**Student Gene Annotation Worksheet**

|  |  |
| --- | --- |
| **Basic Phage Information** | |
| **Phage Name** | **Wardwill** |
| **Gene #** | **53** |
| **Stop Coordinate** | **35347** |
| **Direction (For/Rev)** | **Rev** |
| **Gap (Overlap) with Previous Gene** | **No** |
| **Selected Start Coordinate** | **35511** |
| **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? | *PhagesDB Blast hit: Librie\_52, e-value= 1e-24*  *Hasitha\_52, e-value= 1e-24*  *Fulton\_52, e-value= 1e-24*  *Phamerator: phage: Librie, Hashita, Fulton*  *Similarities found with gene 52 pham: 4548 (5-6-25)* |
| 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 and GeneMark start: 35511* |
| Does the start site have an associated Ribosome Binding Site with a high score? | *Start site 35511*  *FInal score: -6,440*  *Z-value: 1,610*  *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’s the longest ORF and doesn’t result in gene overlap.* |
| Is this start site conserved in other phage genomes as indicated by Starterator? | *The start site is conserve in another phages* [*genomes.It*](http://genomes.it) *is the same predicted by Starterator.* |
| Is this start site conserved in other phage genomes as indicated by BlastP? | ***PhagesDB***  *Best Hit: Librie\_52, function unknown*  *E-value: 1e-24*  *Q#1: MNTAANITKWAILNYKGQQVEVLAVDSDPAADVWVKLLSDPDKSFFVPRSALTA 54*  *MNTAANITKWAILNYKGQQVEVLAVDSDPAADVWVKLLSDPDKSFFVPRSALTA*  *S#1: MNTAANITKWAILNYKGQQVEVLAVDSDPAADVWVKLLSDPDKSFFVPRSALTA 54*  ***BlastP match from NCBI***  *Best Hit: hypothetical protein SEA\_HASITHA\_52 [Microbacterium phage Hasitha]*  *E-value: 6e-30*  *Q#1: MNTAANITKWAILNYKGQQVEVLAVDSDPAADVWVKLLSDPDKSFFVPRSALTA 54*  *MNTAANITKWAILNYKGQQVEVLAVDSDPAADVWVKLLSDPDKSFFVPRSALTA*  *S#1: MNTAANITKWAILNYKGQQVEVLAVDSDPAADVWVKLLSDPDKSFFVPRSALTA 54*  ***DNA Master***  *Best Hit: hypothetical protein SEA\_HASITHA\_52 [Microbacterium phage Hasitha]*  *E-value: 6,5E-30*  *Q#1: MNTAANITKW*  *S#1: MNTAANITKW* |
| **DECISION:** | *This gene start 35511* |

**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: Librie\_52, function unknown*  *NCBI: hypothetical protein SEA\_HASITHA\_52 [Microbacterium phage Hasitha]*  *DNA Master: hypothetical protein SEA\_HASITHA\_52 [Microbacterium phage Hasitha]* |
| 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? | *2HEQ\_A YorP protein; SH3-like, BSU2030, YorP, NESG, Structural Genomics, PSI-2, Protein Structure Initiative, Northeast Structural Genomics Consortium, UNKNOWN FUNCTION; NMR {Bacillus subtilis} SCOP: b.34.20.1, l.1.1.1*  *Probability: 76.09%, E-value: 9.2* |
| 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 4548 is flanked upstream by Pham 7069 and downstream by Pham 8815, 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* |