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. TinyTimmy

Your Name. Allison Johnson

Your Institution. Virginia Commonwealth University

Your email. aajohnson@vcu.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.

Gp5 (2689-3153) function terminase?

There is [an HHpred hit](https://www.rcsb.org/pdb/explore.do?structureId=6Z6E) with 97% probability and Evalue 0.012 that don’t meet our criteria but I guess are being used for functional annotation in the cluster?

Matched tail assembly chaperone regions to Mulciber.

gp34, 26087-26647 and gp 35, 26640-26915: ParA and ParB-like dsDNA partitioning proteins. Confirmed homology to RedRock 37/38.

Gp60 (38936-38439) please check start position 38804 genemark vs 38936 glimmer. Equivalent RBS scores, longer ORF introduces 50 bp overlap, but seems to be preferred start.~ ⅔ of genomes don’t have the most called start if I am reading the [starterator](http://phages.wustl.edu/starterator/Pham166330Report.pdf) correctly.

Gp 70 (42225-41899) has good CDD and HHPred hits for a [pfam](http://pfam-legacy.xfam.org/family/PF14373.10) for T4 super infection immunity protein. Phagesdb function frequency on Pecaan isn’t helping, there are a bunch of related names being used. Please suggest proper name for this protein. We used “Imm-like superinfection immunity protein”.

Gp75 (44421-44044) has [HHPred support](https://www.rcsb.org/pdb/explore.do?structureId=7TZ1) for immunity repressor, and this pham is named this in 19 A11 phages.

Gp76 (44607-44422) doesn’t have HHpred (looks like a transcription factor but E values are fairly high) but has [CDD support](https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi?RID=61HXX0K5016&mode=all) for transcription repressor annotation, and this pham is also named immunity repressor in 20 A11 phages. Are there two?

Gp93 (49220-49047) we went with start at 49220 as predicted by both tools, slightly better RBS, [starterator](http://phages.wustl.edu/starterator/Pham131725Report.pdf) support, as this genome doesn’t have most annotated start. Note there is a start @ position 49232 with a -4 overlap that SMART team may prefer (RBS not as good).

 I reset a few -4 annotations to -1 after talking about this with Debbie at the SEA faculty meeting!!

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?

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?

Yes 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:

X PECAAN output

      DNA Master shorthand (previously used format)

      Spreadsheet

      Powerpoint

      Word document (must be easily searchable)

X Other: Describe. Students document annotation in a wiki-based electronic notebook

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.