Genome Annotation Submission Cover Sheet

Preliminary Annotation Review Checklist 5-15-2018

Phage Name: Sixama

Your Name: Marie Fogarty

Your Institution: Durham Technical Community College

Your email: fogartym@durhamtech.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?
- [ ] 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.

Sixama is a DS cluster phage. There was a lack of comparative data for this annotation. Sixama shows ~ 63% gene similarity to two DS phages - Forza and Boopy found at Pitt. Forza and Boopy were annotated this spring and have just been published in NCBI. Although their BLAST data is not yet available through DNA Master or PECAAN (as of May 17), I have referred to similar annotated functions (available on Phagesdb) when applicable. PECAAN, alongside with DNA master, was used for all aspects of annotation and the 'new SEA format' annotation notes were imported to DNA Master. 29 tRNAs were annotated for Sixama using Aragorn and tRNAscan-SE. Two orphams called by Glimmer only (stop 86539 and stop 87997), with no Genemark coding potential were deleted in favor of the tRNAs they overlapped tRNA-15 (infernal score 58.4) and tRNA-24 (infernal score 70.7). There are three