**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **42** |
| **Stop Coordinate** | **29252** |
| **Direction (For/Rev)** | **Rev** |
| **Gap (Overlap) with Previous Gene** | **1bp** |
| **Selected Start Coordinate** | **29527** |
| **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: Zepp, gene 41, function unknown, 91 Score 193 Evalue 2e-49*  *Phamerator: phage: Zepp, gene: 41, pham: 7632* |
| 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: 29527*  *GeneMark Start Coordinate: 29527* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Start site 29527*  *RBS final score: -3.920*  *Zvalue: 2.301*  *Gap with previous gene (downstream): 1bp* |
| Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)? | *The length of the predicted site 29527 is 276bp. It is not the longest ORF predicted, however the longest ORF (29656) results in an excessive gene overlap of 129 bp* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *The start site was found in 12 of 12 ( 100.0% ) of genes in pham.*  *Called 91.7% of time when present*  *(29527)* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | *the best BlastP match from NCBI:*  *QIG58129.1 [Zepp]*  *e-value: 7e-60*  *Q#: 1 MYPITSYTVT*  *S#: 1 MYPITSYTVT* |
| **DECISION:** | *Keep DNA Master predicted site 29527* |

**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, DNA master)*  *BlastP hits:*  *PhagesDB: Zepp\_41, function unknown*    *evalue: 2e-49*  *NCBI: hypothetical protein SEA\_ZEPP\_41*    *evalue: 7e-60*  *DNA Master: SEA\_Zayuliv\_41, hypothetical protein*  *Evalue: 3.5 E-38* |
| 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: 3G80\_A (Protein B2; RNA-binding, viral protein, suppressor of RNAi, RNA interference; 2.5A {Nodamura virus})*  *Probability: 70.58.*  *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? | *As of the July 31, 2025 Phamerator database version, this gene is flanked upstream by a DNA helicase and downstream by a membrane protein belonging to Pham 247197, forming a conserved syntenic arrangement observed 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* |