# Genome Annotation Submission Cover Sheet

**Pre-QC Phage Genome Annotation Checklist**

| | |
|---|---|
| Phage Name: | Lennon |
| Your Name: | Breimann, Jake; Seraly, Paul; Wynn, |
| Your Institution: | University of Pittsburgh |
| Your email: | jsb100@pitt.edu; |
| Additional emails: (For correspondence) | pas142@pitt.edu; mjw120@pitt.edu; rid19@pitt.edu; |

*Please check each box indicating completion of each task. Annotation Guide section #'s indicated*

☑ 1. Does the genome sequence in your final contain the same number of bases and is it the same as the posted sequence on phagesdb.org?
☑ 2. Are all the genes "valid" when you click the "validate" button? *Section 9.3.2*
☑ 3. Have the genes been renumbered such that they go sequentially from 1 to the highest number? *Section 9.3.3*
☑ 4. Have all old BLAST hits been cleared, and all gene features reBLASTed? *Section 9.3.4*
☑ 5. Are the locus tags the phage name? *Section 9.3.3*
☑ 6. Has the Documentation been recreated to match the information in the feature table? *Section 1.4*
☑ 7. Have tRNA ends been adjusted with web-based Aragorn and/or tRNAscan SE? *Section 9.5.3-4*

8. For the items below, generate a genome profile, and review the following. *Section 11.3*

For the YourPhageName_CompleteNotes.dnam5 file:

☑ a. Have any duplicate genes (or any with the same stop coordinate?) been removed?
☑ b. Does every gene have **one and only one** complete set of Notes (see fig 12.2 in the Annotation Guide)?
☑ c. Do the functions in the Notes match the official function list?
☑ d. Is the function field EMPTY for all features?
☑ e. Do the notes contain the initial Glimmer/GeneMark data from the autoannotation?

For the YourPhageName .dnam5 file:

☑ a. Have any duplicate genes (or any with the same stop coordinate?) been removed?
☑ b. Is the Notes field empty for all the features with no known function?
☑ c. Do the function names in the Notes match the official function list, when applicable?
☑ d. Is the function field EMPTY for all features?

9. Describe any issues or specific genes that you were unable to satisfactorily resolve, and warrant further inspection in the Quality Control review.

There is a gap between genes 3 and 4 that contains coding potential. Initially, it was thought that this could be an additionally added gene, however when looking through the Frames window there were no valid stop codons. Our rationale is that this was once a gene, is no longer functional, and over time has lost its stop codon.

Gap between bps (47929-48333) - This is a significant gap in the genome, however there is no coding potential in the GeneMark map. In this Frames region there is one potential start and one potential stop. However, due to the lack of coding potential we can determine that this region does not contain a gene.