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

 **Phage Name:**  *Stultus*

**Your Name:**  Lauren Callewaert, Rebecca Cohen, Phillip Korth, Manisha Kukkillaya, and Jenna Li

**Your Institution:**  University of Pittsburgh

**Your E-mails:**  LAC192@pitt.edu, RLC103@pitt.edu, PAK88@pitt.edu, MAK385@pitt.edu, JEL180@pitt.edu

**Additional E-mails: 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 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 Stultus\_FinalCompleteNotes\_11-28.dnam5 complete notes file:

* a. Have any duplicate genes (or any with the same stop coordinate?) been removed?
* b. Does every gene have 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 auto-annotation?

For the Stultus\_Minimal\_11-28.dnam5 minimalistic 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 large 335 bp gap before gene 68 (53408bp) but there is only a small amount of low coding potential along the gap and open reading frame would only allow for a gene less than 100bp. Therefore there is no reason to believe that there might be an additional gene within the gap.
* Genes 7, 10, 18, 19, 37, 56, 61, 70, and 73 are all short genes with less than 200 base-pairs in length, but they cover all of their respective coding potential.
* Genes 28, 42, 53 and 59 are not annotated at the longest ORFs, nor do they capture all of their respective coding potential, but moving those genes to earlier start codons would result in significant gene overlaps.
* Gene 16 (11,631-11,849bp) and 54 (45,669-46,175bp) had functional calls on HHpred that were not on the approved functional call list (ribosome modulation factor and replisome organizer, respectively) and therefore the genes should be reviewed for the possible gene function to ensure that the functional call is valid for a phage before assigning such calls.