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.  Anilorac

Your Name.  Catalina Dávila-Aguer

Your Institution.  Universiad Ana G. Mendez, Gurabo campus

Your email.  davilac2@uagm.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.

Please record yes/no for each of the questions below. If further explanation is needed, please add this item to the above box.

1. The Starterator report is not informative for genes 24, 35, 45, 51, 53, 55, 56, 60, and 62.
2. HHPRED identified gene 11 product as a minor capsid protein with a 99.5% probability. However, with did not assign this function, since the alignment do not stretches the full length of both proteins (75% coverage)
3. Legitimate large gaps were found at different locations within the genome (i.e. between genes 38-39, 39-40, 40-41, 50-51, 60-61, and 61-62). We researched all these gaps and did not find blast hits.
4. We changed the start in genes 12, 29, 40, 58, and 61, in order to improve the alignment and in most cases have better Z scores and Final scores. We also took into account the Starterator reports for these genes.
5. We did not detect a slippery sequence and did not identify a programmed translational frameshift in the two tail assembly chaperone genes.
6. Genes 25 and 49 were identified by DNA Master as membrane proteins. However, when Deep TmHHm was run only one transmembrane domain was found, so we assigned NKF to those genes. The only gene with three transmembrane domains was gene 26 (holin). Images from the TmHHm results were included in the spreadsheet named Anilorac\_MemProt\_Supporting inf.
7. We are not sure about the function DNA primase/helicase in gene 33, since the function of this gene varies among members of the cluster EA1, even when comparing with the most similar genomes. HHPRED only detected the function helicase.
8. A forward gene overlapping almost entirely with genes 34 and 35 (reverse genes) was detected by the auto-annotation in DNA Master. It was eliminated.
9. We are not sure about the function VRR-Nuc domain protein in gene 34, since it varies among members of the cluster EA1, even when comparing with the most similar genomes. Only GeneMark called this gene.
10. We are not sure about the function of gene 46 (thymydilate kinase). Some phages of the cluster report an AAA-ATPase in this gene, others agree with our result (thymydilate kinase)
11. tRNAs and tmRNAs were not detected in Anilorac’s genome

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](https://seaphagesbioinformatics.helpdocsonline.com/article-84)?

 Yes  3. Are the genes (and matching LocusTag numbers) [sequential](https://seaphagesbioinformatics.helpdocsonline.com/article-77), starting with #1, counting by 1s.

 Yes  4. Are the Locus Tags the “[SEA\_PHAGE NAME](https://seaphagesbioinformatics.helpdocsonline.com/article-77)” format?

 Yes  5. Has the [documentation been recreated](https://seaphagesbioinformatics.helpdocsonline.com/article-86) from the Feature Table to match the latest file version?

 N/A  6. Have tRNAs followed the [tRNA protocol](https://seaphagesbioinformatics.helpdocsonline.com/undefined), **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](https://seaphagesbioinformatics.helpdocsonline.com/article-54) been annotated correctly (if applicable)?

 Yes  8. Have you cleared your Draft\_Blast data and have you [re-Blasted](https://seaphagesbioinformatics.helpdocsonline.com/article-57) the submitted DNA Master file?

 Yes  9. Has every gene been [described and supported in your Supporting Data file](https://seaphagesbioinformatics.helpdocsonline.com/article-44)?

 Yes  10. Did you investigate ‘[gaps](https://seaphagesbioinformatics.helpdocsonline.com/article-31)’?

 Yes  11. Did you [delete the genes](https://seaphagesbioinformatics.helpdocsonline.com/article-65) that you meant to delete?

Now, [make a profile of the file](https://seaphagesbioinformatics.helpdocsonline.com/article-64) you plan to send. (And you can save this file for [Review to Improve!)](https://seaphagesbioinformatics.helpdocsonline.com/untitled-18)

 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

      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.:

      PECAAN output

      DNA Master shorthand (previously used format)

  X   Spreadsheet

      Powerpoint

      Word document (must be easily searchable)

      Other: Describe.