

Genome Annotation Submission Cover Sheet

Pre-QC Phage Genome Annotation Checklist

Phage Name: Cobra
Your Name: Deborah Tobiason
Your Institution: Carthage College
Your email: dtobiason@carthage.edu
Additional emails:
(For correspondence)

Please check each box indicating completion of each task. Annotation Guide section #'s indicated

- 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? *Section 9.3.2*
- 3. Have the genes been renumbered such that they go sequentially from 1 to the highest number? *Section 9.3.3*
- 4. Have all old BLAST hits been cleared, and all gene features reBLASTed? *Section 9.3.4*
- 5. Are the locus tags the phage name? *Section 9.3.3*
- 6. Has the Documentation been recreated to match the information in the feature table? *Section 1.4*
- 7. Have tRNA ends been adjusted with web-based Aragorn and/or tRNAscan SE? *Section 9.5.3-4*

8. For the items below, generate a genome profile, and review the following. *Section 11.3*

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 (see fig 12.2 in the Annotation Guide)?
- c. Do the functions in the Notes match the official function list?
- d. Is the function field EMPTY for all features?
- 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?
- c. Do the function names in the Notes match the official function list, when applicable?
- d. Is the function field EMPTY for all features?

9. Describe any issues or specific genes that you were unable to satisfactorily resolve, and warrant further inspection in the Quality Control review.

I am not sure what the last .dnam5 file listed is referring to as we followed the instructions in the manual and sent in the 'complete notes' and 'final' files.

The following genes may warrant closer inspection:

- > Gene 16 has strong hit with small terminase but terminase small subunit already identified at beginning of genome so function not labeled in 'final' file
- > Genes 33, 41 and 42 come up with a strong hit to a structural protein in another phage but that seemed too vague to call as a function in the final file
- > Gene 64 came up with a hit to RDF protein from phage Nacho with BlastP but it is not clear what an RDF protein is and there was no other support so did not give as function on