Genome Annotation Submission Cover Sheet

Preliminary Annotation Review Checklist 5-15-2018

Phage Name: Sonali

Your Name: Tamarah Adair

Your Institution: Baylor University

Your email: tamarah_adair@baylor.edu

Additional emails:
(For correspondence)

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.

Aragorn and tRNA SCAN do not agree. The Aragorn call is trimmed and we chose to report this result.
We did not find the frameshift or the tail assembly chaperone.
We used PECAAN and exported the files from there to make the complete notes. We checked all the evidence, but if nothing was found for SYN, HHPRED, or CDD it leaves a blank instead of indicating NKF. This should be interpreted as NKF or lacking evidence. PECAAN also exported the SIF-BLAST whether or not the evalues were acceptable or whether or not we check the box. We have reviewed all the evidence and made our best calls on all 3 lines of evidence. We erred on the conservative side and did not call genes without acceptable evidence. There were 5 genes that were deleted (10571, 12569, 47482P