

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 that achieve a specific goal in our bodies. These actions involve molecules working together in a precise, controlled way.

Appendix - Sequence Contents

A1. Input Sequence

```
>input_sequence
GAGCTCACATTAACATTTTACAGGGTAACTGCTTAGGACCAGTATTATGAGGAGAATTTA
CCTTTCCCGCCTCTCTTTTCCAAGAAACAAGGAGGGGGTGAAGGTACGGAGAACAGTATTT
CTTCTGTTGAAAGCAACTTAGCTACAAAGATAAAATTACAGCTATGTACACTGAAGGTAGC
TATTTTCATTCCACAAAATAAGAGTTTTTTTTAAAAAGCTATGTATGTATGTGCTGCATATAG
AGCAGATATACAGCCTATTAAGCGTTCGTCCTACTAAAACATAAAACATGTCAGCCTTTCTTA
ACCTTACTCGCCCCAGTCTGTCCCGACGTGACTTCCTCGACCTCTAAAAGACGTACAGAC
CAGACACGGCGGGCGGGCGGGGAGAGGGGATTCCCTGCGCCCCCGGACCTCAGGGCCGCT
CAGATTCCTGGAGAGGAAGCCAAGTGTCTTCTGCCCCCCCCCGGTATCCCATCCAAAGGC
GATCAGTCCAGAACTGGCTCTCGGAAGCGCTCGGGCAAAGACTGCGAAGAAGAAAAAGACA
TCTGGCGGAAACCTGTGCGCCTGGGGCGGTGGAACCTCGGGGAGGAGAGGGAGGGATCAGA
CAGGAGAGTGGGGACTACCCCCCTCTGCTCCCAAATTGGGGCAGCTTCCTGGGTTTCCGAT
TTTCTCATTTCCGTGGGTAAAAAACCTGCCCCCACC GGCTTACGCAATTTTTTTAAGG
GGAGAGGAGGGAATAATTTGTGGGGGGGTACGAAAAGGCGGAAAGAAACAGTCATTAC
ATGGGCTTGGTTTTTCAGTCTTATAAAAAAGGAAGTTCTCTCGGTTAGCGACCAATTGTCA
TACGACTTGCAGTGAGCGTCAGGAGCACGTCCAGGAACCTCCTCAGCAGCGCCTCCTTCAG
CTCCACAGCCAGACGCCCTCAGACAGCAAAGCCTACCCCCGCGCCGCGCCCTGCCCCCGG
CTCGGATGCTCGCCCCGCGCCCTGCTGCTGTGCGCGGTCTCGGCGCTCAGCCATACAGGTG
AGTACCTGGCGCCGCGCACCCGGGACTCCGGTTCACGCACCCGGGCAGAGTTTCCGCTC
TGACCTCCTGGGTCTATCCAGTACTCCGACTTCTCTCCGAATAGAGAAGCTACGTGACT
TGGGAAAGAGCTTGGACCGCTAGAGTCCGAAAGAACTCCGTGGATATTCCAGCTTTCCCA
CAAGCACTGATCATTATGAGCCAGTTACTTAAACCGATCTGAGACACTCTCACCTCCTAAA
TAGGGATAGATGATACTAATTTGCAGGTTGTCAATTATGATAAGACAGGATCTGATCAATA
TATGTGAATTGTTTATATTTGGAACCTTTTTATTGAGTGGAAAGAAGTTGTTTTAAATATT
CTAGTCAGTTCTTTCTGCTCCCAGGAAAGCCCGGATTATGTTTTAAGATAAGCAAAATG
TCTTAAAAGTAAGCTGTTTTACTTTGAATTTTTCCCTAAATGTTGATTAGTGTACTAGAT
CCATTTTTAATTTGGAAAGTGAAGTGCTACTTATTTGAACCTTCTTAAAAATGCTAATTTTA
ACATCTAAAGAGTTAACTAAGAAAAGCTTAGTAACATGATGTACCAAGTTGAATATGCTG
TTATCCTTATTTAGAATAGAAAATTGGTATTTCTACGTTTTATCCATTCTAAGGCAGGTT
AAAAAATTGATTTCCATGACTACCTATATATTTCTTGAATTTATTATTGTAAAGTTGAT
TCATAGTCAAACAATTAATGTTTAAATTAAGATTAAGACACTAGAGAATGATTTATTTG
CTGTCTTTAATTGCAGCAAATCCTTGCTGTTCCACCCATGTCAAAACCGAGGTGTATG
TATGAGTGTGGGATTTGACCAGTATAAGTGCGATTGTACCCGGACAGGATTCATATGGAGA
AAACTGCTCAACACGTAAGTTTGTCTTTGGTTGCCTCATTAGGAGTGGGGCTGGATACA
GTTATCATTGTATAGATTTGTGTCTTATAATGAGTCCCATTAATTTCTCCCTCCCTTTCT
TCGTCTTCTTGACGCGGAATTTTTGACAAGAATAAAATTATTTCTGAAACCCACTCCAAA
CACAGTGCCTACATACTTACCCACTTCAAGGGATTTTGGAAACGTTGTGAATAACATTCC
CTTCCTTCGAAATGCAATTATGAGTTATGTCTTGACATGTAAGTACAAGTGTCTTTCTAA
GGTTTTTTAGCCTTCTCAAAGAAAAATATGCTTTTATAATACTGTAAGCCTAATCTAAAAAC
ATATTTCCAAGCTTATCAAAAAGACTTTAAGATAGCTTTTAAAGTTTGCCTTCCATCTTAA
TCGCCAAAAATATTGACATTTAGTCCCATCCAGTTTATACAGTCTGCTCACAACCTCTGTA
TACCTCTTCTAACCTTTACTGTTTGGTCAGTTTGTGGAGGTAGCATGGTCCAGCTGTTTA
TTGAATGCCCATGGGCCACAGAATTGTTCTGAACATGTAGCACCCATTAAAAATAAATTTG
GATTTGGATCAGCAAGAAAAATACTTTCCATGATTCTAAAGTGGGTGCCATACTCAGCCA
TTCTTTTTCATAGGCCTCTGGATAGTGAGCAGATGGCTACCTGAAAAATCAATATTGCCA
GATTATAATGTGCAGAGTATATGTATTTTATTAAAGATGTATTTCAAGTGGCCATTAGAC
TATAAAGTGTAGTTGTTTAAAAATAGATTTTTTTTTATTTTGGAGTTACATTCAACCTCAG
GTGCCACTTTCCACATTTTACAATAAAAAATAATGGTTGATTTACTTAACAAATGAGAATA
AATAAAACATTTTTTTCTTTGAAAAATTCAGCCAGATCACATTTGATTGACAGTCCACCA
ACTTACAATGCTGACTATGGCTACAAAAGCTGGGAAGCCTTCTCTAACCTCTCCTATTAT
ACTAGAGCCCTTCTCCTGTGCTGATGATTGCCCGACTCCCTTGGGTGTCAAAGGTGAG
TAAGAAGAATCCATTAGAGATGTATTAACATAAGACGGGCTGCATTGTGCCAAAAAAA
AAAATTGACCTTAGACTACCATTTATTTATTAACAAAAGCAGTTTTTACTTTTAGCATGG
TTATCTATGGGTATTTTTTAAAGTATGAGTCTATATAAACTATTATGTAAAAGCAAATGA
GCGTCTTGGTATAATGTCTTAATATTTTCAAATTATTTCTTTAGAAATGAAATAATTTCTA
ATTAATAATGATAAAATCATTTCAGTAAGAAGTTGTTCCACCATACTTAGAACTGTTGTT
TATATTATGATCCTAATTCACAATTGTAATTCATATAAATGAAGAATTCTTGGTAGATT
GACAGTCACCATCTCCTTTCTTGAATACATAGATGGATTCTTACCTTAGCTTTCTCATTTT
TTCAGGTAAAAAGCAGCTTCTGATTCAAATGAGATTGTGGAAAAATTGCTTCTAAGAAG
AAAGTTTCATCCCTGATCCCCAGGGCTCAAACATGATGTTGCATTCTTTGCCCAGCACTT
CACGCATCAGTTTTTCAAGACAGATCATAAGCGAGGGCCAGCTTTACCAACGGGCTGGG
CCATGGGGTAAAGATAGAGTTAATATCTTAGAGTTAGTAAAAATTATACCAATCATAGTCA
AGGGCTAACATTAAAGGATATACAGTAGATAGATCCAAATAAATTTATCCACTTTTTTT
```

TAAAAAGAAGTCTTATCTATAAAAAACCTTAAAGGAATTTTCCATTTACTTCACTGGTCTA
GTAAAATTATACACACACACAGACATGCACACACATATATAAACATTCACACACATACAT
ATGTACAGGTATTGTTATTTGTAATTTGACCTTGTATTTTTTAGTTTAAAAATGTTAGTA
CTGCAAAATGTTATGTCCTCAAAAACACATTGTACCATGATTATGCCGCTTTCAATATTG
TAAAGTGAGGTTTTTGCCGCATTATTATTTTTTGGATTTCATAGCATAGCTTCAAGTTA
TTCGTAAGAATTTTTTATAAATAATACATTTTTTATACTTTTTTATAATTACCATATCATC
ATAGTGAAGTATATAATATATATGATATAAGCTCAATATAGTATATTAATTCGGTTAAAC
ACAAAGACATATCAGTTTGTAGCTTTGGTGGATAAAACAAATTAATTTAGCAATTCATGGC
TATGAAAAATGTATATTTTTATTTAAAAATTTTAAAGAAAAGCTAAATGATCAAATTTATTTA
ATGATGAATTATATGATAGACACTTTATATAAGAAAAACTTCAACAGCAACAAATTTAAAA
TTTTTTCATCATTTTTCTAGGTGGACTTAAATCATATTTACGGTGAAACTCTGGCTAGACA
GCGTAAACTGCGCCTTTTTCAAGGATGGAAAAATGAAATATCAGGTATGCTTCCTTTGACT
ATTAAGACTTAGTTATTACCGCTTATACCCATATTTTTAAAAATCCCTAAAAATGTGTTCCCT
TAACTTTTTAACTGATGTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTT
TATCCTCCCACTGATCAAGATACTCAGGCAGAGATGATCTACCTCCTCAAGTCCCTGAG
CATCTACGGTTTGCTGTGGGGCAGGAGGTCTTTGGTCTGGTGCCTGGTCTGATGATGTAT
GCCACAATCTGGCTGCGGGAACACAACAGAGTATGCGATGTGCTTAAACAGGAGCATCCT
GAATGGGGTGATGAGCAGTTGTTCCAGACAAGCAGGCTAATACTGATAGGTAAACAAGAA
AATGATTTATATAAAACCTCTTCCCAAGGAAAAATAGTGTGCTATCTTTGTTATGTTT
TGAGTAAATGACAAGATGTGGTAAATGAAAACTCACACATTCTATATACATTAAATATGT
AAGCATGACTGATAAAATAGCTATCTTTTGATACTGACAAGGAAGAAAAACAGAAATGAAG
GAATAGCAAATTTTAAAAATTGCATTCCAGTTGCTTGAAAGCTTGTGATCAGATGCAATA
AATGTTTTTATTATTTATTTTGTGCAAATAGGAGAGACTATTAAGATTGTGATTGAAGAT
TATGTGCAACACTTGAGTGGCTATCACTTCAAACCTGAAATTTGACCCAGAACTACTTTTC
AACAAACAATTCAGTACCAAAATCGTATTGCTGCTGAATTTAACACCTCTATCACTGG
CATCCCTTCTGCTGACACCTTTCAAATTCATGACCAGAAATACAACATCAACAGTTT
ATCTACAACAACCTCTATATTGCTGGAACATGGAATTACCCAGTTTGTGTAATCATTACC
AGGCAAATTGCTGGCAGGGTAAGCATTATTATTGAAAAACAAAAACAAAAAGACTAGTCAGT
AACTTTTAGAATTTCTGCCACGGAAATTATTTTTCTTAAACTTACTAAAAGAGTAGTTAGT
TATATTGCTAGTAAAATTATTTTATTGATATAAGAAAGCCTAACTTTGTTTGAAAAGTCTA
AACTTTTAGTCTAGTCTACAGTTGTGAGACAAATAGCAAATTTGTACCCCTACCTTAAAAA
TATTTTCAAAAAGTATCTATAATCTTATAGGAATAAATATTTTAGGCTTGAATACTAGTG
TTATTTTTGAAATGTAAAAAGGCAAATTAGTTCTAGGCTGGTGTCCCATTGAATTTTAAG
CAGAGCTCCTGTTGAAATGTAGGTAAGCATCTTTCAGCAAATAAAAAATTGTCTCCGCTG
GGAGTTTCAGTTTTACCTGATTTGTACCTAAGGCAAGCTGAATACAAACAGTAAATATGC
CTAAAATTCTTGTTTTACAACATAATTTTACTTTCCACAGGTTGCTGGTGGTAGGAATGTT
CCACCCGCAGTACAGAAAGTATCACAGGCTTCCATTGACCAGAGCAGGCAGATGAAATAC
CAGTCTTTTAAATGAGTACCGCAAACGCTTTATGCTGAAGCCCTATGAATCATTTGAAGAA
CTTACAGGTAAGAAACAGTTTCTAAACTTCTTCGTTTTTTGTTTTGTTTTGTTTTGTTTTGT
TGTTTTTTGGTTTTCTTTTCGAGATGGAGCCGCCCTCTGTCACCCAGGCTGGAGTGCAGTG
GCGCCATCTCGGCTCACTGCAACCTCCGCCTCCTGGGTTCAGCAATTTCTCCTGCCTCAA
CTTCCTGAGTAGCTGGGACTACAGGCTCACGTGCGACGCATGGATAATTTTTTGTTATTTT
CAGTATAGACGGGTTTTACCGTGTTAGCCAGGCTGGTCTCAAACCTCTGACCTAGTGAT
CCGCCGGCTTCGGCCTCCCGAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCTGGCCC
CTAAACTTCTTAAAAGAATCAGGGGTCAAATGGAAACAGAGAAGTTGGCAGCAAATTGAG
CAAAAGAATCAAACCTGTTTTTTATTTTGTGAAGTTTGACATTGGTTGTATCTCTGTCTTC
ATCGCCTTCACAGGAGAAAAGGAAATGTCTGCAGAGTTGGAAGCACTCTATGGTGACATC
GATGCTGTGGAGCTGTATCCTGCCCTTCTGGTAGAAAAAGCCTCGGCCAGATGCCATCTTT
GGTGAATACCATTGGTAGAAGTTGGAGCACCATTCTCCTTGAAAGGACTTATGGGTAATGTT
ATATGTTCTCCTCGCTACTGGAAGCCAAGCACTTTTGGTGGAGAAGTGGGTTTTTCAAATC
ATCAACACTGCCTCAATTCAGTCTCTCATCTGCAATAACGTGAAGGGCTGTCCCTTTACT
TCATTCAGTGTTCCAGATCCAGAGCTCATTAACAGTCACCATCAATGCAAGTTCTTCC
CGCTCCGACTAGATGATATCAATCCCACAGTACTACTAAAAGAAGCTTCGACTGAACTG
TAGAAGTCTAATGATCATATTTATTTATTTATATGAACCATGTCTATTAATTTAATTATT
TAATAATATTTATATTAACCTCCTTATGTTACTTAAACATCTTCTGTAAACAGAAAGTCAGTA
TCCTGTGTGCGGAGAAAGGAGTCATACTTGTGAAGACTTTTATGTCACTACTCTAAAGAT
TTTGCTGTTGCTGTAAAGTTTGGAAAAACAGTTTTTTATTTCTGTTTTTATAAAACAGAGAGAA
ATGAGTTTTGACGTCTTTTTACTTGAATTTCAACTTATATTATAAGAACGAAAGTAAAGA
TGTTTTGAATACTTAAACACTATCACAAAGATGGCAAAATGCTGAAAGTTTTTACACTGTCTG
ATGTTTCCAATGCATCTTCCATGATGCATTAGAAGTAACTAATGTTTGAAAATTTTAAAGT
ACTTTTGGTTATTTTTCTGTCTATCAACAAAAACAGGTATCAGTGCATTATTAATGAAT
ATTTAAATTAGACATTACCAGTAATTTCTATGTCTACTTTTTTAAAAATCAGCAATGAAACAA
TAATTTGAAATTTCTAAATTCATAGGGTAGAATCACCTGTAAAAAGCTTGTTTGATTTCTT
AAAGTTATTAACTTGTACATATAACAAAAAGAAGCTGTCTTGGATTAAATCTGTAAAA
TCAGATGAAATTTTACTACAATTGCTTGTAAAAATATTTTATAAGTGATGTTCCTTTTTC
ACCAAGAGTATAAACCTTTTTTAGTGTGACTGTTAAAACTTCTTTTTAAATCAAAATGCCA
AATTTATTAAGGTGGTGGAGCCACTGCAGTGTTATCTCAAAATAAGAATATTTTGTGTGAG
ATATTCAGAAATTTGTTTATATGGCTGGTAACATGTAAAAATCTATATCAGCAAAAGGGTC

TACCTTTTAAAAATAAGCAATAACAAAGAAGAAAAACCAAATTATTGTTCAAATTTAGGTTTAA
AACTTTTGAAGCAAACCTTTTTTTTATCCTTGTGCACTGCAGGCCTGGTACTCAGATTTTGG
CTATGAGGTTAATGAAGTACCAAGCTGTGCTTGAATAACGATATGTTTTCTCAGATTTTTC
TGTTGTACAGTTTAAATTTAGCAGTCCATATCACATTGCAAAAAGTAGCAATGACCTCATAA
AATACCTCTTCAAATGCTTAAATTCATTTACACATTAATTTTATCTCAGTCTTGAAGC
CAATTCAGTAGGTGCATTGGAATCAAGCCTGGCTACCTGCATGCTGTTCCCTTTTCTTTTC
TTCTTTTAGCCATTTTGCTAAGAGACACAGTCTTCTCATCACTTCGTTTCTCCTATTTTGG
TTTTACTAGTTTTTAAAGATCAGAGTTCACCTTCTTTGGACTCTGCCTATATTTTCTTACCT
GAACCTTTTGCAAGTTTTTCAGGTAAACCTCAGCTCAGGACTGCTATTTAGCTCCTCTTAAAG
AAGATTAAAAGAGAAAAAAAAGGCCCTTTTAAAAATAGTATACACTTATTTTAAAGTGAA
AAGCAGAGAATTTTATTTATAGCTAATTTTAGCTATCTGTAACCAAGATGGATGCAAAGA
GGCTAGTGCCTCAGAGAGAACTGTACGGGGTTTGTGACTGGAAAAAGTTACGTTCCCATTT
CTAATTAATGCCCTTTCTTATTTAAAAACAAAACCAAATGATATCTAAGTAGTTCTCAGC
AATAATAATAATGACGATAATACTTCTTTCCACATCTCATTGTCACTGACATTTAATGG
TACTGTATATTACTTAATTTTATTGAAGATTATTATTTATGTCTTATTAGGACACTATGGT
TATAAACTGTGTTTAAAGCCTACAATCATTGATTTTTTTTTTGTATGTCACAATCAGTATA
TTTTCTTTGGGGTTACCTCTCTGAATATTATGTAAACAATCCAAAGAAATGATTGTATTA
AGATTTGTGAATAAAATTTTAGAAATCTGATTGGCATATTGAGATATTTAAGGTTGAATG
TTTGTCCTTAGGATAGGCCTATGTGCTAGCCACAAAGAATATTGTCTCATTAGCCTGAA
TGTGCCATAAGACTGACCTTTTAAATGTTTTGAGGGATCTGTGGATGCTTCGTTAATTT
GTTTCAGCCACAATTTATTGAGAAAATATTCTGTGTCAAGCACTGTGGGTTTTAATATTTT
TAAATCAAACGCTGATTACAGATAATAGTATTTATATAAAATAATTGAAAAAAATTTTCTT
TTGGGAAGAGGGAGAAAATGAAATAAATATCATTAAAGATAACTCAGGAGAATCTTCTTT
ACAATTTTACGTTTAGAATGTTTAAAGGTTAAGAAAAGAAATAGTCAATATGCTTGATATAA
ACACTGTTCACTGTTTTTTTTTAAAAAAAACCTTGATTTGTTATTAACATTGATCTGCTG
ACAAAACCTGGGAATTTGGGTTGTGTATGCGAATGTTTCAGTGCCTCAGACAAATGTGTA
TTTTAACTTATGTAAAAGATAAGTCTGGAATAAATGTCTGTTTATTTTGTACTATTTAA
AAAAAAAATCGATGTGCACTCGAGTC

A2. Predicted Genes

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

A3. Protein Sequences

>gene1

MLARALLLCAVLALSHTANPCCSHPCQNRGVCMSVGFDQYKCDCTRTRGFYGENCSTPEFL
TRIKLFLKPTPNTVHYILTHFKGFWNVNNIPFLRNAIMSYVLTSRSHLIDSPPTYNADY
GYKSWEAFSNLSYYTRALPPVPDDCPTPLGVKGKKQLPDSNEIVEKLLLRRKFIPDPQGS
NMMFAFFAQHFTHQFFKTDHKRGPAFTNGLGHGVDLNHIYGETLARQRKLRLFKDGKMKY
QIIDGEMYPPTVKDTQAEMIYPPQVPEHLRFVAVGQEVFGLVPGLMMYATIWLREHNRVCD
VLKQEHPEWGDEQLFQTSRLILIGETIKIVIEDYVQHLSGYHFKLKFDPELLFNKQFQYQ
NRIAAEFNTLYHWHPLLPDTFQIHDQKYNYYQFIYNNSILLEHGITQFVESFTRQIAGRV
AGGRNVPPAVQKVSQASIDQSRQMKYQSFNEYRKRFMLKPYESFEELTGEKEMSAELEAL
YGDIDAVELYPALIVEKPRPDAIFGETMVEVGAPFSLKGLMGNVICSPAYWKPSTFGGEV
GFQIINTASIQSLICNNVKGCPFTSFVSPDPELIKTVTINASSSRSGLDDINPTVLLKER
STEL