Genome Annotation Submission Cover Sheet

Preliminary Annotation Review Checklist 5-15-2018

Phage Name: Shelob

Your Name: Ann Powell

Your Institution: University of Evansville

Your email: ap96@evansville.edu

Additional emails: (For correspondence) js383@evansville.edu, am621@evansville.edu

Please check each box indicating completion of each task. If you are not sure how to do something, please see the Online Bioinformatics manual page "How to Pass Preliminary Review".

- ☑ 1. Does the genome sequence in your final contain the same number of bases and is it the same as the posted sequence on phagesdb.org?
- ☑ 2. Are all the genes "valid" when you click the "validate" button?
- ☑ 3. Have the genes been renumbered such that they go sequentially from 1 to the highest number?
- ☑ 4. Have all old BLAST hits been cleared, and all gene features reBLASTed?
- ☑ 5. Are the locus tags the"SEA_ PHAGENAME"?
- ☑ 6. Has the Documentation been recreated to match the information in the feature table?
- ☑ 7. Have tRNA ends been adjusted with web-based Aragorn and/or tRNAscan SE?
- ☑ 8. Has the frameshift in the tail assembly chaperone been annotated (where applicable?)

9. For the items below, generate a genome profile, and review the following. For the

YourPhageName_CompleteNotes.dnam5 file:

- ☑ a. Have any duplicate genes (or any with the same stop coordinate?) been removed?
- ☑ b. Does every gene have **one and only one** complete set of Notes
- ☑ c. Do the functions in the Notes match the official function list?
- ☑ d. Are all three lines of functional evidence described for EVERY gene?
- ☑ e. Do the notes contain the initial Glimmer/GeneMark data from the autoannotation?

For the YourPhageName .dnam5 file:

- ☑ a. Have any duplicate genes (or any with the same stop coordinate?) been removed?
- ☑ b. Is the Notes field empty (including hidden marks?)
- ☑ c. Do the function names in the Product field either match the official function list or say "Hypothetical Protein"?
- ☑ d. Is the Function field empty (including hidden marks?)

☑ 10. Did you use PECAAN to annotate your phage?

If, so please describe how in the text field after question 11.

11. Describe any issues or specific genes that you were unable to satisfactorily resolve, and warrant further inspection in the Quality Control review.

At gene 71 (Stop 23102), the best start without regard to the previous gene is at 22713. However, this leads to a 34 base pair overlap. We have included both genes, but wonder if the previous gene should be removed. Gene 100 (Stop 52551) has a longer start that would include all coding potential, but it has no one-to-one blast matches, while the later start has one match. We believe that the first part of the previous gene may have been lost in this phage. We believe gene 108 (Stop 58125) to be a membrane protein related to iron permeases found in mycobacterium. Gene 115 (Stop 60796) to be a capsid decoration protein. It is found in the corresponding location in Phox, but the only other evidence available was found on the Phagesdb BLAST. We are unsure of gene 155 (Stop 95354) as it has a small number of BLAST matches. The query cover for gene 161 (Stop 103136) is only 50%. The second half of the gene appears to be unique to our phage. We are unsure of the start for gene 198 (Stop 129186). The earlier stop gives a four base pair overlap, but the second start is the second of two identical start sequences. Both choices have matches that are one-to-one. We feel strongly that gene 215 (Stop 137359) is a nicotinamide ribosyltransferase, but the only approved name is a nicotinate rybosyltransferase.
PECAAN was used for most of the primary annotation. The first half of the genome was completed before we