**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **54** |
| **Stop Coordinate** | **35521** |
| **Direction (For/Rev)** | **Rev** |
| **Gap (Overlap) with Previous Gene** | **No** |
| **Selected Start Coordinate** | ***35676*** |
| **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)? | *Only GeneMark* |
| Is there evidence for coding potential? | *GeneMarkS andr GeneMark-M. folio show coding potential.* |
| Is this gene present in other annotated genomes? | *PhagesDB Blast hit: zepp\_53; e-value= 1e-25*  *Phamerator: phage: zepp; gene: 53; pham: 8815 (06-05-25)* |
| 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:NA*  *GeneMark Start Coordinate: 35676* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Start site 35589*  *FInal score: -3,391*  *Z-value: 2,587*  *This is the best score.* |
| Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)? | *It’s not the longest ORF and doesn’t result in gene overlap.* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *The start site is conserve in another phages* [*genomes. It*](http://genomes.it) *is the same predicted by Starterator.* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | ***PhagesDB***  *Best Hit: Zepp\_53, function unknown*  *E-value: 1e-25*    ***BlastP match from NCBI***  *Best Hit:* hypothetical protein SEA\_ZEPP\_53 [Microbacterium phage Zepp]  *E-value: 6e-29*    ***DNA Master***  *Best Hit: hypothetical protein SEA\_ZEPP\_53 [Microbacterium phage Zepp]*  *E-value: 5,8E-29*  *Q#1: MSDGVRDTQW*  *S#1: MSDGVRDTQW* |
| **DECISION:** | *The gene should start at 35676* |

**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? | *List the most informative BlastP match from each source PhagesDB: Zepp\_53, function unknown*  *NCBI:* hypothetical protein SEA\_ZEPP\_53 [Microbacterium phage Zepp]  *DNA Master: hypothetical protein SEA\_ZEPP\_53 [Microbacterium phage Zepp]* |
| 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? | *8X24\_D AcrIF25; type I-F CRISPR-Cas system, anti-CRISPR (Acr) proteins, VIRAL PROTEIN; 2.834A {Acinetobacter baumannii}*  *Probability: 70.2%,*  *E-value: 12* |
| Is this gene located adjacent to genes of known function and in a region of the genome that shows high conservation of gene order? | *According to the July 30, 2025 Phamerator database version, Pham 8815 is flanked upstream by Pham 4548 and downstream by an aGPT-Pplase1 domain-containing protein, forming a conserved syntenic arrangement among related phages* |
| Is this gene a possible transmembrane protein? | *NO* |
| Is the proposed function found on the SEA-PHAGES approved function list? | *NO* |
| **DECISION:** | *Hypothetical Protein* |