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. Shatter
Your Name. Nicholas Klotz
Your Institution. Webster University
Your email. nicholasklotz@webster.edu
Additional emails. (for correspondence). marypreuss34@webster.edu, vbrownkennerly64@webster.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.

Gene #36: I deliberated on the function for this gene, and ended up calling the function minor tail protein. Looking at the pham number in phagesdb, 3/19 members have called this as a minor tail protein. One member that called that function, Axumite, had an e-value of 1e-160 when compared to Shatter's gene in phagesdb BLAST. Mariokart, who also called that function, had an e-value of 4e-86. Many other phages had no function listed for this gene, though. Neither HHPRED nor CDD have any good hits for function. NCBI BLAST shows that Axumite and Shatter are exactly the same for this gene (100% coverage, 100% identity, 100% aligned), which led me to select minor tail protein as the function.

Genes #60-61: These two genes had a particularly large gap between them (462 bp). There seems to be no coding potential between them, but it seems like a rather large gap. Glimmer/GM do not call any genes in this gap, and many phages in Phamerator do not call any genes between this gap. There were no long ORFs found in the gap and the ORFs present were put into NCBI BLAST with no good hits for similar phages calling a gene. The gap is likely because gene 60 is reverse and gene 61 is forward, so there's probably a gap so that there's enough room for ribosomes to fit that go in both directions.

There also seemed to be larger gaps (>100bp) closer to the end of the genome between operons. No significant coding potential was discovered between the gaps (based on automated GM/Glimmer, Phamerator, and NCBI BLAST).

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:
PECAAN output

What is the file type (sort) submitted for QC to document your gene calls? Choose only one.:
PECAAN output