**Cluster AK Arthrobacteriophage Nubia**

Tamarah L. Adair, Caroline M. Addison, Christina A. Gaw, Catherine Corinne N. Hughes, Roshni T. Jaffery, Pranav Kapoor, Navya Katragadda, Alex A. Muñoz, Christian N. Sessa, Lathan G. Lucas, Micheal C. Munson, Jennifer A. Wilson, Ashley M. Young,

60 genes, 5 reverse genes, 3 deletions, 0 insertions, 7 extensions, 0 shortenings

**Gene start choices:**

**Gene 4** was extended to cover coding potential, improve SD and ORF. However, the original start agrees with Starterator, Glimmer, GeneMark, and results in Q1:S1 for PhagesDB and NCBI. **Gene 5** was extended to improve SD and ORF. **Gene 16 and 16 (originally 17)** were called separately and modified for a frameshift mutation. **Gene 19 (originally 20)** was extended to improve Z-score and SD value. **Gene 26 (originally 27)** was extended to agree with Starterator, shorten gap, and have better SD and Z-score. **Gene 27 (originally 28)** was extended to to have smaller gap and better blast results. **Gene 33 (originally 34)** was extended to improve ORF. **Gene 57 (originally 60)** was extended to agree with Starterator and have longest ORF.

**Functions:**

**Gene 5:** terminase, large subunit, **Gene 6:** portal protein, **Gene 7:** capsid maturation protease, **Gene 8:** major capsid protein, **Gene 9:** capsid decoration protein, **Gene 11:** head-to-tail connector protein, **Gene 16:** tail assembly chaperone, **Gene 17:** tape-measure protein, **Gene 18:** minor tail protein, **Gene 19:** minor tail protein, **Gene 20:** minor tail protein, **Gene 21:** minor tail protein, **Gene 26:** HTH DNA binding protein. MerR-like, **Gene 28:** DNA helicase, **Gene 41:** DNA helicase, **Gene 43:** RecA-like exonuclease, **Gene 45:** DNA primase/polymerase, **Gene 47:** DNA primase/polymerase, **Gene 48:** adenylosuccinate synthetase, PurA like, **Gene 51:** hydrolase, **Gene 53:** helix-turn-helix DNA binding domain, **Gene 60:** HNH endonuclease

**Gene 43 (originally gene 44)** called for a RecA-like exonuclease through PhagesDB (E-value=0.0) and NCBI (E-Value=0.0) and RecA-like exonuclease is called in Nubia’s homologs in Phamerator. This function is not listed.

**Gene inclusion/exclusion:**

**Gene 45 (originally gene 47)** was interesting because it completely overlapped original genes 46 and 48. Gene 46 was a reverse gene and gene 48 was a forward gene. Neither gene had good hits from NCBI, PhagesDB, and HHPred. Both were deleted. The remaining Gene 45 was extended to cover coding potential and agree with Starterator.

At **Gene 16**, it is most probable that a frameshift occurred. By comparing Phamerator’s results for Nubia’s Gene 17 to other AK phages at Gene 17, it was determined that the slippage in Nubia occurred at base pair 411 after the start of Gene 16 (which began at base pair 11830). By subtracting 1 from the start of Gene 16 (11830) and then adding 411 determined the 12240 bp “spot” for where the frameshift occurred. Therefore, it was decided unanimously to merge Gene 16 and 17 into one continuous Gene 17 (renaming it as Gene 16.1 and Gene 16.2 on DNA Master, essentially deleting Gene 16).

The rest of the genes were confidently annotated.