCATATATTTAGTCAACCATGCGGAGACGTGCGCCGTACAGAAATGATGCCAGTGGAG
*****
```

```
left_jct_BGBL2 TTCGTTCTTACACTCACGTCCAGGTGAAACCACCTTTATGGGTTTTTTCTCGCAGGAGCA
Hd16_DRJ1L TTCGTTCTTACACTCACGTCCAGGTGAAACCACCTTTATGGGTTTTTTCTCGCAGGAGCA
Contig_BGBL2-19 TTCGTTCTTACACTCACGTCCAGGTGAAACCACCTTTATGGGTTTTTTCTCGCAGGAGCA
Contig_BGBL2-12 TTCGTTCTTACACTCACGTCCAGGTGAAACCACCTTTATGGGTTTTTTCTCGCAGGAGCA
```

Contig_BGBL2-16 TTCGTTCTTGCACTCACGTCCAGGTGAAACCACTTTATTTGGGTTTTTTCTCGCAGGAGCA
SH_BGBL2 TTCGTTCTTGCACTCACGTCCAGGTGAAACCACTTTATTTGGGTTTTTTCTCGCAGGAGCA
right_jct_BGBL2 TTCGTTCTTGCACTCACGTCCAGGTGAAACCACTTTATTTGGGTTTTTTCTCGCAGGAGCA
Hd16_DRJ1R TTCGTTCTTGCACTCACGTCCAGGTGAAACCACTTTATTTGGGTTTTTTCTCGCAGGAGCA

left_jct_BGBL2 GGTGTCGCTGGCGACCGCCTTGCACTTCTGGAAGCGAATCTTCTTCACGAACGTAGTAT
Hd16_DRJ1L GGTGTCGCTGGCGACCGCCTTGCACTTCTGGAAGCGAATCTTCTTCACGAACGTAGTAT
Contig_BGBL2-19 GGTGTCGCTGGCGACCGCCTTGCACTTCTGGAAGCGAATCTTCTTCACGAACGTAGTAT
Contig_BGBL2-12 GGTGTCGCTGGCGACCGCCTTGCACTTCTGGAAGCGAATGTTCTTCACGAACGTAGTAT
Contig_BGBL2-16 GGTGTCGCTGGCGACCGCCTTGCACTTCTGGAAGCGAATGTTCTTCACGAACGTAGTAT
SH_BGBL2 GGTGTCGCTGGCGACCGCCTTGCACTTCTGGAAGCGAATGTTCTTCACGAACGTAGTAT
right_jct_BGBL2 GGTGTCGCTGGCGACCGCCTTGCACTTCTGGAAGCGAATGTTCTTCACGAACGTAGTAT
Hd16_DRJ1R GGTGTCGCTGGCGACCGCCTTGCACTTCTGGAAGCGAATGTTCTTCACGAACGTAGTAT

left_jct_BGBL2 TACGGGCAGCCTCACAAAATCTCTCTCACCCGGCGCCGCTCCAAAATAGAGGCATGAG
Hd16_DRJ1L TACGGGCAGCCTCACAAAATCTCTCTCACCCGGCGCCGCTCCAAAATAGAGGCATGAG
Contig_BGBL2-19 TACGGGCAGCCTCACAAAATCTCTCTCACCCGGCGCCGCTCCAAAATAGAGGCATGAG
Contig_BGBL2-12 TACGGGCAGCCTCACAAAATCTCTCTCACCCGGCGCCGCTCCAAAATAGAGGCATGAG
Contig_BGBL2-16 TATGGGCAGCCTCACAAAATCTCTCTCACCCGGCGCCGCTCCAAAATAGAGGCATGAG
SH_BGBL2 TACGGGCAGCCTCACAAAATCTCTCTCACCCGGCGCCGCTCCAAAATAGAGGCATGAG
right_jct_BGBL2 TACGGGCAGCCTCACAAAATCTCTCTCACCCGGCGCCGCTCCAAAATAGAGGCATGAG
Hd16_DRJ1R TACGGGCAGCCTCACAAAATCTCTCTCACCCGGCGCCGCTCCAAAATAGAGGCATGAG
** *****

left_jct_BGBL2 CCAGTCCCAATTTATCAGCAGGCGATCTTCTTCTGGCCTTGTTGCATCGAAGCTGTAGCG
Hd16_DRJ1L CCAGTCCCAATTTATCAGCAGGCGATCTTCTTCTGGCCTTGTTGCATCGAAGCTGTAGCG
Contig_BGBL2-19 CCAGTCCCAATTTATCAGCAGGCGATCTTCTTCTGGCCTTGTTGCATCGAAGCTGTAGCG
Contig_BGBL2-12 CCAGTCCCAATTTATCAGCAGGCGATCTTCTTCTGGCCTTGTTGCATCGAAGCTGTAGCG
Contig_BGBL2-16 CCAGTCCCAATTTATCAGCAGGCGATCTTCTTCTGGCCTTGTTGCATCGAAGCTGTAGCG
SH_BGBL2 CCAGTCCCAATTTATCAGCAGGCGATCTTCTTCTGGCCTTGTTGCATCGAAGCTGTAGCG
right_jct_BGBL2 CCAGTCCCAATTTATCAGCAGGCGATCTTCTTCTGGCCTTGTTGCATCGAAGCTGTAGCG
Hd16_DRJ1R CCAGTCCCAATTTATCAGCAGGCGATCTTCTTCTGGCCTTGTTGCATCGAAGCTGTAGCG

left_jct_BGBL2 GACGGGGCATGCCTTCCCGTTCAGGAAAGTCAAGTCCAGGTTTCGGATCGATGTAAGAAA
Hd16_DRJ1L GACGGGGCATGCCTTCCCGTTCAGGAAAGTCAAGTCCAGGTTTCGGATCGATGTAAGAAA
Contig_BGBL2-19 GACGGGGCATGCCTTCCCGTTCAGGAAAGTCAAGTCCAGGTTTCGGATCGATGTAAGAAA
Contig_BGBL2-12 GACGGGGCATGCCTTCCCGTTCAGGAAAGTCAAGTCCAGGTTTCGGATCGATGTAAGAAA
Contig_BGBL2-16 GACGGGGCATGCCTTCCCGTTCAGGAAAGTCAAGTCCAGGTTTCGGATCGATGTAAGAAA
SH_BGBL2 GACGGGGCATGCCTTCCCGTTCAGGAAAGTCAAGTCCAGGTTTCGGATCGATGTAAGAAA
right_jct_BGBL2 GACGGGGCATGCCTTCCCGTTCAGGAAAGTCAAGTCCAGGTTTCGGATCGATGTAAGAAA
Hd16_DRJ1R GACGGGGCATGCCTTCCCGTTCAGGAAAGTCAAGTCCAGGTTTCGGATCGATGTAAGAAA

left_jct_BGBL2 CTGTTGATATGCTACGTTGAAACGTATAAATTCGATGAAGCTGAAAGGGGCGCCATTCCGG
Hd16_DRJ1L CTGTTGATATGCTACGTTGAAACGTATAAATTCGATGAAGCTGAAAGGGGCGCCATTCCGG
Contig_BGBL2-19 CTGTTGATATGCTACGTTGAAACGTATAAATTCGATGAAGCTGAAAGGGGCGCCATTCCGG
Contig_BGBL2-12 CTGTTGATATGCTACGTTGAAACGTATAAATTCGATGAAGCTGAAAGGGGCGCCATTCCGG
Contig_BGBL2-16 CTGTTGATATGCTACGTTGAAACGTATAAATTCGATGAAGCTGAAAGGGGCGCCATTCCGG
SH_BGBL2 CTGTTGATATGCTACGTTGAAACGTATAAATTCGATGAAGCTGAAAGGGGCGCCATTCCGG
right_jct_BGBL2 CTGTTGATATGCTACGTTGAAACGTATAAATTCGATGAAGCTGAAAGGGGCGCCATTCCGG
Hd16_DRJ1R CTGTTGATATGCTACGTTGAAACGTATAAATTCGATGAAGCTGAAAGGGGCGCCATTCCGG

left_jct_BGBL2 CATGGCAGGCCCGCGGAACGTGAAGGTGTCCACCGGGTAAGTATGTGTTCTGTCTGTGATC
Hd16_DRJ1L CATGGCAGGCCCGCGGAACGTGAAGGTGTCCACCGGGTAAGTATGTGTTCTGTCTGTGATC
Contig_BGBL2-19 CATGGCAGGCCCGCGGAACGTGAAGGTGTCCACCGGGTAAGTATGTGTTCTGTCTGTGATC
Contig_BGBL2-12 CATGGCAGGCCCGCGGAACGTGAAGGTGTCCACCGGGTAAGTATGTGTTCTGTCTGTGATC
Contig_BGBL2-16 CATGGCAGGCCCGCGGAACGTGAAGGTGTCCACCGGGTAAGTATGTGTTCTGTCTGTGATC
SH_BGBL2 CATGGCAGGCCCGCGGAACGTGAAGGTGTCCACCGGGTAAGTATGTGTTCTGTCTGTGATC
right_jct_BGBL2 CATGGCAGGCCCGCGGAACGTGAAGGTGTCCACCGGGTAAGTATGTGTTCTGTCTGTGATC
Hd16_DRJ1R CATGGCAGGCCCGCGGAACGTGAAGGTGTCCACCGGGTAAGTATGTGTTCTGTCTGTGATC

left_jct_BGBL2 TTCGTTGAACGGCATGTTGAGTCTGGATATCACCTGCGTCTGGACGATGTTCCCTTTGTG
Hd16_DRJ1L TTCGTTGAACGGCATGTTGAGTCTGGATATCACCTGCGTCTGGACGATGTTCCCTTTGTG
Contig_BGBL2-19 TTCGTTGAACGGCATGTTGAGTCTGGATATCACCTGCGTCTGGACGATGTTCCCTTTGTG
Contig_BGBL2-12 TTCGTTGAACGGCATGTTGAGTCTGGATATCACCTGCGTCTGGACGATGTTCCCTTTGTG
Contig_BGBL2-16 TTCGTTGAACGGCATGTTGAGTCTGGATATCACCTGCGTCTGGACGATGTTCCCTTTGTG
SH_BGBL2 TTCGTTGAACGGCATGTTGAGTCTGGATATCACCTGCGTCTGGACGATGTTCCCTTTGTG
right_jct_BGBL2 TTCGTTGAACGGCATGTTGAGTCTGGATATCACCTGCGTCTGGACGATGTTCCCTTTGTG
Hd16_DRJ1R TTCGTTGAACGGCATGTTGAGTCTGGATATCACCTGCGTCTGGACGATGTTCCCTTTGTG

left_jct_BGBL2 AATGTTACGGTTGAGTGTCTAGCCGTCGACTGCGCGCAAATCCATTTATACCCTCCA

```

Hd16_DRJ1L      AATGTTACGGTTGAGTGTCTAGCCGTCGACTGCGCGCAAATCCATTTATACCACTCCAA
Contig_BGBL2-19 AATGTTACGGTTGAGTGTCTAGCCGTCGACTGCGCGCAAATCCATTTATACCACTCCAA
Contig_BGBL2-12 AATGTTACGGTTGAGTGTCTAGCCGTCGACTGCGCGCAAATCCATTTATACCACTCCAA
Contig_BGBL2-16 AATGTTACGGTTGAGTGTCTAGCCGTCGACTGCGCGCAAATCCATTTATACCACTCCAA
SH_BGBL2        AATGTTACGGTTGAGTGTCTAGCCGTCGACTGCGCGCAAATCCATTTATACCACTCCAA
right_jct_BGBL2 AATGTTACGGTTGAGTGTCTAGCCGTCGACTGCGCGCAAATCCATTTATACCACTCCAA
Hd16_DRJ1R      AATGTTACGGTTGAGTGTCTAGCCGTCGACTGCGCGCAAATCCATTTATACCACTCCAG
*****

```

```

left_jct_BGBL2  TAGCTTATCAT
Hd16_DRJ1L      TAGCTTATCAT
Contig_BGBL2-19 TAGCTTATCAT
Contig_BGBL2-12 TAGCTTATCAT
Contig_BGBL2-16 TAGCTTATCAT
SH_BGBL2        TAGCTTATCAT
right_jct_BGBL2 TAATTTATCAT
Hd16_DRJ1R      TAATTTATCAT
** *****

```

Segment Hd29 (present in BAC clone # BR-08001)

Sequences used for alignment:

- (i) DRJs from genome: Hd29_DRJ1R, Hd29_DRJ1L
- (ii) DRJs from BAC clone # BR-08001: left_jct_BR, right_jct_BR
- (iii) HdIV segments: Contig_BR-1, Contig_BR-4, Contig_BR-7, Contig_BR-10, SH_BR

Alignment

```

left_jct_BR      -----TAGCGATCGCAGTGCCTCGCTGCAT
Hd29_DRJ1L      ATTTTCGTGGCAGGCTTAGTT-TTGATGGAGGGACGTAGCGATCGCAGTGCCTCGCTGCAT
Contig_BR-10    -----TGGCCACCGCTGTGCGTTGTTGAAC
Contig_BR-1     -----TGGCCACCGCTGTGCGTTGTTGAAC
Contig_BR-7     -----TGGCCACCGCTGTGCGTTGTTGAAC
SH_BR           -----TGGCCACCGCTGTGCGTTGTTGAAC
Contig_BR-4     -----TGGCCACCGCTGTGCGTTGTTGAAC
right_jct_BR    -----TGGCCACCGCTGTGCGTTGTTGAAC
Hd29_DRJ1R      ATTTTCGTGGCAAGCGACTTTTCATCGTAGAACGGATGGCCACCGCTGTGCGTTGTTGAAC
* * * * *

```

```

left_jct_BR      CAATGATGCAGCTTTGTGTTCTTCAGAACGGTGAACATTCGTGCAACCAACTTGACACG
Hd29_DRJ1L      CAATGATGCAGCTTTGTGTTCTTCAGAACGGTGAACATTCGTGCAACCAACTTGACACG
Contig_BR-10    CAATGATCCAACCTTTGTGTTCTTCAGGAACGGTGAACATTCGTGCAATCAACTTGACACG
Contig_BR-1     CAATGATCCAACCTTTGTGTTCTTCAGGAACGGTGAACATTCGTGCAATCAACTTGACACG
Contig_BR-7     CAATGATCCAACCTTTGTGTTCTTCAGGAACGGTGAACATTCGTGCAATCAACTTGACACG
SH_BR           CAATGATCCAACCTTTGTGTTCTTCAGGAACGGTGAACATTCGTGCAATCAACTTGACACG
Contig_BR-4     CAATGATCCAACCTTTGTGTTCTTCAGGAACGGTGAACATTCGTGCAATCAACTTGACACG
right_jct_BR    CAATGATCCAACCTTTGTGTTCTTCAGAACGGTGAACATTCGTGCAATCAACTTGACACG
Hd29_DRJ1R      CAATGATCCAACCTTTGTGTTCTTCAGAACGGTGAACATTCGTGCAATCAACTTGACACG
***** ** *****

```

```

left_jct_BR      TCGCCATTCGGAGAGAGGTTTCGATAGACCATGAGGAATATCCAACGTCATTCATTTGTGC
Hd29_DRJ1L      TCGCCATTCGGAGAGAGGTTTCGATAGACCATGAGGAATATCCAACGTCATTCATTTGTGC
Contig_BR-10    TCGTCATGCGGCAAGAAATTCGATAGACCATGAGGAATATCGTAACGTCATTCATTCGTGC
Contig_BR-1     TCGTCATGCGGCAAGAAATTCGATAGACCATGAGGAATATCGTAACGTCATTCATTCGTGC
Contig_BR-7     TCGTCATGCGGCAAGAAATTCGATAGACCATGAGGAATATCGTAACGTCATTCATTCGTGC
SH_BR           TCGTCATGCGGCAAGAAATTCGATAGACCATGAGGAATATCGTAACGTCATTCATTCGTGC
Contig_BR-4     TCGTCATGCGGCAAGAAATTCGATAGACCATGAGGAATATCGTAACGTCATTCATTCGTGC
right_jct_BR    TCGTCATGCGGCAAGAAATTCGATAGACCATGAGGAATATCGTAACGTCATTCATTCGTGC
Hd29_DRJ1R      TCGTCATGCGGCAAGAAATTCGATAGACCATGAGGAATATCGTAACGTCATTCATTCGTGC
*** ** *

```

```

left_jct_BR      GCACT--TCTAGGCTTTTGCAAT---CGAATCATTGCAGATACGCATTTTGAAGCAACGT
Hd29_DRJ1L      GCACT--TCTAGGCTATTGCAAT---CGAATCATTGCAGATACGCATTTTGAAGCAACGT
Contig_BR-10    ACGCATTCCTAGGACAAGTCTCTGCGCTCGAATCGTTGCGGATGCATTTGGAAGCAACGT
Contig_BR-1     ACGCATTCCTAGGACAAGTCTCTGCGCTCGAATCGTTGCGGATGCATTTGGAAGCAATGT
Contig_BR-7     ACGCATTCCTAGGACAAGTCTCTGCGCTCGAATCGTTGCGGATGCATTTGGAAGCAATGT
SH_BR           ACGCATTCCTAGGACAAGTCTCTGCGCTCGAATCGTTGCGGATGCATTTGGAAGCAATGT
Contig_BR-4     ACGCATTCCTAGGACAAGTCTCTGCGCTCGAATCGTTGCGGATGCATTTGGAAGCAATGT
right_jct_BR    ACGCATTCCTAGGACAAGTCTCTGCGCTCGAATCGTTGCGGATGCATTTGGAAGCAATGT
Hd29_DRJ1R      ACGCATTCCTAGGACAAGTCTCTGCGCTCGAATCGTTGCGGATGCATTTGGAAGCAATGT
* * * * *

```

```

left_jct_BR      GGTTCAAATGATGTTGGGAAAAGTGTGCAAAATTTATTAGATTGACCGAATGTTTCATT

```

Hd29_DRJ1L	GGTTC CA AATGATGTTGGGGAAAAGTGTG CA AAACATTTATTAGATTGACCGAATGTT C ATT
Contig_BR-10	GGTTC CA AATGATGTTGGGGAAAAGTGTG CA AAACATTTATTAGATTGACCGAATGTT C ATT
Contig_BR-1	GTCTCCAAT T ATGTTGGGAAAAGTGTCC AA AAATTTATTAGATTGACCGAATGTT A AATT
Contig_BR-7	GTCTCCAAT T ATGTTGGGAAAAGTGTCC AA AAATTTATTAGATTGACCGAATGTT A AATT
SH_BR	GTCTCCAAT T ATGTTGGGAAAAGTGTCC AA AAATTTATTAGATTGACCGAATGTT A AATT
Contig_BR-4	GTCTCCAAT T ATGTTGGGAAAAGTGTCC AA AAATTTATTAGATTGACCGAATGTT A AATT
right_jct_BR	G TCTC CAAT T ATGTTGGGAAAAGTGT C CAAA A ATTTATTAGATTGACCGAATGTT A AATT
Hd29_DRJ1R	G TCTC CAAT T ATGTTGGGAAAAGTGT C CAAA A ATTTATTAGATTGACCGAATGTT A AATT

* * * * *

left_jct_BR	TG AA CTCTTG AA AA CT CACAGCTCGAGTT AA GGAGAGTT GA CGGGATA GA CGT G AC
Hd29_DRJ1L	TG AA CTCTTG AA AA CT CACAGCTCGAGTT G--- AGAGTT GA CGGGATA GA CGT G AC
Contig_BR-10	TG AA CTCTTG AA AA CT CACAGCTCGAGTT AA GGAGAGTT GA CGGGATA GA CGT G AC
Contig_BR-1	TGGAGCTCTTG AA AA CT CACAGCTCGAGTT AA CGAGAGTT GA AGGGATA AG CGT C AC
Contig_BR-7	TGGAGCTCTTG AA AA CT CACAGCTCGAGTT AA CGAGAGTT GA AGGGATA AG CGT C AC
SH_BR	TGGAGCTCTTG AA AA CT CACAGCTCGAGTT AA CGAGAGTT GA AGGGATA AG CGT C AC
Contig_BR-4	TGGAGCTCTTG AA AA CT CACAGCTCGAGTT AA CGAGAGTT GA AGGGATA AG CGT C AC
right_jct_BR	TG GA CTCTTG AA AA CT CACAGCTCGAGTT AA CGAGAGTT GA AGGGATA AG CGT C AC
Hd29_DRJ1R	TG GA CTCTTG AA AA CT CACAGCTCGAGTT AA CGAGAGTT GA AGGGATA AG CGT C AC

* * * * *

left_jct_BR	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA
Hd29_DRJ1L	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA
Contig_BR-10	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA
Contig_BR-1	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA
Contig_BR-7	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA
SH_BR	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA
Contig_BR-4	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA
right_jct_BR	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA
Hd29_DRJ1R	GTTCTTCATGG AA GGT G GGCTCG CA AA CG TCGCAT G CAGCTTGC G AGTTCC AA AT G AAA

* * * * *

left_jct_BR	ACA ACTGCGCGGGGTATTGTT T ATTCT GG GAG AC ACCG
Hd29_DRJ1L	ACA ACTGCGCGGGGTATTGTT T ATTCT GG GAG AC ACCG
Contig_BR-10	ACA ACTGCGCGGGGTATTGTT T ATTCT GG GAG AC ACCG
Contig_BR-1	ACA ACTGCGCGGGGTATTGTT T ATTCT GG GAG AC ACCG
Contig_BR-7	ACA ACTGCGCGGGGTATTGTT T ATTCT GG GAG AC ACCG
SH_BR	ACA ACTGCGCGGGGTATTGTT T ATTCT GG GAG AC ACCG
Contig_BR-4	C -AACTGCGTGGGGTATTGTT T ATTCT T GGAGAG AC CGG
right_jct_BR	C -AACTGCGTGGGGTATTGTT T ATTCT T GGAGAG AC CGG
Hd29_DRJ1R	C -AACTGCGTGGGGTATTGTT T ATTCT T GGAGAG AC CGG

* * * * * no switch ??? *

Segment Hd22

Sequences used for alignment:

- (i) DRJs from genome: Hd22_DRJ1R, Hd22_DRJ1L
- (ii) DRJs from BAC clone # CN-29M18: left_junction, right_junction
- (iii) HdIV segments: Contig_CN-31, Contig_CN-33, Contig_CN-36, Contig_CN-39

Alignment

left_junction	CAATAGGGTGAAT T CA AA CGCT GGT CC AA TGGCAGGAGCTT GGC AG G ACAGAT G ACAT C
Hd22_DRJ1L	-----AA T CA AA CGCT GGT CC AA TGGCAGGAGCTT GGC AG G ACAGAT G ACAT C
Contig_CN-31	CAATAGGGTGAAT T CA AA CGCT GGT CC AA TGGCAGGAGCTT GGC AG G ACAGAT G ACAT C
Contig_CN-33	CAATAGGGTGAAT T CA AA CGCT GGT CC AA TGGCAGGAGCTT GGC AG G ACAGAT G ACAT C
Contig_CN-39	CAATAGGGTGAAT T CA AA CGCT GGT CC AA TGGCAGGAGCTT GGC AG G ACAGAT G ACAT C
Contig_CN-36	CAATAGGGTGAAT T CA AA CGCT GGT CC AA TGGCAGGAGCTT G CA AG GCAGAT G AT C
right_junction	CATTGAGAAAA C TA AG CG CG CA T AG CA AATGGCAGGAGCTT GA CA AG GCAGAT G AT C
Hd22_DRJ1R	-----AA C TA AG CG CG CA T AG CA AATGGCAGGAGCTT GA CA AG GCAGAT G AT C

* * * * *

left_junction	GAGTGA A TGCGTAGG AG TCT GA AA AG CG CG TCGTGAACAATATCGTGACCATTGATCG C
Hd22_DRJ1L	GAGTGA A TGCGTAGG AG TCT GA AA AG CG CG TCGTGAACAATATCGTGACCATTGATCG C
Contig_CN-31	GAGTGA A TGCGTAGG AG TCT GA AA AG CG CG TCGTGAACAATATCGTGACCATTGATCG C
Contig_CN-33	GAGTGA A TGCGTAGG AG TCT GA AA AG CG CG TCGTGAACAATATCGTGACCATTGATCG C
Contig_CN-39	GAGTGA A TGCGTAGG AG TCT GA AA AG CG CG TCGTGAACAATATCGTGACCATTGATCG C
Contig_CN-36	GAGTGA A TGCGTAGG AG TCT GA AA AG CG CG TCGTGAACAATATCGTGACCATTGATCG C
right_junction	GAGTGA A TGCGTAGG AG TCT GA AA AG CG CG TCGTGAACAATATCGTGACCATTGATCG C
Hd22_DRJ1R	GAGTGA A TGCGTAGG AG TCT GA AA AG CG CG TCGTGAACAATATCGTGACCATTGATCG C

* * * * *

left_junction	CATAGCGACTTGCCGCATCTCAGGATCGTGATCTAAAA-----CGTTTCACGTGA
Hd22_DRJ1L	CATAGCGACTTGCCGCATCTCAGGATCGTGATCTAAAA-----CGTTTCACGTGA
Contig_CN-31	CAAGCGACTTGGCGACTTGGCACACGTCAGCGTCAAGTTCCACAAAATCTGACGCGA
Contig_CN-33	CAAGCGACTTGGCGACTTGGCACACGTCAGCGTCAAGTTCCACAAAATCTGACGCGA
Contig_CN-39	CATAGCGACTTGGCGACTTGGCACACGTCAGCGTCAAGTTCCACAAAATCTGACGCGA
Contig_CN-36	CACAGCGACTTGGCGACTTGGCACACGTCAGCGTCAAGTTCCACAAAATCTGACGCGA
right_junction	CACAGCGACTTGGCGACTTGGCACACGTCAGCGTCAAGTTCCACAAAATCTGACGCGA
Hd22_DRJ1R	CACAGCGACTTGGCGACTTGGCACACGTCAGCGTCAAGTTCCACAAAATCTGACGCGA

** ***** ** *

Segment Hd19

Sequences used for alignment:

- (i) DRJs from genome: Hd19_DRJ1R, Hd19_DRJ1L
- (ii) DRJs from BAC clone # CP-38F08+CQ-40J21: CPCQ_right, CPCQ_left
- (iii) HdIV segments: jct_CPCQ-28, jct_CPCQ-27, jct_CPCQ-23, jct_CPCQ-22

Alignment

CPCQ_left	TTTTGCTGTTTCATGTTTCAGAGCAGCGGTCTATGTGATACACGCTGCTTAGCCTGGT
Hd19_DRJ1L	TTTTGCTGTTTCATGTTTCAGAGCAGCGGTCTATGTGATACACGCTGCTTAGCCTGGT
jct_CPCQ-28	TTTTGCTGCTTCATGTTTCAGAGCAGCTTGTCTACATGATACACGCTGATTTGTC-TGT
jct_CPCQ-27	TTTTGCTGCTTCATGTTTCAGAGCAGCTTGTCTACATGATACACGCTGATTTGTC-TGT
jct_CPCQ-23	TTTTGCTGCTTCATGTTTCAGAGCAGCTTGTCTACATGATACACGCTGATTTGTC-TGT
jct_CPCQ-22	TTTTGCTGCTTCATGTTTCAGAGCAGCTTGTCTACATGATACACGCTGATTTGTC-TGT
CPCQ_right	TTTTGCTGCTTCATGTTTCAGAGCAGCTTGTCTACATGATACACGCTGATTTGTC-TGT
Hd19_DRJ1R	TTTTGCTGCTTCATGTTTCAGAGCAGCTTGTCTACATGATACACGCTGATTTGTC-TGT

***** ***** ***** ***** ***** ***** ***** ***** ***** *****

CPCQ_left	GACGGCAGCATGGTTCCTCCTCATGCTGCACGAGCTCTGTGAGCTTT
Hd19_DRJ1L	GACGGCAGCATGGTTCCTCCTCATGCTGCACGAGCTCTGTGAGCTTT
jct_CPCQ-28	GACGGCAGCATGGTTCCTCCTCATGCTGCACGAGCTCTGTGAGCTTT
jct_CPCQ-27	GACGGCAGCATGGTTCCTCCTCATGCTGCACGAGCTCTGTGAGCTTT
jct_CPCQ-23	GACGGCAGCATGGTTCCTCCTCATGCTGCACGAGCTCTGTGAGCTTT
jct_CPCQ-22	GACGGCAGCATGGTTCCTCCTCATGCTGCACGAGCTCTGTGAGCTTT
CPCQ_right	GACGGCAGCATGGTTCCTCCTCATGCTGCACGAGCTCTGTGAGCTTT
Hd19_DRJ1R	GACGGCAGCATGGTTCCTCCTCATGCTGCACGAGCTCTGTGAGCTTT

***** ** *

Segments Hd28

Sequences used for alignment:

- (i) DRJs from genome: Hd28_DRJ1R, Hd28_DRJ1L, Hd24_DRJ1R (reverse), Hd24_DRJ1L (reverse)
- (ii) DRJs from BAC clone # CR-47M17: CR_left, CR_right
- (iii) HdIV segments: Contig_CR-05, Contig_CR-06, Contig_CR-09, Contig_CR-10

Alignment

CR_left	GCCTTTGAGCTGAGCT-GCATGAGGTTTTCATGTCCGATAACTGGATTCATTGTCA
Hd28_DRJ1L	GCCTTTGAGCTGAGCT-GCATGAGGTTTTCATGTCCGATAACTGGATTCATTGTCA
Contig_CR-05	GCCTTTGGAGCTAAACTTGCACGAGATTTCCGGTGTCCGTGATAACTGAAATGGCATTGTCA
Contig_CR-06	GCCTTTGGAGCTAAACTTGCACGAGATTTCCGGTGTCCGTGATAACTGAAATGGCATTGTCA
CR_right	GCCTTTGAGCTAAACTTGCACGAGATTTCCGGTGTCCGTGATAACTGAAATGGCATTGTCA
Hd28_DRJ1R	GCCTTTGAGCTAAACTTGCACGAGATTTCCGGTGTCCGTGATAACTGAAATGGCATTGTCA

***** *

CR_left	GGCGTAAGCATTATGACGGCGGTGGTGTAGGCGTTGCTTGGGAAGGTTGGTGAGCCACG
Hd28_DRJ1L	GGCGTAAGCATTATGACGGCGGTGGTGTAGGCGTTGCTTGGGAAGGTTGGTGAGCCACG
Contig_CR-05	GGCGTAATATTTATGACGGC-----GGTGATG
Contig_CR-06	GGCGTAATATTTATGACGGC-----GGTGATG
CR_right	GGCGTAAATTTATGACGGC-----GGTGATG
Hd28_DRJ1R	GGCGTAAATTTATGACGGC-----GGTGATG

***** ***** * * * * *

CR_left TCGACGTAGATTTCAGGCGTGAACGGAGAGCTGAAAGGTATAGTAGTCGTTTCGACCAACT
Hd28_DRJ1L TCGACGTAGATTTCAGGCGTGAACGGAGAGCTGAAAGGTATAGTAGTCGTTTCGACCAACT
Contig_CR-05 TTAACGTAGGTTTCAGGCGTTAGACTGAGAGCCGGAAGGTATGGTAGTCGTTTCGACCGTCT
Contig_CR-06 TTAACGTAGGTTTCAGGCGTTAGACTGAGAGCCGGAAGGTATGGTAGTCGTTTCGACCGTCT
CR_right TTAACGTAGGTTTCAGGCGTTAGACTGAGAGCCGGAAGGTATGGTAGTCGTTTCGACCGTCT
Hd28_DRJ1R TTAACGTAGGTTTCAGGCGTTAGACTGAGAGCCGGAAGGTATGGTAGTCGTTTCGACCGTCT
* * * * *

CR_left CGGAATCACACCGACGAAACTATTGTCCATTGCTCAGGCTATGAGGACTTGAGGCAAT
Hd28_DRJ1L CGGAATCACACCGACGAAACTATTGTCCATTGCTCAGGCTATGAGGACTTGAGGCAAT
Contig_CR-05 CCGAATCACGCTGACGAGAACTCTTCGTTTCATCGCTTAAGCTCTGAGGACTTGACGCAAT
Contig_CR-06 CCGAATCACGCTGACGAGAACTCTTCGTTTCATCGCTTAAGCTCTGAGGACTTGACGCAAT
CR_right CCGAATCACGCTGACGAGAACTCTTCGTTTCATCGCTTAAGCTCTGAGGACTTGACGCAAT
Hd28_DRJ1R CCGAATCACGCTGACGAGAACTCTTCGTTTCATCGCTTAAGCTCTGAGGACTTGACGCAAT
* * * * *

CR_left CAAGTTCGGACGCAAAGT-----TTTGTTCGGCCATAGATCACAAAGTAAGAA
Hd28_DRJ1L CAAGTTCGGACGCAAAGT-----TTTGTTCGGCCATAGATCACAAAGTAAGAA
Contig_CR-05 CAAGTTCGGATGATAAACGCCAGACGCAAAGTTTGTTCATAGATCACAAAGTAAGAA
Contig_CR-06 CAAGTTCGGATGATAAACGCCAGACGCAAAGTTTGTTCATAGATCACAAAGTAAGAA
CR_right CAAGTTCGGATGATAAACGCCAGACGCAAAGTTTGTTCATAGATCACAAAGTAAGAA
Hd28_DRJ1R CAAGTTCGGATGATAAACGCCAGACGCAAAGTTTGTTCATAGATCACAAAGTAAGAA
* * * * *

CR_left CAGTTCTGGCAGACTGGACAGTCCACCGATACCTCACAAGACTGTAAA-ACCTCTGGAC
Hd28_DRJ1L CAGTTCTGGCAGACTGGACAGTCCACCGATACCTCACAAGACTGTAAA-ACCTCTGGAC
Contig_CR-05 CAGTTCTTACGGACCTGACAGTCCAACATTAATCTCAGACACTGTAAA-ACCTATGGAC
Contig_CR-06 CAGTTCTTACGGACCTGACAGTCCAACATTAATCTCAGACACTGTAAA-ACCTATGGAC
CR_right CAGTTCTTACGGACCTGACAGTCCAACATTAATCTCAGACACTGTAAA-ACCTATGGAC
Hd28_DRJ1R CAGTTCTTACGGACCTGACAGTCCAACATTAATCTCAGACACTGTAAA-ACCTATGGAC
* * * * *

CR_left GTACTGCGCACGTACCCAGTCCACGAAACACTCGTGCCACATCCATTTGATGTTGGCTT
Hd28_DRJ1L GTACTGCGCACGTACCCAGTCCACGAAACACTCGTGCCACATCCATTTGATGTTGGCTT
Contig_CR-05 GTACTTTCGCACGTACCCGAGTCCACGAAACACTCGTGCTACATCCATTTGACGTTGGCCT
Contig_CR-06 GTACTTTCGCACGTACCCGAGTCCACGAAACACTCGTGCTACATCCATTTGACGTTGGCCT
CR_right GTACTTTCGCACGTACCCAGTCCACGAAACACTCGTGCTACATCCATTTGACGTTGGCCT
Hd28_DRJ1R GTACTTTCGCACGTACCCAGTCCACGAAACACTCGTGCTACATCCATTTGACGTTGGCCT
* * * * *

CR_left CTGCCAGGCATGTCGCTAT--AGATTTCCAGCGTGTGCCCGTGTCCAGAAAAGCTTT
Hd28_DRJ1L CTGCCAGGCATGTCGCTAT--AGATTTCCAGCGTGTGCCCGTGTCCAGAAAAGCTTT
Contig_CR-05 TCT---GACAAGGCCACTATAGAGTCCAGCGTGTGCCCGTGTCCAGAAAAGCTTT
Contig_CR-06 TCT---GACAAGGCCACTATAGAGTCCAGCGTGTGCCCGTGTCCAGAAAAGCTTT
CR_right TCT---GACAAGGCCACTATAGAGTCCAGCGTGTGCCCGTGTCCAGAAAAGCTTT
Hd28_DRJ1R TCT---GACAAGGCCACTATAGAGTCCAGCGTGTGCCCGTGTCCAGAAAAGCTTT
* * * * *

CR_left GTTCACAACCGGACTTTTACTGTACCTTTGTGTTACAGTTGCGTTACGTAACACTACT
Hd28_DRJ1L GTTCACAACCGGACTTTTACTGTACCTTTGTGTTACAGTTGCGTTACGTAACACTACT
Contig_CR-05 T-ACTCCAACCGGACTTTTACTGTACCTTTGTGTTACAGTTGCGTTACGTAACACTACT
Contig_CR-06 T-ACTCCAACCGGACTTTTACTGTACCTTTGTGTTACAGTTGCGTTACGTAACACTACT
CR_right T-ACTCCAACCGGACTTTTACTGTACCTTTGTGTTACAGTTGCGTTACGTAACACTACT
Hd28_DRJ1R T-ACTCCAACCGGACTTTTACTGTACCTTTGTGTTACAGTTGCGTTACGTAACACTACT
* * * * *

CR_left ACTGTGGAACCTTCTACTGCCACAGTACAACAGCTATCGACAACGTGTTCTTTCAGTAT
Hd28_DRJ1L ACTGTGGAACCTTCTACTGCCACAGTACAACAGCTATCGACAACGTGTTCTTTCAGTAT
Contig_CR-05 ACTGTGGAACCTTCTACTGCCACAGTACAACAGCTATCGACAACAGTCTTTCAGTAT
Contig_CR-06 ACTGTGGAACCTTCTACTGCCACAGTACAACAGCTATCGACAACAGTCTTTCAGTAT
CR_right ACTGTGGAACCTTCTACTGCCACAGTACAACAGCTATCGACAACAGTCTTTCAGTAT
Hd28_DRJ1R ACTGTGGAACCTTCTACTGCCACAGTACAACAGCTATCGACAACAGTCTTTCAGTAT
* * * * *

CR_left CAAGAAGTCATTATGCAAGTAGAGAATGGCCGGGTTTTAAGTTGCATGAAA-GCACTCAA
Hd28_DRJ1L CAAGAAGTCATTATGCAAGTAGAGAATGGCCGGGTTTTAAGTTGCATGAAA-GCACTCAA
Contig_CR-05 CAAGAAGTAATTATCAAGTGGAGAATGGCCGGGTTTTAAGTTGCATGAAA-GCACTCAA
Contig_CR-06 CAAGAAGTAATTATCAAGTGGAGAATGGCCGGGTTCGG-AGTTGCATGAAA-GCACTCAA
CR_right CAAGAAGTAATTATCAAGTGGAGAATGGCCGGGTTCGG-AGTTGCATGAAA-GCACTCAA

```
Hd28_DRJ1R CAAGAAGTAATTATCAAGTGGAGAATGGCCGGGTGG-AGTTGCATGAAAAGCACGGAA
***** ***** ***** ***** ***** ***** ***** **

CR_left TTGTCTACAATGCAGACGAG
Hd28_DRJ1L TTGTCTACAATGCAGACGAG
Contig_CR-05 TTGTCTACAATGCAGACGAG
Contig_CR-06 TTGTCTACAATGCAGACGAG
CR_right GTGTGTACAATACAGACGGG
Hd28_DRJ1R GTGTGTACAATACAGACGGG
*** ***** ***** *
```