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. Avvas

Your Name. Lori West

Your Institution. Lee University

Your email. lwest@leeuniversity.edu

Additional emails. (for correspondence). jdaft@leeuniversity.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.

1. Gene 16 on phagesdb (original start = 12681) – We have called the function as minor tail protein based on recommendations from the CAT leader. Our only hesitation was that the gene length is 0.5kb rather than 1-3kb.
2. Gene 20 on phagesdb (original start = 14494) – We originally called this holin. It has 2TMDs, has a small region of alignment in HHPred to holin (pfam) and is near endolysin (one gene in between). The gene adjacent to the gene encoding endolysin is predicted to code for a protein with a TMD, however, it is only predicted to have a single TMD and therefore would not be classified as a holin under the current criteria. SEA-PHAGES information on calling holin vs membrane protein was the following: “If you have a transmembrane domain (or domains) found by DeepTMHMM for your protein and it has a pfam hit to a holin, please call it a holin.” However, our CAT leader recommended that we call this a membrane protein until more research is performed. The concern is that both genes likely encode proteins with holin associated function.
3. Gene 27 on phagesdb (original start = 17012) – We discussed this in the CAT. We have called an HNH endonuclease. Ours has an HNN sequence. This type of sequence is described in the forum and reference literature that supports inclusion of HNN in the HNH family.
4. Gene 22 on phagesdb (original start = 15439) The call we have selected goes against both Glimmer and GeneMark (which disagree). We are concerned with possible loss of coding potential. The call we have chosen decreases a significant gap (but creates a 4bp overlap).
5. Gene 25 on phagesdb (original start = 16529) – The call we have selected goes against the Glimmer/GeneMark call as well as 23MAs. However, we are concerned with possible loss of coding potential with that call. This call also decreases the gap.

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)’?

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)

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:

      PECAAN output

      DNA Master shorthand (previously used format)

      Spreadsheet

      Powerpoint

      Word document (must be easily searchable)

Yes Other: Describe. Google doc

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

      Word document (must be easily searchable)

X Other: Describe. Google slides