

GeneVision Analysis Report

File name: genevision_report_2025-05-14.pdf

Generation date: May 14, 2025

User: Mariem_Gaaliche

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	A biological process is how a cell carries out a specific task based on its genes. It's a series of steps, done by specific molecules in a certain order, to achieve a biological goal.

Appendix - Sequence Contents

A1. Input Sequence

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

TAAAAAGAAGTCTTATCTATAAAAAACCTTAAAGGAATTTTCCATTTACTTCACTGGTCTA
GTAAAATTATACACACACACAGACATGCACACACATATATAAACATTCACACACATACAT
ATGTACAGGTATTGTTATTTGTAATTTGACCTTGTATTTTTTAGTTTAAAAATGTTAGTA
CTGCAAAATGTTATGTCCTCAAAAACACATTGTACCATGATTATGCCGCTTTCAATATTG
TAAAGTGAGGTTTTTGGCCGATTATTATTTTTTGGATTTCATAGCATAGCTTCAAGTTA
TTCGTAAGAATTTTTTATAAATAATACATTTTTTATACTTTTTTATAATTACCATATCATC
ATAGTGAAGTATATAATATATATGATATAAGCTCAATATAGTATATTAATTCGGTTAAAC
ACAAAGACATATCAGTTTGTAGCTTTGGTGGATAAAACAAATTAATTTAGCAATTCATGGC
TATGAAAAATGTATATTTTTATTTAAAAATTTTAAAGAAAAGCTAAATGATCAAATTTATTTA
ATGATGAATTATATGATAGACACTTTATATAAGAAAAACTTCAACAGCAACAAATTTAAAA
TTTTTTCATCATTTTTCTAGGTGGACTTAAATCATATTTACGGTGAAACTCTGGCTAGACA
GCGTAAACTGCGCCTTTTTCAAGGATGGAAAAATGAAATATCAGGTATGCTTCCTTTGACT
ATTAAGACTTAGTTATTACCGCTTATACCCATATTTTTAAAAATCCCTAAAAATGTGTTCCCT
TAACTTTTTAACTGATGTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTT
TATCCTCCCACTGATCAAGATACTCAGGCAGAGATGATCTACCTCCTCAAGTCCCTGAG
CATCTACGGTTTGTCTGTGGGGCAGGAGGTCTTTGGTCTGGTGCCTGGTCTGATGATGTAT
GCCACAATCTGGCTGCGGGAACACAACAGAGTATGCGATGTGCTTAAACAGGAGCATCCT
GAATGGGGTGATGAGCAGTTGTTCCAGACAAGCAGGCTAATACTGATAGGTAAACAAGAA
AATGATTTATATAAAACCTCTTCCCAAGGAAAAATAGTGTGCTATCTTTGTTATGTTT
TGAGTAAATGACAAGATGTGGTAAATGAAAACTCACACATTCATATACATTAAATATGT
AAGCATGACTGATAAAATAGCTATCTTTTGATACTGACAAGGAAGAAAAACAGAAATGAAG
GAATAGCAAATTTTAAAAATTGCATTCCAGTTGCTTGAAAGCTTGTGATCAGATGCAATA
AATGTTTTTATTATTTATTTTGTGCAAATAGGAGAGACTATTAAGATTGTGATTGAAGAT
TATGTGCAACACTTGAGTGGCTATCACTTCAAACCTGAAATTTGACCCAGAACTACTTTTC
AACAAACAATTCAGTACCAAAATCGTATTGCTGCTGAATTTAACACCTCTATCACTGG
CATCCCTTCTGCTGACACCTTTCAAATTCATGACCAGAAATACAACATCAACAGTTT
ATCTACAACAACCTCTATATTGCTGGAACATGGAATTACCCAGTTTGTGTAATCATTACC
AGGCAAATTGCTGGCAGGGTAAGCATTATTATTGAAAAACAAAAACAAAAAGACTAGTCAGT
AACTTTTAGAATTTCTGCCACGGAAATTATTTTTCTTAAACTTACTAAAAGAGTAGTTAGT
TATATTGCTAGTAAAATTATTTTATTGATATAAGAAAGCCTAACTTTGTTTGAAAAGTCTA
AACTTTTAGTCTAGTCTACAGTTGTGTCAGACAAATAGCAAATTTGTACCCCTACCTTAAAAA
TATTTTCAAAAAGTATCTATAATCTTATAGGAATAAATATTTTAGGCTTGAATACTAGTG
TTATTTTTGAAATGTAAAAAGGCAAATTAGTTCTAGGCTGGTGTCCCATTGAATTTTAAG
CAGAGCTCCTGTTGAAATGTAGGTAAGCATCTTTCAGCAAATAAAAAATTGTCTCCGCTG
GGAGTTTCAGTTTTACCTGATTTGTACCTAAGGCAAGCTGAATACAAACAGTAAATATGC
CTAAAATTCTTGTTTTACAACATAATTTTACTTTCCACAGGTTGCTGGTGGTAGGAATGTT
CCACCCGCAGTACGAAAGTATCACAGGCTTCCATTGACCAGAGCAGGCAGATGAAATAC
CAGTCTTTTAAATGAGTACCGCAAACGCTTTATGCTGAAGCCCTATGAATCATTTGAAGAA
CTTACAGGTAAGAAACAGTTTCTAAACTTCTTCGTTTTTTGTTTGTGTTGTTGTTGTTTGT
TGTTTTTGGTTTTCTTTTCGAGATGGAGCCGCCCTCTGTCACCCAGGCTGGAGTGCAGTG
GCGCCATCTCGGCTCACTGCAACCTCCGCCTCCTGGGTTCAGCAATTCCTCTGCCTCAA
CTTCCTGAGTAGCTGGGACTACAGGCTCACGTGCGACGCATGGATAATTTTTTGTTATTTT
CAGTATAGACGGGTTTTACCGTGTTAGCCAGGCTGGTCTCAAACCTCTGACCTAGTGAT
CCGCCGGCTTCGGCCTCCGAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCTGGCCC
CTAAACTTCTTAAAAGAATCAGGGGTCAAATGGAAAACAGAGAAGTTGGCAGCAAATTGAG
CAAAAGAATCAAACCTGTTTTTTATTTTGTGAAGTTTGACATTGGTTGTATCTCTGTCTTC
ATCGCCTTCACAGGAGAAAAGGAAATGTCTGCAGAGTTGGAAGCACTCTATGGTGACATC
GATGCTGTGGAGCTGTATCCTGCCCTTCTGGTAGAAAAAGCCTCGGCCAGATGCCATCTTT
GGTGAATACCATTGTTAGTAAAGTTGGAGCACCATTCTCCTTGAAAGGACTTATGGGTAAATGTT
ATATGTTCTCCTCGCTACTGGAAGCCAAGCACTTTTGGTGGAGAAGTGGGTTTTTCAAATC
ATCAACACTGCCTCAATTCAGTCTCTCATCTGCAATAACGTGAAGGGCTGTCCCTTTACT
TCATTAGTGTTCCAGATCCAGAGCTCATTAACAGTCACCATCAATGCAAGTTCTTCC
CGCTCCGACTAGATGATATCAATCCCACAGTACTACTAAAAGAAGCTTCGACTGAACTG
TAGAAGTCTAATGATCATATTTATTTATTTATATGAACCATGTCTATTAATTTAATTATT
TAATAATATTTATATTAACCTCCTTATGTTACTTAAACATCTTCTGTAAACAGAAAGTCAGTA
TCCTGTTGCGGAGAAAGGAGTCATACTTGTGAAGACTTTTATGTCACTACTCTAAAGAT
TTTGCTGTTGCTGTAAAGTTTGGAAAAACAGTTTTTTATTTCTGTTTTTATAAAACCAGAGAGAA
ATGAGTTTTGACGTCTTTTTACTTGAATTTCAACTTATATTATAAGAACGAAAGTAAAGA
TGTTTTGAATACTTAAACACTATCACAAGATGGCAAAATGCTGAAAGTTTTTACACTGTCTG
ATGTTTCCAATGCATCTTCCATGATGCATTAGAAGTAACTAATGTTTGAAATTTTAAAGT
ACTTTTGGTTATTTTTCTGTCTATCAACAAAAACAGGTATCAGTGCATTATTAATGAAT
ATTTAAATTAGACATTACCAGTAATTTCTATGTCTACTTTTTTAAAAATCAGCAATGAAACAA
TAATTTGAAATTTCTAAATTCATAGGGTAGAATCACCTGTAAAAAGCTTGTGTTGATTTCTT
AAAGTTATTAACTTGTACATATACCAAAAAGAAGCTGTCTTGGATTAAATCTGTAAAA
TCAGATGAAATTTTACTACAATTGCTTGTGTTAAAAATATTTTATAAGTGATGTTCCTTTTTC
ACCAAGAGTATAAACCTTTTTTAGTGTGACTGTTAAAACTTCTTTTTAAATCAAAATGCCA
AATTTATTAAGGTGGTGGAGCCACTGCAGTGTTATCTCAAAATAAGAATATTTTGTGAG
ATATTCAGAAATTTGTTTATATGGCTGGTAACATGTAAAAATCTATATCAGCAAAAGGGTC

TACCTTTTAAAAATAAGCAATAACAAAGAAGAAAAACCAAATTATTGTTCAAATTTAGGTTTAA
AACTTTTGAAGCAAACCTTTTTTTTATCCTTGTGCACTGCAGGCCTGGTACTCAGATTTTGG
CTATGAGGTTAATGAAGTACCAAGCTGTGCTTGAATAACGATATGTTTTCTCAGATTTTTC
TGTTGTACAGTTTAAATTTAGCAGTCCATATCACATTGCAAAAAGTAGCAATGACCTCATAA
AATACCTCTTCAAATGCTTAAATTCATTTACACATTAATTTTATCTCAGTCTTGAAGC
CAATTCAGTAGGTGCATTGGAATCAAGCCTGGCTACCTGCATGCTGTTCCCTTTTCTTTTC
TTCTTTTAGCCATTTTGCTAAGAGACACAGTCTTCTCATCACTTCGTTTTCTCCTATTTTTG
TTTTACTAGTTTTTAAAGATCAGAGTTCACCTTCTTTGGACTCTGCCTATATTTTCTTACCT
GAACCTTTTGCAAGTTTTTCAGGTAAACCTCAGCTCAGGACTGCTATTTAGCTCCTCTTAAAG
AAGATTAAAAGAGAAAAAAAAGGCCCTTTTAAAAATAGTATACACTTATTTTAAAGTGAA
AAGCAGAGAATTTTATTTATAGCTAATTTTAGCTATCTGTAACCAAGATGGATGCAAAGA
GGCTAGTGCCTCAGAGAGAACTGTACGGGGTTTTGTGACTGGAAAAAGTTACGTTCCCAT
CTAATTAATGCCCTTTCTTATTTAAAAACAAAACCAAATGATATCTAAGTAGTTCTCAGC
AATAATAATAATGACGATAATACTTCTTTCCACATCTCATTGTCACTGACATTTAATGG
TACTGTATATTACTTAATTTTATTGAAGATTATTATTTATGTCTTATTAGGACACTATGGT
TATAAACTGTGTTTAAAGCCTACAATCATTGATTTTTTTTTTGTATGTCACAATCAGTATA
TTTTCTTTGGGGTTACCTCTCTGAATATTATGTAAACAATCCAAAGAAATGATTGTATTA
AGATTTGTGAATAAATTTTTTAGAAATCTGATTGGCATATTGAGATATTTAAGGTTGAATG
TTTGTCCTTAGGATAGGCCTATGTGCTAGCCACAAAGAATATTGTCTCATTAGCCTGAA
TGTGCCATAAGACTGACCTTTTTAAATGTTTTGAGGGATCTGTGGATGCTTCGTTAATTT
GTTTCAGCCACAATTTATTGAGAAAATATTCTGTGTCAAGCACTGTGGGTTTTAATATTTT
TAAATCAAACGCTGATTACAGATAATAGTATTTATATAAAATAATTGAAAAAAATTTTCTT
TTGGGAAGAGGGAGAAAATGAAATAAATATCATTAAAGATAACTCAGGAGAATCTTCTTT
ACAATTTTACGTTTAGAATGTTTAAAGGTTAAGAAAAGAAATAGTCAATATGCTTGATATAA
ACACTGTTCACTGTTTTTTTTTAAAAAAAACCTTGATTTGTTATTAACATTGATCTGCTG
ACAAAACCTGGGAATTTGGGTTGTGTATGCGAATGTTTCAGTGCCTCAGACAAATGTGTA
TTTTAACTTATGTAAAAGATAAGTCTGGAAATAAATGTCTGTTTATTTTTGTACTATTTAA
AAAAAAAATCGATGTGCTGACTCGAGTC

A2. Predicted Genes

>gene1
ATGCTCGCCCGCGCCCTGCTGCTGTGCGCGGTCTGCGGCTCAGCCATACAGCAAATCCT
TGCTGTTCCACCCATGTCAAAACCGAGGTGTATGTATGAGTGTGGGATTTGACCAGTAT
AAGTGCATTTGTACCCGGACAGGATTCTATGGAGAAAACCTGCTCAACACCGGAATTTTTG
ACAAGAATAAAATTTATTTCTGAAACCCACTCCAAACACAGTGCCTACATACTTACCCAC
TTCAAGGGATTTTGAACGTTGTGAATAACATTCCCTTCCCTTCGAAATGCAATTATGAGT
TATGTCTTGACATCCAGATCACATTTGATTGACAGTCCACCAACTTACAATGCTGACTAT
GGCTACAAAAGCTGGGAAGCCTTCTCTAACCTCTCCTATTATATACTAGAGCCCTTCTCCT
GTGCTGATGATTGCCCCGACTCCCTTGGGTGTCAAAGGTAAAAAGCAGCTTCTCTGATTCA
AATGAGATTGTGGAATAATGCTTCTAAGAAGAAAGTTTATCCCTGATCCCCAGGGCTCA
AACATGATGTTTGCATTCTTTGCCCAGCACTTCACGCATCAGTTTTTCAAGACAGATCAT
AAGCGAGGGCCAGCTTTACCAACGGGCTGGGCCATGGGGTGGACTTAAATCATATTTAC
GGTGAAACTCTGGCTAGACAGCGTAAACTGCGCCTTTTCAAGGATGGAATAATGAAATAT
CAGATAAATTGATGGAGAGATGTATCCTCCACAGTCAAAGATACTCAGGCAGAGATGATC
TACCCTCCTCAAGTCCCTGAGCATCTACGGTTTGTCTGTGGGGCAGGAGGTCTTTGGTCTG
GTGCTGCTGATGATGATGTATGCCACAATCTGGCTGCGGGAACACAACAGAGTATGCGAT
GTGCTTAAACAGGAGCATCCTGAATGGGGTGATGAGCAGTTGTTCCAGACAAGCAGGCTA
ATACTGATAGGAGAGACTATTAAGATTGTGATTGAAGATTATGTGCAACACTTGAGTGGC
TATCACTTCAAACCTGAAATTTGACCCAGAACTACTTTTCAACAAACAATTCAGTACCAA
AATCGTATTGCTGCTGAATTTAACACCTCTATCACTGGCATCCCCCTTCTGCCTGACACC
TTTTCAAATTCATGACCAGAAATACAACATATCAACAGTTTATCTACAACAACCTCTATATTG
CTGGAACATGGAATTACCCAGTTTGTGTAATCATTACCAGGCAAATTGCTGGCAGGGTT
GCTGGTGGTAGGAATGTTCCACCCGAGTACAGAAAAGTATCACAGGCTTCCATTGACCAG
AGCAGGCAGATGAAATACCAGTCTTTTAAATGAGTACCGCAAAACGCTTTATGCTGAAGCCC
TATGAATCATTTGAAGAACTTACAGGAGAAAAGGAAATGTCTGCAGAGTTGGAAGCACTC
TATGGTGACATCGATGCTGTGGAGCTGTATCCTGCCCCTTCTGGTAGAAAAGCCTCGGCCA
GATGCCATCTTTGGTGAAACCATGGTAGAAGTTGGAGCACCATTCTCCTTGAAAGGACTT
ATGGGTAAATGTTATATGTTCTCCTGCCTACTGGAAGCCAAGCACTTTTGGTGGAGAAGTG
GGTTTTCAAATCATCAACACTGCCTCAATTCAGTCTCTCATCTGCAATAACGTGAAGGGC
TGTCCCTTTACTTCATTTCAGTGTTCAGATCCAGAGCTCATTTAAACAGTCACCATCAAT
GCAAGTTCTTCCCGCTCCGGACTAGATGATATCAATCCCACAGTACTACTAAAAGAACGT
TCGACTGAACTGTAG

A3. Protein Sequences

>gene1

MLARALLLCAVLALSHTANPCCSHPCQNRGVCMSVGFDQYKCDCTRTRTFYGENCSTPEFL
TRIKLFLKPTPNTVHYILTHFKGFVNVMNIPFLRNAIMSYVLTSRSHLIDSPPTYNADY
GYKSWEAFSNLSYYTRALPPVPDDCPTPLGVKGKKQLPDSNEIVEKLLLRRKFIPDPQGS
NMMFAFFAQHFTHQFFKTDHKRGPAFTNGLGHGVDLNHIYGETLARQRKLRLFKDGKMKY
QIIDGEMYPPTVKDTQAEMIYPPQVPEHLRFAVGQEVFGLVPGLMMYATIWLREHNRVCD
VLKQEHPEWGDEQLFQTSRLILIGETIKIVIEDYVQHLSGYHFKLKFDPELLFNKQFQYQ
NRIAAEFNTLYHWHPLLPDTFQIHDQKYNYYQFIYNNSILLEHGITQFVESFTRQIAGRV
AGGRNVPPAVQKVSQASIDQSRQMKYQSFNEYRKRFMLKPYESFEELTGEKEMSAELEAL
YGDIDAVELYPALLVEKPRPDAIFGETMVEVGAPFSLKGLMGNVICSPAYWKPSTFGGEV
GFQIINTASIQSLICNNVKGCPFTSFVSPDPELIKTVTINASSSRSGLDDINPTVLLKER
STEL