Genome Annotation Submission Cover Sheet

Pre-SM*ART QC Phage Genome Annotation Checklist

Phage Name: Avle17

Your Name: JoAnn Whitefleet-Smith

Your Institution: Worcester Polytechnic Institute

Your email: jaws@wpi.edu

Additional emails:
(For correspondence)

Please check each box indicating completion of each task.

☑ 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?

no 9. Did you use PECAAN to annotate your phage?
   a. If, so please describe how in the text field after question 10.

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

-Added gp 59, NKF; start is not called by either gl or gm; "GeneMark self" is only coding potential analysis that has a positive result, and only in the middle of the region; gp is in Cocoabery's GenBank annotation.

-Gene 60: assigned function as "hydrolase", however there were an approximately equal number of equally well-supported BLAST hits to both hydrolases and esterases. HHPred also contained hits to both hydrolases and esterases, and of the two most closely-related phages, one lists the function as "hydrolase" and the other as "esterase".

-Gene 62: little support for earlier start at 40491 (although equivalent start is used in the