

# GeneVision Analysis Report

**File name:** genevision\_report\_2025-04-26.pdf

**Generation date:** April 26, 2025

**User:** Mariem

## 1. General Information

**Analyzed sequence:** Available in appendix

**Predicted genes:** Available in appendix

**Protein sequences:** Available in appendix

**Tools used:**

- Gene prediction: AUGUSTUS
- Functional annotation: DeepGOPlus + QuickGO
- Structural modeling: ESMAtlas

## 2. Sequence Data

- Number of predicted genes: 1
- Number of protein sequences: 1
- Total input sequence length: 9453 bp

## 3. Results Summary

Gene ID	Position	Score	GO Term	Function	Description
gene1	966 - 6903	0.94	GO:0008150	biological_process	Living things follow instructions in their genes to carry out tasks, like building proteins or breaking down food. These step-by-step tasks are called biological processes.

# Appendix - Sequence Contents

## A1. Input Sequence

```
>input_sequence
GAGCTCACATTAACATTTTACAGGGTAACTGCTTAGGACCAGTATTATGAGGAGAATTTA
CCTTTCCCGCCTCTCTTTTCCAAGAAACAAGGAGGGGGTGAAGGTACGGAGAACAGTATTT
CTTCTGTTGAAAGCAACTTAGCTACAAAGATAAAATTACAGCTATGTACACTGAAGGTAGC
TATTTTCATTCCACAAAATAAGAGTTTTTTTAAAAAGCTATGTATGTATGTGCTGCATATAG
AGCAGATATACAGCCTATTAAGCGTTCGTCCTACTAAAACATAAAACATGTCAGCCTTCTTA
ACCTTACTCGCCCCAGTCTGTCCCGACGTGACTTCCTCGACCTCTAAAAGACGTACAGAC
CAGACACGGCGGGCGGGCGGGGAGAGGGGATTCCCTGCGCCCCCGGACCTCAGGGCCGCT
CAGATTCCTGGAGAGGAAGCCAAGTGTCTTCTGCCCCCCCCCGGTATCCCATCCAAGGC
GATCAGTCCAGAACTGGCTCTCGGAAGCGCTCGGGCAAAGACTGCGAAGAAGAAAAAGACA
TCTGGCGGAAACCTGTGCGCCTGGGGCGGTGGAACCTCGGGGAGGAGAGGGAGGGATCAGA
CAGGAGAGTGGGGACTACCCCCCTCTGCTCCCAAATTGGGGCAGCTTCCTGGGTTTCCGAT
TTTCTCATTTCCGTGGGTAAAAAACCTGCCCCACCGGGCTTACGCAATTTTTTTAAGG
GGAGAGGAGGGAATAATTTGTGGGGGGGTACGAAAAGGCGGAAAGAAACAGTCATTAC
ATGGGCTTGGTTTTTCAGTCTTATAAAAAAGGAAGTTCTCTCGGTTAGCGACCAATTGTCA
TACGACTTGCAGTGAGCGTCAGGAGCACGTCCAGGAACCTCCTCAGCAGCGCCTCCTTCAG
CTCCACAGCCAGACGCCCTCAGACAGCAAAGCCTACCCCCGCGCCGCGCCCTGCCCCCGG
CTCGGATGCTCGCCCCGCGCCCTGCTGCTGTGCGCGGTCTCGGCGCTCAGCCATACAGGTG
AGTACCTGGCGCCGCGCACCCGGGACTCCGGTTCACGCACCCGGGCAGAGTTCCGCTC
TGACCTCCTGGGTCTATCCAGTACTCCGACTTCTCTCCGAATAGAGAAGCTACGTGACT
TGGGAAAGAGCTTGGACCGCTAGAGTCCGAAAGAACTCCGTGGATATTCCAGCTTTCCCA
CAAGCACTGATCATTATGAGCCAGTTACTTAAACCGATCTGAGACACTCTCACCTCCTAAA
TAGGGATAGATGATACTAATTTGCAGGTTGTCAATTATGATAAGACAGGATCTGATCAATA
TATGTGAATTGTTTATATTTGGAACCTTTTTATTGAGTGGAAAGAAGTTGTTTTAAATATT
CTAGTCAGTTCTTTCCTGCTCCCAGGAAAGCCCGGATTATGTTTTAAGATAAGCAAAATG
TCTTAAAAGTAAGCTGTTTTACTTTGAATTTTTCCCTAAATGTTGATTAGTGTACTAGAT
CCATTTTTAATTTGGAAAGTGAAGTGCTACTTATTTGAACCTTCTTAAAAATGCTAATTTTA
ACATCTAAAGAGTTAACTAAGAAAAGCTTAGTAACATGATGTACCAAGTTGAATATGCTG
TTATCCTTATTTAGAATAGAAAATTGGTATTTCTACGTTTTATCCATTCTAAGGCAGGTT
AAAAAATTGATTTCCATGACTACCTATATATTTCTTGAATTTATTATTGTAAAGTTGAT
TCATAGTCAAACAATTAATGTTTAAATTAAGATTAAGACACTAGAGAATGATTTATTTG
CTGTCCTTTAATTGCAGCAAATCCTTGCTGTTCCACCCATGTCAAAACCGAGGTGTATG
TATGAGTGTGGGATTTGACCAGTATAAGTGCGATTGTACCCGGACAGGATTCATATGGAGA
AAACTGCTCAACACGTAAGTTTGTCTTTGGTTGCCTCATTAGGAGTGGGGCTGGATACA
GTTATCATTGTATAGATTTGTGTCTTATAATGAGTCCCATTAATTTCTCCCTCCCTTTCT
TCGTCTTCTTGACGCGGAATTTTTGACAAGAATAAAATTATTTCTGAAACCCACTCCAAA
CACAGTGCCTACATACTTACCCACTTCAAGGGATTTTGGAAACGTTGTGAATAACATTCC
CTTCCTTCGAAATGCAATTATGAGTTATGTCTTGACATGTAAGTACAAGTGTCTTTCTAA
GGTTTTTTAGCCTTCTCAAAGAAAAATATGCTTTTATAATACTGTAAGCCTAATCTAAAAAC
ATATTTCCAAGCTTATCAAAAAGACTTTAAGATAGCTTTTAAAGTTTGCCTTCCATCTTAA
TCGCCAAAAATATTGACATTTAGTCCCATCCAGTTTATACAGTCTGCTCACAACCTCTGTA
TACCTCTTCTAACCTTTACTGTTTGGTCAGTTTGTGGAGGTAGCATGGTCCAGCTGTTTA
TTGAATGCCCATGGGCCACAGAATTGTTCTGAACATGTAGCACCCATTAAAAATAAATTTG
GATTTGGATCAGCAAGAAAAATACTTTCCATGATTCTAAAGTGGGTGCCATACTCAGCCA
TTCTTTTTCATAGGCCTCTGGATAGTGAGCAGATGGCTACCTGAAAAATCAATATTGCCA
GATTATAATGTGCAGAGTATATGTATTTTATTAAAGATGTATTTCAAGTGGCCATTAGAC
TATAAAGTGTAGTTGTTTAAAAATAGATTTTTTTTATTTTGGAGTTACATTCAACCTCAG
GTGCCACTTTCCACATTTTACAATAAAAAATAATGGTTGATTTACTTAACAAATGAGAATA
AATAAAACATTTTTTTCTTTGAAAAATTCAGCCAGATCACATTTGATTGACAGTCCACCA
ACTTACAATGCTGACTATGGCTACAAAAGCTGGGAAGCCTTCTCTAACCTCTCCTATTAT
ACTAGAGCCCTTCTCCTGTGCTGATGATTGCCCGACTCCCTTGGGTGTCAAAGGTGAG
TAAGAAGAATCCATTAGAGATGTATTAACATAAGACGGGCTGCATTGTGCCAAAAAAA
AAAATTGACCTTAGACTACCATTTATTTATTAACAAAAGCAGTTTTTACTTTTAGCATGG
TTATCTATGGGTATTTTTTAAAGTATGAGTCTATATAAACTATTATGTAAAAGCAAATGA
GCGTCTTGGTATAATGTCTTAATATTTTCAAATTATTTCTTTAGAAATGAAATAATTTCTA
ATTAATAATGATAAAATCATTTCAGTAAGAAGTTGTTCCACCATACTTAGAACTGTTGTT
TATATTATGATCCTAATTCACAATTGTAATCTCATATAAATGAAGAATTCTTGGTAGATT
GACAGTCACCATCTCCTTTCTTGAATACATAGATGGATTCTTACCTTAGCTTTCTCATTTT
TTCAGGTAAAAAGCAGCTTCCTGATTCAAATGAGATTGTGGAAAAATTGCTTCTAAGAAG
AAAGTTCATCCCTGATCCCCAGGGCTCAAACATGATGTTGCATTCTTTGCCCAGCACTT
CACGCATCAGTTTTTCAAGACAGATCATAAGCGAGGGCCAGCTTTACCAACGGGCTGGG
CCATGGGGTAAAGATAGAGTTAATATCTTAGAGTTAGTAAAAATTATACCAATCATAGTCA
AGGGCTAACATTAAAGGATATACAGTAGATAGATCCAAATAAATTTATCCACTTTTTTT
```

TAAAAAGAAGTCTTATCTATAAAAAACCTTAAAGGAATTTTCCATTTACTTCACTGGTCTA  
GTAAAATTATACACACACACAGACATGCACACACATATATAAACATTCACACACATACAT  
ATGTACAGGTATTGTTATTTGTAATTTGACCTTGTATTTTTTAGTTTAAAAATGTTAGTA  
CTGCAAAATGTTATGTCCTCAAAAACACATTGTACCATGATTATGCCGCTTTCAATATTG  
TAAAGTGAGGTTTTTGCCGCATTATTATTTTTTGGATTTCATAGCATAGCTTCAAGTTA  
TTCGTAAGAATTTTTTATAAATAATACATTTTTTATACTTTTTTATAATTACCATATCATC  
ATAGTGAAGTATATAATATATATGATATAAGCTCAATATAGTATATTAATTCGGTTAAAC  
ACAAAGACATATCAGTTTGTAGCTTTGGTGGATAAAACAAATTAATTTAGCAATTCATGGC  
TATGAAAAATGTATATTTTTATTTAAAAATTTTAAAGAAAAGCTAAATGATCAAATTTATTTA  
ATGATGAATTATATGATAGACACTTTATATAAGAAAAACTTCAACAGCAACAAATTTAAAA  
TTTTTTCATCATTTTTCTAGGTGGACTTAAATCATATTTACGGTGAAACTCTGGCTAGACA  
GCGTAAACTGCGCCTTTTTCAAGGATGGAAAAATGAAATATCAGGTATGCTTCCTTTGACT  
ATTAAGACTTAGTTATTACCGCTTATACCCATATTTTTAAAAATCCCTAAAAATGTGTTCCCT  
TAACTTTTTAACTGATGTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTT  
TATCCTCCCACTGATCAAGATACTCAGGCAGAGATGATCTACCTCCTCAAGTCCCTGAG  
CATCTACGGTTTGCTGTGGGGCAGGAGGTCTTTGGTCTGGTGCCTGGTCTGATGATGTAT  
GCCACAATCTGGCTGCGGGAACACAACAGAGTATGCGATGTGCTTAAACAGGAGCATCCT  
GAATGGGGTGATGAGCAGTTGTTCCAGACAAGCAGGCTAATACTGATAGGTAAACAAGAA  
AATGATTTATATAAAACCTCTTCCCAAGGAAAAATAGTGTGCTATCTTTGTTATGTTT  
TGAGTAAATGACAAGATGTGGTAAATGAAAACCTCACACATTCATATACATTAAATATGT  
AAGCATGACTGATAAAATAGCTATCTTTTGATACTGACAAGGAAGAAAAACAGAAATGAAG  
GAATAGCAAATTTAAAAATTGCATTCCAGTTGCTTGAAAGCTTGTGATCAGATGCAATA  
AATGTTTTTATTATTTATTTTGTGCAAATAGGAGAGACTATTAAGATTGTGATTGAAGAT  
TATGTGCAACACTTGAGTGGCTATCACTTCAAACCTGAAATTTGACCCAGAACTACTTTTC  
AACAAACAATTCAGTACCAAAATCGTATTGCTGCTGAATTTAACACCTCTATCACTGG  
CATCCCTTCTGCTGACACCTTTCAAATTCATGACCAGAAATACAACATCAACAGTTT  
ATCTACAACAACCTCTATATTGCTGGAACATGGAATTACCCAGTTTGTGTAATCATTACC  
AGGCAAATTGCTGGCAGGGTAAGCATTATTATTGAAAACCAAAACAAAAGACTAGTCAGT  
AACTTTTGAATTTCTGCCACGGAAATTATTTTTCTTAAACTTACTAAAAGAGTAGTTAGT  
TATATTGCTAGTAAAATTATTTTATTGATATAAGAAAGCCTAACTTTGTTTGAAAAGTCTA  
AACTTTTAGTCTAGTCTACAGTTGTGTCAGACAAATAGCAAATTTGTACCCCTACCTTAAAAA  
TATTTTCAAAAAGTATCTATAATCTTATAGGAATAAATATTTTAGGCTTGAATACTAGTG  
TTATTTTTGAAATGTAAAAAGGCAAATTAGTTCTAGGCTGGTGTCCCATTGAATTTTAAG  
CAGAGCTCCTGTTGAAATGTAGGTAAGCATCTTTCAGCAAATAAAAAATTGTCTCCGCTG  
GGAGTTTCAGTTTTACCTGATTTGTACCTAAGGCAAGCTGAATACAAACAGTAAATATGC  
CTAAAATTCTTGTTTTACAACATAATTTTACTTTCCACAGGTTGCTGGTGGTAGGAATGTT  
CCACCCGCAGTACAGAAAGTATCACAGGCTTCCATTGACCAGAGCAGGCAGATGAAATAC  
CAGTCTTTTAAATGAGTACCGCAAACGCTTTATGCTGAAGCCCTATGAATCATTTGAAGAA  
CTTACAGGTAAGAAACAGTTTCTAAACTTCTTCGTTTTTTGTTTTGTTTTGTTTTGTTTTGT  
TGTTTTTTGGTTTTCTTTTCGAGATGGAGCCGCCCTCTGTCACCCAGGCTGGAGTGCAGTG  
GCGCCATCTCGGCTCACTGCAACCTCCGCCTCCTGGGTTCAGCAATTCCTCTGCCTCAA  
CTTCCTGAGTAGCTGGGACTACAGGCTCACGTGCGACGCATGGATAATTTTTTGTTATTTT  
CAGTATAGACGGGTTTTACCGTGTTAGCCAGGCTGGTCTCAAACCTCTGACCTAGTGAT  
CCGCCGGCTTCGGCCTCCGGAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCTGGCCC  
CTAAACTTCTTAAAAGAATCAGGGGTCAAATGGAAAACAGAGAAGTTGGCAGCAAATTGAG  
CAAAAGAATCAAACCTGTTTTTTATTTTGTGAAGTTTGACATTGGTTGTATCTCTGTCTTC  
ATCGCCTTCACAGGAGAAAAGGAAATGTCTGCAGAGTTGGAAGCACTCTATGGTGACATC  
GATGCTGTGGAGCTGTATCCTGCCCTTCTGGTAGAAAAAGCCTCGGCCAGATGCCATCTTT  
GGTGAATACCATTGTTAGTAAAGTTGGAGCACCATTCTCCTTGAAAGGACTTATGGGTAATGTT  
ATATGTTCTCCTCGCTACTGGAAGCCAAGCACTTTTGGTGGAGAAGTGGGTTTTTCAAATC  
ATCAACACTGCCTCAATTCAGTCTCTCATCTGCAATAACGTGAAGGGCTGTCCCTTTACT  
TCATTCAGTGTTCCAGATCCAGAGCTCATTAACAGTCACCATCAATGCAAGTTCTTCC  
CGCTCCGACTAGATGATATCAATCCCACAGTACTACTAAAAGAAGCTTCGACTGAACTG  
TAGAAGTCTAATGATCATATTTATTTATTTATATGAACCATGTCTATTAATTTAATTATT  
TAATAATATTTATATTAACCTCCTTATGTTACTTAAACATCTTCTGTAAACAGAAAGTCAGTA  
TCCTGTTGCGGAGAAAAGGAGTCATACTTGTGAAGACTTTTATGTCACTACTCTAAAGAT  
TTTGCTGTTGCTGTTAAGTTTGGAAAAACAGTTTTTTATTTCTGTTTTTATAAAACCAGAGAGAA  
ATGAGTTTTGACGTCTTTTTACTTGAATTTCAACTTATATTATAAGAACGAAAGTAAAGA  
TGTTTTGAATACTTAAACACTATCACAAAGATGGCAAAATGCTGAAAGTTTTTACACTGTCTG  
ATGTTTCCAATGCATCTTCCATGATGCATTAGAAGTAACTAATGTTTGAAAATTTTAAAGT  
ACTTTTGGTTATTTTTCTGTCTATCAACAAAAACAGGTATCAGTGCATTATTAATGAAT  
ATTTAAATTAGACATTACCAGTAATTTTCATGTCTACTTTTTTAAAAATCAGCAATGAAACAA  
TAATTTGAAATTTCTAAATTCATAGGGTAGAATCACCTGTAAAAAGCTTGTTTGATTTCTT  
AAAGTTATTAACTTGTACATATACCAAAAAGAAGCTGTCTTGGATTAAATCTGTAAAA  
TCAGATGAAATTTTACTACAATTGCTTGTGTTAAAAATATTTTATAAGTGATGTTCCTTTTTC  
ACCAAGAGTATAAACCTTTTTTAGTGTGACTGTTAAAACTTCTTTTTAAATCAAAATGCCA  
AATTTATTAAGGTGGTGGAGCCACTGCAGTGTTATCTCAAAATAAGAATATTTTGTGAG  
ATATTCAGAAATTTGTTTATATGGCTGGTAACATGTAAAAATCTATATCAGCAAAAGGGTC

TACCTTTTAAAAATAAGCAATAACAAAGAAGAAAAACCAAATTATTGTTCAAATTTAGGTTTA  
AACTTTTGAAGCAAACCTTTTTTTTATCCTTGTGCACTGCAGGCCTGGTACTCAGATTTTGG  
CTATGAGGTTAATGAAGTACCAAGCTGTGCTTGAATAACGATATGTTTTCTCAGATTTTTC  
TGTTGTACAGTTTAATTTAGCAGTCCATATCACATTGCAAAAAGTAGCAATGACCTCATAA  
AATACCTCTTCAAATGCTTAAATTCATTTACACATTAATTTTATCTCAGTCTTGAAGC  
CAATTCAGTAGGTGCATTGGAATCAAGCCTGGCTACCTGCATGCTGTTCCCTTTTCTTTTC  
TTCTTTTAGCCATTTTGCTAAGAGACACAGTCTTCTCATCACTTCGTTTCTCCTATTTTG  
TTTTACTAGTTTTTAAGATCAGAGTTCACCTTCTTTGGACTCTGCCTATATTTTCTTACCT  
GAACCTTTTGCAAGTTTTTCAGGTAAACCTCAGCTCAGGACTGCTATTTAGCTCCTCTTAAAG  
AAGATTAAAAGAGAAAAAAAAGGCCCTTTTAAAAATAGTATACACTTATTTTAAAGTGAA  
AAGCAGAGAATTTTATTTATAGCTAATTTTAGCTATCTGTAACCAAGATGGATGCAAAGA  
GGCTAGTGCCTCAGAGAGAACTGTACGGGGTTTGTGACTGGAAAAAGTTACGTTCCCAT  
CTAATTAATGCCCTTTCTTATTTAAAAACAAAACCAAATGATATCTAAGTAGTTCTCAGC  
AATAATAATAATGACGATAATACTTCTTTCCACATCTCATTGTCACTGACATTTAATGG  
TACTGTATATTACTTAATTTTATTGAAGATTATTATTTATGTCTTATTAGGACACTATGGT  
TATAAACTGTGTTTAAAGCCTACAATCATTGATTTTTTTTTTGTATGTCACAATCAGTATA  
TTTTCTTTGGGGTTACCTCTCTGAATATTATGTAAACAATCCAAAGAAATGATTGTATTA  
AGATTTGTGAATAAAATTTTAGAAATCTGATTGGCATATTGAGATATTTAAGGTTGAATG  
TTTGTCCTTAGGATAGGCCTATGTGCTAGCCACAAAGAATATTGTCTCATTAGCCTGAA  
TGTGCCATAAGACTGACCTTTTAAATGTTTTGAGGGATCTGTGGATGCTTCGTTAATTT  
GTTTCAGCCACAATTTATTGAGAAAATATTCTGTGTCAAGCACTGTGGGTTTTAATATTTT  
TAAATCAAACGCTGATTACAGATAATAGTATTTATATAAAATAATTGAAAAAAATTTTCTT  
TTGGGAAGAGGGAGAAAATGAAATAAATATCATTAAAGATAACTCAGGAGAATCTTCTTT  
ACAATTTTACGTTTAGAATGTTTAAAGGTTAAGAAAAGAAATAGTCAATATGCTTGATATAA  
ACACTGTTCACTGTTTTTTTTTAAAAAAAACCTTGATTTGTTATTAACATTGATCTGCTG  
ACAAAACCTGGGAATTTGGGTTGTGTATGCGAATGTTTCAGTGCCCTCAGACAAATGTGTA  
TTTAACTTATGTAAAAGATAAGTCTGGAAATAAATGTCTGTTTATTTTGTACTATTTAA  
AAAAAAAATCGATGTGCTGACTCGAGTC

## A2. Predicted Genes

>gene1  
ATGCTCGCCCGCGCCCTGCTGCTGTGCGCGGTCCTGGCGCTCAGCCATACAGCAAATCCT  
TGCTGTTCCACCCATGTCAAAACCGAGGTGTATGTATGAGTGTGGGATTTGACCAGTAT  
AAGTGCATTTGTACCCGGACAGGATTCTATGGAGAAAACCTGCTCAACACCGGAATTTTTG  
ACAAGAATAAAATTTATTTCTGAAACCCACTCCAAACACAGTGCCTACATACTTACCCAC  
TTCAAGGGATTTTGAACGTTGTGAATAACATTCCCTTCCCTTCGAAATGCAATTATGAGT  
TATGTCTTGACATCCAGATCACATTTGATTGACAGTCCACCAACTTACAATGCTGACTAT  
GGCTACAAAAGCTGGGAAGCCTTCTCTAACCTCTCCTATTATACTAGAGCCCTTCTCCT  
GTGCCTGATGATTGCCCCGACTCCCTTGGGTGTCAAAGGTAAAAAGCAGCTTCTCTGATTCA  
AATGAGATTGTGGAATAATTTGCTTCTAAGAAGAAAGTTTCATCCCTGATCCCAGGGCTCA  
AACATGATGTTTGCATTCTTTGCCCAGCACTTCACGCATCAGTTTTTCAAGACAGATCAT  
AAGCGAGGGCCAGCTTTACCAACGGGCTGGGCCATGGGGTGGACTTAAATCATATTTAC  
GGTGAAACTCTGGCTAGACAGCGTAAACTGCGCCTTTTCAAGGATGGAATAATGAAATAT  
CAGATAAATTGATGGAGAGATGTATCCTCCACAGTCAAAGATACTCAGGCAGAGATGATC  
TACCCTCCTCAAGTCCCTGAGCATCTACGGTTTGTCTGTGGGGCAGGAGGTCTTTGGTCTG  
GTGCCTGGTCTGATGATGTATGCCACAATCTGGCTGCGGGAACACAACAGAGTATGCGAT  
GTGCTTAAACAGGAGCATCCTGAATGGGGTGATGAGCAGTTGTTCCAGACAAGCAGGCTA  
ATACTGATAGGAGAGACTATTAAGATTGTGATTGAAGATTATGTGCAACACTTGAGTGGC  
TATCACTTCAAACCTGAAATTTGACCCAGAACTACTTTTCAACAAACAATTCAGTACCAA  
AATCGTATTGCTGCTGAATTTAACACCTCTATCACTGGCATCCCTTCTGCCTGACACC  
TTTCAAATTCATGACCAGAAATACAACATCAACAGTTTATCTACAACAACCTCTATATTG  
CTGGAACATGGAATTACCCAGTTTGTGTAATCATTACCAGGCAAATTGCTGGCAGGGTT  
GCTGGTGGTAGGAATGTTCCACCCGAGTACAGAAAAGTATCACAGGCTTCCATTGACCAG  
AGCAGGCAGATGAAATACCAGTCTTTTAAATGAGTACCGCAAACGCTTTATGCTGAAGCCC  
TATGAATCATTTGAAGAACTTACAGGAGAAAAGGAAATGTCTGCAGAGTTGGAAGCACTC  
TATGGTGACATCGATGCTGTGGAGCTGTATCCTGCCCCTTCTGGTAGAAAAGCCTCGGCCA  
GATGCCATCTTTGGTGAAACCATGGTAGAAGTTGGAGCACCATTCTCCTTGAAAGGACTT  
ATGGGTAAATGTTATATGTTCTCCTGCCTACTGGAAGCCAAGCACTTTTGGTGGAGAAGTG  
GGTTTTCAAATCATCAACACTGCCTCAATTCAGTCTCTCATCTGCAATAACGTGAAGGGC  
TGTCCCTTTACTTCATTTCAGTGTTCAGATCCAGAGCTCATTTAAACAGTCACCATCAAT  
GCAAGTTCTTCCCGCTCCGGACTAGATGATATCAATCCCACAGTACTACTAAAAGAACGT  
TCGACTGAACTGTAG

### A3. Protein Sequences

>gene1

MLARALLLCAVLALSHTANPCCSHPCQNRGVCMSVGFDQYKCDCTRTRTFYGENCSTPEFL  
TRIKLFLKPTPNTVHYILTHFKGFVNVMNIPFLRNAIMSYVLTSRSHLIDSPPTYNADY  
GYKSWEAFSNLSYYTRALPPVPDDCPTPLGVKGKKQLPDSNEIVEKLLLRKFIPDPQGS  
NMMFAFFAQHFTHQFFKTDHKRGPAFTNGLGHGVDLNHIYGETLARQRKLRLFKDGKMKY  
QIIDGEMYPPTVKDTQAEMIYPPQVPEHLRFVAVGQEVFGLVPGLMMYATIWLRHNRVCD  
VLKQEHPEWGDEQLFQTSRLILIGETIKIVIEDYVQHLSGYHFKLKFDPELLFNKQFQYQ  
NRIAAEFNTLYHWHPLLPDTFQIHDQKYNYYQFIYNNSILLEHGITQFVESFTRQIAGRV  
AGGRNVPPAVQKVSQASIDQSRQMKYQSFNEYRKRFMLKPYESFEELTGEKEMSAELEAL  
YGDIDAVELYPALLVEKPRPDAIFGETMVEVGAPFSLKGLMGNVICSPAYWKPSTFGGEV  
GFQIINTASIQSLICNNVKGCPFTSFVDPDELIKTVTINASSSRSGLDDINPTVLLKER  
STEL