**Genome Annotation Submission Cover Sheet
Pre-QC Phage Genome Annotation Checklist**

 **Phage Name:**  Eyes

**Your Name:**  Michael Lauer, Orges Alabaku, Hanna Brubaker, Devin Dikec, Madison Hanaway

**Institution:**  University of Pittsburgh

**Emails:**  MJL157@pitt.edu.edu, DAD165@pitt.edu, MAH368@pitt.edu, HAB91@pitt.edu, ORA6@pitt.edu

**Other Emails:**  Instructor - Rebecca Bortz [RLB6@pitt.edu]

* 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 Section 1.4 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 Eyes\_FinalCompleteNotes.dnam5 file:

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

For the Eyes.dnam5 file:

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

Describe any issues or specific genes that you were unable to satisfactorily resolve, and warrant further inspection as the instructor reviews your genome:

* Gene 8: Head-to-tail connector complex protein, HHPred his >90% probability (95.81%) to SPP1 15 Bacillus (6492-6875 bp)
* Gene 29 (24,305-26,248Fwd) Called Minor tail protein, hydrolase domain based on alignment with other phage in phagesdb and strong hits to Hydrolase domain using HHpred and CDD and hit to singleton Finch\_29
* Gene 30 (26251-27021Rev) Many HHpred hits over 90% probability to multiple functions such as esterase, hydrolase, and lipase – Lipase called based on Phamerator hit to Bowser\_30
* Genes 32-33: Over 200 base pair gap but no genes present in range
* Gene 33 (29318-28272Rev): called as HTH DNA Binding Domain from HHPred/CDD, but NCBI BLAST suggests XRE transcriptional regulator
* Gene 35: 30005-30118 gene deleted - Overlap, small short gene. Not called by GeneMark.
* Gene 37: 31605-31916 New gene added – coding potential present and removes a 488 bp gap, called by NCBI BLAST- NKF
* Gene 41: (33167-33919) called NKF but BLAST hits aligned with Smoothie and ClubL to Toxic cation resistance protein not on list of approved functions. HHpred and CDD hits to vWA (von Willebrand factor type A)
* Gene 59: 42425-43786 New gene added – reduces a gap by 1420 bp (large gap), called by NCBI BLAST and Phamerator for function – HNH Endonuclease