**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **36** |
| **Stop Coordinate** | **23285** |
| **Direction (For/Rev)** | **Rev** |
| **Gap (Overlap) with Previous Gene** | **3 bp** |
| **Selected Start Coordinate** | **23980** |
| **Selected Function** | **ASCE ATPase** |

**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\_34 E-value: 0*  *Phamerator: phage Zayuliv (EA5), gene 34, pham 210844(199).* |
| 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: 23980*  *GeneMark Start Coordinate: 23980* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Final RBS score: -5.650*  *Zscore: 1.525*  *It's not the best score, the best is for the start coordinate: 24112 (Final RBS score: -4.108).* |
| Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)? | *Predicted ORF length: 696.*  *Is not the longest ORF, longest ORF length: 844.*  *The longest ORF has an excessive gene overlap (150 bp).* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *The start match the consensus start 20 (23980 bp), found in 28 of 194 ( 14.4% ) of genes in pham with 25 manual annotations* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | *DNA master:*  *Best hit: Q0P65231 (Zayuliv)*  *Evalue: 0*  *Q#: 1 MSVITIYGRP*  *S#: 1 MSVITIYGRP* |
| **DECISION:** | *Keep DNA master predicted site at 23980* |

**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? | *Phages DB hit: Zayuliv\_34, AAA-ATPase, 231 Evalue: 1e-128*    *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: 8RC5\_4I, “ORF016; Annealase, SSAP, Single Strand Annealing, Single Strand Binding, Recombineering, Recombination, SaPI, Bacterioph”*  *Probability: 99.88.*  *Evalue: 4e-20*  *Score: 152.74* |
| Is this gene located adjacent to genes of known function and in a region of the genome that shows high conservation of gene order? | *This gene is flanked upstream by a gene from Pham 246632 and downstream by an exonuclease, forming a conserved syntenic arrangement observed in related phages.* |
| Is this gene a possible transmembrane protein? | *No. According to SOSUI it’s a SOLUBLE protein* |
| Is the proposed function found on the SEA-PHAGES approved function list? | *YES* |
| **DECISION:** | *ASCE ATPase (Official function list doesn’t suggest AAA ATPase in this case)* |