**Student Gene Annotation Worksheet**

| **Basic Phage Information** | |
| --- | --- |
| **Phage Name** | **Wardwill** |
| **Gene #** | **12** |
| **Stop Coordinate** | **7515** |
| **Direction (For/Rev)** | **For** |
| **Gap (Overlap) with Previous Gene** | **Overlap 28pb** |
| **Selected Start Coordinate** | **7105** |
| **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: QuadZero\_11 with e-value = 5e-70*  *Phamerator; phage: QuadZero; gene: 11; pham 632 (05/05/25)* |
| 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:7105*  *GeneMark Start Coordinate:7105* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Final RBS score: -3.812*  *Z-score: 2.418*  *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 is not. The longest ORF results in excessive gene overlap (202 bp)* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *The start number called is not the most often in the published annotations.*  *This start number was called in 48 of 182 non-draft genes in the pham.*  *Called 76.4% of time when present* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | *DNA Master:*  *Best hit: QOP65208 (Zayuliv)*  *E-value: 0*  *Q#1 MVMSTRTSLE*  *S#1 MVMSTRTSLE* |
| **DECISION:** | *Keep DNA Master predicted start site* ***7105*** |

**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: Zayuliv\_11, function unknown 5e-70*  *NCBI: Zayuliv\_11, function unknown 4e-91* |
| 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? | *No Synteny Observed.* |
| Is this gene a possible transmembrane protein? | *NO* |
| Is the proposed function found on the SEA-PHAGES approved function list? | *NO* |
| **DECISION:** | *NKF* |