

Elaine M. Majerus, Xinglong Zheng, Elodee A. Tuley, and J. Evan Sadler  
Cleavage of the ADAMTS13 Propeptide is not Required for Protease Activity

Supplemental Data

Candidate *Mus musculus* ADAMTS13 cDNA sequence

The mouse sequence is from NCBI Entrez GenomeScan model Mm2\_39245\_30\_123\_3, which overlaps with LocusLink LOC279028 and GenBank XP\_205053. The 5' end of the transcript that includes exons 1-4 appears to be correctly spliced. Some later exons appear to have been predicted less successfully.

>lcl|Mm2\_39245\_30\_123\_3

```
ATGAGCCAGCTTTGCCTGTGGTTGACGTGCCAGCCTTGTATGCTGTCAGTGTGTCAGAGGA
ATCCTCACTGGTGCCATCTTCATTCTGGGCTGCTGGGGGCTCTCTGACTTCCAGAAGAGT
CTTCTTCAAGATCTGGAGCCCAAGGATGTGTCTTCTTACTTTGGCCACCATGCTGCTCCA
TTCACAGGCCATCCTCCCTCTCACCTCCAGAGACTGAGACGGAGAAGGACTTTGGAGGAC
ATTCTGCACCTGGAACCTCCTGGTAGCTGTGGGCCCGATGTTTCCCGGGCTCATCAGGAG
GACACAGAACGCTACGTGCTCACTAATCTCAATATCGGGTCAGAACTGTTGAGAAACCCA
TCCCTGGGAGTCCAGTTCAGGTGCACCTGGTGAAGCTAATCACCCCTCTCTGACTCAGAG
AGTACTCCGAATATCACGGCCAACATCACCTCATCCTTGATGAGCGTCTGCGAGTGGAGC
CAGACGATCAACCCCCACGATGACAGGGATCCAAGTCACGCTGACCTGATTCTCTATATC
ACCAGCAACGTGGCTGGTGCCACTGTCCTTGTGATTCATTTTCTCTTATCAAGGTTTGAC
CTGGAGTTGCCTGATGGCAACCAGCAGGTTCCGGGGTGTACCCAGCTGGGAGGTGCCTGC
TCCCTTTCCTGGAGTTGCCTTATCACTGAGGATACTGGCTTTGACCTGGGGGTACCATC
GCCATGAGATTGGGCACAGCTTCGGGCTGGACCATGATGGTGCTCCAGGTAGTGGCAGC
ACCTGCAAGGCCAGTGGCCACGTGATGGCGGCTGATGGCGCAACACCTACTGGAGGGACC
CTGGAGTGGTCTGCCTGCAGCCAAAGGCAGTTGCAGCACCTACTCAGCACAGGGCAAATG
CACTGCTTCCAGGACCCACCTGGGCTGCAGTCAGGACTTACACGGCACCAGCTGATGGCA
CAGCCTGGCCTCTACTACAGTGCAGATGATCAGTGCCGTGTGGCTTTCGGTTCTGGGGCT
GTCGCCTGCACCTTCTCCAGGGAGGGTCTGGATGTATGCCAGGCCCTGTCTGCCACACA
GACCCCTTGGACCAAAGCAGCTGCAGCCGCCTCCTTGTTCCTCTCCTGGATGGGACAGAA
TGTGGTGTGGAGAAGGTTGGAGTAGTGAGAGGGATGAAGCTATGGTGCTCCAAGGCTCGC
TGTCGCTCCCTAGCTGAGCTGGCTCCTGTGGCTGCAGTACATGGACACTGGTCTAGCTGG
GGCCCCATAGTCCCTGCTCCCGATCCTGTGGAGGAGGTGTGATTACCAGGAGGCGGTGG
TGCAACAACCCAGGCCTGCATTTGGGGGACGTGCATGTGTGGGTGAAGACCTCCAGGCT
AAGATGTGCAACACGCAGGCTTGTGAGAAGACTCAGCTGGAGTTCATGTCCGAGCAGTGT
GCCAGACAGACAGACAACCCTGCAACTTTCCCAAGGCACTGCCTCCTTCTACCACTGG
GATGCTGCTGTGCAGTATAGTCAAGGAGATACCCTGTGCAGACACATGTGCTGGGCTGTT
GGAGAAAGCTTCATTGTCAGCCGTGGGGACAGGTTCCCTAGATGGGACCCGTTGTGTGCCA
AGTGGTCCCTCAGGATGATGGGACCCTAAGCCTCTGTTTGTGGGCAGCTGCAGGACCTTT
GGCTGTGATGGCAGGATGGACTCCCAGAAGGTTTGGGATGCGTGCCAGGTGTGTGGAGGA
GACAACAGCACCTGCAGCTCACGGAATGGTTCTTTACAGCTGGGAGAGCCAGAGAATAT
GTCACGTTCCCTGATTGTTACTCCCAACATGACCAACGCACACATTGTCAACCGCAGGCCT
CTCTTACACACTTGGCGGTGAGGATCCAGGGCCACTACATTGTGGCAGGGAAGACTAGC
ATCTCACCCAACACCACCTACCCTTCCCTTCTGGAGGACTACCGTGTGGAATACAGAGTG
ACTCTCACTGAGGACCAGCTGCCCCACTTAGAGGAGATTCACATCCGGGGACCCGTCCGG
```

GATGACATTGAGATTCAGGTGTACAGACGATATGGAGGAGAATATGGGGATCTTACACAC  
CCAGACATCACCTTTTCTACTTTCAACTGAAGCAGCAGGCAGCCTGGGTATGGACCGCT  
AAGCGTGGACCCTGCTCAGTGAGCTGTGGGGCAGGGCTGCGCTGGGTGACCTACAGCTGC  
CAGGATCAAGCTCAAGACAAGTGGGTAAAGAACGCCAGTGCCAAGGGAGCCCACAGCCA  
CCTGCATGGCAAGAGCCTTGTGTCTCTGCCCCCTGCTCCCCATATTGGGTAGCTGGGGAC  
TTCAGCCCATGTAGCGTGTCTTGTGGCGGGGGCCTTCGGGAGCGGTCACTGCGCTGTGTA  
GAGACCCAAGATGGCTTCTTAAAGACACTGCCACCTGCCCGGTGCAGAGCAGTAGCCCAG  
CAGCCAGCAGCAGAAGTGGAAAACCTGCAACTCCCAGCCCTGTCCCACCAGGCGAGTATAT  
CAAGGTCTGGACTCAAGGAATGTGACATGTGTGTCCAGGGCGGGTGACCCGGAGAAGCCA  
GAAACTGCAGGCCCTGCCGCACCGACGAGATGTCAGCTATGCTGGAGCCCTGCTCCAGG  
AGCCTGTGTTCTCCAGGCTTGGGTGAGGTGGACAACACCATGTCTCTGGGAGAGGAGGCT  
CCATCCCCGGTGGGCAGTGACAAGCCAGGGGCTCAGGCTGAGCATGTGTGGACCCCTCTG  
GTGGGGCTGTGCTCCATCTCTTGTGGGAGAGGTCTGAAGGAACGTATTTTCTGTGCATG  
GATTCTGTCCTCAAATGCCTGTCCAGGAAGAGCTATGCGGCTTGGCTAGTAAGCCCCCA  
AGCCGGTGGGAGGTCTGCAGGGCTCGCCCCTGTCCTGCTCGGTGGGAGACTCAAGTCTTG  
GCACCGTGCCCGGTGACCTGTGGTGGGGGGCGAGTGCCACTGTCTGTTGTTGTGTGCAG  
CTAGACCGTGGCCACCCGATATCTGTACCTCACTCCAAGTGCTCGCCAGTGCCTAAGCCA  
GGCTCCTTCGAGGACTGCAGCCCTGAGCCTTGTCTGCTAGATGGAAAGTCCTGTCCCTT  
GGCCCATGCTCAGCCAGCTGTGGCCTTGGCACTGCCACACAAATGGTGGCCTGCATGCAG  
CTTGACCAGGGCCATGACAATGAGGTGAATGAAACTTTCTGTAAGGCTCTAGTGCGGCCT  
CAGGCCAGTGTCCCCTGCCTCATTTGCTGACTGCGCATTCGGGTGGCACATCAGCGCCTGG  
ACAGAGTGCTCTGTCTCCTGTGGGGACGGTATTTCAGCGCCGGCATGACACCTGCCTTGA  
CCCAGGCCCAAGTGCCTGTGCCAGCCAATTTCTGCCAGCACCTGCCCAAGCCAATGACC  
TGGTCCCAGCCCCGAGCCCGACCCCCACCCTTTTCTCAGCCTCCCAGTCTCTGGGGCTC  
CAGGAAAACCTGGAAGAGCATGGTGCCTGTGGAAGGCAGTACCTGGAGCCCACAGGAACC  
ATTCACATGCGAGATCAAGGACGGCTAGACTGTGTGGTGGCCATTGGGAGACCTCTAGGT  
GAAGTGGTGACCTACAAATCCTTGAGAGCTCCCTCAAGTGTAGTGCAGGGGAGCAACTG  
CTGCTCTGGGGTAGGTTACATGGAGGAAGACATGCAGGAAGATGCCTGGCATGACTTTC  
AGCACAAAACCAACACAGTGGTGGTGAAGCAGCATCGTGTGCTGCCAGGGGGTGGAGTG  
CTGCTACGGTACTGGAGTCAACCTGCCCGGGGACCTTCTACAAAGAATGTGACAGGCAG  
CTCTTTGGACCTCGGGGTGAAATTTGTGAGTCCCTCACTGAGCCAGACGGGAGGAAGGCA  
GGGACCTGCAGGGTCTTCATCAGTGTGGCTCCACAGGCCCGCATTGCCATCCGCGCCCTG  
GCCAGTGATATGGGCACTGCTTCTGAGGGGACCAATGCCAACTATGTCTCGATCAGGGAC  
ATCCACAGTCTGAGGACCACAACATTTTGGGGGCAGCAGGTGCTCTACTGGGAGTCCGAA  
GGCAGCGAGGCTGAACTGGAGTTTAGCCCTGGCTTCCCTGGAGGCACATGCCAGTCTCCAG  
GGCGAGTATTGGACCATCTCACCTAGGACATCGGAACAGGACGACAGCCTGGCTCTGTCC  
TAG

Candidate *Rattus norvegicus* ADAMTS13 cDNA sequence

The rat sequence is from NCBI Entrez GenomeScan model NCBI Entrez GenomeScan model Rn3\_990\_1\_5\_5. The 5' end of the transcript that includes exons 1-4 appears to be correctly spliced. Some portions of later exons deviate significantly from the structure of human ADAMTS13.

>lcl|Rn3\_990\_1\_5\_5  
ATGAGCCAGCTTTGCTTATGGTTGCCGTGCCAGCCTTTTCTTGCTGCCAGTATCAGAGGA

ATCCTCACGGGCGCCATCTTCATTCTGGGCTGCTGGGGGCTCTCTGATTTCCAGAAGAGT  
TTTCTTCAAGATCTGGAGCCCAAGGACTTATCTTCCTCCTTCAGCCGCCATACTGCTCCA  
CTCACAGGCCATCCTTCCCTCTCACCTCCAGAGACTGAGACAGAGGAGGACTTCGGAGGAC  
ATTCTGCACCTAGAACTCCTGGTAGCTGTGGGCCCGATGTTTACCAGGCTCACCAGGAG  
GACACAGAGCGCTACGTGCTCACCAACCTCAATATCGGGTCAGAACTGTTGAGAAACCCG  
TCCCTAGGGGCTCAGCTCCAAGTGCACCTGGTGAAGCTGATCATTCTCTCGGACTCAGAG  
AATACTCCGAATATCACGACCAACATCACCTCATCCTTGATGAGTGTCTGCGAGTGGAGC  
CAGACGATCAATCCCCATGATGACAGGGACCCAAGTCATGCTGACCTGATCATCTATATA  
ACCAGGTTTGACCTGGAGCTGCCGGATGGTAACCAGCAGGTTTCGGGGAGTCACCCAACAG  
GGTGGTGCCTGCTCTCCTTCCCTGGAGTTGCCTTATCACTGAGGACACTGGCTTCGATCTG  
GGGGTCACCATTGCCACGAGATTGGACACAGCTTTGGGCTGGACCATGATGGTGTCTCCA  
GGCAGTGGCAGCACCTGCGAGGCCCGTGGCCATGTGATGGCATCTGACGGCGCAGTGCCT  
GCTGGAGGGGCCCTGGAGTGGTCCACCTGCAGCCAAAGGCAGCTGCAGCACCTGCTCAGC  
ACAGGGCAAATGCACTGCCTCTGGGACCCGCCGGGCTGCAGTCAGGACTCACATGGCAC  
CAGCTGGTGGCACAGCCTGGCCTCTTCTACAGTGCAGATGACCAGTGCCGTGTGGCTTTC  
GGCGCCGGGGCTGTGCGCTGCACCTTCTCCAGGGAGGGTCTGGATGTATGCCAGGCTCTG  
TCCGTGCACACAGACCCCCCTGGACCAAAGCAGCTGCAGCCGCCCTCCTTGTTCTCTCCTG  
GACGGGACAGAATGTGGTGTAGAGAAGTGGTGTCTAAGGCTCGCTGTGCTCCCTAGCT  
GAGCTGGCTCCTGTGGCTGTGGTACATGGACACTGGTCTAGCTGGGGCCCCACAGTCTT  
TGCTCTCGATCCTGTGGAGGAGGTGTGATCACCAGGAGACGGTGGTGAACAACCCAGG  
CCTGCATTTGGGGGACGTGCATGTGTGGGCGAGGACCTCCAGGCTAAGATGTGCAACACG  
CAGGCTTGTGAGAAGACTCAGCTGGAGTTCATGTGACAGCAGTGTGCCCAGACAGACAGA  
CATCCACTGCAACTTTCCCAAGGCAGTGCCTCCTTCTACCCTGGGATGCTGCTGTGCAG  
TATAGTCAAGGGGATACTCTGTGCAGACACATGTGCTGGGCTGTTGGAGAAAGTTTCATT  
GTCAGCCGTGGGGACAGTTTCCCTCGATGGGACCCGTTGTGTGCCAAATGGTCTCGGGAG  
GATGGGACCACCTTTGGCTGTGATGGCAAGATGGACTCCCAGCAGGTTTGGGATGCGTGC  
CAGGTGTGTGGAGGAGACAACAGCACCTGCACCTCCCAAATGGCTCTTTACAGCTGGG  
AGAGCCAGAGAGTATGTCACGTTCCCTGATAGTCACTCCCAACATGACCAGCGCACACGTC  
ATCAACCACAGGCCTCTCTTACACACTTGGCGGTGAGGATCCAGGACCACTACATTGTG  
TCAGGGAAGAGCAGCATGTGCGCTAACGTACGTACCCCTTCCCTTCTGGAGGACTACCGC  
GTGGAGTACAGAGTAACTCTCACTGAGGACCAGCTGCCCCGCTTAGAGGAGATTACATC  
CGGGGACCTGTCCGGGATGACATTGAGATCCAGGTGTACAGACGATATGGAGGAGAATAT  
GGGGGTCTTACACACCCAGACATCACCTTTACCTACTTCCAACGGAAGCAGCAGGCAGCC  
TGGGTATGGGCCGCTAAGCGTGGGCCCTGCTCAGTGAAGTGTGGGGCAGGGCTGCGCTGG  
GTGACCTACAGCTGCCAGGATCAAGCTCAAAACAAGTGGGTGAGGAACGCCAGTGCCAA  
GGGAGCCCGCAGCCACCTGCATGGCAGGAGCCTTGTGTCTCTGCCCCGTGCCCTCCACAC  
TGGGTAGCTGACGACTTTGGCCCATGCAGCGTGTCTTGTGGCGGAGGCCCTTCGGGAGCGG  
TCACTGCGCTGTGTAGAGGCCACAATGGCTTCTTAAAGACGTTGCCGCCAGCCCCATGC  
AGAGCAGTAGCCCAGCAGCCAGCAGGAGAGGTAGAAAAGTGCAACTCCCAGCCCTGCCCC  
ACTGGATGGGAGGTGTGACACCTAGCCCTTGCCTGTGATCTGCGTGTGAACCCGGTCTG  
GAGTCAGGGAGTGGGACATGTGTGCCAGGGTGGGTGACTCAGAGAAGCCAGCGATGGCA  
GGCCCTGCCGCCGGATGACATGTACCTGTGCCGGAGCCCTGCGTTAGGACCACATGT  
TCTCCAGGCTCGGGTCCGGTGGACACCATGTCTCTGGGAGAGGAGGCTTCATCCCAGTG  
GGCAGTGGCAAGCCAAGCTCTCAGGCTGAACATGTGTGGACTCCTCTGGTGGGGCTGTGC  
TCCATTTCTTGTGGGAAAGGTCTGAAGGAGCTGTATTTCTGTGCATGGATTCTGTCTC  
AAGATGCCGTGTCAGGAAGAGCTATGTGGCTTGGCCAGTAAGCCCCAAGCCGGTGGGAG  
GTCTGCAGGGCTCGCCATGTCTGCTCGGTGGGAGACTCGAGTCTTGGCACCGTGTCCG  
GTGACCTGTGGCGGGGGCGAGTGCCTGGCCATTGATGTGTGCAGCTGGACCACGGG

CGCCCTGTGTCTGTACCTCACTCCAAGTGCTGGCCAGCACCTAGGCCAGGCTCCTTCGAG  
 GACTGCAGCCCTGAGCCTTGCTCTGCTAGATGGAAAGTCCTGCCCTTGCCCGTGCTCA  
 GCCAGCTGTGGCCTTGGCACTGCCACACAAATGGTGGCCTGCGTGCAGCTTGACCAAGGC  
 CATGACAATGAGGTGGATGAAACTTTCTGTAAGGCTCTGGTGCGGCCCCCGCCAGTGTT  
 CCCTGCCTCATTGCTGACTGTGCATACCGGTGGCACATCAGCACCTGGACAGAGTGCTCC  
 GTCTCCTGTGGGATGGCATCCAGCGCCGGCATGGCACCTGCCCTCGGACCCAGGCTCAA  
 GTGCCTGTGCCAGCCAACTTCTGCCAGCACCTGCCAAGCCAGTGACTGTGAGAGGCTGC  
 TGGGCTGGACCCTGTGCAGGGCAGAAGATCTCCAGCTCGCTGCCCTCACAAGGAAGCTACT  
 CTTCCCAGCCAGACCCTGGCTGCGGTTACTGCTGCCTCTCTGCAATGGTCCCAGCCCCAA  
 GCCAGGCCACAGCCTTTTCCCAGACTCCCAGTCTCTGGGGCTCCAGGAAACCCTGGAA  
 GAGCCTGGAGCCTGTGGAAGGCAATACCTGGAGCCCACAGGAACCATTGACATGCGAGGT  
 CAAGGACAGCTAGACTGTGTGGTAGCCATCGGGAGACCTCTAGGTGAAGTGGTGACCCTC  
 CAAATCCTCGAGAGCTCCCTCAAGTGCAGAGCAGGGGAGCTGCTGCTGCTCTGGGGTAAG  
 TTCACATGGCGGAAGATGTGCTGGAAGATGTCTGGCAGGACTTTCAGCACAAAAGCCAAC  
 ACGGTGGTGGTGAGGCAGCATCGTGTGCTGCCGGGAGGTGGAGTGCTGCTACTGTACCGG  
 AGCCAACCTGCCCCAGGGACCTTCTACAAAGAATGTGACAGGCAGCTCTTTGGACCTCGG  
 GGTGAAATCACGAGTCCCTCACTGAGCCAGACGGGAGGAAGGCAGGAACCTGCAGGGTC  
 TTTGTGAGTGTGGCCCCACAGGCCCGAATTGCCATCCGTGCCCTGGCCAGTGACATGGGC  
 ACTGCTTCTGAGGGGACCAGTGCCAACATATGTCTCGATCAGGGATATCCACGGTCTGAAG  
 ACTACAACATTTTGGGGGACAGGCTCTACTGGGAGTCCGAAGGCAGTGAGGCTGAA  
 CTGGAGTTTAGCCATGGCTTCCCTGGAGGCGCATGCCAGTTTCCAGGGCCAGTATTGGACC  
 ATCCACCGAGGACACCGAACAGCTGGCTCTGCTGAACAGCACAGTGGAGCAGCGCTGC  
 AGGTGGGATGCTGATGACAGGGCTGCAGGCCCTCGGCCTGTGCCATCTCTGGACTCTTC  
 AACTGCGTCACGATCCACCCTCTGAACATCGCAGCTGGGGTGTGGATGATCATGAACGCC  
 TTTATCCTGTTGCTGTGTGAGGCCCCCTTCTGCTGCCAGTTTGTGGAGTTTGCAAACACA  
 GTAGCTGAGAAGGTTGACCGGCTGCGCTCCTGGCAGAAGGCCGTCTTCTACTGCGGGATG  
 GCCATCGTCCCCATCGTCATGAGCCTGACCCTGACTACACTGCTGGGCAACGCCATCGCC  
 TTCGCCACGGGGTGTGTACGGACTGTCTGCCCTGGGCAAAAAGTGCCTCTACAGCCCA  
 GTCCCGGGGGGAGATGCCATTTCTTATGCTCGGATCCAGCAGCAGAGGCAGCAGGCGGAT  
 GAGGAGAAGCTGGCTGAGACATTCGAGGGGGAAGTGTGA

Candidate *Fugu rubripes* ADAMTS13 cDNA sequence

Sequence derived from Joint Genome Institute gene model FRUP00000165554 was obtained from <http://genome.jgi-psf.org/fugu6/fugu6.home.html> and is shown in uppercase characters. 168 nucleotides for candidate exon 1, exon 2 and exon 3 (partial) are shown in lowercase characters.

>Extended FRUP00000165554 transcript sequence

atgttttactgattcagctctgtgttctcttgctgccgctggcctggctggttctcctgagttctccttc  
 tgatgagtccttctgtatccattaagctctccgatctcaggcaagagaacatttttccctcttttgatt  
 ctgtccctgatgactctggatttatt  
 CCACTTTTCAGCAGAGGACGCGCGTCCGGTTCGGTTCGCTGATGTTACACATCTGGAGCTTCTCG  
 TTGTGGTCCGACCTGATGTGCATCAGGTCCACAAGCAGGACACGGAACGATACATCCTCACCAACCTAAA  
 TATTGCCTCAGAGCTGTTGAGAGACATTAGCCTGGGCGGCAACATGAGGGTGCACCTGGTCCGCATGATC  
 ATCCTGTCAGAGCCAGAGCCGGAGATCCAGCTGTGACCAACATCACGTCTTCCCTCACCAGTGTGTGTG  
 AGTGGGGCAGGAAGATAAACCCCGCAATGATTCAGACCCCTGCATGCTGATCTTCTCTTGTACATCAC

GAGGTATGACCTGGTGTTCCTGATGGGAATAACCAGGTCAGGGGTGTGGCACAGCTGGGCGGGCCTGC  
AGTAGCGAGTGGAGCTGCGTGATCACCAGGACACGGGCTTTGACCTGGGGATTACCATCACTCACGAGA  
TTGGTCACAGTTTTTGAATCAACCACGATGGAGTGAGAAACACCTGCAGCCGATCCGGCTTCATCATGGC  
CTCTGATGGAGGCTACAACAGCGTGGATCTGACCTGGTCCCCGTGCAGCCGACAGCAGCTGCTCGCCTTC  
TTCAGTGACGAAAAGCGGAATGTGTGAAAGACCTTCCGTACTGGAAGGCTCTCTGCAGGACTGGAAAC  
CTGGTCTGTATTATGGTGTGGATGACCAGTGTGCGATAGCGTTTGGTAGCGCCGCACGAGCCTGTTCCCT  
CACCAACAGCGACTTGGTACCTGTGCGACCCTGTCTTGTACAGTCAACCCTGGGGATGACAGCTCATGC  
AAACGCCTCTTAGTTCCTACTGCTGGACGGGACAGAATGTGCACCTCATCGGTGGTGGCTTGAGGGGTCATT  
GCGTGTCTACGGAGGAGCTGAGCTCCTCTGTGGTGGTGCACGGCGCCTGGTCCAGCTGGTCCGGGTTCTC  
GCCTTGTTCACGTACTTGCGGTGGAGGAATTTCTCATCGCACGCGTAAATGCACCAATCCGAGACCTGCC  
TTTGGAGGGAGCGATTGTGAGGGAGCAGATGTTGAAGCTGAACTTTGTACCAGCAGCGATGTGAAAACA  
GAACTCAGCTGGACTTCATAGCAGAACAGTGTTCAGAACAGACTCCCAGGCTCTCTACCTGCTACCAA  
CGCCGCCTCCTTCTACACGTGGATCCCTGCTGTAGGATTTACACAAGGAGAGGAACAGTGCCAATACATG  
TGCCAGTCAGATGGGAGAACTTCCCTCGTGAGACGTGGTTCTCAGTTCCTGGACGGAACCTGCTGCGAGT  
CGGACGCTTACCTCCGTTTGGCTCCACGGCGGCTTGTGTTGAGAGGCATATGCCAGCTGTTTGGCTGTGA  
TGGAACACTGCATCTGGGAAAGTGAAGGATGTGTGTGGCGTGTGCGGGGGAAACGCATCATCCTGCAGC  
TTGAGGTCAGGATCTTACACCGCCGGTCAAGCCAGAGAGTACACCACCTTCCCTCTCCCTGCCAGTGAACG  
CCACACAAGTCCACGTCATCAACAGGGCGCCCTCTTCACTCACATGGCTGTGGTATTGAGGAGTCAGTA  
CATTGTGGCTGGACGCGGTGCCATGTCTTTGAATGTGACCTACCCGTCTGTTCTGGACGAGGACCGCCTT  
GAGTACCGACTGTATCTGACCCCTGATCTCCTGCCTGAAATGGAGGAGCTGCTGCTGCCTGGACCACTGC  
AGGAAGAGATAAACGTAGAGGTCTATCGTAAATATGGAAAGGAATATGGAGAGCAAACGAATCCAAACAT  
TAGCTTTCAGTTTTATGTGCCTGGGCTCCCAAAGATGTGACAAACACCAAACCTTAAAGGCAAATGGACA  
TTTGTCAACAACCCCTGCTCCGTGTCTTGTGGATCAGGTCTGTGTACCAGAAATCTTATCAGACATCACC  
AATGCTTGAAGGTCTTGTGCTTACAGGTGGGATACGGGCCGTTTGGACCCTGCAGTGCTTCCTGTGG  
TGGAGGAGAGCGTAAACGCCTCCTGAAATGCGTTCAAGTACTTGGGGGCACCGTGACGCAGGTTCAAAAT  
TCAGACTGTCCAGCAGATAAAGCTCCTACGTCTGTTGAAAGATGTAACCTGCAACACTGCCCTGCCAGGT  
GGCAGGTCTCAGAACCAGGAGAATGTTCTGCTGTTTTGCGGGCCAGGAGAAGCCAAGCAGGTTGTGTCTG  
TGTGCGTCCAGAAAACGGACAGGATGTCAAAGTGGACGAACGCTTTTGTGTCAGAGACTGAAGAGACCAGCC  
GATTCTGTGCCCTGTGTGGTCAATGTCTGTCTTATTGGATGGGAATCAACGGGGCAGGACCAGGTCTGT  
TGAAGTCGCTGCCAGACACAAAACAAGATCCTGTGTATGTGTGGAGCCCCATTATCAGCAAGTGTCTCAA  
GACCTGTGGAATGGAACCCTACAGCTGTGGTTTTTCTGTGTGGACCACCAGACCAGACTCGGGGTGAA  
GACTTGTACTGCGACGCTTCCACCAAACCGCCTCCTCAGTCTGAGAGCTGCAGCACGTCTCCCTGTCTC  
CTACGTGGCGCTTTAAGCAAGGAGCGTGCAGCGTAACGTGTGGAGGCGGGTTGGCCAACAGAGTCTGT  
CTGTGGAAGGGAGGCCGGAGGAGAGGAGGAGGAGGTGCTGGAGGATTACAGACTGCAGCCACTCGGTCAA  
CCCACGGCCGTGGTTTTCTGTGCAACGGCCACAGCTGCCAGCGAGATGGAAAGTGTGCGGGACGTCGCCCT  
GCTCGGCGTCTGTGATCTCGGCATCGCTCAGAGGGCCGTGCTTGTGTCCGGTTCATCAACGGCACAGA  
GAGCGCAGTGGCGGAGGAAAAGTGCCACGCTGGGGTCAAACCAGCCACCAGCGTTCCCTGTCTGGTGCAG  
GTCTGCACCTTCAGATGGGAGGTGAAACCTTGGAGCCAGTGTTCAGTCTCCTGTGGATACGGGATCCAGT  
CCAGAGCCGTGTCTTGCATGGGTCCCTCGCATCCGCAGCCTCTAAGCCCCATGCTGTGCATGCACATGCC  
CAAGCCCATCACCATCCAGAGCTGCAGCACGGGGAGCTGCAGCGAGGAGCGACCTTCCACGCCGGCAGC  
CCGCTGCAGCCCACGGCTCGCACCGCTCCTCCAGGTCCAACAGAGGCCAGCACCGTCTTCCAGAACCTCA  
CAACCCCTATTCCCCTACCAAGATGAGTTTTATTCTCTTGGCGTTGGGTCTCACCAGATCCATGTGGGCA  
CCTTCTCCTGCAGCCATCAGGCACTGTGGACCTGAGGGACGTAAGTGGACGCTGCACAGTGTCCATTGGT  
CGGCCACTGGATGAGGTCACTACGTCAAAGTGAATCTGGCACCTGAACTGCAAACCTGCAAAGAGAGT  
ACGTGGCGTTTTTTGACCGGCTGGTGTGTTGTGAGGAAGTGAAGCAGGTGACAGGAACTGAGCTGACCAC  
GCGAACCAACGTCCTGCTGGTTGCGCAAAGCCTGCTGACCCTGGGAATGGGGTCGTCTTGACTTTCAAC  
TCCATGAAGAACACGAAGAGGAGCCATCATCAGGATTGTGACGTTTACAGCTGTTTTACCGAGTGGCATCT  
TTGAGAATCCAACAACCTTCCACACTAACCACACCTGCAGGGTCTCCTCAATCCCCCTTCGGTGAAGAT

CCAAATCCAAGCGCTGCACATCGGACTCAACGCCACCAACTCTCAGTCTACATACATCATGGTGAACACA  
CACTTTTCATGGGACATGGACATCTTAAAGACCACCGTGTTTAAAGGCCGCCAGCTGTTCCCTGTGGCACT  
CGTCCGGAAACATGGCTGAGGTGGAATTCATGGCGATTTCCCTGCGATCCAAAGGAAGCTTCAGAGCAGA  
ATAT