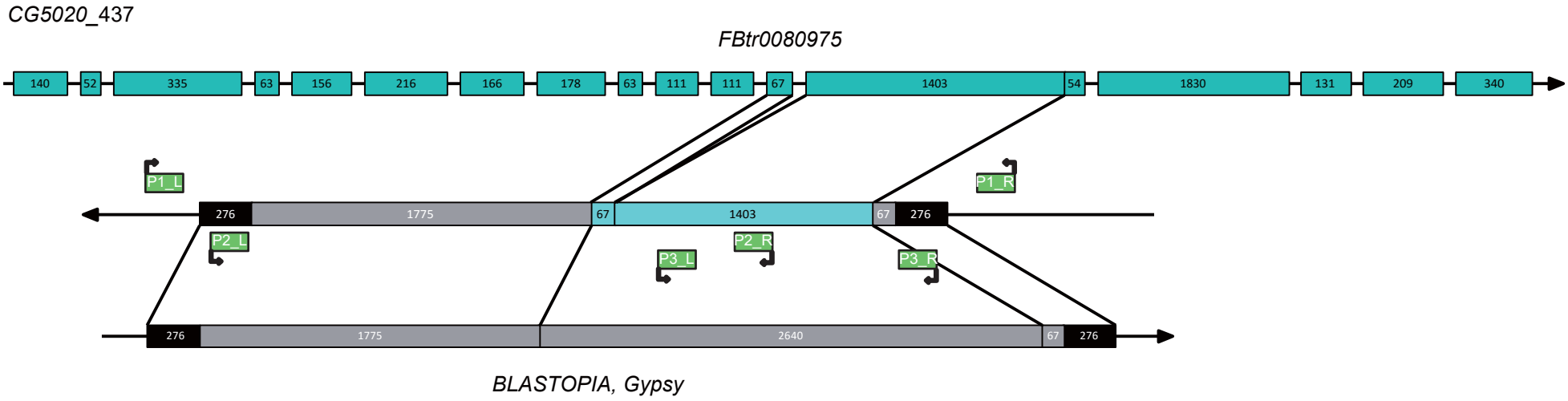
Sanger sequencing result of the retroposed locus, *CG5020_r*.

On the basis of the location information of *CG5020_r* revealed by mate-pair sequencing data, we used the following primers to do PCR and sequence the whole retroposed locus.
P1_L “TTTGGGCTTTGTCTATCCGCAC”;
P1_R “TAACCCTTCACCTCCCCCTTCG”.

Supposedly, driven by the 5’ regulatory context of the LTR retrotransposon, this chimeric locus is transcribed as a long transcript just like the canonical LTR retrotransposon. In order to test this possibility, we designed two sets of primers to RT-PCR the whole locus with the corresponding products overlapping each other while avoiding amplifying the corresponding LTR elements. The exact primers are as follows:
P2_L “GCAACATGATTGTCGGAAATCCA”;
P2_R “TCAAGTTCGTTTTGTTTTAAAGCCA”;
P3_L “CTGAGTTCACCCGAAAACTAGC”;
P3_R “CGATGACTGATTGTCGATAACTTC”.

The position of all these three sets of primers relative to the retroposed locus is shown below.
The figure convention follows Supplemental Dataset 1.



We aligned the *de novo* assembled fragment together with flanking regions predicted based on consensus *BLASTOPIA* sequence, DNA sequence and two RNA sequences and show that these sequences are consistent with each other. In other words, the theoretical model in Fig. 4 holds and the chimeric retroposed locus is transcribed as a standard LTR retrotransposon.

The alignment is color-coded as follows:
5' and 3' LTR retrotransposon repeat zone
Coding region of LTR retrotransposon *BLASTOPIA*
Microsimilarity shared by LTR retrotransposon and retrocopy
Retrocopy
Assembled sequence
Primer

```
>CG5020_predicted&assembled   ----------------------------------------TGTAACAT-
>CG5020_sequenced_DNA         TTGTAAATTCTTTTGTTAATTTGTAGAAACAGAGACATACATGTAACATG
>CG5020_sequenced_RNA_1       ------------------------------------------------
>CG5020_sequenced_RNA_2       ------------------------------------------------
                                                                      ********

>CG5020_predicted&assembled   GAGTAAGGCTGAAGGCTGGCAACAACCCGGTTGGCAGCGCTGTTGAGCAG
>CG5020_sequenced_DNA         GAGTAAGGCTGAAGGCTGGCAACAACCCGGTTGGCAGCGCTGTTGAGCAG
>CG5020_sequenced_RNA_1       ------------------------------------------------
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************************************

>CG5020_predicted&assembled   CAACATGATTGTCGGAAATCCAAGTTATCGACAATCAGTCATCGAAGG-A
>CG5020_sequenced_DNA         CAACATGATTGTCGGAAATCGAAGTTATCGACAATCAGTCATCGAAGGAA
>CG5020_sequenced_RNA_1       ---------------------------------TCAGTCATCGA-GG-A
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ******************* ********************** ** *

>CG5020_predicted&assembled   CGATCGC-AGGCAGCAGTAGAGGCGAGTGGAAGTCAGCGTTGCAGTCAGT
>CG5020_sequenced_DNA         CGATCGCAAGGCAGCAGTGGAGTCGAGTGGAAGTCAGCGTTGCAGTCAGT
>CG5020_sequenced_RNA_1       CGATCGCAAGGCAGCAGTGGAGTCGAGTGGAAGTCAGCGTTGCAGTCAGT
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ******* ********** *** ***************************

>CG5020_predicted&assembled   CGAGTTCTCAGCAGCAGTCGTTCGGTCCACAAACTAAGAAATACTTTATA
>CG5020_sequenced_DNA         CGAGTTCTCAGCAGCAGTCGTTCGGT-CACAAACTAAG-AATACTTCATA
>CG5020_sequenced_RNA_1       CGAGTTCTCAGCAGCAGTCGTTCGGT-CACAAACTAAG-AATACTTCATA
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************* *********** ******* ***
```

```
>CG5020_predicted&assembled   TAATTACCGCATTTAGAATTAAACTAATAATTAAATTAATAATAAACAAT
>CG5020_sequenced_DNA         TAATTACCGCATTTAGAATTAAACTAATAATTAAATTAATAATAAACAAT
>CG5020_sequenced_RNA_1       TAATTACCGCATTTAGAATTAAACTAATAATTAAATTAATAATAAACAAT
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   AATAATAAACAATCTTACATGGGGGCTCGTCCAGTCCTAAATCGGTTATA
>CG5020_sequenced_DNA         AATAATAAACAATCTTACATGGGGGCTCGTCCAGTCCTAAATCGGTTATA
>CG5020_sequenced_RNA_1       AATAATAAACAATCTTACATGGGGGCTCGTCCAGTCCTAAATCGGTTATA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   TGAAGGTGCAGTTGTTTAAAGAAAAAAGACATTGTTGTGTGCGTGGGTAT
>CG5020_sequenced_DNA         TGAAGGTGCAGTTGTTTAAAGAAAAAAGACATTGTTGTGTGCGTGGGTAT
>CG5020_sequenced_RNA_1       TGAAGGTGCAGTTGTTTAAAGAAAAAAGACATTGTTGTGTGCGTGGGTAT
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   AGTCTTTAAAACGTTGTAAAGTTGTGGCTATATCTATTGCATTTAAAGTT
>CG5020_sequenced_DNA         AGTCTTTAAAA-GTTGTAAAGTTGTGGCTATATCTATTGCATTTAAAGTT
>CG5020_sequenced_RNA_1       AGTCTTTAAAA-GTTGTAAAGTTGTGGCTATATCTATTGCATTTAAAGTT
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              ********** **********************************

>CG5020_predicted&assembled   GGAAAAATCAGTTGTACAGATTTTGTTTGAACACAAGTCGGTAAAAGTCG
>CG5020_sequenced_DNA         GGAAAAATCAGTTGTACAGATTTTGTTTGAACACAAGTCGGTAAAAGTCG
>CG5020_sequenced_RNA_1       GGAAAAATCAGTTGTACAGATTTTGTTTGAACACAAGTCGGTAAAAGTCG
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   GGAAAGCTGCTAGAGAGAACTGATAAAGTTGAAATTGTCGTGTGCGTGGA
>CG5020_sequenced_DNA         GGAAAGCTGCTAGAGAGAACTGATAAAGTTGAAATTGTCGTGTGCGTGGA
>CG5020_sequenced_RNA_1       GGAAAGCTGCTAGAGAGAACTGATAAAGTTGAAATTGTCGTGTGCGTGGA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   TTTAGTCTTTAAAGTTGTAAAGTTATGGCTACGTCTACTGCATTGAAAGT
>CG5020_sequenced_DNA         TTTAGTCTTTAAAGTTGTAAAGTTATGGCTACGTCTACTGCATTGAAAGT
>CG5020_sequenced_RNA_1       TTTAGTCTTTAAAGTTGTAAAGTTATGGCTACGTCTACTGCATTGAAAGT
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   TGAAAAAATCGATTGAACTCATACAGACTCAAGTCGTTTTGCTGTTGTGG
>CG5020_sequenced_DNA         TGAAAAAATCGATTGAACTCATACAGACTCAAGTCGTTTTGCTGTTGTGG
```

```
>CG5020_sequenced_RNA_1       TGAAAAAATCGATTGAACTCATACAGACTCAAGTCGTTTTGCTGTTGTGG
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   AATTTAAAACAATTAAATTGCAAAGGTGGTGAAATTCGTTTCTAACGAAA
>CG5020_sequenced_DNA         AATTTAAAACAATTAAATTGCAAAGGTGGTGAAATTCGTTTCTAACGAAA
>CG5020_sequenced_RNA_1       AATTTAAAACAATTAAATTGCAAAGGTGGTGAAATTCGTTTCTAACGAAA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   ATCAAAATTTGTCTTTTAACCGGTGGCGCCGTCTGCAAAATCGACTACCG
>CG5020_sequenced_DNA         ATCAAAATTTGTCTTTTAACCGGTGGCGCCGTCTGCAAAATCGACTACCG
>CG5020_sequenced_RNA_1       ATCAAAATTTGTCTTTTAACCGGTGGCGCCGTCTGCAAAATCGACTACCG
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   TCGCGCCGTTAGAACATTGTCGTTGTTTGCTGGTGTTAGTGCCTTGTCGC
>CG5020_sequenced_DNA         TCGCGCCGTTAGAACATTGACGTTGTTTGCTGGT-TTAGTGCCTTGTCGC
>CG5020_sequenced_RNA_1       TCGCGCCGTTAGAACATTGTCGTTGTTTGCTGGTGTTAGTGCCTTGTCGC
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              ****************** ************** ***************

>CG5020_predicted&assembled   GGAATGTTCACACGTACACCACCTACAAATAAAAAACTTAACACCGACCA
>CG5020_sequenced_DNA         GGAATGTTCACACGTACACCACCTACAAATAAAAAAC-TAACACCGACC-
>CG5020_sequenced_RNA_1       GGAATGTTCACACGTACACCACCTACAAATAAAAAACTTAACACCGACCA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              ********************************** ***********

>CG5020_predicted&assembled   AATACAAGCAATTCTAGAGAACGAAAGCGAGGACGAAAGCAGAAAAGAAA
>CG5020_sequenced_DNA         ATTACAAGCAATTCTAGAGAACGAAAGCGAGGACGAAAGCATAAAAGAAT
>CG5020_sequenced_RNA_1       AATACAAGCAATTCTAGAGAACGAAAGCGAGGACGAAAGCAGAAAAGAAA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              * ************************************ *******

>CG5020_predicted&assembled   AAATGAACGAAGAAGATCAAAAGTTGGCGCCTGTAGGAGAAGCAGAGGCA
>CG5020_sequenced_DNA         AAATGAACGAATATGATCACAAGTTGGCGCCTGTATGATAAGCCGAGGTA
>CG5020_sequenced_RNA_1       AAATGAACGAAGAAGATCAAAAGTTGGCGCCTGTAGGAGAAGCAGAGGCA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              *********** * ***** *************** ** **** **** *

>CG5020_predicted&assembled   AAGAAGCAGAATAAAGACGCTAGTGCTAAAGTCGAAGAGAAATTTGAACA
>CG5020_sequenced_DNA         AAGAAGCAGAATAAAGACGCTAGT-TTACAAGCGAAGAG-AATTTGAAC-
>CG5020_sequenced_RNA_1       AAGAAGCAGAATAAAGACGCTAGTGCTAAAGTCGAAGAGAAATTTGAACA
>CG5020_sequenced_RNA_2       --------------------------------------------------
```

```
                          ************************  ** *  ******* *********

>CG5020_predicted&assembled   AATGATGAATACTCTAACCCAGAGCATGTTGGCAAAATCTAAACAAGAGG
>CG5020_sequenced_DNA         AATGATGAATACTCTAACCC-GAGCTTGTTGGC-AAATCT-AACCAGAGG
>CG5020_sequenced_RNA_1       AATGATGAATACTCTAACCCAGAGCATGTTGGCAAAATCTAAACAAGAGG
>CG5020_sequenced_RNA_2       -------------------------------------------------
                          ******************* **** ******* ****** *** *****

>CG5020_predicted&assembled   GGCAAGTAATTATCGCTGCAGAAAAATTTGAAAAAGTTGTAAGTGACTGT
>CG5020_sequenced_DNA         GTCATGC-ATTATCGCTGCAGAAAAA-TTGAAAAAGTCGCAAGCGACTGG
>CG5020_sequenced_RNA_1       GGCAAGTAATTATCGCTGCAGAAAAATTTGAAAAAGTTGTAAGTGACTGT
>CG5020_sequenced_RNA_2       -------------------------------------------------
                          * ** *  ****************** ********** * *** *****

>CG5020_predicted&assembled   GATGGCAAATCAATTCCTATTAAAAAATGGTTTGAAATTTTTGAGAAAAA
>CG5020_sequenced_DNA         GA-AGCAATTCA--TCCTCTTAAATACTGCCTTGAAACTTTCGTGAATAA
>CG5020_sequenced_RNA_1       GATGGCAAATCAATTCCTATTAAAAAATGGTTTGAAATTTTTGAGAAAAA
>CG5020_sequenced_RNA_2       -------------------------------------------------
                          **  **** ***  **** ***** * **  ****** *** * *** **

>CG5020_predicted&assembled   T-GCCGAGGCATATGAACTTTCGGAGAAACAAAAATATGTTCAAGCCAGA
>CG5020_sequenced_DNA         TGGCCTAGGCCTATGGATACTCTGGC------------------------
>CG5020_sequenced_RNA_1       T-GCCGAGGCATATGAACTTTCGGAGAAACAAAAATATGTTCAAGCCAGA
>CG5020_sequenced_RNA_2       -------------------------------------------------
                          * *** **** **** *   ** *  ************************

>CG5020_predicted&assembled   AGTAAGATGATTGGATCAGCAGAACTTTTCTTAGAATCTGAATGTGTCAG
>CG5020_sequenced_DNA         -------------------------------------------------
>CG5020_sequenced_RNA_1       AGTAAGATGATTGGATCAGCAGAACTTTTCTTAGAATCTGAATGTGTCAG
>CG5020_sequenced_RNA_2       -------------------------------------------------
                          *************************************************

>CG5020_predicted&assembled   TGGATACACTGAACTCAAAGAGTTACTAATTGAAGAATTTTCAGGCAGCT
>CG5020_sequenced_DNA         -------------------------------------------------
>CG5020_sequenced_RNA_1       TGGATACACTGAACTCAAAGAGTTACTAATTGAAGAATTTTCAGGCAGCT
>CG5020_sequenced_RNA_2       -------------------------------------------------
                          *************************************************

>CG5020_predicted&assembled   ATAATAGCGCCGTTATTCACAAAAAGTTGCAAGACAGGAAGAAGAAGAGG
>CG5020_sequenced_DNA         -------------------------------------------------
>CG5020_sequenced_RNA_1       ATAATAGCGCCGTTATTCACAAAAAGTTGCAAGACAGGAAGAAGAAGAGG
>CG5020_sequenced_RNA_2       -------------------------------------------------
                          *************************************************
```

```
>CG5020_predicted&assembled   GAGGAAACTCTACACGACTATTTGTTACAAATGAAGAAAATAGCAGCCTT
>CG5020_sequenced_DNA         ------------------------------------------------
>CG5020_sequenced_RNA_1       GAGGAAACTCTACACGACTATTTGTTACAAATGAAGAAAATAGCAGCCTT
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************************************

>CG5020_predicted&assembled   AGGTGAAGTTGAAACAGTTGCTTTGATAACTCATATCGTAAACGGCCTCG
>CG5020_sequenced_DNA         ------------------------------------------------
>CG5020_sequenced_RNA_1       AGGTGAAGTTGAAACAGTTGCTTTGATAACTCATATCGTAAACGGCCTCG
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************************************

>CG5020_predicted&assembled   ACATTAAAAAGGAGTATAAGGGTGCTATGCTCCGTTGTAAAACTCTTAAG
>CG5020_sequenced_DNA         ------------------------------------------------
>CG5020_sequenced_RNA_1       ACATTAAAAAGGAGTATAAGGGTGCTATGCTCCGTTGTAAAACTCTTAAG
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************************************

>CG5020_predicted&assembled   GAATTAAAGCAAGAATTCGAAATCTACGAGAGTCTGAATATTGTTGACAA
>CG5020_sequenced_DNA         ------------------------------------------------
>CG5020_sequenced_RNA_1       GAATTAAAGCAAGAATTCGAAATCTACGAGAGTCTGAATATTGTTGACAA
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************************************

>CG5020_predicted&assembled   GCCGAATATTCAACCAAAACCAAAGCAAATTACACAAGGTGTAAAAGCAG
>CG5020_sequenced_DNA         ------------------------------------------------
>CG5020_sequenced_RNA_1       GCCGAATATTCAACCAAAACCAAAGCAAATTACACAAGGTGTAAAAGCAG
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************************************

>CG5020_predicted&assembled   ATCACTGCTTCAACTGTGGTTCGAGGGAACACAAACGAAAGGATTGTACA
>CG5020_sequenced_DNA         ------------------------------------------------
>CG5020_sequenced_RNA_1       ATCACTGCTTCAACTGTGGTTCGAGGGAACACAAACGAAAGGATTGTACA
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************************************

>CG5020_predicted&assembled   CTTCCTACCAAATGTTTCAGCTGTAATCAAGAGGGCCATATCTCAAGCAA
>CG5020_sequenced_DNA         ------------------------------------------------
>CG5020_sequenced_RNA_1       CTTCCTACCAAATGTTTCAGCTGTAATCAAGAGGGCCATATCTCAAGCAA
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              ************************************************

>CG5020_predicted&assembled   GTGTCCGGAAAAAGTAAACAGCATGCGCATTCACGTTGATAGTGCACGAA
>CG5020_sequenced_DNA         ------------------------------------------------
```

```
>CG5020_sequenced_RNA_1       GTGTCCGGAAAAAGTAAACAGCATGCGCATTCACGTTGATAGTGCACGAA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   CAAAGCCAGTAATCATAAATGGGATTATCATCAACTGTCTGGTGGACACA
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       CAAAGCCAGTAATCATAAATGGGATTATCATCAACTGTCTGGTGGACACA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   GGATCAGATGTGACCATAATTAAAGAAGCTATATTCAAGAAGATGAAAGA
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       GGATCAGATGTGACCATAATTAAAGAAGCTATATTCAAGAAGATGAAAGA
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   TGTTGATTTAAACCGCACTGCAACAGTATTGCGAGGTTTGGGAAATGCCT
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       TGTTGATTTAAACCGCACTGCAACAGTATTGCGAGGTTTGGGAAATGCCT
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   CAACACAGCCGATTGGATGCTTCAGAGCATTAATCAAGACCGACCAGGTG
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       CAACACAGCCGATTGGATGCTTCAGAGCATTAATCAAGACCGACCAGGTG
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   GAAGCAAGCCACAACGTTTTTAGTCGTCCACGATTCTAAATTCAGTTGCG
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       GAAGCAAGCCACAACGT-TTTAGTCGTCCACGATTCTAAATTCAGTTGCG
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              ***************** ********************************

>CG5020_predicted&assembled   ATGGAATAGTGGGACACGATTTTTATCAGCAAGTTTCGTCTTATCTGTAG
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       ATGGAATAGTGGGACACGA-TTTTATCAGCAAGTTTCGTCTTATCTGTAG
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              ****************** ******************************

>CG5020_predicted&assembled   TGCAGAAGGCTATACTTTTCTTGACCTGGAAGCAGATAAAAAACAAGCGG
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       TGCAGAAGGCTATACTTTTCTTGACCTGGAAGCAGATAAAAAACAAGCGG
>CG5020_sequenced_RNA_2       --------------------------------------------------
```

```
                                                  **************************************************

>CG5020_predicted&assembled    CTCAGTCACAGGTTTACAAGGAAAAAATCCATGATCTGGAGTCAAAAATC
>CG5020_sequenced_DNA          --------------------------------------------------
>CG5020_sequenced_RNA_1        CTCAGTCACAGGTTTACAAGGAAAAAATCCATGATCTGGAGTCAAAAATC
>CG5020_sequenced_RNA_2        --------------------------------------------------
                                                  **************************************************

>CG5020_predicted&assembled    ACAAAACTGGTGTCCGCCACGCCAAGCCTACAAAGTATACTACCGCCCGA
>CG5020_sequenced_DNA          --------------------------------------------------
>CG5020_sequenced_RNA_1        ACAAAACTGGTGTCCGCCACGCCAAGCCTACAAAGTATACTACCGCCCGA
>CG5020_sequenced_RNA_2        --------------------------------------------------
                                                  **************************************************

>CG5020_predicted&assembled    TCTACCTTCAGACGATGGTGCTTTGCAGGAGGAAATTGCCCAGCTGCAGG
>CG5020_sequenced_DNA          --------------------------------------------------
>CG5020_sequenced_RNA_1        TCTACCTTCAGACGATGGTGCTTTGCAGGAGGAAATTGCCCAGCTGCAGG
>CG5020_sequenced_RNA_2        --------------------------------------------------
                                                  **************************************************

>CG5020_predicted&assembled    AAAAGATGACCATTCAGCAGAAGGAGGTTGAATCTCGGATTGCGGAACAG
>CG5020_sequenced_DNA          --------------------------------------------------
>CG5020_sequenced_RNA_1        AAAAGATGACCATTCAGCAGAAGGAGGTTGAATCTCGGATTGCGGAACAG
>CG5020_sequenced_RNA_2        --------------------------------------------------
                                                  **************************************************

>CG5020_predicted&assembled    CTGGAGGAGGAGCAGCGATTGAGGGAAAATGTGAAGTACCTTAATGAGCA
>CG5020_sequenced_DNA          --------------------------------------------------
>CG5020_sequenced_RNA_1        CTGGAGGAGGAGCAGCGATTGAGGGAAAATGTGAAGTACCTTAATGAGCA
>CG5020_sequenced_RNA_2        --------------------------------------------------
                                                  **************************************************

>CG5020_predicted&assembled    AATCGCCACTCTACAGTCCGAGTTGGTGTCCAAAGATGAGGCCCTGGAGA
>CG5020_sequenced_DNA          --------------------------------------------------
>CG5020_sequenced_RNA_1        AATCGCCACTCTACAGTCCGAGTTGGTGTCCAAAGATGAGGCCCTGGAGA
>CG5020_sequenced_RNA_2        --------------------------------------------------
                                                  **************************************************

>CG5020_predicted&assembled    AATTCTCCCTCTCGGAATGTGGCATCGAGAATCTCCGAAGGGAACTCGCA
>CG5020_sequenced_DNA          --------------------------------------------------
>CG5020_sequenced_RNA_1        AATTCTCCCTCTCGGAATGTGGCATCGAGAATCTCCGAAGGGAACTCGCA
>CG5020_sequenced_RNA_2        --------------------------------------------------
                                                  **************************************************
```

```
>CG5020_predicted&assembled   CTTCTCAAGGAGGAGAACGAAAAGCAAGCTCAGGAGGCTCAGGCTGAGTT
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       CTTCTCAAGGAGGAGAACGAAAAGCAAGCTCAGGAGGCTCAGGCTGAGTT
>CG5020_sequenced_RNA_2       --------------------------------------------------
                              **************************************************

>CG5020_predicted&assembled   CACCCGAAAACTAGCCGAAAAATCCGTAGAGGTGTTAAGATTAAGCTCTG
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       CACCCGAAAACTAGCCGAAAAATCCGTAGAGGTGTTAAGATTAAGCTCTG
>CG5020_sequenced_RNA_2       ------------------------GTAGAGGTGTT-AGATT-AGCTCTG
                              ********************************** ***** *******

>CG5020_predicted&assembled   AATTGCAAAACTTGAAGGCAACATCCGATTCCCTGGAAAGCGAAAGGGTT
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       AATTGCAAAACTTGAAGGCAACATCCGATTCCCTGGAAAGCGAAAGGGTT
>CG5020_sequenced_RNA_2       -ATTGCAAAACTTGAAGGCAACATCCGATTCCCTGGAAAGCGAAAGGGTT
                               **************************************************

>CG5020_predicted&assembled   AACAA-TCCGACGAATGTGAAATTCTTCAAACCGAGTACCGAATGCGG-A
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       AACAAATCCGACGAATGTGAAATTCTTCAAACCGAAGTCCGAATGCGGGA
>CG5020_sequenced_RNA_2       AACAAATCCGACGAATGTGAAATTCTTCAAACCGAAGTCCGAATGCGGGA
                              ***** ****************************   ********** *

>CG5020_predicted&assembled   TGAGCAAATTAGAGAGCTAAACCAACAACTCGATGAGGTTACCACACAAC
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       TGAGCAAATTAGAGAGCTAAACCAACAACTCGATGAGGTTACCACACAAC
>CG5020_sequenced_RNA_2       TGAGCAAATTAGAGAGCTAAACCAACAACTCGATGAGGTTACCACACAAC
                              **************************************************

>CG5020_predicted&assembled   TAAATGTACAAAAAGCGGATAGTTCTGCTCTGGATGATATGCTTCGATTG
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       TAAATGTACAAAAAGCGGATAGTTCTGCTCTGGATGATATGCTTCGATTG
>CG5020_sequenced_RNA_2       TAAATGTACAAAAAGCGGATAGTTCTGCTCTGGATGATATGCTTCGATTG
                              **************************************************

>CG5020_predicted&assembled   CAAAAGGAGGGTACTGAAGAAAAATCTACTCTTTTAGAGAAAACCGAAAA
>CG5020_sequenced_DNA         --------------------------------------------------
>CG5020_sequenced_RNA_1       CAAAAGGAGGGTACTGAAGAAAAATCTACTCTTTTAGAGAAAACCGAAAA
>CG5020_sequenced_RNA_2       CAAAAGGAGGGTACTGAAGAAAAATCTACTCTTTTAGAGAAAACCGAAAA
                              **************************************************

>CG5020_predicted&assembled   GGAGCTAGTTCAAATCAAAGAACAAGCTGCGAAAACTCTACAGGATAAGG
>CG5020_sequenced_DNA         --------------------------------------------------
```

```
>CG5020_sequenced_RNA_1       GGAGCTAGTTCAAATCAAAGAACAAGCTGCGAAAACTCTACAGGATAAGG
>CG5020_sequenced_RNA_2       GGAGCTAGTTCAAATCAAAGAACAAGCTGCGAAAACTCTACAGGATAAGG
                              **************************************************

>CG5020_predicted&assembled   AACAACTTGAAAAACAGATATCAGATTTAAAGCAATTGGCGGAACAAGAA
>CG5020_sequenced_DNA         ---------------------------------------CG--ACCAAGG
>CG5020_sequenced_RNA_1       AACAACTTGAAAAACAGATATCAGATTTAAAGCAATTGGCGGAACAAGAA
>CG5020_sequenced_RNA_2       AACAACTTGAAAAACAGATATCAGATTTAAAGCAATTGGCGGAACAAGAA
                              ****************************************  ** *

>CG5020_predicted&assembled   AAACTAGTCAGGGAAAAGACTGAAAATGCAATCAATCAAATACAACTAGA
>CG5020_sequenced_DNA         AAACTAGTCAGG-AAAAG-CTG-AAATGC-ATCAATCAATTCCAACTAG-
>CG5020_sequenced_RNA_1       AAACTAGTCAGGGAAAAGACTGAAAATGCAATCAATCAAAT-CAACTAGA
>CG5020_sequenced_RNA_2       AAACTAGTCAGGGAAAAGACTGAAAATGCAATCAATCAAATACAACTAGA
                              ************ ***** *** ****** ********* * *******

>CG5020_predicted&assembled   AAAAGAATCCATAGAACAGCAATTGGCTTTAAAACAAAACGAACTTGAGG
>CG5020_sequenced_DNA         AAAAGA--TCATAG-ACAGCA---TGCTTT-AAACAAAACG-ACTTGAGG
>CG5020_sequenced_RNA_1       AAAAGGATCCA-----------------------------------------
>CG5020_sequenced_RNA_2       AAAAGAATCCATAGAACAGCAATTGGCTTTAAAACAAAACGAACTTGAGG
                              *****   ****** ******   ****** ********** ********

>CG5020_predicted&assembled   ACTTCCAAAAGAAACAGTCAGAATCGGAAGTTCATCTTCAGGAAATCAAA
>CG5020_sequenced_DNA         ACTGC--AAAAGAACAGTCAG-ATCGGAAGTTCATCTTCAGGAAATCAAA
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       ACTTCCAAAAGAAACAGTCAGAATCGGAAGTTCATCTTCAGGAAATCAAA
                              *** *  ************** ****************************

>CG5020_predicted&assembled   GCTCAGAACACGCAGAAGGATTTAGAATTAGTTGAATCTGGTGAGTCCCT
>CG5020_sequenced_DNA         GCTCAG-ACACGCAGAAGGATTTAGAATTAGTTG-ATCTGGTGAGTCCCT
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       GCTCAGAACACGCAGAAGGATTTAGAATTAGTTGAATCTGGTGAGTCCCT
                              ****** ************************** ***************

>CG5020_predicted&assembled   TAAAAAACTGCAACAGCAATTAGAGGAGAAAACTTTGGGTCATGAAAAAC
>CG5020_sequenced_DNA         TAAAAAACTGCAACAGCAATTAGAGGAGAAAACTTTGGGTCATGAAAAAC
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       TAAAAAACTGCAACAGCAATTAGAGGAGAAAACTTTGGGTCATGAAAAAC
                              **************************************************

>CG5020_predicted&assembled   TGCAAGCTGCTTTGGAAGAACTAAAGAAAGAAAAGGAAACGATTATAAAG
>CG5020_sequenced_DNA         TGCAAGCTGCTTTGGAAGAACTAAAGAAAGAAAAGGAAACGATTATAAAG
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       TGCAAGCTGCTTTGGAAGAACTAAAGAAAGAAAAGGAAACGATTATAAAG
```

```
                                **************************************************

>CG5020_predicted&assembled    GAAAAGGAGCAGGAGCTCCAGCAGCTCCAAAGCAAGTCAGCTGAATCGGA
>CG5020_sequenced_DNA          GAAAAGGAGCAGGAGCTCCAGCAGCTCCAAAGCAAGTCAGCTGAATCGGA
>CG5020_sequenced_RNA_1        --------------------------------------------------
>CG5020_sequenced_RNA_2        GAAAAGGAGCAGGAGCTCCAGCAGCTCCAAAGCAAGTCAGCTGAATCGGA
                                **************************************************

>CG5020_predicted&assembled    AAGTGCTTTAAAGGTCGTACAAGTACAACTAGAGCAACTCCAGCAACAGG
>CG5020_sequenced_DNA          AAGTGCTTTAAAGGTCGTACAAGTACAACTAGAGCAACTCCAGCAACAGG
>CG5020_sequenced_RNA_1        --------------------------------------------------
>CG5020_sequenced_RNA_2        AAGTGCTTTAAAGGTCGTACAAGTACAACTAGAGCAACTCCAGCAACAGG
                                **************************************************

>CG5020_predicted&assembled    CTGCCGCATCTGGAGAAGAGGGATCCAAAACTGTGGCCAAATTGCACGAT
>CG5020_sequenced_DNA          CTGCCGCATCTGGAGAAGAGGGATCCAAAACTGTGGCCAAATTGCACGAT
>CG5020_sequenced_RNA_1        --------------------------------------------------
>CG5020_sequenced_RNA_2        CTGCCGCATCTGGAGAAGAGGGATCCAAAACTGTGGCCAAATTGCACGAT
                                **************************************************

>CG5020_predicted&assembled    GAGATTAGTCAGCTTAAGTCCCAAGCGGAAGAAACTCAGTCCGAGTTAAA
>CG5020_sequenced_DNA          GAGATTAGTCAGCTTAAGTCCCAAGCGGAAGAAACTCAGTCCGAGTTAAA
>CG5020_sequenced_RNA_1        --------------------------------------------------
>CG5020_sequenced_RNA_2        GAGATTAGTCAGCTTAAGTCCCAAGCGGAAGAAACTCAGTCCGAGTTAAA
                                **************************************************

>CG5020_predicted&assembled    ATCCACCCAATCGAACTTGGAAGCCAAAAGCAAACAATTGGAAGCAGCAA
>CG5020_sequenced_DNA          ATCCACCCAATCGAACTTGGAAGCCAAAAGCAAACAATTGGAAGCAGCAA
>CG5020_sequenced_RNA_1        --------------------------------------------------
>CG5020_sequenced_RNA_2        ATCCACCCAATCGAACTTGGAAGCCAAAAGCAAACAATTGGAAGCAGCAA
                                **************************************************

>CG5020_predicted&assembled    ATGGCAGCCTAGAAGAGGAAGCCAAGAAGTCAGGCCAACTGCAGGAACAG
>CG5020_sequenced_DNA          ATGGCAGCCTAGAAGAGGAAGCCAAGAAGTCAGGCCAACTGCAGGAACAG
>CG5020_sequenced_RNA_1        --------------------------------------------------
>CG5020_sequenced_RNA_2        ATGGCAGCCTAGAAGAGGAAGCCAAGAAGTCAGGCCAACTGCAGGAACAG
                                **************************************************

>CG5020_predicted&assembled    ATTACCAAACTTAAATCTGAAGTGGAGGAGACGCAGGCAGCTCTCAGTTC
>CG5020_sequenced_DNA          ATTACCAAACTTAAATCTGAAGTGGAGGAGACGCAGGCAGCTCTCAGTTC
>CG5020_sequenced_RNA_1        --------------------------------------------------
>CG5020_sequenced_RNA_2        ATTACCAAACTTAAATCTGAAGTGGAGGAGACGCAGGCAGCTCTCAGTTC
                                **************************************************
```

```
>CG5020_predicted&assembled   ATATCATACGGATGTGGAAGTACATAGCCGAAAATGCAGACCTATTGTCA
>CG5020_sequenced_DNA         ATATCATACGGATGTGGAAGTACATAGCCGAAAATGCAGACCTATTGTCA
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       ATATCATACGGATGTGGAAGTACATAGCCGAAAATGCAGACCTATTGTCA
                              *************************************************

>CG5020_predicted&assembled   TCCGGGTCGGATGATGATGATCAGGAGGGCCGAATGTAACATGGAGTAAG
>CG5020_sequenced_DNA         TCCGGGTCGGATGATGATGATCAGGAGGGCCGAATGTAACATGGAGTAAG
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       TCCGGGTCGGATGATGATGATCAGGAGGGCCGAATGTAACATGGAGTAAG
                              *************************************************

>CG5020_predicted&assembled   GCTGAAGGCTGGCAACAACCCGGTTGGCAGCGCTGTTGAGCAGCAACATG
>CG5020_sequenced_DNA         GCTGAAGGCTGGCAACAACCCGGTTGGCAGCGCTGTTGAGCAGCAACATG
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       GCTGAAGGCTGGCAACAACCCGGTTGGCAGCGCTG-TGAGCAGCAACATG
                              ********************************** **************

>CG5020_predicted&assembled   ATTGTCGGAAATCGAAGTTATCGACAATCAGTCATCGAAGGAACGATCGC
>CG5020_sequenced_DNA         ATTGTCGGAAATCGAAGTTATCGACAATCAGTCATCGAAGGAACGATCGC
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       AT-----------------------------------------------
                              *************************************************

>CG5020_predicted&assembled   AAGGCAGCAGTGGAGTCGAGTGGAAGTCAGCGTTGCAGTCAGTCGAGTTC
>CG5020_sequenced_DNA         AAGGCAGCAGTGGAGTCGAGTGGAAGTCAGCGTTGCAGTCAGTCGAGTTC
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       -------------------------------------------------
                              *************************************************

>CG5020_predicted&assembled   TCAGCAGCAGTCGTTCGGTCACAAACTAAGAATACTTTATATAATTACCG
>CG5020_sequenced_DNA         TCAGCAGCAGTCGTTCGGTCACAAACTAAGAATACTTCATATAATTACCG
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       -------------------------------------------------
                              *********************************** ************

>CG5020_predicted&assembled   CATTTAGAATTAAACTAATAATTAAATTAATAATAAACAATAATAATAAA
>CG5020_sequenced_DNA         CATTTAGAATTAAACTAATAATTAAATTAATAATAAACAATAATAATAAA
>CG5020_sequenced_RNA_1       -------------------------------------------------
>CG5020_sequenced_RNA_2       -------------------------------------------------
                              *************************************************

>CG5020_predicted&assembled   CAATCTTACA----------------------------------------
>CG5020_sequenced_DNA         CAATCTTACATACACAAACACGCCGGCTAATAGATGCACGCCCACGATCC
```

```
>CG5020_sequenced_RNA_1       ------------------------------------------------
>CG5020_sequenced_RNA_2       ------------------------------------------------
                              **********

>CG5020_predicted&assembled   ------------------------------------------------
>CG5020_sequenced_DNA         TCAGGGGCATTTATGATGATGATGATGATCGGATCGCAAACGAGTTTTGC
>CG5020_sequenced_RNA_1       ------------------------------------------------
>CG5020_sequenced_RNA_2       ------------------------------------------------


>CG5020_predicted&assembled   ------------------------------------------------
>CG5020_sequenced_DNA         CATAAAAATAAAGATGATTTTGGCCAGATCACTGTCAGCAACTAAACGCA
>CG5020_sequenced_RNA_1       ------------------------------------------------
>CG5020_sequenced_RNA_2       ------------------------------------------------


>CG5020_predicted&assembled   ---------------
>CG5020_sequenced_DNA         ATAGCGGTAGTCCCC
>CG5020_sequenced_RNA_1       ---------------
>CG5020_sequenced_RNA_2       ---------------
```