

Name _____ Class Period _____ Date _____

Gene Annotation

Activity Three: Defining Protein Products for a Gene Model

Use the information from the corresponding website to complete the questions on this worksheet.

1. Briefly describe an Open Reading Frame.

2. What is the function of the ORF Finder portal in the Annotation Tool? Describe what it does and how it finds an ORF.

3. What is the first amino acid produced during translation of mRNA to Protein by a ribosome?

4. List two criteria you could use in selecting an ORF for a gene model?

5. Describe the difference between a BLASTp and BLASTn query.

1. The protein coding region of a gene, it is a sequence of codons without any stop codons ideally containing a start codon and possibly a stop codon at the end
2. This portal links to the NCBI website and accesses the data stored there about each reading frame in order to create a list of potential open reading frames
3. Methionine
4. Length, Presence of Start/Stop codons, and the relative size compared to the gene model
5. BLASTp searches protein databases for similar amino acid sequences while BLASTn does the same for nucleotide sequences (DNA and mRNA).
- 6.

7. The Evidence Plot and the Evidence Table
8. The Evidence Plot
9. The entire structure (all of the exons) are added to the (green) gene model
10. First click and it is selected (not the other exons on the same structure) on the second click it is deselected.
11. The coordinates of the exon correspond to nucleotide bases in the genome, basically defining where the exon starts and stops in the sequence a chromosome
12. The Max Sim Score and the number of cDNA and EST structures that contain the exons in question
13. A color block represents a groups of variants for the same exon, so only one should be selected for each gene model
14. Exons