**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **56** |
| **Stop Coordinate** | **36536** |
| **Direction (For/Rev)** | **Rev** |
| **Gap (Overlap) with Previous Gene** | **Yes, with gene 55 (3 pb)** |
| **Selected Start Coordinate** | **36787** |
| **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 (different start site)* |
| Is there evidence for coding potential? | *GeneMarkS andr GeneMark-M. folio show coding potential.* |
| Is this gene present in other annotated genomes? | PhagesDB Blast hit: Zayuliv\_55; e-value=4e-22  Phamerator: phage: Zayuliv; gene: 55; pham: 4886 (06-05-25) |
| 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:36724*  *GeneMark Start Coordinate: 36787* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Start site 36724*  *FInal score: -3,958*  *Z-value: 2,282*  *Start site 36787*  *FInal score: -3,368*  *Z- value: 2,589*  *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)? | *Start site 36724: It’s not the longest ORF and doesn’t result in gene overlap.*  *Start site 36787: It’s not the longest ORF and doesn’t result in gene overlap.* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *There are not as well conserved indicated by Starterator* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | ***PhagesDB***  *Best Hit: TwoBits\_Draft\_52, function unknown*  *E-value: 4e-23*    ***BlastP match from NCBI***  *Best Hit:* hypothetical protein SEA\_ZAYULIV\_55 [Microbacterium phage Zayuliv]  *E-value: 3e-26*    ***DNA Master***  *Best Hit: hypothetical protein SEA\_ZAYULIV\_55 [Microbacterium phage Zayuliv]*  *E-value: 2,7E-26*  *Q#1: EEFRISEEEW*  *S#1: EEFRITREDW* |
| **DECISION:** | *The best posible start is 36787* |

**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: TwoBits\_Draft\_52, function unknown*  *NCBI:* hypothetical protein SEA\_ZAYULIV\_55 [Microbacterium phage Zayuliv]  *DNA Master: hypothetical protein SEA\_ZAYULIV\_55 [Microbacterium phage Zayuliv]* |
| 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? | *5OHG\_J Ribonuclease E; enolase, RNase E, Lyase; HET: PO4; 1.997A {Escherichia coli (strain K12)}*  *Probability: 75.5%,*  *E-value: 3.4* |
| 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 the July 30, 2025 Phamerator database version, Pham 4886 is flanked upstream by an aGPT-Pplase1 domain-containing protein and downstream by Pham 197878, forming a conserved syntenic arrangement among related phages* |
| Is this gene a possible transmembrane protein? | *NO* |
| Is the proposed function found on the SEA-PHAGES approved function list? | *NO* |
| **DECISION:** | *Hypothetical Protein* |