**Biol. 433 Assignment:**

**Obtaining Information from a Cloned Gene**

You have obtained the following sequence information for the gene that you have just cloned:

catcttcaaaaataccctaatcacattttgtaacaataatacaattatacattaaaactctccgacgATGCCTCAGGCACCGATGCCAGAGTTCTCTAGCTCGGTGAAGCTCAAGTACGTGAAACTTGGTTACCAATATTTGGTTAACCATTTCTTGAGTTTTCTTTTGATCCCGATCATGGCTATTGTCGCCGTTGAGCTTCTTCGGATGGGTCCTGAAGAGATCCTTAATGTTTGGAATTCACTCCAGTTTGACCTAGTTCAGGTTCTATGTTCTTCCTTCTTTGTCATCTTCATCTCCACTGTTTACTTCATGTCCAAGCCACGCACCATCTACCTCGTTGACTATTCTTGTTACAAGCCACCTGTCACGTGTCGTGTCCCCTTCGCAACTTTCATGGAACACTCTCGTTTGATCCTCAAGGACAAGCCTAAGAGCGTCGAGTTCCAAATGAGAATCCTTGAACGTTCTGGCCTCGGTGAGGAGACTTGTCTCCCTCCGGCTATTCATTATATTCCTCCCACACCAACCATGGACGCGGCTAGAAGCGAGGCTCAGATGGTTATCTTCGAGGCCATGGACGATCTTTTCAAGAAAACCGGTCTTAAACCTAAAGACGTCGACATCCTTATCGTCAACTGCTCTCTTTTCTCTCCCACACCATCGCTCTCAGCTATGGTCATCAACAAATATAAGCTTAGGAGTAATATCAAGAGCTTCAATCTTTCGGGGATGGGCTGCAGCGCGGGCCTGATCTCAGTTGATCTAGCCCGCGACTTGCTCCAAGTTCATCCCAATTCAAATGCAATCATCGTCAGCACGGAGATCATAACGCCTAATTACTATCAAGGCAACGAGAGAGCCATGTTGTTACCCAATTGTCTCTTCCGCATGGGTGCGGCAGCCATACACATGTCAAACCGCCGGTCTGACCGGTGGCGAGCCAAATACAAGCTTTCCCACCTCGTCCGGACACACCGTGGCGCTGACGACAAGTCTTTCTACTGTGTCTACGAACAGGAAGACAAAGAAGGACACGTTGGCATCAACTTGTCCAAAGATCTCATGGCCATCGCCGGTGAAGCCCTCAAGGCAAACATCACCACAATAGgtaataaacacacatcattttatttcacctacatatatacactaacaatgatatacacggtgaaaactatatcaaggggtgggcattaaatattattggtcggctccgttttgacttcaattcggttttaaataacagaataattactctatctctatatatgttttgttattatggttaaatgtactttcttggtatatgtctcacttctatggaaatttaggttactctattcggttgcattcaattcggtttgttcgatataaattttattatgatttaaagcaaagtttggtttggcttaggtttaaggaattatggtacgtttttggttttgatagatgaatttgtttgaacagGTCCTTTGGTCCTACCGGCGTCAGAACAACTTCTCTTCCTCACGTCCCTAATCGGACGTAAAATCTTCAACCCGAAATGGAAACCATACATACCGGATTTCAAGCTGGCCTTCGAACACTTTTGCATTCACGCAGGAGGCAGAGCGGTGATCGACGAGCTCCAAAAGAATCTACAACTATCAGGAGAACACGTTGAGGCCTCAAGAATGACACTACATCGTTTTGGTAACACGTCATCTTCATCGTTATGGTACGAGCTTAGCTACATCGAGTCTAAAGGGAGAATGAGGAGAGGCGATCGCGTTTGGCAAATCGCGTTTGGGAGTGGTTTCAAGTGTAACTCTGCCGTGTGGAAATGTAACCGTACGATTAAGACACCTAAGGACGGACCATGGTCCGATTGTATCGACCGTTACCCTGTCTTTATTCCCGAAGTTGTCAAACTCTAAactgaaaacgtctttgaacggtttagtaacggtttgattttgtgttacggtttaggtttatttggtctcgggatttggtttaaaggggattgagaaatgggaagttagagagaagaaaaagcaaagcataaatgtttgtatttaattgctctgcatatacttaaatctctgcttttcatttggggtattttttagtttctcgtgctgtaattaataacttgtggtgtactcaaataagaatatttctctctgttttattatttttaaattggagaataaacttctttgttttcatcgctgccttatatatattgcttctgatttgggttgaccatccatccagtttgcaggatccgatagcagt

Using this sequence information and The Arabidopsis Information Resource Website (<http://www.arabidopsis.org/>) as your starting points you are trying to get as much information about this gene. Your research supervisor asked you to try to answer the following questions, and be as specific as possible by using not only predictions, but also **experimental evidence** **whenever it is available**:

1. What is the identity of the gene that you have cloned (Gene name, MIPS designation-eg At5g38340)? Using the tools available find the location of exons, introns, 5’ and 3’ untranslated regions (UTRs), and translational START and STOP codons for this gene and indicate them using the format shown on the next page.
2. What is the predicted/confirmed biochemical function of this gene/protein?
3. Are there any other related sequences in other organisms? Are there any other related sequences in the Arabidopsis genome? Is the biochemical activity known for any of them?
4. When and where is this gene expressed (predicted/confirmed)? Where in the cell is the protein encoded by this gene localized (predicted/confirmed)? For predictions use prediction tools eFP and Genevestigator; experimental evidence if available!
5. What are the phenotypes of the mutant(s) disrupted in this gene? Based on this phenotype what is the biological function of this gene/protein?
6. Are genes with high sequence similarity/identity present in genomes of other plant species? What does this suggest about the function of this protein?

Example of a gene with indicated parts for question 1:

|  |  |  |
| --- | --- | --- |
| ATG,TAA = Translational Start/StopATGC = Exonatgc = Intronatgc = UTR  |   |  |
|  |  |  |

cacctcatatttctctctctctctccatttacc**ATG**ACTTTCAATAAGAGGCAAGTCAAGATCAATCACTGGCCGGAGAAGAACGACAAAGAGAAGCAGAAATACTCCAAAAACAGAGAAACCGTCAAGTTGACACTTCTCACTCTTCTTCTCCTTTGCTCAATTTGTTTCCTCTTCCTAACTCTCAATTTCCCCTTCACAATAGAATTCACAGCCTCAATCCCTCGCACATGCGACCACAATTTCACCGTCTACGTGTACGATCTCCCCAAAGAGTTCAACATCGGTCTCCTCCAAAATTGCCGACACTTAAACATCTACACCAACATGTGCCCTCACGTGGCCAACAATGGCCTTGGACAGCCTCTTCACCGCGGACGAACTTCGTGGTTCTCTACGCACCAGTTCATAGCCGAGATGATCTTCCACGCACGTGTCGAAAACCATCCATGCCGCACGTATGAGCCAGACACTGCCGACATCTTCTACGTCCCTTTCTACGGTGGTCTCTACGCTTCAAGCGTGTTCCGAGAGCAAAACCTGACCAAGCGCGACGAGCTAGCTGTCAGATTAGTCAACTACATTAGTGGCCAACGATGGTGGAAGAGAAGCAACGGTCGTGATCATTTCTTGGCCATAGGGAGGACAGCTTGGGATTTCATGCGCTCCTCCGACACTGACTTTGGAGCTAACATGCTCATGCAAATGCCACGTGTCATGAACATGTCGGTGCTGACCGTGGAGAGGCAGCCTTGGAATGGTGACAATCACTTTGGTATACCGTATCCTTCTTATTTTCATCCGTACACGTCAGCAGAAATGGTGACGTGGCAGGACAAGATGAAAAATGTCGAGAGACCAAACTTGTTTAGTTTTGTCGGTGGGCCGAGAAAGGGCTTGGAGAAAGCTGCCATTAGAGACGAGCTGATCAAGCAATGCGCTGAGTCGAGCCATTGTGAGCTTCTCAAGTGTGAAAATGGAGGGTCCAGgttttttacttttatcatgaatttgacaatttcgcaattaagtattatttaatacactcactagtcaccaaaatttcttgattggaggggcccataatcataaactttatttaatttttattgacaactggcacattgggttatactttattgggtagattacattttttgtggttaactctctctcaacttatttgactaagaggtgtaacaacaattcaactaggagacttggttgttgaactatattttaagcgttttgtgtatccttaactaagaggagtaacaataattcaacttattttttgtcttcatcaggaaatattgtaaccttgattatttttactcaatgtagGTGTCACAACCCAATGACGGTATTAGGAGTGATGGCTAGGTCACGGTTTTGCCTACAAGCACCGGGGGACTCATTCACGCGGAGATCAACATTCGATGCGATGTTAGCCGGGTGCATACCGGTCTTTTTTTCACCCCACACTATGTATACACAGTACATGTGGTATCTTCCGGACGATAAAAGAAGTTACTCAGTGTTCATGGACGAAAAGAACAACACTCATATAGAACAAGAGCTCTTGAGGATCTCGGAGAACGAAGTGGTTCAAATGAGAGAAATAGTCATCGATTTGATCCCGAGACTGACTTATGCGCACCCAAACTCTACTAACTATGATTTACCAGATGCGGTTGATATAGCGTTAGAGGCACTAGCTAAGCAAGCAAGAGACAATGTTGTGGTTTCATTGTAAgacctatagttttgtgctactcagatgtaactaatcccaaagaataaagaattgtagtgttcagacatcaagccaagcaactattgaactatggacattggcaatatgaaacatgcaacttaagacaaatgacagttatagaaacacgcttgtctatggttatgaagtttccctaat