**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **22** |
| **Stop Coordinate** | **15890bp** |
| **Direction (For/Rev)** | **For** |
| **Gap (Overlap) with Previous Gene** | **Overlap 1bp** |
| **Selected Start Coordinate** | **14718bp** |
| **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? | *GeneMarkS and GeneMark-host show coding potencial.* |
| Is this gene present in other annotated genomes? | *PhagesDB blast hit Quadzero\_20 with e value= 0*  *Phamerator phage: pham 87190 (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:8203*  *GeneMark Start Coordinate:8203* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *final RBS score: -3.739*  *Z-score: 2.452*  *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)? | *No, in gene 22 the start codon of the ORF is number 2.*  *Yes, the longest ORF results in an overlap with gene 21 that exceeds 30 bp* |
| 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 15 non-draft genes in the pham.* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | *DNA Master: QWY84465 (QuadZero)*  *E-value: 0*  *Q#1: MPEPIIPIRN*  *S#1: MPEPIIPIRN* |
| **DECISION:** | *Keep DNA Master pre start site 14718* |

**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: LilTerminator\_Draft\_21, function unknown*  *e-value:0.0*  *NCBI: minor tail protein [Microbacterium phage QuadZero]*  *Sequence ID: QWY84465.1*  *e-value: 0.0*  *DNA Master: minor tail protein [Microbacterium phage QuadZero]*  *Score:2033* |
| 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:Prophage MuSo2, 43 kDa tail protein; MuSo2, Shewanella oneidensis MR-1, Structural Genomics, PSI-2, Protein Structure In*  *Probability: 99.95*  *e-value: 9.4e-24* |
| 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:7779 (17) (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”* |