**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **35** |
| **Stop Coordinate** | **22393** |
| **Direction (For/Rev)** | **Rev** |
| **Gap (Overlap) with Previous Gene** | **No** |
| **Selected Start Coordinate** | **23253** |
| **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? | *Blast hit: phage Zayuliv\_33 E-value: 0*  *Phamerator: phage Zayuliv (EA5), gene 33, pham 231101 (198).* |
| 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: 23253*  *GeneMark Start Coordinate: 23253* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Final RBS Score: -2.507*  *Zvalue: 3.130*  *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)? | *This is the longest ORF predicted (861 bp)* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *The start site is conserved in other phage genomes indicated by Starterator. This start was found in 193 of 201 ( 96.0% ) of genes in pham.*  *Called in 98.4% of time when present.* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | *DNA master:*  *Best hit: QWY84478 (Quadzero)*  *Evalue: 0*  *Q#: 1 MGKKITLDFS*  *S#:1 MGKKITLDFS* |
| **DECISION:** | *Keep DNA master predicted start site: 23253* |

**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:*  *QuadZero\_33, function unknown, 286*  *Evalue: 1e-163*    *NCBI and DNA master show the same result* |
| 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? | *HHPred match: 8S4T\_B “PrgE; SSB, DNA BINDING PROTEIN; HET: PGE; 2.67A {Enterococcus faecalis}”*  *Probability: 99.76.*  *Evalue: 2.7e-1.*  *Score: 141.22* |
| 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 Actino\_Draft 600, in the ZAYULIV phage, the most closely related phage, the adjacent genes are a nuclease upstream and an AAA-ATPase downstream.* |
| Is this gene a possible transmembrane protein? | *No. According to SOSUI this gene corresponds to a SOLUBLE protein* |
| Is the proposed function found on the SEA-PHAGES approved function list? | *Yes* |
| **DECISION:** | *Hypothetical Protein* |