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. Goodgraces
Your Name. Suparna Bhalla
Your Institution. Mount Saint Mary College
Your email. Suparna.bhalla@msmc.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.

Goodgraces is a podovirus with a small genome and a few orphams. There are some podovirus specific genes that are not yet in the official function list. I spoke to David Bollivar who was a CAT leader for one of the last CAT meetings who consulted with Debbie and he sent us references to the Phi 29 phage which we have done at times when tagging functions.

Additional issues

1. Gene 14: Function is a collar protein. HHPRED hits to a collar protein in phi 29, staph virus with a 100% probability and low e values. This is not an official function so labelled it as hypothetical protein and would request it be changed
2. Gene 16 – Labelled as Endolysin but data supports Endolysin M23 domain as there is no lysin B. This was not an option on the function list and would request a change
3. Gene 22 – Labelled as Endolysin but data supports Endolysin with N-acetylmuramoyl-L-alanine amidase domain as there is no lysin B. This was not an option on the function list and would request a change
4. Gene 21 – Really small, only 120bp. There is only one call (Glimmer) and we had several discussion on whether to delete as we evaluated it against the guiding principles. We kept it but would appreciate another opinion
5. We annotated 5 orphams. Genes 3, 11, 27, 29, 30. Besides coding potential we some poor alignment we did not have much to go on and would appreciate the smart team to evaluate our choices

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 and No 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
DNA Master shorthand (previously used format)
Spreadsheet
Powerpoint
Yes 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.