**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **40** |
| **Stop Coordinate** | **27341** |
| **Direction (For/Rev)** | **Rev** |
| **Gap (Overlap) with Previous Gene** | **89 pb gap** |
| **Selected Start Coordinate** | **27739** |
| **Selected Function** | **Hypothetical Protein** |

**Annotation Decision #1: Is this a Gene?**

|  |  |
| --- | --- |
| **Gathering Evidence** | **Explain Your Rationale** |
| Was the gene called by an auto-annotation program (Glimmer, GeneMark)? | *YES BOTH* |
| Is there evidence for coding potential? | *GeneMarkS and GeneMark-M show coding potential.* |
| Is this gene present in other annotated genomes? | *PhagesDB Blast hit: Zayuliv, gene 39, function unknown, 132 Score 252 Evalue 3e-67*  *Phamerator: phage: Zayuliv, gene: 39, pham: 652* |
| Does the gene violate any major guiding principles? | *No* |
| **DECISION:** | *YES* |

**Annotation Decision #2: What is the best possible start site for this gene?**

|  |  |
| --- | --- |
| **Gathering Evidence** | **Explain Your Rationale** |
| What start site do Glimmer and GeneMark suggest? | *Glimmer Start Coordinate: 27739*  *GeneMark Start Coordinate: 27739* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Start site: 27739*  *RBS final score: -2.443*  *Zvalue: 2.995* |
| Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)? | *This start codon is the longest ORF (399bp)* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *The start site 27739 is conserved in other phage genomes indicated by Starterator. This start was found in 199 of 199 (100%) of genes in pham.*  *Called 99.0% of time when present* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | *the best BlastP match from NCBI:*  *QOP65235.1 [Zayuliv]*  *e-value: 2e-82*  *Q#: 1 MAARTRKTA*  *S#: 1 MAARTRKTA* |
| **DECISION:** | *Keep DNA master predicted site, 27739* |

**Annotation Decision #3: What is the Function of the Putative Protein?**

|  |  |
| --- | --- |
| **Gathering Evidence** | **Explain Your Rationale** |
| Does this protein align with a protein having a functional assignment in BlastP (phagesDB and/or GenBank) with an alignment of 10-4 or smaller with appropriate coverage? | *The protein doesn’t align with proteins having a functional assignment in BlastP (NCBI, PhagesDB)*  *BlastP hits:*  *PhagesDB: Zayuliv, function unknown*    *evalue: 3e-67*  *NCBI: hypothetical protein ZAYULIV*    *evalue: 2e-82*  *DNA Master:*  *SEA\_GREENIVY\_39, hypothetical protein*    *Evalue: 4.0 E-18* |
| Does this protein align with a protein having a functional assignment in the PDB or other database in HHPred with a probability of 90% or greater with appropriate coverage? | *No quality match*  *HHpred match: 2LBF\_B (60S acidic ribosomal protein P2; ribosome, stalk, P1/P2, RIBOSOMAL PROTEIN; NMR {Homo sapiens})*  *Probability 80.06*  *Database: PDB* |
| Is this gene located adjacent to genes of known function and in a region of the genome that shows high conservation of gene order? | *This gene is flanked upstream by a tRNA annotated as 'tRNA-OTHER' and downstream by a DNA helicase, forming a syntenic arrangement conserved in related phages.* |
| Is this gene a possible transmembrane protein? | *No, according to DeepTMHMM is a globular protein* |
| Is the proposed function found on the SEA-PHAGES approved function list? | *Hypothetical Protein* |
| **DECISION:** | *Hypothetical Protein* |