CONSIDER FOR TALK

2021 SEA Symposium Abstract

Hope College

Holland MI

Corresponding Faculty Member: Joseph Stukey (stukey@hope.edu)

Mycobacteriophage PenguinLover67: A New Member of Subcluster B9

Quinn Baar, Natnael Belay, Kayla Carpenter, Alexis Erickson, Therese Joffre, Sophie Klanseck, Lydia Konings, Ellie Marion, Rachel Mast, Jenna Mustapha, Madeleine O'Donnell, Tasia Parsons, Allison Ziemba

Fourteen new mycobacteriophages were isolated from soil samples collected around the state of Michigan and parts of the United States. All phages were capable of infecting Mycobacterium smegmatis and were isolated through either enrichment or direct plating at 37°C. A variety of plaque morphologies were produced based on size, shape, and clarity. The mycobacteriophage PenguinLover67 was chosen as one of two phages for complete genome sequencing and comparative genomic analyses. The predominant plaque produced by PenguinLover67 after 24-48 hours at 37°C was clear with well-rounded edges and approximately 1 mm in diameter. The complete genome sequence for PenguinLover67 showed it was similar to mycobacteriophages of cluster B, subcluster B9, which now contains 4 sequenced members. PenguinLover67 is most similar to the B9 phage CRB2. It has a genome size of 70,299 bp, making it the smallest of the B9 mycobacteriophage genomes, and it contains 97 protein-encoding genes but no tRNA or tmRNA genes. PenguinLover67 contains several interesting genomic features including repeats of a conserved 17-bp sequence (consensus: GCGGGA(A/C/T)(C/T)G(C/G)CCCGCCG). The repeated sequence is found upstream to 8 different genes on the right arm of the genome and at distances of 14-25 bp to the translation initiation sites and 3-8 bp to the likely ribosome binding sites (AGGAG). Five of those repeats also have an additional 26-bp partially conserved sequence nearby. All repeat sequences are located in intergenic sequence space. The conserved genomic positions and specific orientations of the repeat sequences relative to the translation initiation sites of adjacent genes suggests these may have a role in gene expression. The 17-bp sequence element also happens to be found as direct repeats flanking PenguinLover67 gene 90. The gene 90 protein is just the third member of Pham 5667; the other two Pham 5667 genes are found in Quesadilla (B9) and Thonko (B8). Interestingly, all three Pham 5667 genes are found in different locations in the three phage genomes and all are flanked by direct repeats of the conserved 17-bp sequence element. HHpred analysis of all three gene products showed a strong match (full protein lengths, probability ≥ 96.9) to the conserved domain of unknown function, DUF1840. PenguinLover67 also contains 4 ORPHam genes. Comparative analysis suggests their presence may impact expression of adjacent genes. In all 4 cases, the quality of the Ribosome Binding Site (RBS) for the adjacent gene is much stronger in comparison to the RBS scores for the corresponding genes in the similar B9 phage, Quesadilla, which lacks all 4 ORPHam genes. We continue to explore these and other genomic features of the new B9 mycobacteriophage, PenguinLover67.