Actinobacteriophage Genome Annotation Submission Cover Sheet

This Cover Sheet will accompany each genome's annotation file(s) submission and succinctly describe the work that your students and you have done. This document ensures that the work done was as complete and thorough as it could be. Most important to the QC reviewer, denote where the trouble spots were in your annotation and how they were resolved.

Phage Name. Enochoraptor
Your Name. Rocky Ng
Your Institution. Skyline College
Your email. ngrocky@smccd.edu
Additional emails. (for correspondence). kapp@smccd.edu

Describe any issues or specific genes that you would like to highlight for the QC reviewer. This includes any genes that you had questions about or received help with or that warrant further inspection in the QC review process. Include those genes that you deliberated on and/or want to strongly advocate for. If you contacted SMART, workshop facilitator, or a buddy school for help, please document.

Cluster BU phages is conserved and lacks genetic diversity ([Marinelli et al. 2012](#)). The last time any cluster BU phages was annotated on 2016. So some of the previously annotated BU phages function would be different name under the updated Official Function List. Please take a look at these genes:

Gene 4 – MuF-like minor capsid protein to now called **capsid maturation protease**
Gene 9 – minor tail protein to **hypothetical protein** due to HK97-gp10 hit on HHPred (see PECAAN notes)
Gene 10 – minor tail protein to **tail terminator**
Gene 13 – **tail assembly chaperone**, can you help me double-check the slippery sequence?
Gene 20 – endolysin to **endolysin, N-acetylmuramoyl-L-alanine amidase domain**
Gene 23 – helix-turn-helix DNA-binding protein to **helix-turn-helix DNA-binding domain**
Gene 25 – helix-turn-helix DNA-binding protein to **helix-turn-helix DNA-binding domain**
Gene 26 – exonuclease to **MRE11 double-strand break endo/exonuclease**
Gene 27 – helix-turn-helix DNA-binding protein to **helix-turn-helix DNA-binding domain**
Gene 36 – exonuclease to **RecB-like exonuclease/helicase**
Gene 39 – hypothetical protein to **membrane protein**

Please record yes/no for each of the questions below. If further explanation is needed, please add this item to the above box.

In the submitted DNA Master file (Yes/No):

    Yes 1. Does the genome sequence in your submitted DNA Master file match the nucleotide fasta file posted on phagesDB (same number of bases, no N bases, etc.)?
    Yes 2. Are all the genes 'Valid" when you click the Validation button?
    Yes 3. Are the genes (and matching LocusTag numbers) sequential, starting with #1, counting by 1s.
    Yes 4. Are the Locus Tags the "SEA_PHAGE NAME" format?
    Yes 5. Has the documentation been recreated from the Feature Table to match the latest file version?
    Yes (no tRNAs) 6. Have tRNAs followed the tRNA protocol, **COPYING** tRNA-AMINOACID type (DNA equivalent of the anti-codon) from Aragorn output - tRNA-Gln(ctg) - AND the ends been adjusted to match the Aragorn output?
    Yes 7. Has the frameshift in the tail assembly chaperone been annotated correctly (if applicable)?

Yes 8. Have you <u>cleared your Draft </u>Blast data and have you <u>re-Blasted</u> the submitted DNA Master file?

Yes 9. Has every gene been <u>described and supported in your Supporting Data file</u>?

Yes 10. Did you investigate '<u>gaps</u>'?

Yes 11. Did you <u>delete the genes</u> that you meant to delete?

Now, <u>make a profile of the file</u> you plan to send. (And you can save this file for <u>Review to Improve!)</u>

Yes 1. Have any duplicate genes been deleted?

Yes 2. Has the Notes field been cleared (using the automated buttons)?

Yes 3. Do the gene numbers and locus tags match?

Yes 4. Are the correct Feature_Types correctly selected (most will be ORFs, but check that tRNAs and tmRNAs are correctly labeled)?

Yes 5. Do the function names in the Product field either match the official function list or say "Hypothetical Protein"?

Yes 6. Has the Function field been cleared (using the automated buttons)?

How are you documenting your gene calls in class? Choose any/all that apply:
- X PECAAN output
- DNA Master shorthand (previously used format)
- Spreadsheet
- Powerpoint
- Word document (must be easily searchable)
- Other: Describe.

What is the file type (sort) submitted for QC <u>to document your gene calls</u>? Choose only one.:
- X PECAAN output
- DNA Master shorthand (previously used format)
- Spreadsheet
- Powerpoint
- Word document (must be easily searchable)
- Other: Describe.