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

 **Phage Name:**  Kiko

**Your Name:**  Mira Patel, Cierra Jarmon, Leigh Coonelly, Natalia Zakachi,
Danny (Chen-Jung) Yu, Judy Kim

**Institution:**  University of Pittsburgh

**Emails:**  MKP32@pitt.edu, CPJ7@pitt.edu, LEC116@pitt.edu, NAJ42@pitt.edu, CHY75@pitt.edu, JSK76@pitt.edu

**Other Emails:**  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 Kiko\_CompleteNotes.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 Kiko.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 34 (25895-26950 reverse): The predicted function by BLAST is **acetyltransferase**. This function, however, is not on the approved list of functional calls. When BLASTed this gene has many functional hits to acetyltransferase in many bacterial species and *Gordonia* phage Schwebeltier\_30. HHpred search also has hits to acetyltransferase with probability greater than 90% and evalue=-22.
* Gene 36 (27509-28771 reverse): The predicted function by BLAST is **fic family protein** but this is not on the approved list of functions. The functional calls made by HHpred had probability greater than 90%, coverage greater than 85%, and evalue=-35. BLAST hits also support fic family protein with similarity to many bacterial species.
* Gene 40 (bp 31778-31443 reverse) was added in a large gap that contained an open reading frame and a low amount of coding potential from GeneMark host model.
* Gene 42 (32442-32807): The predicted function by HHpred is **chromosome partitioning protein PARB** and is on the approved list of functions but PARA was not also discovered.
* Gene 50 (bp 37007-37117) was deleted due to overlap and lack of supporting evidence.
* Gene 56 (39215-39607): The predicted function by HHpred is **ATP-dependent DNA helicase** which is not on the approved list of functions. It is suggested that this gene be reviewed in greater detail in order to confirm a functional call.
* Gene 59 (bp 40835-41173) was deleted due to overlap and lack of supporting evidence.
* Gene 61 (42578-43237): The predicted function by HHpred is **transcription initiation domain** and is not on the approved list of functions. P>90%, cover=19%, e=0.69