

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. Kabocha
Your Name. Bonnie Madden
Your Institution. George Mason University
Your email. Bmadden1@gmu.edu
Additional emails. (for correspondence). Aschere2@gmu.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.

GP_13

Blast has NKF, HHPred hits, near a terminase, official FX on list is oxidoreductase. I think it codes for a subunit of the pump that drives the terminase

GP_17

Chose GL. Other start stats were worse. Should we choose the longest start with a -4 overlap? Chose to call it HNH endonuclease

GP_29

Possible portal protein based on HHPred High probability of 98.4% but low coverage of 37.156%. Not calling it to be conservative.

GP_40

Possible capsid decoaration protein but the % identity was less than 90% on NCBI Blast

GP_41

Possible membrane protein. HHPred only had a 73.1% probability. TmHmm was 0, but SOSUI predicton is a membrane protein.

GP_42

Called function as head-to-tail adapter based on ChisanaKitsune synteny.

GP_43

No function called. No hits on NCBI Blast, but HHPred called it a Phage tail S with a 99.8% probability and 89.404% coverage. Not on official function list. In the region of other tail proteins

GP_45

No function called No hits on NCBI Blast, but HHPred called it a Minor tail protein U with a 97.9% probability. In the region of other tail proteins.

GP_65

No function called. No hits on NCBI Blast, but HHPred calls it a Cell division protein FtsB with 95.9% probability and 89.2857% coverage. Not on official function list.

GP_82

No function called. No hits on NCBI Blast, but HHPred calls it as a Host-nuclease inhibitor protein Gam with a 91.2% probability and 83.5443% coverage. Not on official function list.

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?

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:

Yes PECAAN output

Yes 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

Yes DNA Master shorthand (previously used format)

Spreadsheet

Powerpoint

Word document (must be easily searchable)

Other: Describe.

