CONSIDER FOR TALK

7th Annual SEA-PHAGES Symposium Abstract

Gettysburg College

Gettysburg PA

Corresponding Faculty Member: Veronique Delesalle (delesall@gettysburg.edu)



Alex Agesen



Maddi Strine

Phamished and the Cluster B Phages

Alex Agesen, Sarah Brantley, Elizabeth Burton, Jenna DeCurzio, Jenny Giannini, Isabella Katsanos, Aphra Murray, Julia Palmucci, Amanda Richman, Sarah Rivera, Maddi Strine, Breanna Titchen, Bobby Weisensee, Rachel Wigmore, Veronique A Delesalle, Greg P Krukonis

The unique mycobacteriophage Phamished was isolated from the Gettysburg College campus, sequenced at the University of Pittsburgh, and assigned to cluster B1. Phamished, with a genome of 68,515 bp, 66.5% GC, 101 putative genes with an average length of 649.7 bp, and no tRNA genes, is typical of its subcluster. Phamished has no unique genes; all of its genes belong to gene families (“phams”) found in other B1 phages.

A gene-by-gene analysis of amino acid composition in Phamished revealed six ORFs with disproportionate use of a particular amino acid. Five ORFs (genes 11, 15, 19, 74, and 79) had more than 20% alanine. ORF 32, a minor tail protein, contained 24.8% glycine and was the only one of these ORFs with a putative function. Given its steric flexibility, glycine may be important in tail fiber attachment. In support of that hypothesis, we found that glycine was the most prevalent amino acid for at least one minor tail protein in other mycobacteriophages (*e.g*., C1 Gizmo, D2 Hawkeye) and unrelated phages (*e.g*., phage lambda, *Salmonella* phage Chi).

The protein product of the cluster B-specific gene 52 (pham 380) fell within the PD-(D/E)XK nuclease superfamily (HHPred probability of 99.0%). This conclusion was supported by Vilnius University’s Institute of Biotechnology PDEXK server, which assigned this protein to the AddAB-type helicase-nuclease complex with at least a 97.9% probability, a complex with similar functions to *E. coli*’s REC BCD.

As in other B1 phages, we added an HNH endonuclease gene, not called by Glimmer or GeneMark, at position 10. Two forms of the pham (2895) exist: a 435 bp ORF found in the majority of sequenced B1 phages and a 159 bp segment of the longer ORF found in four B1 phages. The longer ORF has a GC content of 60.5% which differed from the short version (67.9%) and the surrounding genes (68.2%). We explored whether the short version arose via a deletion in the front half of the long ORF with a loss of endonuclease function, or whether a homing event with the HNH endonuclease motif cleaved into the DNA producing the long ORF.

Start codon analysis of all B phages with available DNAMaster files showed that the frequencies of ATG, GTG, and TTG starts differ significantly between subclusters. Although ATG was the most prevalent start codon across all five subclusters, its frequency ranged from 57% to 71%. GTG was used less frequently in subcluster B1 and TTG was always the least used. Other studies have shown that modifying the start codon in a given mRNA affects protein expression levels, with ATG starts conferring the highest expression levels. Interestingly, all B tapemeasure genes start with GTG while the capsid genes in pham 9976 use both ATG and GTG codons. The variation in start codon usage within a pham suggests that factors besides control of gene expression may affect for start codon selection. The way selection operates on start codons is an area worthy of further exploration.