**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **52** |
| **Stop Coordinate** | **35158** |
| **Direction (For/Rev)** | **Rev** |
| **Gap (Overlap) with Previous Gene** | **No** |
| **Selected Start Coordinate** | ***35337*** |
| **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 andr GeneMark-M. folio show coding potential.* |
| Is this gene present in other annotated genomes? | *PhagesDB Blast hit: Zepp\_51, e-value=1e-24*  *Librie\_51, E-value=1e-24*  *Hasitha\_51, e-value= 1e-24*  *Fulton\_51, e-value=1e-24*  *Phamerator: phage: zepp, Librie, Hashita, Fulton*  *Similarities found with gen 51; pham: 7069* (5-6-25) |
| 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 and GeneMark Start Coordinate: 35337* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Start site 35373*  *Final score: -3,020*  *Z-value: 2,761*  *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 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\_51, function unknown*  *E-value: 1e-24*    ***BlastP match from NCBI***  *Best Hit:* hypothetical protein SEA\_ZEPP\_51 [Microbacterium phage Zepp]  *E-value: 5e-32*    ***DNA Master***  *Best Hit: hypothetical protein SEA\_ZEPP\_51 [Microbacterium phage Zepp]*  *E-value: 3,8E-24*  *Q#1: MATITLDQQI*  *S#1: MATITLDQQI* |
| **DECISION:** | *The gene should start at 35337* |

**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\_51, function unknown*  *NCBI:* hypothetical protein SEA\_ZEPP\_51 *[Microbacterium phage Zepp]*  *DNA Master: hypothetical protein SEA\_ZEPP\_51 [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? | *3SHG\_B VbhA; AMPylation, adenylylation, toxin-antitoxin complex, Fic fold, AMP transfer, TRANSFERASE-PROTEIN BINDING complex; HET: TLA; 1.5A {Bartonella schoenbuchensis R1}*  *Probability: 61.3%,*  *E-value: 31* |
| 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, this gene is flanked upstream by a thymidylate synthase and downstream by a gene from Pham 4548, 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* |