Genome Annotation Submission Cover Sheet

Preliminary Annotation Review Checklist 5-15-2018

Phage Name:	Naji
Your Name:	Alexandra Huang-Queiroz
Your Institution:	College of William and Mary
Your email:	aahuangqueiroz@email.wm.ed
Additional emails: (For correspondence)	mssaha@wm.edu, mhfors@wm.edu, spaudel@email.wm.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".

- [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?)

- [ ] 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.

BLAST results for gene that starts from 14205 to 14540 are matching with Q1 to S14, it is possible for an extension but extension creates a big gap
Genes (SSC: 41301: 41065 REV, to 41454, SSC: 41850-41464 REV to 41880) can be extended in order to obtain LO, however there is no BLAST match in that extended region and RBS is not the best score either.
Although, there is no consistency in start site, gene (SSC: 45602-45333 REV) that potentially codes for immunity repressor, can be extended to obtain better BLAST hits but produces a huge overlap. Also,it does not match with L5 immunity repressor.