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. Skitty
Your Name. Julia Simler
Your Institution. Purdue University
Your email. jsimler@purdue.edu
Additional emails. (for correspondence). klclase@purdue.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.

Skitty had many gene deletions, possibly due to its lack of similarity with other phages. The beginning portion of its genome is the only portion with strong conservation with fellow FQ phage MrSmee.

The gene from 20618-20797 bp is a reverse gene called by both GeneMark and Glimmer. GeneMark showed heavy coding potential and there were decent BLAST results. However, it is a reverse gene after a forward gene and does not have the usual 50 bp gap. Despite violating this rule, there was not enough evidence to delete the gene.

The gene from 38085-38489 bp is a reverse gene that overlaps the upstream forward gene by 91 bp. It has no coding potential, has no BLAST results, is not called by GeneMark, is an orpham, and has no similarity with other phages. It is called by Glimmer with strength of 3.01. The gene from 38642-38836 bp is a reverse gene following the upstream reverse gene with a gap of 152 bp. It has no coding potential, has no BLAST results, is not called by GeneMark, is an orpham, and has no similarity with other phages. It is called by Glimmer with strength of 1.56. Both genes were deleted due to the overwhelming lack of evidence. However, this leaves a gap of 659 bp. The gap was investigated for new genes, but none were found.

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)
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)
Spreadsheet
Powerpoint
Word document (must be easily searchable)
Other: Describe.