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. BearBQ

Your Name. Dane Bowder and Erin Doyle

Your Institution. Doane University

Your email. dane.bowder@doane.edu

Additional emails. (for correspondence). erin.doyle@doane.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.

We had a lot of fun annotating BearBQ, it is quite the oddball phage! We ultimately were able to identify and annotate the tail assembly chaperone programmed translational frameshift in genes 14 and 15, and eliminated the autoannotated tRNA due to a low Infernal score (only detected by tRNAScan.SE.) Notably, BearBQ almost seems like a mosaic phage, composed of sections from several of the DN subclusters, which made it difficult to identify close relatives, it was almost like there were close relatives for specific chunks of the genome. After autoanotation, we ended up deleting three genes and adding two based on coding potential and the rules of annotation. As with our other genes we tried to follow the guidelines, paying particular attention to HHPRED when calling functions. There were a number of genes we felt were important to mention in the cover letter to explain our process.

Gene 4 (Stop 4232)- The function for this gene was difficult to call. Many other phages call this MuF like (no longer on the approved list), and some call this a capsid maturation protease, but we could not identify any HHPRED hits that would support such a function. Ultimately we chose to call this an NKF despite many other phages calling a function.

Gene 14/15 (Stop 8984 and 9373 respectively)- This is the Tail Assembly Chaperone with the frameshift. We believe that we correctly identified a -1 slippery base.

Gene 21 (Stop 20055)- We identified this as one of the two lysin A proteins. We were torn in our call between an M15 domain lysin A and a L-ala D-glu domain. We chose the latter because in the conserved domains database it explains that the M15 is a larger family of proteins, implying the L/D function is more specific.

Gene 26 (Stop 22341)- We believe this can be called a membrane protein because it has 4 transmembrane domains. No other phages have identified this function.

Gene 28 (Stop 24742) and 32 (Stop 27110)- Though some phages call this a minor tail, and it is in the region of tail proteins, we did not identify any structural support for this, so we called it an NKF.

Gene 40 (Stop 30396)- We really struggled with the start call on this gene. Starterator suggested one start, but there was another start with a higher RBS score that we called. However, the issue is that both of these starts result in some cut off of coding potential. We chose the higher RBS, but want to acknowledge that we noticed the coding potential, but chose to align with another phage. We would perhaps suggest moving this start further back to accommodate all of the coding potential, but didn’t have other phages to compare to- so we welcome a second set of eyes.

Gene 48 (Stop 34848)- We didn’t actually have any issues calling this gene, but noticed that there were some hits in HHPRED to interesting antigenic proteins from M. tuberculosis that seem to play an immunoprotective role. No major issues, but may be an interesting gene for someone to look further into!

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 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)

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:

XPECAAN output

      DNA Master shorthand (previously used format)

      Spreadsheet

X Powerpoint

      Word document (must be easily searchable)

X Other: Describe. Students make their documentation in a set of google slides that we grade. We transfer and annotate the genome for submission in PECAAN based on their calls and evidence.

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

      PECAAN output

X DNA Master shorthand (previously used format)

      Spreadsheet

      Powerpoint

      Word document (must be easily searchable)

      Other: Describe.