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. Dalilpop
Your Name. Ann Koga
Your Institution. College of Idaho
Your email. akoga@collegeofidaho.edu
Additional emails. (for correspondence). NA

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 1 (stop 685): Some of the related phages call this gene Terminase, small subunit. There is HHPred evidence to support this (PF05119.15, 96.9% probability), but it is very far from the Terminase, large subunit, which is about 28 genes downstream from this gene. I left NKF as Flapper and Turuncu annotations did.

Gene 29 (Stop 16.257) This is the large subunit mentioned above. Have left as Terminase.

Gene 4 (Stop 4266) Strong hit to HHPred for zinc finger DNA binding protein, but no hits in Conserved domains, so left as NKF

Gene 50 (stop 38,589) 25 bp overlap, but start is consistent with other CR1 phages

Gene 51 (stop 38,935) Most annotated start not in this genome. Leaves large gap, but previous start would create 29 bp overlap.

Gene 56 (stop 41.870) 26 bp overlap. Consistent with many other phages. Next start would omit coding potential

Added gene, which filled a gap but created a huge overlap with the preceding gene. Flapper and Turuncu have kept both genes with the big overlap. New gene stop is 44,726 (Gene 59) which has ample evidence for a function, WhiB family transcription factor. Previous gene, no function, but there are many members of Pham and very good coding potential. There is another start codon that could reduce the overlap, but would be inconsistent with starts of other phages.

Gene 73 (stop 54,868) There are 2 equally conserved start sites, one with a 4 bp overlap and one with a 50 bp gap. For some reason, other annotations have chosen the second start. However, choosing the first start captures more coding potential and has the advantage of the -4bp gap.

Added gene, which filled a gap but created a huge overlap. This is another WhiB Transcription factor gene and creates a large overlap just as mentioned above. New gene stop is 57,295 (Gene 76) which has ample evidence for a function. Previous gene is a DNA Helicase.

Added a gene at the tail end of the reverse genes (stop 60,156). There is only a small amount of atypical coding potential. A Blastn resulted in very close alignment with regions of 3 of the most closely related phages (GRU-1, Flapper, and Turuncu). 96% identity and e= -99 No Blastp hits. The gap between the reverse and forward genes is pretty large, so adding this gene still leaves a 346 bp space for ribosomes to bind for translation in both directions. I really don't know if I should add this gene, but figured it would be easier for you to delete it than to add it.

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?
NA 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"?
yesyes 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
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.