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. ClamChowder

Your Name. Dane Bowder; Erin Doyle

Your Institution. Doane University

Your email. dane.bowder@doane.edu

Additional emails. (for correspondence). erin.doyle@doane.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.

This genome was fairly straightforward to annotate due to high sequence similarity to several published genomes (Barb, Fugax and PinkCoffee). We ultimately deleted two genes that were predicted by autoannotation: 41 (stop codon 33730) and 92 (57319). These were both eliminated because they overlapped other genes running in the opposite direction and have been eliminated in other phages. In general, we made sure to prioritize 1 bp overlaps over 4 bp overlaps and to ensure we had reasonable structural evidence when calling functions. Also, when selecting helix-turn-helix DNA binding domain fucntions, we always checked for the required structural pattern.

We were not able to identify a slippery sequence in the tail assembly chaperone, this appears to be consistent with other DC1 phages. Additionally we were unable to identify any tRNAs or tmRNAs in this genome.

There were a few function calls we want to highlight prior to review:

Gene 50 we called an immunity repressor, though it was classified as an orpham. There appears to be a mutation that introduces a stop codon, shortening the gene substantially and pushing it out of the pham with other similar phages. We debated whether or not to call the immunity repressor function, but ultimately decided to call it due to the high sequence similarity to other phages and HHPRED hits. We were however unsure whether this truncated version of this gene actually serves an immunity repressor function. (This may contribute to our expectation that this phage was lytic during phage discovery, and warrants some further investigation in the lab).

Genes 1 and 5 we called as the terminase- small subunit and large subunit, respectively. Many phages in the subcluster do not call in this way, but we ultimately chose to go with the small/llarge calls because of relevant HHPRED hits to both.

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)?

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

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?

Yes 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?

NA 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)’?

Yes 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)

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:

      PECAAN output

      DNA Master shorthand (previously used format)

      Spreadsheet

      Powerpoint

      Word document (must be easily searchable)

X Other: Describe. Students have been annotating in DNAMaster and using a google sheet to document their work. As we check their work, we input the data into PECAAN and use that to generate all notes and documentation.

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

      PECAAN output

X DNA Master shorthand (previously used format)

      Spreadsheet

      Powerpoint

      Word document (must be easily searchable)

      Other: Describe.