Genome Annotation Submission Cover Sheet

Preliminary Annotation Review Checklist 5-15-2018

Phage Name: Josuke

Your Name: Victoria Brown-Kennerly

Your Institution: Webster University

Your email: vbrownkennerly64@webster.e

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 (including hidden marks?)
- ☑ c. Do the function names in the Product field either match the official function list or say "Hypothetical Protein"?
- ☑ d. Is the Function field empty (including hidden marks?)

☑ 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.

ORF12: GeneMark showed upstream coding potential in Josuke, maybe unique to Josuke, but not available in the list of choices for START sites; however, this is a well-known protein, tapemeasure, and blasted 1:1 in NCBI with other tapemeasures, so would seem bold to change it drastically.
ORF 16: Had a few similarly named hits in different data sources, encouraging, but all with weak p-values.
ORF 18: Several hits all over the place, different proteins. No difinitive consensus.
ORF 19: Called this as holin; there were many lines of evidence and though each itself was not super strong, there were many different data sources pointing to holin. Called it.
TECHNICAL: This is a tiny genome, annotated in PECAAN to transfer info into DNAmaster,