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. Alucard

Your Name. Emily Savage

Your Institution. Southern Maine Community College

Your email. esavage@smccme.edu

Additional emails. (for correspondence). btarbox@smccme.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.

     - Contacted Debbie Jacobs-Sera and Karen Klyczek for help with the frame shift

* Called GP\_19 as a holin instead of a membrane protein. Documented evidence is attached. Would love to talk with someone about this call and if it is or isn’t a holin, and why!

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):

X 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.)?

X 2. Are all the genes ‘Valid” when you click the [Validation button](https://seaphagesbioinformatics.helpdocsonline.com/article-84)?

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

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

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

x 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?

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

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

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

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

x 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)

x 1. Have any duplicate genes been deleted?

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

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

x 4. Are the correct Feature\_Types correctly selected (most will be ORFs, but check that tRNAs and tmRNAs are correctly labeled)?

x 5. Do the function names in the Product field either match the official function list or say “Hypothetical Protein”?

x 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)

x Spreadsheet

      Powerpoint

      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)

x Spreadsheet

      Powerpoint

      Word document (must be easily searchable)

      Other: Describe.

Called Holin with evidence collected below:

1. 2 transmembrane regions identified by Sosui (1) and Deep tmhmm (2). Results below

A picture containing timeline

Description automatically generatedChart

Description automatically generated

1. HHPred identifies as a transmembrane domain. Second HHpred hits to holins in D29-11, specifically in the region that Sosui and TMHMM identify the domain.
2. Gene 19 is not directly adjacent to the endolysin (gene 17), but gene 18 has weak hits in hhpred to a Spanin protein (not strong enough to call other than a membrane protein.), which has been shown to be an important part in lysis in Lambda (link below)
   1. <https://pubmed.ncbi.nlm.nih.gov/22904283/> “
3. Sequence similarity is not the best identifying tool for holins, and ours has somewhat conflicting Sosui and Deep TMHMM results. This looks like it could be either a class 2 or class 3 holin. This paper goes a little beyond my understanding. Looking at the net charge density from Sosui, it seems that there is a net positive charge at the N terminus.
   1. <https://virologyj.biomedcentral.com/articles/10.1186/1743-422X-9-70>
4. Based on the “Official names list” there is one criteria we may not meet to call Holin.

“(4), and the abscence of additional transmembrane domains in the area.”

I’m not sure what the “area” is defined as. The upstream adjacent gene (gp 18) is a membrane protein identified by Sosui and Deep TMHMM.