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. Hermeonysus

Your Name. Breonna Martin

Your Institution. Harford Community College

Your email. brmartin@harford.edu

Additional emails. (for correspondence). Jaclyn Madden – jmadden@harford.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.

Gp 26 Stop@15543 We investigated the preliminary function of "Major Capsid Hexamer Protein", it lacked evidence on HHPred. We decided to go with the call, "Major Capsid Hexamer Protein" based on the notes for the EC cluster. This call was discussed with Debbie.

Gp 27 Stop@16406 We investigated the preliminary function of "Major Capsid Pentamer Protein", there was a lack of evidence for this call on HHPred. We decided to go with the call "Major Capsid Pentamer Protein" based on the notes for the EC cluster. This call was discussed with Debbie.

GP 29 Stop@17879 We investigated the preliminary function of "major tail protein." However, we decided to go with the conservative call "Hypothetical Protein" due to the lack of useful evidence on HHPred. This call was discussed with Debbie.

GP 30 Stop@18803 Conservative call "Hypothetical Protein" made due to lack of appropriate hits on HHPred that are described in the approved functions list. This call was discussed with Debbie.

GP 42 Stop@30804 Preliminary function called this gene as a Hypothetical protein, however, when looking at both CDD and HHPred, significant results pointed to this being an endolysin. No lysin b was found, so this was called endolysin as opposed to lysin a.

Gap between gp53 and gp54 36151-36387 We investigated this gap as a potential gene, but concluded that it was not likely a gene. There is no CP on GMSelf. The very small CP on GMHost only spans the second half of the potential ORF. While there are two hits on BLASTP (PiperSansNom and Pulchra) we didn't think we had sufficient evidence to insert a gene

GP 54 Stop@36593 There is a large gap, but the RBS scores for alternative starts are poor, there is no CP on GeneMark to support extending this gene. This start also aligns with all other MAs on Starterator.

GP 58 Stop@38369 Gene inserted (114 bp). small CP on GM self and GM Host. BLASTP Hit with Scumberland and Quhwah. stop has a -4bp overlap with the downstream gene. We decided to insert this even though it is small

GP 61 Stop@39863 There were many great hits on HHPred (integrins, anthrax toxin receptor, collagen, etc.), but none of the proteins align with phage functions or the approved functions list. We were looking to investigate DNA ligase but couldn't identify evidence to support this call.

GP43, GP 44, GP 45, GP 46, GP65, GP70, GP 71, and GP 81 - Each of these genes had hits for transmembrane domains on DeepTMHMM. We kept the call “hypothetical protein” since “membrane protein” did not come up in our preliminary survey of protein functions, but wanted to mention that transmembrane domains were present.

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](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?

Yes (no tRNAs found) 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?

Yes (not located) 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)

N/A 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.