Genome Annotation Submission Cover Sheet

Pre-QC Phage Genome Annotation Checklist

Phage Name: Nanodon

Your Name: Steven Caruso

Your Institution: UMBC

Your email: scaruso@umbc.edu

Additional emails:
(For correspondence)

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

- [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? *Section 9.3.2*
- [x] 3. Have the genes been renumbered such that they go sequentially from 1 to the highest number? *Section 9.3.3*
- [x] 4. Have all old BLAST hits been cleared, and all gene features reBLASTed? *Section 9.3.4*
- [x] 5. Are the locus tags the phage name? *Section 9.3.3*
- [x] 6. Has the Documentation been recreated to match the information in the feature table? *Section 1.4*
- [x] 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:

- [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 (see fig 12.2 in the Annotation Guide)?
- [x] c. Do the functions in the Notes match the official function list?
- [x] d. Is the function field EMPTY for all features?
- [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 for all the features with no known function?
- [x] c. Do the function names in the Notes match the official function list, when applicable?
- [x] 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.

Good afternoon,

Nanodon is a cluster BD1 phage with some interesting features. It is interesting that it encodes two primases, as do all of the members of the cluster. The genes were originally annotated with a large overlap by the students (as are Lika, Sujidade, Zamlya) which made me consider if we weren't looking at a frameshift of a single primase gene. I don't see evidence of that, but it is interesting that there are two. I have chosen to annotate without the large overlap (as per Caliburn, Danzina, and Aaronocolus).

Another interesting feature is that all of the phages carry a serine integrase (gp52 on Nanodon in Phamerator, gp51 on the final version). In all fourteen BD1 phages, there are only three different serine integrases, with thirteen of the phages split into two groups, and