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. MrGreen

Your Name. Kari Clifton

Your Institution. University of West Florida

Your email. kclifton@uwf.edu

Additional emails. (for correspondence).

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.

We used PECAAN again, only our second genome. Previously I lacked proper understanding about making the start SS/NA/NI note. I think I corrected that, although I will say that we may have gotten some NA/NI mixed up. We checked starts for 4 bp overlap. This time around we did a better job with checking for membrane proteins, checking all with SOSUI and Deep TMHMM. TMPred offline?? All gaps were checked.

Gene 2. Glimmer suggested start 3, genemark suggested start 4 (most annotated). We chose start 4, as it resulted in a 4 bp overlap (rather than 19 bp overlap with start 3), and had better Z and RBS scores. Also, re-blast resulted in Q:S 1:1, which is better than that produced with start 3.

Genes 10-11 translational frameshift of tail assembly. This is my first time annotating a frameshift. A “standard” -1 frameshift. GGGAAA sequence. Checked the six-frame translation to verify. To help myself visualize, I selected Bri160, with essentially identical sequence. I located the slippery sequence, highlighted the gene 10 codons through the sequence. Then I marked the codons starting from the 3’ end of the fusion protein, and was able to ID the spot where they overlapped, which was indeed the first A. start of Gene 11 was moved up to 7001. I believe I located the exact position of the slippery A (7285), but you might want to verify please.

Gene 18. We called this as a membrane protein rather than hypothetical protein, as it only has 1 but was found on both SOSUI and TMHMM. Also, start was changed to one that resulted in 4 bp overlap. Gives bst RBS and Zvalue scores.

Gene 19. Also called this as a membrane protein as shown by Deep TNHMM and SOSUI.

We have reviewed and blasted the gap between Genes 21 and 22 and have decided that a gene with coding potential is NOT present. However, some EE phages (i.e., Kaijohn) have a small gene inserted in between two of the reverse genes (usually 21 and 22). We looked at some of those with the insertion and could not determine why the call was made to insert. If you feel that a gene should be inserted there, I would appreciate some feedback as to how to make that call.

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](https://seaphagesbioinformatics.helpdocsonline.com/article-84)?

Yes 3. Are the genes (and matching LocusTag numbers) [sequential](https://seaphagesbioinformatics.helpdocsonline.com/article-77), starting with #1, counting by 1s.

Yes 4. Are the Locus Tags the “[SEA\_PHAGE NAME](https://seaphagesbioinformatics.helpdocsonline.com/article-77)” format?

Yes 5. Has the [documentation been recreated](https://seaphagesbioinformatics.helpdocsonline.com/article-86) from the Feature Table to match the latest file version?

N/A 6. Have tRNAs followed the [tRNA protocol](https://seaphagesbioinformatics.helpdocsonline.com/undefined), **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](https://seaphagesbioinformatics.helpdocsonline.com/article-54) been annotated correctly (if applicable)?

yes 8. Have you cleared your Draft\_Blast data and have you [re-Blasted](https://seaphagesbioinformatics.helpdocsonline.com/article-57) the submitted DNA Master file?

yes 9. Has every gene been [described and supported in your Supporting Data file](https://seaphagesbioinformatics.helpdocsonline.com/article-44)?

yes 10. Did you investigate ‘[gaps](https://seaphagesbioinformatics.helpdocsonline.com/article-31)’?

N/A 11. Did you [delete the genes](https://seaphagesbioinformatics.helpdocsonline.com/article-65) that you meant to delete?

Now, [make a profile of the file](https://seaphagesbioinformatics.helpdocsonline.com/article-64) you plan to send. (And you can save this file for [Review to Improve!)](https://seaphagesbioinformatics.helpdocsonline.com/untitled-18)

N/A 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

X Word document (must be easily searchable)

      Other: Describe.

What is the file type (sort) submitted for QC to document your gene calls? Choose only one.:

      PECAAN output

      DNA Master shorthand (previously used format)

      Spreadsheet

      Powerpoint

 X Word document (must be easily searchable)

      Other: Describe.