

1021 tccaggtccatggagagcagccctaggaaggaagtaaaaagaagccagtc aaagttgag
292 ·S·R·S·M·E·S·S·P·R·K·G·S·K·K·K·P·V·K·V·E·

1081 gctccggaatacatcccataagtgatgaccctaaggcctccgcaaagaaaagatgaag
312 ·A·P·E·Y·I·P·I·S·D·D·P·K·A·S·A·K·K·K·M·K·

1141 tccaaaaagaaggtagagcagccagtc atcgaggagccagctctgaaaaggaagaaaaag
332 ·S·K·K·K·V·E·Q·P·V·I·E·E·P·A·L·K·R·K·K·K·

1201 aagaagaggaaagagagtggggtagcaggagacccttggaaggaggaaacagacacggac
352 ·K·K·R·K·E·S·G·V·A·G·D·P·W·K·E·E·T·D·T·D·

Ex3 | Ex4
Region DUF5594

1261 ttagaggtgggtggttggaaaaaaaggcaacatggatgaggcgcatagaccagggtgagg
372 L·E·V·V·L·E·K·K·G·N·M·D·E·A·H·I·D·Q·V·R

Ex4 | Ex5

1321 cgaaaggccttgcaagaagagatcgatcgcgagtcaggcaaaacggaagcttctgaaacc
392 R·K·A·L·Q·E·E·I·D·R·E·S·G·K·T·E·A·S·E·T·

Ex5 | Ex6

1381 aggaagtggacgggaaccagtttggccagtgaggatactgctggttttgagaacgaggac
412 ·R·K·W·T·G·T·Q·F·G·Q·W·D·T·A·G·F·E·N·E·D·

Region SMAP

1441 caaaaactgaaatctcagacttatgggtggcttcaaaaacctgtccccttcggttcagc
432 Q·K·L·K·F·L·R·L·M·G·G·F·K·N·L·S·P·S·F·S·

1501 cgccccgcagcagcagattgcaaggcccaacatggcctcggcaagaaggcggctgacagc
452 R·P·A·S·T·I·A·R·P·N·M·A·L·G·K·K·A·A·D·S·

1561 ctgcagcagaatctgcagcgggactacgaccgggcatgagctggaagtacagccgggga
472 L·Q·Q·N·L·Q·R·D·Y·D·R·A·M·S·W·K·Y·S·R·G·

1621 gccggcctcggcttctccaccgcccccaacaagatcttttacattgacaggaacgcttcc
492 A·G·L·G·F·S·T·A·P·N·K·I·F·Y·I·D·R·N·A·S·

1681 aagtcagtc aagctggaagat taaactctagagttttgtcccccaaaactgccacaatt
512 ·K·S·V·K·L·E·D·*

1741 gctttgattattccatttatgctggagattacaaatTTTTTTTTgtgaaaaaatcagatct
1801 tgggtgaggacctcgagcagtaagatataaaataactcccataagcttagcgttccagtaat
1861 ggaacactaggcataaatgggtttattcagttgtgcaaatgaaagccatctgacagttggc
1921 tcacattgaacacctgtggagattaaggacgaggacaactatattgatgggcttggatga
1981 actggggcagggcagctcatatttcggggagccaggagaacgagtgagtgctaaaacctcc
2041 tgttttctgtgttaaaccattccgtccctgtttgagacatcagtatgtacagttactttt
2101 gttgagtgtttagcaggtactagggacatactagtgtttcccttaatgtatttaactttc
2161 ataattatgaaatgggtgctattattagccccatcttatagatgaggcaactgaggttca
2221 gggataaagt aataaaaattgctggggtcaccagccacta agtgtcagggtgttgtactt
2281 ttgtaccogaagccctaagttcactattcgccactctgaatgtccccttttagggaatttc
2341 caccagaatcctcgttggggattgaaatgtcttttagatggaggaaaaagttttatgacaa
2401 gtctgcatctctgataaaaaagtgaggatgaatgaggaaacggagaatcgaaaactccttttc
2461 ctctccttttctttccctgtcatagagcagctgtaggcagagaggtgtctgagttgta
2521 ccaaacacatgtgactgagctgctgctgctgggaaatgtgcctctaaaccagtcctgcct
2581 tttccactgggtgtcccttcggccctaagtgtgaggtaggatTTTTgtgctttgactgga
2641 ttctgaacaaggcatttttaggaagagtggttagtccgcaaataatcaatggcagatcc
2701 cagggacgtgtgtgtgtgttttgagataggctctcactctgtcaccaggtggagtgta
2761 gtggcatatacggccttgttgcaaccactacctatccagcttaagtgatcctcccacctc
2821 agcctcccagtagctgggactgtgggtgtgcaccacaacaaccagctagtgtttaaaat
2881 tttttgtagagatggggctcactatattgtccaggctagtcccaaactcctggactcaa
2941 ctcatgctcctgcctcagcccccaagagctgggatcacaggtgtgagccactgtgccacg

polyA signal sequence
major polyA site