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. JeriBeth

Your Name. Jennifer Broderick

Your Institution. Thiel College

Your email. jbroderick@thiel.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.

----------------

This is my first time submitting a genome, so I apologize for any errors in documentation. After submitting JeriBeth, we are trying to finish up some annotations started previously so any advice is appreciated. Because our previous annotation instructor left, I did not get to attend the formal genome training and have been working off of the bioinformatic guide.

**Gene 4**

Matches “MuF-like minor capsid protein” in other phages but Approved function list says to keep it as a Hypothetical Protein? Doesn’t match the HHPRed alignment to be a capsid maturation protein?

**Genes 12-19 Tape Measure / Tail protein Synteny Calls**

The genes (12-19) require names based on evidence and synteny. Based on our calls, gene 12 is a Tail Terminator Protein; Gene 13 is a Hypothetical Protein-Unknown Function, Gene 14 is a Major Tail Protein, Gene 15 and 16 are Tail Assembly Chaperones, Gene 17 is a Tape Measure Protein, and lastly Gene 18 and 19 are a Minor Tail Proteins. We believe their proximity and synteny support their calls, in addition to their other evidence.

**Gene Product 20**

Aragorn did not pick up a tRNA at this coordinate when the whole genome was run (only picked up one that would be in the middle of Gene 19). tRNA scan also reported no tRNA’s found.

We didn’t delete Gene 20 officially due to uncertainty on if this was correct and protocol.

**Gene Product 26** –unsure on how much evidence is needed.

Gene product 26 comes up as a match for a Holin in other bacteriophages (ilzat, aubergine, etc) on BLAST NCBI (100% Identity, alignment, coverage) and on HHPRED (Phage\_holin\_5\_1 ; Bacteriophage A118-like holin, Hol118----99.7% probability, 75% coverage). It appears to have three transmembrane regions. It has an endolysin several downstream.

For the approved function list it requires: evidence needed to call a holin can include biochemical data (1), sequence similarity to genes with biochemical data (2), at least 2 transmembrane domians found and the gene be adjacent to the endolysins (s), conserved domain hits (4), and the abscence of additional transmembrane domains in the area. The literature suggests that some phages have more than one holin, for now when we seem multiple possibilities for a holin gene, let's call them membrane proteins.

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?

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?

Yes 7. Has the [frameshift in the tail assembly chaperone](https://seaphagesbioinformatics.helpdocsonline.com/article-54) been annotated correctly (if applicable)?

No 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:

X PECAAN output

      DNA Master shorthand (previously used format)

      Spreadsheet

X Powerpoint

      Word document (must be easily searchable)

      Other: Describe. 3 Page Worksheet

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

X Powerpoint

      Word document (must be easily searchable)

      Other: Describe.