Genome Annotation Submission Cover Sheet

Preliminary Annotation Review Checklist 5-15-2018

Phage Name: Rizwana

Your Name: Dr.Hari Kotturi

Your Institution: University of Central Oklahoma

Your email: hkotturi@uco.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?
- [x] 2. Are all the genes "valid" when you click the "validate" button?
- [x] 3. Have the genes been renumbered such that they go sequentially from 1 to the highest number?
- [x] 4. Have all old BLAST hits been cleared, and all gene features reBLASTed?
- [x] 5. Are the locus tags the"SEA_ PHAGENAME"?
- [x] 6. Has the Documentation been recreated to match the information in the feature table?
- [x] 7. Have tRNA ends been adjusted with web-based Aragorn and/or tRNAscan SE?
- [x] 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:

- [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
- [x] c. Do the functions in the Notes match the official function list?
- [x] d. Are all three lines of functional evidence described for EVERY gene?
- [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 (including hidden marks?)
- [x] c. Do the function names in the Product field either match the official function list or say "Hypothetical Protein"?
- [x] d. Is the Function field empty (including hidden marks?)

- [x] 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.

Gene 1 - SMART Members can look at this gene. This was not called by Glimmer or GeneMark. We found this ORF in the 644 bp gap between 1st and 2nd genes at the beginning of the genome. It is in the reverse direction as gene 1. The Z score is 2.233 and Final score is -4.312. There are no matches with Phagesdb or NCBI. HHPRED has 58.1% probability and 41.4% coverage with DNA repair protein. This could be a gene as there are actions in Rizwana that do not match with any member of cluster.

Gene 4 - SMART members should check this gene as we are not sure. We think that it should be deleted as it is 96 bp long with a product os 32 aminoacids. The BLAST matches have very low E-value. The reason for keeping this gene is it has a transmembrane domain predicted by