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. Bolt007
Your Name. Amanda Freise
Your Institution. UCLA
Your email. afreise@ucla.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.

Stop @ **5522**: interesting strong hit for Imm-like superinfection immunity protein. But left as NKF because remaining evidence did not seem to suggest correct function.
Stop @ **9970**: Top HHpred hits have some connection to tail terminator, but 3FZ2 chain I is not listed as an acceptable piece of evidence for tail terminator (must be A-F), so leaving as NKF for now.
Stop @ **20917**: Called as minor tail, but may be possible tailspike protein. Top HHpred hit (w/ only 18% coverage) is 3GW6, a chaperone involved in folding triple-beta coils which are in the tailspike. Other HHpred hits to tailspike. 2 other FH phages called this a tailspike protein.
Stop @ **29398**: tricky function call. Closest PhagesDB hits call esterase, but top HHPred hit is for glycoside hydrolase. Pham includes both calls.
Stop @ **29036**: some HHPred evidence to suggest minor tail protein, but not sure if sufficient.
stop @ **30487**: tricky function call. Exonuclease at minimum, but could it be Cas4 or RecE? None of the SP-recommended HHpred hits present to call Cas4, but there is one good hit to 3FHD_A which mentions PD-(D/E)XK Superfamily. It is just upstream of a RecT, but no other evidence to suggest RecE (except many PhagesDB hits). Unsure of whether to call something more specific than exonuclease.
Stop @ **34071**: for function, a couple high e-value/probability HHpred hits suggest RepA-like protein rather than just HTH DNA binding protein, but HHpred hits don't have very high identity.
Stop @ **34668**: HHpred hits suggest helicase loader, though no one else in this pham (only other FH phages) called that function.
Stop @ **41995**: Added this gene to fill gap and due to CP on GeneMark-Host, but it may not be real. Some poor, but not *terrible*, hits on HHpred for some kind of DNA-related function.

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)?
NO 8. Have you cleared your Draft Blast data and have you re-Blasted the submitted DNA Master file?

I was unable to re-BLAST the genome within DNA master and have yet to figure out what the issue is. Debbie said it would be fine for now to note this here and submit. I did re-blast for the few location calls I wasn't sure about.

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!)

1. Have any duplicate genes been deleted?
2. Has the Notes field been cleared (using the automated buttons)?
3. Do the gene numbers and locus tags match?
4. Are the correct Feature_Types correctly selected (most will be ORFs, but check that tRNAs and tmRNAs are correctly labeled)?
5. Do the function names in the Product field either match the official function list or say "Hypothetical Protein"?
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:

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

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