Genome Annotation Submission Cover Sheet

Phage Name: **Niza**

Your Name: **Claire Rinehart**

Your Institution: **Western Kentucky University**

Your email: **claire.rinehart@wku.edu**

Additional emails: (For correspondence) **bobby.gaffney@wku.edu**

Pre-QC Phage Genome Annotation Checklist

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? **Yes**

2. Are all the genes “valid” when you click the “validate” button? Section 9.3.2 **Yes**

3. Have the genes been renumbered such that they go sequentially from 1 to the highest **Yes**

number? Section 9.3.3

4. Have all old BLAST hits been cleared, and all gene features reBLASTed? Section 9.3.4 **Yes**

5. Are the locus tags the phage name? Section 9.3.3 **Yes**

6. Has the Documentation been recreated to match the information in the feature table? Section 1.4 **Yes**

7. Have tRNA ends been adjusted with web-based Aragorn and/or tRNAscan SE? Section 9.5.3-4 NA

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? **Yes**

b. Does every gene have one and only one complete set of Notes (see fig 12.2 in the

Annotation Guide)? **Yes**

c. Do the functions in the Notes match the official function list? **Yes**

d. Is the function field EMPTY for all features? **Yes**

e. Do the notes contain the initial Glimmer/GeneMark data from the autoannotation? **Yes**

For the YourPhageName .dnam5 file:

a. Have any duplicate genes (or any with the same stop coordinate?) been removed? **Yes**

b. Is the Notes field empty for all the features with no known function? **Yes**

c. Do the function names in the Notes match the official function list, when applicable? **Yes**

d. Is the function field EMPTY for all features? **Yes**

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

CDS 17038 - 17409

 /note=ribosomal slippage

 /note=join(16610..16963:16963..17409)

CDS 28623 - 28943

 /note=Matches PhrostyMug\_33 (BYU, 2016) and Lesedi\_32 (KwaZulu-Natal, 2012). No supporting evidence for MuL-like function.

CDS complement (32888 - 32983)

 /note= PhrostMug\_41: RDF questionable? No support. (BYU direct submission)

CDS complement (37855 - 37995)

 /note=Questionable RDF sources? Matches to PhrostyMug\_49 (BYU), Turj99\_49 (NCBI), BPBiebs32\_51 (KwaZulu-Natal).

CDS complement (41147 - 41311)

 /note=Single match to Switzer\_57 (KwaZulu-Natal, 2012). No other evidence to support.

CDS complement (41902 - 41988)

 /note=TMHMM predicted one transmembrane domain, 7-24

CDS complement (44054 - 44182)

 /note=PhrostyMug\_65: RDF. Questionable function. Lack of other evidence.

CDS complement (46294 - 46689)

 /note=Matches Repressor from BillKnuckles\_68 and TheloniousMonk\_76 at 96% identity but has no other support similar to the next repressor protein. Is it possible to have two repressors or is one perhaps a pseudogene?

CDS complement (47505 - 47804)

 /note=Questionable source for RDF, PhrostyMug\_79 (BYU). TMHMM does predict a transmembrane domain 5-27 at 93% probability.

CDS complement (47797 - 47976)

 /note= PhrostyMug\_80: RDF protein. Lack of support, questionable source (BYU).

CDS complement (48439 - 48627)

 /note=TMHMM predicts a transmembrane domain between 33-52 with a 71% probability.

CDS complement (48627 - 48794)

 /note= Lack of support: RDF, PhrostyMug\_85, and questionable source (BYU).

CDS complement (48813 - 48962)

 /note=Questionable support: HNH endonuclease, PhrostyMug\_88 (BYU)

CDS complement (50890 - 50973)

 /note=TMHMM predicts one membrane domain, 5-24.

CDS complement (51083 - 51286)

 /note= TMHMM predicts one membrane domain, 7-29.