# 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. Fresco
Your Name. Nicholas Klotz
Your Institution. Webster University
Your email. nicholasklotz@webster.edu
Additional emails. (for correspondence). marypreuss34@webster.edu, vbrownkennerly64@webster.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.

Gene 41 had a larger gap (104bp) which was BLASTED and no evidence of a protein was found.

There was a gene deleted (bp 43,848 to 44,159) found in the forward direction among a group of reverse genes. It was found by GeneMark, but not Glimmer. There was no evidence supporting that this was a gene. It was found that this was actually a piece of HNH endonuclease, which was added into the genome (gene 50, from bp 44,095 to 43,628 in the reverse direction); it seems like GM/Glimmer tends to not find HNH endonuclease.

There was a large gap between genes 61 and 60, of 462bp. There was no protein coding potential found between these genes in NCBI BLAST (blastx).

There was another large gap between genes 62 and 61, and no protein coding potential was found between these genes in NCBI BLAST.

Gene 72 had an gene call that may not be accurate- please review. In both PhagesDB BLAST and NCBI BLAST, the protein function "hica-like toxin" was called, but this doesn't appear to be in the approved function list in SEA PHAGES anymore. HHPRED had hits for "toxin HicA; toxin-antitoxin, TA, protein complex, DNA-binding", which was the most similar to the approved function "toxin in toxin/antitoxin system, HicA-like", so that function was selected. However, since the BLAST results called a different function, I wasn't sure if this function represented the same thing or if "Hypothetical Protein" should be selected instead, given the evidence provided.

In gene 79, I had trouble figuring out which start to select. GeneMark and Glimmer didn't agree (GM: 58,225 and Glimmer: 58,243). Starterator, however, wasn't super informative (one start had 2 MA's, and the other had one MA). Looking at the Z-score and final score, the results looked better for 58,243 than the other start, so this start was selected. However, that was really the only evidence supporting that start over the other.

**This phage didn't seem to have a frameshift, and almost all of the phages in this cluster did not call a frameshift as well.

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?
N/A 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?
N/A 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
Spreadsheet

What is the file type (sort) submitted for QC to document your gene calls? Choose only one.:
PECAAN output