**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **23** |
| **Stop Coordinate** | **16654bp** |
| **Direction (For/Rev)** | **For** |
| **Gap (Overlap) with Previous Gene** | **Overlap 3bp** |
| **Selected Start Coordinate** | **15887bp** |
| **Selected Function** | **Minor tail 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, Glimmer and GeneMark call it gene.* |
| Is there evidence for coding potential? | *GeneMark-host show coding potencial, but GeneMarkS does not show promising results.* |
| Is this gene present in other annotated genomes? | *PhagesDB blast hit Librie\_21 with e value= e-42*  *Phamerator phage: pham 7779 (25/4/25)* |
| Does the gene violate any major guiding principles? | *No, the gene does not violate any key guiding principles* |
| **DECISION:** | *Yes, it is a gene.* |

**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: 15887*  *GeneMark Start Coordinate: 15887* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *final RBS score: -3.509*  *Z-score:2.920*  *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)? | *Yes* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *Yes, is the start number called the most often in the published annotations. It was called in 12 of the 12 non-draft genes in the pham.* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | *DNA Master: UVK59179 (Librie)*  *E-value: 0*  *Q#1: MTIQDIRDIT*  *S#1: MTIQDIRDIT* |
| **DECISION:** | *Keep DNA Master pre start site 15887* |

**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? | *PhagesDB: Librie\_21, minor tail protein*  *e-value: e-142*  *NCBI: minor tail protein [Microbacterium phage Librie]*  *Sequence ID: UVK59179.1*  *e-value: 0.0*  *DNA Master: minor tail protein [Microbacterium phage Librie]*  *Score: 1255* |
| 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? | *Yes,*  *Name: Spike (CD1369); Phage tail-like, bacteriocin, baseplate, pre-contraction, VIRUS LIKE PARTICLE;{Clostridioides difficile}*  *Probability: 97.3*  *e-value: 0.00095* |
| Is this gene located adjacent to genes of known function and in a region of the genome that shows high conservation of gene order? | *YES, it is between “minor tail protein” and “Phame:235269 (103) (May 16, 2025)”* |
| Is this gene a possible transmembrane protein? | *No, TMHMM result does not show posible transmembrane subunit.* |
| Is the proposed function found on the SEA-PHAGES approved function list? | *Yes, “minor tail protein”* |
| **DECISION:** | *“minor tail protein”* |