Genome Annotation Submission Cover Sheet

Preliminary Annotation Review Checklist 4-4-2018

Phage Name: Ali17
Your Name: Roy Coomans
Your Institution: North Carolina A&T State University
Your email: coomansr@ncat.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 for all the features with no known function (including hidden marks?
   - ☑ c. Do the function names in the Notes match the official function list?
   - ☑ d. Is the function field EMPTY for all features?

- ☑ 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 by the students during annotation but all information was entered into DNA Master. The PECAAN information is not necessarily updated and correct. Genes that caused concern when determining start site or function have a note entered at the end of the Notes field in the Complete Notes file. Gene 3 we were uncertain whether to list DNA methylase as the function. It is a near perfect match for the last 128 amino acids of Phinally gene 5, identified as DNA methylase, but the first 380 amino acids of the Phinally gene are not present, so good E-value but low coverage of the Phinally gene by the much shorter Ali17 gene. We were conservative and put it as NKF rather than list the function.