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. Rialto

Your Name. Kylie Marks

Your Institution. Montana Technological University

Your email. kmarks@mtech.edu

Additional emails. (for correspondence). mpedulla@mtech.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.

Please investigate our purported attP site between genes 45 and 46. It shows 26 identical DNA bases with the *M. smegmatis* mc2155 tRNA-Ala (gene MSMEG\_RS10375, at 2213142-2213214) and the same identity to Mycobacteriophage Zizzle at its bases 31,913-31938.

We identified BLASTP hits with Transposon at genes 31 and 66. Attempting to follow the official function assignment instructions, we called 31 as transposase because it had 5-base inverted repeated flanking the gene (GAAAC) but found no direct repeats. Please verify this assignment.

We deleted the very first gene called by GeneMarkS and Glimmer. It was not on the phagesdb draft but was on our autoannotation and called in our GenemarkS program.

Original reverse frame gene 58 (final draft replaced with added forward gene 57). There was insufficient space for divergent promotors. Final Gene 57 had coding potential on the GeneMarkS graph but was not called by GeneMark or Glimmer. The reverse original 58 was in phagesdb draft (#57) and our autoannotation.

New inserted forward gene 70 was in a gap on our original autoannoated DNA Master file and not called in phagesdb draft. It was not called in our GeneMarkS but had coding potential on the graph. It had BlastP matches at E value at 7e-67 and HNH endonuclease functional matches.

Our original auto-annotated reverse gene 76, called by GeneMark, and in phagesdb draft as gene 75 was deleted and the final gene 76 start codon was moved upstream to fill in the resultant gap. The longer start encodes a protein with 1:1 BlastP matches with multiple phages’ annotated proteins, although an alternative gene encoded in the first frame appears to have better coding potential..

Final forward gene 85 was added. Not called by Glimmer or GeneMarkS or phagesdb draft. It fills a gap and has 1:1 BlastP matches with annotated mycobacteriophages E value 1e-33 as well as coding potential on the GeneMarkS graph.

Added small forward gene 96 to fill gap. It was not called by Glimmer or GeneMarkS or phagesdb draft. BlastP matches with mycobacteriophage annotated proteins 8e-18 with no function assignment.

Deleted our original auto-annotated reverse gene 99. It was called by Glimmer but not GeneMarkS. It was in phagesdb draft annotation. No significant BlastP matches were identified. A small resultant gap exists between final forward genes 99 and 100.

Our original autoannotated starts were modified for final gene numbers 1, 3, 10, 14, 15, 19, 27, 39, 41, 42, 43, 47, 50, 54, 58, 67,

The regions between genes 76 and 78 were not clear based on GeneMarkS coding potential. We made our best attempt to call genes with BlastP matches and coding potential.

Gene 106 had several conflicting possible functional matches on BlastP and HHPred results.

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)?

No...BLAST FAILED 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

XXXX 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)

      Spreadsheet

      Powerpoint

XXXX Word document (must be easily searchable)

      Other: Describe.