

# Diastata ussurica

Taxonomic lineage: Brachycera > Muscomorpha > Ephydroidea > Diastatidae > Diastata

Assembly: GCA\_035078745.1\_ASM3507874v1\_genomic

Dia\_uss1 | JAWNWP010044231.1:859-1194 (-) | 336 nt | lncRNA: noe consensus e-value: 5.0e-04

5' motif: GCGGT 3' motif: ATCGC Internal max. Poly-T: 4nt Trailing-T: 7nt PSE: 0.94

GGCGGTACGACGCTCTCAGTTGAAATGGTTGGTCGCACACCTTTAATTGCGGCCAAAAAATAAGTTGACAAAGAAAAAACAAAGAAATATTTCCAAATTACAAAAAGAAAAAAGAAAAAGAAAAATCAATTAACAAAAACAATTTCAAAAATAAATTCGAGATCACATCATTCTTCCCTTTGTTAAAGGTTTGAATGCCGATTAGAAAGCAATTATTTCTCCCAATTCGGGATTAGATTGCGATTAGAAAAATCTCTCGATAGTTTTATCGAGGAGTTTAAATTGATAAGATCTGCGAGATCGTAACGTGATCGCTCTCTTTTTTT

```
Dia_uss1      -----ttccctttgttctacttcgctgcattagcggctgaTATAATTCGCAACGCTTCTcgaattttgttaagtacaaatagcgcatttaacacatatgtt
Arthropod_7SK -----ctcattgtttttccttcggtgattacttgatagcGAATAATTCGCAAAACTTTTtcggttatgtgcacagcataaatacaacatgcatttcttagtatttt
U6           ----aataaaatatttttctcdaatatacatgataaATAAATTCCAACTACTTCTcaacgtatgtacaaatfaaatatacacttcgaatgaatcacttc-
U6           ----actgaaattattttacttcaaatgttcactcgataaaATAATTCCAACCACTTTTcgttgttgggacaagttaaataccatatttgatgaacttcatg-
U6atac       ----attgccttcacttcatttcatatcatttaatttcaaTCAATTCGCAACCACCTTCTcgtgcttgtgtagttttaagtacttggttaaaaaatttcaattt-
tRNAsec      ----cdaaaccttgtttttcttcacgccttccttttTGAATTCATACATTTTTTcccttatocctttatttcaaatatgtacaaatcaatcdaatcctt-
U1           ----aaacgttttccattttttcttcgttgagtaacattttTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U1           ----cagttgaagaattatcfcaccttcgttttgcctcgtGAATAATTCCTATTAGGTTTAtatatgcattagataaagcagtggttagtttgcggagttaaagc-
U1           ----gttcctctctattatctacttacttctcttcacaaATAATTCCAACTGGTTTAtgaatgottttgtatatataaagcgaatgaagcgaatgaagc-
U1           ----gtgaccaggtagccttaatatattgggaaaaaaaatgtTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U1           ----tgaccaggtagccttaatatattgggaaaaaaaatgtTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U1           ----aaacaacaggtgacgggcaaaaattttgtcaaaaaatgtTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U2           ----ttcaaacctcatttacttctctacatatttccattttTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U2           ----gtcaacctcatttgcctccgcatatatttgaattttTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U2           ----tgagatataattacatatttgccttgcgttgataaagtTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U2           ----tattttacatttctattacatataatttccctttTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U2           ----gtacatacaaatagcatgagaaagtgaataataaTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U3           ----gtttgcttcgctgtttcattttgtgaaattttatgtTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U4           ----tttatgtctcttttcttcttctccttcttctcdaatTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U4           ----aatatatttttcttcttgcgttaatactgcctgtTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U4           ----atgcaatttgcagacacttacttcacacttttgaTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U4           ----aatatatttttcttcttgcgttaatactgcctgtTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U4           ----tttatgtctcttttgcgtgttgcgtgttgcgtgtTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U4atac       ttaacaatttgaatttttcttcgcagcacattttattttTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U5           -atottttttcttatttttttcaatttcaaatdaattcaaaTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U5           -cgttgaattttctacacttctgttgcgtgttatgtgcTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U5           -gttatatttgaagtttcaaatgaatgatttaccatcgaaTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U11          ----cacaaaaacatttcttcttttadcttgaacacgaaatTAATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
U12          ----acaaaaagcttttgcctgatttgcgtgtatttgcTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
OrCD1       ----ttttttcatatagcttgcgtgttgcgtgttgcTATAATTCCTCAATTCCTTAgctagattttgtataaagcgcggttgagagcgaagaaaagc-
```

Upstream promoter alignment showing conserved regions. Sequences include Pol III promoters (noeCR34335, Arthropod\_7SK, 7SL, RNase\_MRP, RNaseP\_nuc, U6, U6atac, tRNAsec, snoRNAme18SA1806) and Pol II promoters (U1, U2, U3, U4, U4atac, U5, U7, U8, U11, U12, OrCD1). Conserved positions are highlighted in the alignment.