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. Argan

Your Name. Bryan Gibb

Your Institution. NYIT

Your email. bgibb@nyit.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.

**The region between 45,000 and 45,500 bp is a rather large gap with low coding potential. After We opted to call three genes, which are similar to what we did in the previous year for Uzumaki. Coding potential is there, but very weak, but the gap is also very big. The genes we call fill the gap well, but have no homology to anything other than those called from Uzumaki.**

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?

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

YES 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?

NO 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”?

NO 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

      Word document (must be easily searchable)

      **Other: Onenote Notebook**

**Describe. Class split into groups of two. Each pair of students were assigned a subset of genes from Argan and asked to document start sites for each gene in a Onenote notebook. The class shared a single notebook, and each group was given a tap, and each gene a page. Genes were assigned so that two pairs of students evaluated each gene. After structural annotation, groups were shuffled and each pair was assigned a new set of genes to annotate for function. Students again annotated for function using a Onenote notebook, but here they were also expected to review the calls from the each of the two groups that were assigned to annotate start sites as they then evaluated the function. Once the students had evaluated all of their genes in Onenote, they used that information to work on the genome in PECAAN. In this case, PECAAN is reporting the calls based on evidence collected in their notebook rather than being used to collect and evaluate the evidence directly (or so I hope).**

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

      Word document (must be easily searchable)

      Other: Describe.