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. ViaConlectus
Your Name. Dane Bowder and Erin Doyle
Your Institution. Doane University
Your email. dane.bowder@doane.edu; erin.doyle@doane.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.

ViaConlectus was an interesting genome to annotate, and we are proud of our final product. In our annotation section, the students annotated and took notes using DNA Master and a spreadsheet, and we (the faculty members) chose to keep track of annotations, notes and export from PECAAN. As a general rule, many of our genes did not agree with the most common start in Starterator, so we chose NI, but made a note, as we typically chose the start most common for other DE4 phages. There were a few genes of note we wanted to point out specifically:

Gene 16 we had trouble calling a function- many phages called this a MuF-like minor capsid protein, however, after posting on the forum and the approved function list, Debbie informed us that was no longer an approved function, and that there was not experimental evidence for that gene in the capsid of the virus, so we ultimately called this NKF. Of note, gene 16 also contains a hedgehog like intein that was interesting, but we were unable to assign to a function.

Gene 18 was an added gene, though it is quite small. We added this gene because it filled a large gap and accommodated some coding potential. Of note, after we annotated, we noticed that the newly completed Tardus also added this gene.

We did not call a function for gene 37, though many other phages call this a minor tail protein. The reason for this was a lack of HHPRED evidence, and it did not meet the additional requirements (collagen or glycine rich HHPRED hits in the syntenic minor tail region), so we chose to stick with NKF.

We chose to call gene 44 and 45 membrane proteins due to the presence of two TMH domains found by tmhmm. Very few other phages have done this, but we thought we would go with membrane protein, as it is outlined in the approved functions list. Gene

Genes 54 and 55 we ultimately chose to call the function HTH DNA binding protein, rather than including a MerR like qualifier. This is aligned with what other phages have done, though there was some minimal evidence for the MerR like, but it wasn't consistent.

We added gene 64, because it filled a gap and accommodated coding potential and had 100% similarity to Zipp_65, along with a 4 bp overlap. We were convinced this was a gene.

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:

Yes PECAAN output
Yes DNA Master shorthand (previously used format)
Yes 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.:

Yes PECAAN output
DNA Master shorthand (previously used format)
Spreadsheet
Powerpoint
Word document (must be easily searchable)
Other: Describe.