

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	Biological processes are series of actions driven by our genes to achieve a specific goal in the body, like building a protein or breaking down food. These actions are carried out by molecules in a specific order.

Appendix - Sequence Contents

A1. Input Sequence

```
>input_sequence
GAGCTCACATTAACATTTTACAGGGTAACTGCTTAGGACCAGTATTATGAGGAGAATTTA
CCTTTCCCGCCTCTCTTTTCCAAGAAACAAGGAGGGGGTGAAGGTACGGAGAACAGTATTT
CTTCTGTTGAAAGCAACTTAGCTACAAAGATAAAATTACAGCTATGTACACTGAAGGTAGC
TATTTTCATTCCACAAAATAAGAGTTTTTTTAAAAAGCTATGTATGTATGTGCTGCATATAG
AGCAGATATACAGCCTATTAAGCGTTCGTCCTACTAAAACATAAAACATGTCAGCCTTCTTA
ACCTTACTCGCCCCAGTCTGTCCCGACGTGACTTCCTCGACCTCTAAAAGACGTACAGAC
CAGACACGGCGGGCGGGCGGGGAGAGGGGATTCCCTGCGCCCCCGGACCTCAGGGCCGCT
CAGATTCCTGGAGAGGAAGCCAAGTGTCTTCTGCCCCCCCCCGGTATCCCATCCAAGGC
GATCAGTCCAGAACTGGCTCTCGGAAGCGCTCGGGCAAAGACTGCGAAGAAGAAAAAGACA
TCTGGCGGAAACCTGTGCGCCTGGGGCGGTGGAACCTCGGGGAGGAGAGGGAGGGATCAGA
CAGGAGAGTGGGGACTACCCCCCTCTGCTCCCAAATTGGGGCAGCTTCCTGGGTTTCCGAT
TTTCTCATTTCCGTGGGTAAAAAACCTGCCCCCACC GGCTTACGCAATTTTTTTAAGG
GGAGAGGAGGGAAAAAATTTGTGGGGGGGTACGAAAAGGCGGAAAGAAACAGTCATTAC
ATGGGCTTGGTTTTTCAGTCTTATAAAAAAGGAAGTTCTCTCGGTTAGCGACCAATTGTCA
TACGACTTGCAGTGAGCGTCAGGAGCACGTCCAGGAACCTCCTCAGCAGCGCCTCCTTCAG
CTCCACAGCCAGACGCCCTCAGACAGCAAAGCCTACCCCCGCGCCGCGCCCTGCCCCCGG
CTCGGATGCTCGCCCCGCGCCCTGCTGCTGTGCGCGGTCTCGGCGCTCAGCCATACAGGTG
AGTACCTGGCGCCGCGCACCCGGGACTCCGGTTCACGCACCCGGGCAGAGTTCCGCTC
TGACCTCCTGGGTCTATCCAGTACTCCGACTTCTCTCCGAATAGAGAAGCTACGTGACT
TGGGAAAGAGCTTGGACCGCTAGAGTCCGAAAGAACTCCGTGGATATTCCAGCTTTCCCA
CAAGCACTGATCATTATGAGCCAGTTACTTAAACCGATCTGAGACACTCTCACCTCCTAAA
TAGGGATAGATGATACTAATTTGCAGGTTGTCAATTATGATAAGACAGGATCTGATCAATA
TATGTGAATTGTTTATATTTGGAACCTTTTTATTGAGTGGAAAGAGTTGTTTTAAATATT
CTAGTCAGTTCTTTCCTGCTCCCAGGAAAGCCCGGATTATGTTTTAAGATAAGCAAAATG
TCTTAAAAGTAAGCTGTTTTACTTTGAATTTTTCCCTAAATGTTGATTAGTGTACTAGAT
CCATTTTTAATTTGGAAAGTGAAGTGCTACTTATTTGAACCTTCTTAAAAATGCTAATTTTA
ACATCTAAAGAGTTAACTAAGAAAAGCTTAGTAACATGATGTACCAAGTTGAATATGCTG
TTATCCTTATTTAGAATAGAAAATTGGTATTTCTACGTTTTATCCATTCTAAGGCAGGTT
AAAAAATTGATTTCCATGACTACCTATATATTTCTTGAATTTATTATTGTAAAGTTGAT
TCATAGTCAAACAATTAATGTTTAAATTAAGATTAAGACACTAGAGAATGATTTATTTG
CTGTCCTTTAATTGCAGCAAATCCTTGCTGTTCCACCCATGTCAAAACCGAGGTGTATG
TATGAGTGTGGGATTTGACCAGTATAAGTGCGATTGTACCCGGACAGGATTCATATGGAGA
AAACTGCTCAACACGTAAGTTTGTCTTTGGTTGCCTCATTAGGAGTGGGGCTGGATACA
GTTATCATTGTATAGATTTGTGTCTTATAATGAGTCCCATTAATTTCTCCCTCCCTTTCT
TCGTCTTCTTGACGCGGAATTTTTGACAAGAATAAAATTTATTTCTGAAACCCACTCCAAA
CACAGTGCCTACATACTTACCCACTTCAAGGGATTTTGGAAACGTTGTGAATAACATTCC
CTTCCTTCGAAATGCAATTATGAGTTATGTCTTGACATGTAAGTACAAGTGTCTTTCTAA
GGTTTTTAGCCTTCTCAAAGAAAAATATGCTTTATAATACTGTAAGCCTAATCTAAAAAC
ATATTTCCAAGCTTATCAAAAAGACTTTAAGATAGCTTTTAAAGTTTGCCTTCCATCTTAA
TCGCCAAAAATATTGACATTTAGTCCCATCCAGTTTATACAGTCTGCTCACAACTCTGTA
TACCTCTTCTAACCTTTACTGTTTGGTCAGTTTGTGGAGGTAGCATGGTCCAGCTGTTTA
TTGAATGCCCATGGGCCACAGAATTGTTCTGAACATGTAGCACCCATTAAAAATAAATTTG
GATTTGGATCAGCAAGAAAAATACTTTCCATGATTCTAAAGTGGGTGCCATACTCAGCCA
TTCTTTTTCATAGGCCTCTGGATAGTGAGCAGATGGCTACCTGAAAAATCAATATTGCCA
GATTATAATGTGCAGAGTATATGTATTTTATTAAAGATGTATTTCAAGTGGCCATTAGAC
TATAAAGTGTAGTTGTTTAAAAATAGATTTTTTTTATTTTGGAGTTACATTCAACCTCAG
GTGCCACTTTCCACATTTTACAATAAAAAATAATGGTTGATTTACTTAACAAATGAGAATA
AATAAAACATTTTTTTCTTTGAAAAATTCAGCCAGATCACATTTGATTGACAGTCCACCA
ACTTACAATGCTGACTATGGCTACAAAAGCTGGGAAGCCTTCTCTAACCTCTCCTATTAT
ACTAGAGCCCTTCTCCTGTGCTGATGATTGCCCGACTCCCTTGGGTGTCAAAGGTGAG
TAAGAAGAATCCATTAGAGATGTATTAACATAAGACGGGCTGCATTGTGCCAAAAAAA
AAAATTGACCTTAGACTACCATTTATTTATTAACAAAAGCAGTTTTTACTTTTAGCATGG
TTATCTATGGGTATTTTTTAAAGTATGAGTCTATATAAACTATTATGTAAAAGCAAATGA
GCGTCTTGGTATAATGTCTTAATATTTTCAAATTTTCTTTAGAAATGAAATAATTTCTA
ATTAATAATAGATAAAATCATTTCAGTAAGAAGTTGTTCCACCATACTTAGAACTGTTGTT
TATATTATGATCCTAATTCACAATTGTAATTCATATAAAATGAAGAATTCTTGGTAGATT
GACAGTCACCATCTCCTTTCTTGAATACATAGATGGATTCTTACCTTAGCTTTCTCATTTT
TTCAGGTAAAAAGCAGCTTCTGATTCAAATGAGATTGTGGAAAAATTGCTTCTAAGAAG
AAAGTTTCATCCCTGATCCCCAGGGCTCAAACATGATGTTGCATTCTTTGCCCAGCACTT
CACGCATCAGTTTTTCAAGACAGATCATAAGCGAGGGCCAGCTTTACCAACGGGCTGGG
CCATGGGGTAAAGATAGAGTTAATATCTTAGAGTTAGTAAAAATTATACCAATCATAGTCA
AGGGCTAACATTAAAGGATATACAGTAGATAGATCCAAATAACTTATCCACTTTTTTT
```

TAAAAAGAAGTCTTATCTATAAAAAACCTTAAAGGAATTTTCCATTTACTTCACTGGTCTA
GTAAAATTATACACACACACAGACATGCACACACATATATAAACATTCACACACATACAT
ATGTACAGGTATTGTTATTTGTAAATTTGACCCTGTATTTTTTAGTTTAAAAATGTTAGTA
CTGCAAAATGTTATGTCCTCAAAAACACATTGTACCATGATTATGCCGCTTTCAATATTG
TAAAGTGAGGTTTTTGGCCGATTATTATTTTTTGGATTTCATAGCATAGCTTCAAGTTA
TTCGTAAGAATTTTTTATAAATAATACATTTTTTATACTTTTTTATAATTACCATATCATC
ATAGTGAAGTATATAATATATATGATATAAGCTCAATATAGTATATTAATTCGGTTAAAC
ACAAAGACATATCAGTTTGTAGCTTTGGTGGATAAAACAAATTAATTTAGCAATTCATGGC
TATGAAAAATGTATATTTTTATTTAAAAATTTTAAAGAAAGCTAAATGATCAAATTTATTTA
ATGATGAATTATATGATAGACACTTTATATAAGAAAAACTTCAACAGCAACAAATTTAAAA
TTTTTTCATCATTTTTCTAGGTGGACTTAAATCATATTTACGGTGAAACTCTGGCTAGACA
GCGTAAACTGCGCCTTTTTCAAGGATGGAAAAATGAAATATCAGGTATGCTTCCCTTTGACT
ATTAAGACTTAGTTATTACCGCTTATACCCATATTTTTAAAAATCCCTAAAAATGTGTTCCCT
TAACTTTTTAACTGATGTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTT
TATCCTCCCACTGATCAAGATACTCAGGCAGAGATGATCTACCTCCTCAAGTCCCTGAG
CATCTACGGTTTGTCTGTGGGGCAGGAGGTCTTTGGTCTGGTGCCTGGTCTGATGATGTAT
GCCACAATCTGGCTGCGGGAACACAACAGAGTATGCGATGTGCTTAAACAGGAGCATCCT
GAATGGGGTGATGAGCAGTTGTTCCAGACAAGCAGGCTAATACTGATAGGTAAACAAGAA
AATGATTTATATAAAACCTCTTCCCAAGGAAAAATAGTGTGCTATCTTTGTTATGTTT
TGAGTAAATGACAAGATGTGGTAAATGAAAACTCACACATTCATATACATTAAATATGT
AAGCATGACTGATAAAATAGCTATCTTTTGATACTGACAAGGAAGAAAAACAGAAATGAAG
GAATAGCAAATTTTAAAAATTGCATTCCAGTTGCTTGAAAGCTTGTGATCAGATGCAATA
AATGTTTTTATTATTTATTTTGTGCAAATAGGAGAGACTATTAAGATTGTGATTGAAGAT
TATGTGCAACACTTGAGTGGCTATCACTTCAAACCTGAAATTTGACCCAGAACTACTTTTC
AACAAACAATTCAGTACCAAAATCGTATTGCTGCTGAATTTAACACCCCTCTATCACTGG
CATCCCTTCTGCTGACACCTTTCAAATTCATGACCAGAAATACAACATCAACAGTTT
ATCTACAACAACCTCTATATTGCTGGAACATGGAATTACCCAGTTTGTGTAATCATTACC
AGGCAAATTGCTGGCAGGGTAAGCATTATTATTGAAAAACAAAAACAAAAAGACTAGTCAGT
AACTTTTAGAATTTCTGCCACGGAAATTATTTTTCTTAAACTTACTAAAAGAGTAGTTAGT
TATATTGCTAGTAAAATTATTTTATTGATATAAGAAAGCCTAACTTTGTTTGAAAAAGTCTA
AACTTTTAGTCTAGTCTACAGTTGTGTCAGACAAATAGCAAATTTGTACCCCTACCTTAAAAA
TATTTTCAAAAAGTATCTATAATCTTATAGGAATAAATATTTTAGGCTTGAATACTAGTG
TTATTTTTGAAATGTAAAAAGGCAAATTAGTTCTAGGCTGGTGTCCCATTGAATTTTAAG
CAGAGCTCCTGTTGAAATGTAGGTAAGCATCTTTCAGCAAATAAAAAATTGTCTCCGCTG
GGAGTTTCAGTTTTACCTGATTTGTACCTAAGGCAAGCTGAATACAAACAGTAAATATGC
CTAAAATTCTTGTTTTACAACATAATTTTACTTTCCACAGGTTGCTGGTGGTAGGAATGTT
CCACCCGCAGTACGAAAGTATCACAGGCTTCCATTGACCAGAGCAGGCAGATGAAATAC
CAGTCTTTTAAATGAGTACCGCAAACGCTTTATGCTGAAGCCCTATGAATCATTTGAAGAA
CTTACAGGTAAGAAACAGTTTCTAAACTTCTTCGTTTTTTGTTTGTGTTGTTGTTGTTTGT
TGTTTTTTGGTTTTCTTTTCGAGATGGAGCCGCCCTCTGTCACCCAGGCTGGAGTGCAGTG
GCGCCATCTCGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCAATTCCTCTGCCTCAA
CTTCTGAGTAGCTGGGACTACAGGCTCACGTCGCACGCATGGATAATTTTTTGTTATTTT
CAGTATAGACGGGTTTTACCGTGTTAGCCAGGCTGGTCTCAAACCTCTGACCTAGTGAT
CCGCCGGCTTCGGCCTCCGAAAGTCTGGGATTACAGGCGTGAGCCACCGCGCTGGCCCC
CTAAACTTCTTAAAAGAATCAGGGGTCAAATGGAAACAGAGAAGTTGGCAGCAAATTGAG
CAAAAGAATCAAACCTGTTTTTTATTTTGTGAAGTTTGACATTGGTTGTATCTCTGTCTTC
ATCGCCTTCACAGGAGAAAAGGAAATGTCTGCAGAGTTGGAAGCACTCTATGGTGACATC
GATGCTGTGGAGCTGTATCCTGCCCTTCTGGTAGAAAAAGCCTCGGCCAGATGCCATCTTT
GGTGAATACCATTGTTAGTAAAGTTGGAGCACCATTCTCCTTGAAAGGACTTATGGGTAATGTT
ATATGTTCTCCTCGCTACTGGAAGCCAAGCACTTTTGGTGGAGAAGTGGGTTTTTCAAATC
ATCAACACTGCCTCAATTCAGTCTCTCATCTGCAATAACGTGAAGGGCTGTCCCTTTACT
TCATTAGTGTTCAGATCCAGAGCTCATTAACAGTCACCATCAATGCAAGTTCTTCC
CGCTCCGACTAGATGATATCAATCCCACAGTACTACTAAAAGAAGCTTCGACTGAACTG
TAGAAGTCTAATGATCATATTTATTTATTTATATGAACCATGTCTATTAATTTAATTATT
TAATAATATTTATATTAACCTCCTTATGTTACTTAAACATCTTCTGTAAACAGAAAGTCAGTA
TCCTGTGTGCGGAGAAAGGAGTCATACTTGTGAAGACTTTTATGTCACTACTCTAAAGAT
TTTGCTGTTGCTGTAAAGTTTGGAAAAACAGTTTTTTATTTCTGTTTTTATAAAACAGAGAGAA
ATGAGTTTTGACGTCTTTTTACTTGAATTTCAACTTATATTATAAGAACGAAAGTAAAGA
TGTTTTGAATACTTAAACACTATCACAAAGATGGCAAAATGCTGAAAGTTTTTACACTGTCTG
ATGTTTCCAATGCATCTTCCATGATGCATTAGAAGTAACTAATGTTTGAAATTTTAAAGT
ACTTTTGGTTATTTTCTGTCTATCAACAAAAACAGGTATCAGTGCATTATTAATGAAT
ATTTAAATTAGACATTACCAGTAATTTCTATGTCTACTTTTTTAAAAATCAGCAATGAAACAA
TAATTTGAAATTTCTAAATTCATAGGGTAGAATCACCTGTAAAAAGCTTGTTTGATTTCTT
AAAGTTATTAACTTGTACATATAACAAAAAGAAGCTGTCTTGGATTAAATCTGTAAAA
TCAGATGAAATTTTACTACAATTGCTTGTGTTAAAAATATTTTATAAGTGATGTTCCCTTTTTC
ACCAAGAGTATAAACCTTTTTTAGTGTGACTGTTAAAACTTCTTTTTAAATCAAAATGCCA
AATTTATTAAGGTGGTGGAGCCACTGCAGTGTTATCTCAAAATAAGAATATTTTGTGAG
ATATTCAGAAATTTGTTTATATGGCTGGTAACATGTAAAAATCTATATCAGCAAAAGGGTC

TACCTTTTAAAAATAAGCAATAACAAAGAAGAAAAACCAAATTATTGTTCAAATTTAGGTTTA
AACTTTTGAAGCAAACCTTTTTTTTATCCTTGTGCACTGCAGGCCTGGTACTCAGATTTTGG
CTATGAGGTTAATGAAGTACCAAGCTGTGCTTGAATAACGATATGTTTTCTCAGATTTTTC
TGTTGTACAGTTTAAATTTAGCAGTCCATATCACATTGCAAAAAGTAGCAATGACCTCATAA
AATACCTCTTCAAATGCTTAAATTCATTTACACATTAATTTTATCTCAGTCTTGAAGC
CAATTCAGTAGGTGCATTGGAATCAAGCCTGGCTACCTGCATGCTGTTCCCTTTTCTTTTC
TTCTTTTAGCCATTTTGCTAAGAGACACAGTCTTCTCATCACTTCGTTTTCTCCTATTTTG
TTTTACTAGTTTTTAAAGATCAGAGTTCACCTTCTTTGGACTCTGCCTATATTTTCTTACCT
GAACCTTTTGCAAGTTTTTCAGGTAAACCTCAGCTCAGGACTGCTATTTAGCTCCTCTTAAAG
AAGATTAAAAGAGAAAAAAAAGGCCCTTTTAAAAATAGTATACACTTATTTTAAAGTGAA
AAGCAGAGAATTTTATTTATAGCTAATTTTAGCTATCTGTAACCAAGATGGATGCAAAGA
GGCTAGTGCCTCAGAGAGAACTGTACGGGGTTTGTGACTGGAAAAAGTTACGTTCCCAT
CTAATTAATGCCCTTTCTTATTTAAAAACAAAACCAAATGATATCTAAGTAGTTCTCAGC
AATAATAATAATGACGATAATACTTCTTTCCACATCTCATTGTCACTGACATTTAATGG
TACTGTATATTACTTAAATTTATTGAAGATTATTATTTATGTCTTATTAGGACACTATGGT
TATAAACTGTGTTTAAAGCCTACAATCATTGATTTTTTTTTTGTATGTCACAATCAGTATA
TTTTCTTTGGGGTTACCTCTCTGAATATTATGTAAACAATCCAAAGAAATGATTGTATTA
AGATTTGTGAATAAAATTTTAGAAATCTGATTGGCATATTGAGATATTTAAGGTTGAATG
TTTGTCCTTAGGATAGGCCTATGTGCTAGCCACAAAGAATATTGTCTCATTAGCCTGAA
TGTGCCATAAGACTGACCTTTTAAATGTTTTGAGGGATCTGTGGATGCTTCGTTAATTT
GTTTCAGCCACAATTTATTGAGAAAATATTCTGTGTCAAGCACTGTGGGTTTTAATATTTT
TAAATCAAACGCTGATTACAGATAATAGTATTTATATAAAATAATTGAAAAAAATTTTCTT
TTGGGAAGAGGGAGAAAATGAAATAAATATCATTAAAGATAACTCAGGAGAATCTTCTTT
ACAATTTTACGTTTAGAATGTTTAAAGGTTAAGAAAAGAAATAGTCAATATGCTTGTATAAA
ACACTGTTCACTGTTTTTTTTTAAAAAAAACCTTGATTTGTTATTAACATTGATCTGCTG
ACAAAACCTGGGAATTTGGGTTGTGTATGCGAATGTTTCAGTGCCTCAGACAAATGTGTA
TTTTAACTTATGTAAAAGATAAGTCTGGAAATAAATGTCTGTTTATTTTTGTACTATTTAA
AAAAAAAATCGATGTGCTGACTCGAGTC

A2. Predicted Genes

>gene1
ATGCTCGCCCGCGCCCTGCTGCTGTGCGCGGTCCTGGCGCTCAGCCATACAGCAAATCCT
TGCTGTTCCACCCATGTCAAAACCGAGGTGTATGTATGAGTGTGGGATTTGACCAGTAT
AAGTGCATTTGTACCCGGACAGGATTCTATGGAGAAAACCTGCTCAACACCGGAATTTTTG
ACAAGAATAAAATTTATTTCTGAAACCCACTCCAAACACAGTGCCTACATACTTACCCAC
TTCAAGGGATTTTGAACGTTGTGAATAACATTCCCTTCCCTTCGAAATGCAATTATGAGT
TATGTCTTGACATCCAGATCACATTTGATTGACAGTCCACCAACTTACAATGCTGACTAT
GGCTACAAAAGCTGGGAAGCCTTCTCTAACCTCTCCTATTATACTAGAGCCCTTCTCCT
GTGCCTGATGATTGCCCCACTCCCTTGGGTGTCAAAGGTAAAAAGCAGCTTCTCTGATTCA
AATGAGATTGTGGAATAATTTGCTTCTAAGAAGAAAGTTTATCCCTGATCCCAGGGCTCA
AACATGATGTTTGCATTCTTTGCCAGCACTTCACGCATCAGTTTTTCAAGACAGATCAT
AAGCGAGGGCCAGCTTTACCAACGGGCTGGGCCATGGGGTGGACTTAAATCATATTTAC
GGTGAAACTCTGGCTAGACAGCGTAAACTGCGCCTTTTCAAGGATGGAATAATGAAATAT
CAGATAAATTGATGGAGAGATGTATCCTCCACAGTCAAAGATACTCAGGCAGAGATGATC
TACCCTCCTCAAGTCCCTGAGCATCTACGGTTTGTCTGTGGGGCAGGAGGTCTTTGGTCTG
GTGCCTGGTCTGATGATGTATGCCACAATCTGGCTGCGGGAACACAACAGAGTATGCGAT
GTGCTTAAACAGGAGCATCCTGAATGGGGTGATGAGCAGTTGTTCCAGACAAGCAGGCTA
ATACTGATAGGAGAGACTATTAAGATTGTGATTGAAGATTATGTGCAACACTTGAGTGGC
TATCACTTCAAACCTGAAATTTGACCCAGAACTACTTTTCAACAAACAATTCAGTACCAA
AATCGTATTGCTGCTGAATTTAACACCTCTATCACTGGCATCCCCCTTCTGCCTGACACC
TTTTCAATTCATGACCAGAAATACAACATATCAACAGTTTATCTACAACAACCTCTATATTG
CTGGAACATGGAATTACCCAGTTTGTGTAATCATTACCAGGCAAATTGCTGGCAGGGTT
GCTGGTGGTAGGAATGTTCCACCCGAGTACAGAAAAGTATCACAGGCTTCCATTGACCAG
AGCAGGCAGATGAAATACCAGTCTTTTAAATGAGTACCGCAAACGCTTTATGCTGAAGCCC
TATGAATCATTTGAAGAACTTACAGGAGAAAAGGAAATGTCTGCAGAGTTGGAAGCACTC
TATGGTGACATCGATGCTGTGGAGCTGTATCCTGCCCCTTCTGGTAGAAAAGCCTCGGCCA
GATGCCATCTTTGGTGAAACCATGGTAGAAGTTGGAGCACCATTCTCCTTGAAAGGACTT
ATGGGTAAATGTTATATGTTCTCCTGCCTACTGGAAGCCAAGCACTTTTGGTGGAGAAGTG
GGTTTTCAAATCATCAACACTGCCTCAATTCAGTCTCTCATCTGCAATAACGTGAAGGGC
TGTCCCTTTACTTCATTTCAGTGTTCAGATCCAGAGCTCATTTAAACAGTCACCATCAAT
GCAAGTTCTTCCCGCTCCGGACTAGATGATATCAATCCCACAGTACTACTAAAAGAACGT
TCGACTGAACTGTAG

A3. Protein Sequences

>gene1

MLARALLLCAVLALSHTANPCCSHPCQNRGVCMSVGFDQYKCDCTRTRTFYGENCSTPEFL
TRIKLFLKPTPNTVHYILTHFKGFWNVNNIPFLRNAIMSYVLTSRSHLIDSPPTYNADY
GYKSWEAFSNLSYYTRALPPVPDDCPTPLGVKGKKQLPDSNEIVEKLLLRRKFIPDPQGS
NMMFAFFAQHFTHQFFKTDHKRGPAFTNGLGHGVDLNHIYGETLARQRKLRLFKDGKMKY
QIIDGEMYPPTVKDTQAEMIYPPQVPEHLRFVAVGQEVFGLVPGLMMYATIWLRHNRVCD
VLKQEHPEWGDEQLFQTSRLILIGETIKIVIEDYVQHLSGYHFKLKFDPELLFNKQFQYQ
NRIAAEFNTLYHWHPLLPDTFQIHDQKYNYYQFIYNNSILLEHGITQFVESFTRQIAGRV
AGGRNVPPAVQKVSQASIDQSRQMKYQSFNEYRKRFMLKPYESFEELTGEKEMSAELEAL
YGDIDAVELYPALLVEKPRPDAIFGETMVEVGAPFSLKGLMGNVICSPAYWKPSTFGGEV
GFQIINTASIQSLICNNVKGCPFTSFVSPDPELIKTVTINASSSRSGLDDINPTVLLKER
STEL