

## 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. Arielagos  
Your Name. Eric Sakowski  
Your Institution. Mount St. Mary's University  
Your email. e.g.sakowski@msmary.edu  
Additional emails. (for correspondence). dye@msmary.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.

-We changed the start sites for genes 2 and 19. Changing gene 2 to begin at 424 instead of 363 creates a gap. We investigated this gap and found no strong evidence for a gene in this gap. Changing gene 19 to begin at 13063 makes a gap with no compelling evidence for a gene. There is a gap between genes 10 and 11 (10 stops at 6807 and 11 starts at 6936). There appears to be a potential gene here which shows high sequence similarity (88% identity over 84% of the sequence) with gene 14 in Arthrobacter phage Laila. However, there is no predicted start site that we could find, although there is a predicted leucine that could serve as a potential start site. We have been conservative and not called this a gene. There are no tRNAs, so we have stated yes for Q6 below. There is no tail assembly chaperone, so we have stated yes for Q7 below. We did not delete any genes, so we stated yes for Q11 below.

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](#)?
- Yes 3. Are the genes (and matching LocusTag numbers) [sequential](#), starting with #1, counting by 1s.
- Yes 4. Are the Locus Tags the "[SEA PHAGE NAME](#)" format?
- Yes 5. Has the [documentation been recreated](#) from the Feature Table to match the latest file version?
- Yes 6. Have tRNAs followed the [tRNA protocol](#), **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](#) been annotated correctly (if applicable)?
- Yes 8. Have you [cleared your Draft\\_Blast](#) data and have you [re-Blasted](#) the submitted DNA Master file?
- Yes 9. Has every gene been [described and supported in your Supporting Data file](#)?
- Yes 10. Did you investigate [gaps](#)?
- Yes 11. Did you [delete the genes](#) that you meant to delete?

Now, [make a profile of the file](#) you plan to send. (And you can save this file for [Review to Improve!](#))

- 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:

- PECAAN output
- DNA Master shorthand (previously used format)
- Yes 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)
- Yes Spreadsheet
- Powerpoint
- Word document (must be easily searchable)
- Other: Describe.