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. **Usavi**
Your Name. Tejas Bouklas
Your Institution. UCLA
Your email. tbouklas@mednet.ucla.edu
Additional emails. (for correspondence). afreise@ucla.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.

stop @ 19105: Many calls for TAC, but no convincing evidence. Debbie agreed in this forum post: https://seaphages.org/forums/topic/4518/. Calling as NKF for now. See also gene downstream, which interestingly, does have a 94% probability/41% coverage hit to tail assembly chaperone in HHpred (though the e-value is 1.1).
We just annotated two very similar B1 draft phages. This is stop @ 18600 in Bluehusk and stop @ 19086 in Miniboss

stop @ 49093: needs more digging; some HHpred hits also include polymerase. Other members of pham call all three functions. Asked on forums:
https://seaphages.org/forums/topic/4391/
https://seaphages.org/forums/topic/5493/
We just annotated two very similar B1 draft phages. This is stop @ 48513 in Bluehusk and stop @ 48015 in Miniboss

stop @ 55929: added based on good CP in Self-trained Genemark. Not found in all B1 phages, but Usavi seems to have an insertion in this area. About 15 hits from PhagesDB with good stats; potentially HNH? unclear based on HHPred

stop @ 56609: went with start site 56481. Difficult start site; 3 start sites often MA'ed in pham. Went with this one based on best final score and covers all coding potential.

stop @ 57734: Start site: chose start site based on Starterator. Other start sites are better RBS-wise, but produce far longer genes than other genes in the pham. Function tricky. forum response: ribbon helix helix IS NOT supported for Usavi70 and should not be called. https://seaphages.org/forums/topic/5437/
We just annotated two very similar B1 draft phages. This is stop @ 57128 in Bluehusk and stop @ 56429 in Miniboss

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 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.
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:

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.