Genome Annotation Submission Cover Sheet

Preliminary Annotation Review Checklist 5-15-2018

Phage Name: Emma1919

Your Name: Joyce Stamm

Your Institution: University of Evansville

Your email: js383@evansville.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".

- [x] 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?
- [x] 2. Are all the genes “valid” when you click the “validate” button?
- [x] 3. Have the genes been renumbered such that they go sequentially from 1 to the highest number?
- [x] 4. Have all old BLAST hits been cleared, and all gene features reBLASTed?
- [x] 5. Are the locus tags the"SEA_ PHAGENAME"?
- [x] 6. Has the Documentation been recreated to match the information in the feature table?
- [x] 7. Have tRNA ends been adjusted with web-based Aragorn and/or tRNAscan SE?
- [x] 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:

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

For the YourPhageName .dnam5 file:

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

- [x] 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.

- PECAAN was used to organize the initial annotation and to help in functional annotation. Results were exported to DNA master for final quality control checks.
- Not sure about the function for gp20 (5752-5354) - can this be a DNA binding protein?
- gp113: The chosen start at 72822 was manually called 100% of the time, but the start at 72813 gives a -4 gap.
- gp165: The chosen start at 86362 was most commonly called in other BK1 phage, but the start at 86311 gives a -4 gap.
- gp181, used start at 93942 which was the most annotated start that was present; the Glimmer called start gives a large gap.
- gp231 (111185-111670) is possibly a peptidase based on BLAST results.
- gp 260 (122901-123620) could be a kinase based on multiple HHPred hits, but there are no similar genes labeled as kinases.