

Additional File 4 – Nucleotide alignments for isogroups 08478 and 03233

A Isogroup08478 – putative *piwi-like* gene

```
isotig33900   ATAATCCAGGCAAATCAACGATCATACACTTACAACCAGGACACGGCTGATTGGTCGAAA
isotig33901   ATAATCCAGGCAAATCAACGATCATACACTTACAACCAGGACACGGCTGATTGGTCGAAA
*****

isotig33900   GAGATGAGAGGACTACCCCTGCTGAACCCAGCCCATCTAACCAACTGGGTACTCATCTAC
isotig33901   GAGATGAGAGGACTACCCCTGCTGAACCCAGCCCATCTAACCAACTGGGTACTCATCTAC
*****

isotig33900   ACGAGTCGAGATGCTGGCAATGCAAGTGACCTGAGTTCAACACTAGTGAAAGTCTGTCAG
isotig33901   ACGAGTCGAGATGCTGGCAATGCAAGTGACCTGAGTTCAACACTAGTGAAAGTCTGTCAG
*****

isotig33900   CCAATGGGAATTCGTGTCGACCAACCACGAACATGCCAGTTGCCTGATGATAGTACAAC
isotig33901   CCAATGGGAATTCGTGTCGACCAACCACGAACATGCCAGTTGCCTGATGATAGTACAAC
*****

isotig33900   ACATTCGTACGAGCACTGAAGGAATCTATTGACGACACGGTTTGCATGGTGGTGATAATA
isotig33901   ACATTCGTACGAGCACTGAAGGAATCTATTGACGACACGGTTTGCATGGTGGTGATAATA
*****

isotig33900   CTACCCACAAACAGGAAGGATCGTTATGATGCCATCAAACCATCTGCTGTGTTGAGCAG
isotig33901   CTACCCACAAACAGGAAGGATCGTTATGATGCCATCAAACCATCTGCTGTGTTGAGCAG
*****

isotig33900   CCAGTCCAAGCCAAGTTGTATTGGCTCGTACTTTATCCAAGAAGCAGATGTTGATGTCCT
isotig33901   CCAGTCCAAGCCAAGTTGTATTGGCTCGTACTTTATCCAAGAAGCAGATGTTGATGTCCT
*****

isotig33900   GTGGCAACCAAATTGCCATTCAATTGAACGTCAAATGGGCGGCGAAGTATGGGCGCTG
isotig33901   GTGGCAACCAAATTGCCATTCAATTGAACGTCAAATGGGCGGCGAAGTATGGGCGCTG
*****

isotig33900   GATATACCTATCAAGAATGTCATGATTATTGGAATTGATACTTATCACGATTCTGCCAAG
isotig33901   GATATACCTATCAAGAATGTCATGATTATTGGAATTGATACTTATCACGATTCTGCCAAG
*****

isotig33900   AAAGGTCGATCTGTTGGAGCATTCAATGCCAGCATGAACTCAACCTTTACACGTTATTAT
isotig33901   AAAGGTCGATCTGTTGGAGCATTCAATGCCAGCATGAACTCAACCTTTACACGTTATTAT
*****

isotig33900   TCACGATGTGCATTCCAATCAAGTCACGAGGAACTACAGAATGCTCTGAAAATATGTATG
isotig33901   TCACGATGTGCATTCCAATCAAGTCACGAGGAACTACAGAATGCTCTGAAAATATGTATG
*****

isotig33900   CAGAGTGCTTTGAAGAATTACCATGCCGTTAATGGAGTTCTGCCTGAGAAAAGTCATAGTC
isotig33901   CAGAGTGCTTTGAAGAATTACCATGCCGTTAATGGAGTTCTGCCTGAGAAAAGTCATAGTC
*****

isotig33900   TTCCGTGATGGAGTTGGGGATGGACAGCTGGCAACTGTTTACGAGCATGAAGTTCGTCAG
isotig33901   TTCCGTGATGGAGTTGGGGATGGACAGCTGGCAACTGTTTACGAGCATGAAGTTCGTCAG
*****

isotig33900   CTGACTGACTGCTTTCAGGCACCTGGACCAGATTACAGTCCAAAGTTTGCAGTAATCGTG
isotig33901   CTGACTGACTGCTTTCAGGCACCTGGACCAGATTACAGTCCAAAGTTTGCAGTAATCGTG
*****

isotig33900   GTCAAGAAGCGCATCAACTCCCGTTTCTTCATGACAGAGAGGGGCCAAACTACAAATCCA
isotig33901   GTCAAGAAGCGCATCAACTCCCGTTTCTTCATGACAGAGAGGGGCCAAACTACAAATCCA
*****

isotig33900   CGTGCAGGAACCATCATTGACACAGAGGCAACGCGGCCAGAATGGTACGATTTCCTCTTG
isotig33901   CGTGCAGGAACCATC-----
*****

isotig33900   GTGAGTCAGTCGGTTCGTCAGGGAACGGTTGCTCCAACCTACTACAAATGTCATCTACGAT
isotig33901   -----

isotig33900   ACAACTGGTTTCAAGCCGGATCACATGCAACGTTTGGCTTACAAACTGTGCCATTTGTAC
isotig33901   -----
```

isotig33900 TACAAC TGGCAGGGAACGATT CGTGTGCCGGCGCCTTGCCAGTATGCTCACAAAATGGCA
isotig33901 -----

isotig33900 TTCCTGGTTGGTCAGTCTCTTCATAGAGATCCATCTCTGCGCCTAGCCGACAAGCTGTTC
isotig33901 -----TCTGCGCCTAGCCGACAAGCTGTTC

isotig33900 TTCCTGTAGTATCTCTGACTCGGCTGCTGAACTGTTACACTGTCGATGCTGACGTGCTGA
isotig33901 TTCCTGTAGTATCTCTGACTCGGCTGCTGAACTGTTACACTGTCGATGCTGACGTGCTGA

isotig33900 ATATCAAGATTTTATTTTCATCTCAGTTAGCAGTTGATTCGCGTTGTTGTGATAATTGAT
isotig33901 ATATCAAGATTTTATTTTCATCTCAGTTAGCAGTTGATTCGCGTTGTTGTGATAATTGAT

isotig33900 CTA CTTTGATGATTAATACTTTTTTTCAATGTGTTCAATCGTAATACAGTGGACTGTAGT
isotig33901 CTA CTTTGATGATTAATACTTTTTTTCAATGTGTTCAATCGTAATACAGTGGACTGTAGT

isotig33900 TTTTAAATGTACCATTGGATTTGTATGGTTTACGGAGAAGATGCTCAGCTTTCTTGCC
isotig33901 TTTTAAATGTACCATTGGATTTGTATGGTTTACGGAGAAGATGCTCAGCTTTCTTGCC

isotig33900 GTTTGAGGAAGTGTTCGGAATGATGGTGTGCGCCATTACAAGAGGCTGACATTCACATC
isotig33901 GTTTGAGGAAGTGTTCGGAATGATGGTGTGCGCCATTACAAGAGGCTGACATTCACATC

isotig33900 TATGTATTTT TAGCTAAGTCGTGCCGAGTTGTTGGACTGTTGCAATACACGTTGTCTT
isotig33901 TATGTATTTT TAGCTAAGTCGTGCCGAGTTGTTGGACTGTTGCAATACACGTTGTCTT

isotig33900 GAATGCACCTATGTGTATTTTCTCAAGTATTGAAGCCATCGGAAAGGAACTTTGATCGAG
isotig33901 GAATGCACCTATGTGTATTTTCTCAAGTATTGAAGCCATCGGAAAGGAACTTTGATCGAG

isotig33900 TTCACGTTGATGTCTGAATATTTTCTACTAGCTGTTCCGACGTTAGTGTAGTTTGT
isotig33901 TTCACGTTGATGTCTGAATATTTTCTACTAGCTGTTCCGACGTTAGTGTAGTTTGT

isotig33900 GATACGTTAACATCCATTTCTGTCAGTATTGTTGATAACGCATCTCTAGTGCACACATT
isotig33901 GATACGTTAACATCCATTTCTGTCAGTATTGTTGATAACGCATCTCTAGTGCACACATT

isotig33900 GCTGCAGTAGCACGTGCGTCTCGCCTGCATTTTTTTGACAAATTTGTTGCGATAGATT
isotig33901 GCTGCAGTAGCACGTGCGTCTCGCCTGCATTTTTTTGACAAATTTGTTGCGATAGATT

isotig33900 CTGTTTGAGATTTGTGAAGTAGCGGATGTTATGCATAAAATAATCTGCATCCTGATGCAT
isotig33901 CTGTTTGAGATTTGTGAAGTAGCGGATGTTATGCATAAAATAATCTGCATCCTGATGCAT

isotig33900 AATTGTTCTTGGTGGTGTGATTTAAGAGATGCCCAACTGATTTAAGAGATGCCCAACTG
isotig33901 AATTGTTCTTGGTGGTGTGATTTAAGAGATGCCCAACTGATTTAAGAGATGCCCAACTG

isotig33900 TTGAAACTATTTTCATCCTGTTTATACCGGGCATTAAACTGCATTTACGTCGCATACGA
isotig33901 TTGAAACTATTTTCATCCTGTTTATACCGGGCATTAAACTGCATTTACGTCGCATACGA

isotig33900 TGAATTGATGATTTGTTTGGAGTTGAAAAA
isotig33901 TGAATTGATGATTTGTTTGGAGTTGAAAAA

B Isogroup03233 – putative *frizzled* gene

```
isotig19223 -----
isotig19224 CCCGTAGTCGACGCGTCTGCAACATGTCCAGCGTTTTCCAAGTACGACAACCTGTTGCAG
isotig19225 -----
isotig19226 -----
isotig19227 CCCGTAGTCGACGCGTCTGCAACATGTCCAGCGTTTTCCAAGTACGACAACCTGTTGCAG
isotig19228 -----

isotig19223 -----
isotig19224 ATGTCAACGTTGATGATTTGACGTTTTTCGTACGACAGCAACACTTGTGCGAGTCGTTTCAT
isotig19225 -----
isotig19226 -----
isotig19227 ATGTCAACGTTGATGATTTGACGTTTTTCGTACGACAGCAACACTTGTGCGAGTCGTTTCAT
isotig19228 -----

isotig19223 -----
isotig19224 GTCGACATGGCTGCATGTTGCCGTATTGATCGTAAACTGCTATCTTAACCGACCGAGCGA
isotig19225 -----
isotig19226 -----
isotig19227 GTCGACATGGCTGCATGTTGCCGTATTGATCGTAAACTGCTATCTTAACCGACCGAGCGA
isotig19228 -----

isotig19223 -----
isotig19224 TATTCGGCGCGAAGTGCCTTTGCTCGATTTTATGTGACGAGCTGCGTTTGACCGATTTT
isotig19225 -----
isotig19226 -----
isotig19227 TATTCGGCGCGAAGTGCCTTTGCTCGATTTTATGTGACGAGCTGCGTTTGACCGATTTT
isotig19228 -----

isotig19223 -----CCGAAGTGTTCAGTTTTATCGCTGTCGCTACAAACG
isotig19224 TGTGCGGGTTGTGTTCCGAAA---AACTGTTTCAAGTTTTATCGCTGTCGCTACAAACG
isotig19225 -----
isotig19226 -----CCGAAGTGTTCAGTTTTATCGCTGTCGCTACAAACG
isotig19227 TGTGCGGGTTGTGTTCCGAAA---AACTGTTTCAAGTTTTATCGCTGTCGCTACAAACG
isotig19228 -----

isotig19223 TAGGCTACGCACGCACGGTCTGGCGGGCGTTTCGCTTGGAGCGTCGCTCGCCTTCGAAACA
isotig19224 TAGGCTACGCACGCACGGTCTGGCGGGCGTTTCGCTTGGAGCGTCGCTCGCCTTCGAAACA
isotig19225 -----
isotig19226 TAGGCTACGCACGCACGGTCTGGCGGGCGTTTCGCTTGGAGCGTCGCTCGCCTTCGAAACA
isotig19227 TAGGCTACGCACGCACGGTCTGGCGGGCGTTTCGCTTGGAGCGTCGCTCGCCTTCGAAACA
isotig19228 -----

isotig19223 TCGAGATAACACATACGTACACAAAGATAGCAGCGAAATAATGCGGTCGTCGTCGTCGT
isotig19224 TCGAGATAACACATACGTACACAAAGATAGCAGCGAAATAATGCGGTCGTCGTCGTCGTCGT
isotig19225 -----
isotig19226 TCGAGATAACACATACGTACACAAAGATAGCAGCGAAATAATGCGGTCGTCGTCGTCGTCGT
isotig19227 TCGAGATAACACATACGTACACAAAGATAGCAGCGAAATAATGCGGTCGTCGTCGTCGTCGT
isotig19228 -----

isotig19223 -GATCAGCTGGAATCCGTGTACTTGCCTGCTTCTTAATCCACTGCCATCCTGCAGTAT
isotig19224 -GATCAGCTGGAATCCGTGTACTTGCCTGCTTCTTAATCCACTGCCATCCTGCAGTAT
isotig19225 -----
isotig19226 -----
isotig19227 -----
isotig19228 -----

isotig19223 CAGCGGACTTGAGTCAATATGGAATCGAGTTATGCAGCACCATCGGACGTTTGGCAGCC
isotig19224 CAGCGGACTTGAGTCAATATGGAATCGAGTTATGCAGCACCATCGGACGTTTGGCAGCC
isotig19225 -----
isotig19226 -----
isotig19227 -----
isotig19228 -----
```

isotig19223 AGCGCTCTTTTGCTTCAGCTGTTACTGACCGGCAGTCACCAGATGCCGTCCGCTGTCGGT
isotig19224 AGCGCTCTTTTGCTTCAGCTGTTACTGACCGGCAGTCACCAGATGCCGTCCGCTGTCGGT
isotig19225 -----
isotig19226 -----
isotig19227 -----
isotig19228 -----

isotig19223 TTCGGTTCGCAGCAACAGCAACAGCAGCAACAACAGCCGCAGCAACAGCAGCAGCATTCCG
isotig19224 TTCGGTTCGCAGCAACAGCAACAGCAGCAACAACAGCCGCAGCAACAGCAGCAGCATTCCG
isotig19225 -----
isotig19226 -----
isotig19227 -----
isotig19228 -----

isotig19223 ATGATGTCGCACCATCACCTGC-----
isotig19224 ATGATGTCGCACCATCACCTGC-----
isotig19225 -----CGTTCCCTCGTGTGTATTTCTTCGGCATGGCTGGATC
isotig19226 -----
isotig19227 -----
isotig19228 -----CGTTCCCTCGTGTGTATTTCTTCGGCATGGCTGGATC

isotig19223 -----
isotig19224 -----
isotig19225 GCTCTGGTGGGTCGCTCTGACGGTCGCATTTACTTGTTCGGCCGGACGGAAGTGGAGTCCG
isotig19226 -----
isotig19227 -----
isotig19228 GCTCTGGTGGGTCGCTCTGACGGTCGCATTTACTTGTTCGGCCGGACGGAAGTGGAGTCCG

isotig19223 -----
isotig19224 -----
isotig19225 AGAAGCGATCGAAGCGCGGTCGGCCTACTTTACGTCGCGCGGCTGGACACTTCCGGCCGT
isotig19226 -----
isotig19227 -----
isotig19228 AGAAGCGATCGAAGCGCGGTCGGCCTACTTTACGTCGCGCGGCTGGACACTTCCGGCCGT

isotig19223 -----
isotig19224 -----
isotig19225 GAAAACGATCGCCATATTGGCGCTTCATCGTGTTCGACGCGGACGAATTGACGGGACTCTG
isotig19226 -----
isotig19227 -----
isotig19228 GAAAACGATCGCCATATTGGCGCTTCATCGTGTTCGACGCGGACGAATTGACGGGACTCTG

isotig19223 -----
isotig19224 -----
isotig19225 TTACGTCGGCAACGTCGACCGAGACGCGCTGGTCGCCTTCGTTGTCGTTCCGCTGTTGGC
isotig19226 -----
isotig19227 -----
isotig19228 TTACGTCGGCAACGTCGACCGAGACGCGCTGGTCGCCTTCGTTGTCGTTCCGCTGTTGGC

isotig19223 -----
isotig19224 -----
isotig19225 CTATCTCATCGTCGGCATAGCGTTTATCGCTGCAGGTTTCACTGCTATGTTCCACATCCG
isotig19226 -----
isotig19227 -----
isotig19228 CTATCTCATCGTCGGCATAGCGTTTATCGCTGCAGGTTTCACTGCTATGTTCCACATCCG

isotig19223 -----
isotig19224 -----
isotig19225 CAAAGACCTTCTGCGTGGCGGCAGCGGCGAGAACGACGACGAAACATAATCGCTCCGGG
isotig19226 -----
isotig19227 -----
isotig19228 CAAAGACCTTCTGCGTGGCGGCAGCGGCGAGAACGACGACGAAACATAATCGCTCCGGG

isotig19223 -----
isotig19224 -----
isotig19225 AGGAACGAAACTATTCGGAAGTTGAAAACTAATGGCAAAAATCGGCGTGTCTCTGTG
isotig19226 -----
isotig19227 -----
isotig19228 AGGAACGAAACTATTCGGAAGTTGAAAACTAATGGCAAAAATCGGCGTGTCTCTGTG

isotig19223 -----
isotig19224 -----
isotig19225 CTCTATAACCGTACCGGCGACGTGTGTCGTTGCGTGTTCCTGTACCGTCGCATGAACATG
isotig19226 -----
isotig19227 -----
isotig19228 CTCTATAACCGTACCGGCGACGTGTGTCGTTGCGTGTTCCTGTACCGTCGCATGAACATG

isotig19223 -----
isotig19224 -----
isotig19225 GACATCTGGCGCAGCGAGCTGCCGCTGGCTTGGACTGTGGCGCCCTCGTAAGCATCTG
isotig19226 -----
isotig19227 -----
isotig19228 GACATCTGGCGCAGCGAGCTGCCGCTGGCTTGGACTGTGGCGCCCTCGTAAGCATCTG

isotig19223 -----
isotig19224 -----
isotig19225 AGCTTAGTAGGGTCTGAAGACGAACCGGACTGCGAGCCTCTGAGGCAATCGATACCGCCG
isotig19226 -----
isotig19227 -----
isotig19228 AGCTTAGTAGGGTCTGAAGACGAACCGGACTGCGAGCCTCTGAGGCAATCGATACCGCCG

isotig19223 -----
isotig19224 -----
isotig19225 ATAGAGGTCGAGATGCTGAGGTTGTTTCATGTTATTGGCGGGCGGAATGACCAGCTTCGTG
isotig19226 -----
isotig19227 -----
isotig19228 ATAGAGGTCGAGATGCTGAGGTTGTTTCATGTTATTGGCGGGCGGAATGACCAGCTTCGTG

isotig19223 -----
isotig19224 -----
isotig19225 TGGATCTGCTCGAGGAAAACATTTTGGAGCTGGTACCGATTTTGGACGCGAAGTGGAGGA
isotig19226 -----
isotig19227 -----
isotig19228 TGGATCTGCTCGAGGAAAACATTTTGGAGCTGGTACCGATTTTGGACGCGAAGTGGAGGA

isotig19223 -----
isotig19224 -----
isotig19225 CGAATTGGGTCCCGACATCTGTGCCAGGCGCCAGTGCTCACCAAAACAGACCGGTGCGA
isotig19226 -----
isotig19227 -----
isotig19228 CGAATTGGGTCCCGACATCTGTGCCAGGCGCCAGTGCTCACCAAAACAGACCGGTGCGA

isotig19223 -----
isotig19224 -----
isotig19225 TCAGATGCGGCAGCGCTACCGTCGGCGGTGGGGGATTGCGTTTAGCTACGCCAGCCACCG
isotig19226 -----
isotig19227 -----
isotig19228 TCAGATGCGGCAGCGCTACCGTCGGCGGTGGGGGATTGCGTTTAGCTACGCCAGCCACCG

isotig19223 -----
isotig19224 -----
isotig19225 GATCACGTGATATTCACACCGGGCTGC-TTCAGCAGCCGTCGTCGTCGTCGT---CTG
isotig19226 -----CTG
isotig19227 -----CTG
isotig19228 GATCACGTGATATTCACACCGGGCTGC-----

isotig19223 -----
isotig19224 -----
isotig19225 CAGCTGGAAAATGTGAACGGATCAGGATACCTCTATGTAAGGATATGAAATACAATATGA
isotig19226 CAGCTGGAAAATGTGAACGGATCAGGATACCTCTATGTAAGGATATGAAATACAATATGA
isotig19227 CAGCTGGAAAATGTGAACGGATCAGGATACCTCTATGTAAGGATATGAAATACAATATGA
isotig19228 -----

isotig19223 -----
isotig19224 -----
isotig19225 CGCGGATGCCGAACCTGGTCGGTCAAACCTAAATCAAAGGACGCAGCGCTTCAGGTGCACG
isotig19226 CGCGGATGCCGAACCTGGTCGGTCAAACCTAAATCAAAGGACGCAGCGCTTCAGGTGCACG
isotig19227 CGCGGATGCCGAACCTGGTCGGTCAAACCTAAATCAAAGGACGCAGCGCTTCAGGTGCACG
isotig19228 -----

isotig19223 -----
isotig19224 -----
isotig19225 AGTCCATACCGCTGATACAGTTCGGTTGTTCAAACCTGTTGAAGTTTTTCCTCTGTCTT
isotig19226 AGTCCATACCGCTGATACAGTTCGGTTGTTCAAACCTGTTGAAGTTTTTCCTCTGTCTT
isotig19227 AGTCCATACCGCTGATACAGTTCGGTTGTTCAAACCTGTTGAAGTTTTTCCTCTGTCTT
isotig19228 -----

isotig19223 -----CTACCGGAAGTGCACACGTCTTATCCTG
isotig19224 -----CTACCGGAAGTGCACACGTCTTATCCTG
isotig19225 TGTACGCACCAATGTGTACCGAACCAAGTCG-----
isotig19226 TGTACGCACCAATGTGTACCGAACCAAGTCG-----
isotig19227 TGTACGCACCAATGTGTACCGAACCAAGTCG-----
isotig19228 -----CTACCGGAAGTGCACACGTCTTA-CCTG

isotig19223 CATTCGAACAGTATCAGACTCGAATTTGATTGGTGTATTTATGTATTTTGTCCTATAA
isotig19224 CATTCGAACAGTATCAGACTCGAATTTGATTGGTGTATTTATGTATTTTGTCCTATAA
isotig19225 -----
isotig19226 -----
isotig19227 -----
isotig19228 CATTCGAACAGTATCAGACTCGAATTTGATTGGTGTATTTATGTATTTTGTCCTATAA

isotig19223 TTGTACGTAGCACGGGATTGATCATGGAGTTGTTACGG
isotig19224 TTGTACGTAGCACGGGATTGATCATGGAGTTGTTACGG
isotig19225 -----
isotig19226 -----
isotig19227 -----
isotig19228 TTGTACGTAGCACGGGATTGATCATGGAGTTGTTACGG