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. KUKLA

Your Name. Tamarah Adair

Your Institution. Baylor University

Your email. Tamarah\_Adair@baylor.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.

1. CDS 3064 – 4104: gene 3 hits to Muf-like capsid protein but according to the functional assignment document we use hypothetical protein instead.
2. CDS complement (4101 - 4235) Potentially delete this gene. There is a bit of reverse CP and there is a BLAST hit to Arthrobacter sp, but not strong evidence.
3. CDS 17420 – 18367 Most likely the endolysin; HHPred evidence for peptidase
4. CDS 18379 – 18687 Pham is conserved; hints of holin; upstream is potential lysin; HHpred is weak but the TMHMM predicts 2 transmembrane
5. CDS 22239 - 23315

 /note=best evidence- HHPred lysophospholipase L1-like esterase [Arthrobacter sp. PvP023].

 /note=ACCESSION MBP1135119

 /note=Use hydrolase as general classification

1. CDS complement (23319 - 24104)

 /note=Evidence for esterase in several clusters; chose to go with broader group

1. CDS complement (24156 - 25310)

 /note=PF01757 This family includes a range of acyltransferase enzymes. This domain is found in a wide range of acyltransferase enzymes, including, mainly, bacterial proteins which catalyse the transfer of acyl groups, other than amino-acyl, from one compound to another, such as Glucans biosynthesis protein C (OPGC)

1. CDS 26468 - 26929

 /note=Changes start to cover the CP. strange forward among the reverse.

1. CDS 31650 - 31868

 /note=Small amount of CP not covered. Better scores/acceptable overlap

 /note=some hits to HTH; not consistent or strong

1. CDS 38865 - 39524

 /note=PF06319 This family includes Caulobacter MmcB (CCNA\_03580), which is involved in DNA repair. It has been proposed to be an endonuclease that creates the substrate for translesion synthesis [1].

 /note=Went with nuclease since there is not function for endonuclease and the specific nucleases on the list are not supported.

1. CDS 43182 - 43334

 /note=Changed to -4 overlap

1. CDS 43753 - 43965

 /note=No CP in the other reading frame; overlap is the only available start

1. CDS 44062 - 44535

 /note=Gap is LORF

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?

NA 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/. All BALSTED in Pecaan 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

x DNA Master shorthand (previously used format)

x Spreadsheet

x 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.:

x PECAAN output

      DNA Master shorthand (previously used format)

      Spreadsheet

      Powerpoint

      Word document (must be easily searchable)

      Other: Describe.