

Student Gene Annotation Worksheet

Name: Zion Dewerff

Date: 2/21/21

Basic Phage Information	
Phage Name	DesireeRose
Gene #	1
Stop Coordinate	538
Direction (For/Rev)	Forward
Gap (Overlap) with Previous Gene	NA
Selected Start Coordinate	308
Selected Function	unknown

Annotation Decision #1: Is this a Gene?

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program (Glimmer, GeneMark)?	<i>YES, Glimmer only</i>
Is there evidence for coding potential?	
Is this gene present in other annotated genomes?	<i>There is only one related, annotated phage which is the DesireeRose_Draft, protein #2, e-value of 2e-39. This has the same Pham of 7375 and in the same cluster GE.</i>
Does the gene violate any major guiding principles?	<i>There is a significant overlap with the gene after it because the gap is over 30 bp. and these genes are also in different directions. The gene is long enough as it is over 150 bp and 200 bp.</i>
DECISION:	<i>Respond here with YES or NO after reviewing the evidence gathered above.</i>

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?	<i>Glimmer Start Coordinate: 308 GeneMark Start Coordinate: NA</i>
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	<i>The proposed start site is not the longest and does not have and overlap with the nearest upstream gene.</i>
Is this start site conserved in other phage genomes as indicated by Starterator?	<i>Starterator was not informative and did not have a report.</i>
DECISION:	<i>I think the gene should start at 308 considering that is the only recommendation. GeneMark did not have a start coordinate and Starterator did not have have a report either.</i>

1)DUPLICATE worksheet (table 1 and 2) for each gene you are assigned

2)Add the information to PECAAN

Basic Phage Information	
Phage Name	DesireeRose
Gene #	2
Stop Coordinate	121
Direction (For/Rev)	Reverse
Gap (Overlap) with Previous Gene	161
Selected Start Coordinate	336
Selected Function	unknown

Annotation Decision #1: Is this a Gene?

Gathering Evidence	Explain Your Rationale
Was the gene called by an auto-annotation program (Glimmer, GeneMark)?	<i>YES BOTH</i>
Is there evidence for coding potential?	
Is this gene present in other annotated genomes?	<i>The best match for related, annotated phage is MuffinTheCat, gene protein #1, with an e-value of 3e-38. This gene was in the same cluster GE and same Pham number of 7589.</i>
Does the gene violate any major guiding principles?	<i>This gene did have a significant overlap over 30 bp. It was long enough but was still on the low side of length being at 216. The gene before was a different direction but the gene after was the same, both being reverse.</i>
DECISION:	<i>Respond here with YES or NO after reviewing the evidence gathered above.</i>

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?	<p><i>Glimmer Start Coordinate: 336</i> <i>GeneMark Start Coordinate: 336</i></p>
Is the predicted start codon the longest ORF? If not, does the longest ORF result in excessive gene overlap (>30bp)?	<p><i>The predicted start codon was the longest ORF with a length of 216.</i></p>
Is this start site conserved in other phage genomes as indicated by Starterator?	<p><i>You will also need to provide the following information from Starterator: does the start match the consensus start site predicted from Starterator? If no, is the consensus start site not found in this ORF? If no, is there a better option for the consensus start site instead of the one predicted by Starterator? If Starterator doesn't reveal a consensus start site, you can record that Starterator was not informative.</i></p> <p><i>For the start match from the starterator number 4 one being 336 and one being 337.</i></p>
DECISION:	<p><i>I would start at 336 since both the glimmer and GeneMark start predictions match at 336 and one of the Starterator's were at 336 as well.</i></p>