**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **WardWill** |
| **Gene #** | **13** |
| **Stop Coordinate** | **7856** |
| **Direction (For/Rev)** | **For** |
| **Gap (Overlap) with Previous Gene** | **Overlap 3pb** |
| **Selected Start Coordinate** | **7512** |
| **Selected Function** | **NKF** |

**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. folio show coding potential.* |
| Is this gene present in other annotated genomes? | *PhagesDB Blast hit: Zepp\_13 with e-value=1e-58*  *Phamerator: phage: Zepp; gene: 13; pham: 166474 (05/05/2025)* |
| 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: 7512*  *GeneMark Start Coordinate: 7512* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Final RBS score: -2,584*  *Z-score: 2.995*  *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? | *The start number called is the most often in the published annotations.*  *This start number was called in 180 of 182 non-draft genes in the pham.*  *Called 100% of time when present* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | *DNA Master:*  *Best hit: QIG58102 (Zepp)*  *E-value: 0*  *Q#1 MSITLGDLAS*  *S#1 MSITLGDLAS* |
| **DECISION:** | *Keep DNA Master predicted start site* ***7512*** |

**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: Zepp\_13, function unknown E value 1e-58*  *NCBI: Zepp\_13, hypothetical protein E value 1e-76*  *DNA Master: Zepp\_13, hypothetical protein E value 0.0* |
| 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* |
| 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 Phamerator, this gene is flanked upstream by a hypothetical protein belonging to a pham family conserved among related phages and a tail terminator.* |
| Is this gene a possible transmembrane protein? | *NO* |
| Is the proposed function found on the SEA-PHAGES approved function list? | *NO* |
| **DECISION:** | *NKF* |