CONSIDER FOR TALK

8th Annual SEA-PHAGES Symposium Abstract

University of Maine, Fort Kent

Fort Kent ME

Corresponding Faculty Member: Peter Nelson (peter.nelson@maine.edu)

New graphical perspectives on genomic mosaicism using multivariate analysis of Cluster A

Zach Beal, Joseph Enerva, Peter R Nelson

Bacteriophages contain an abundance of information within their respective genomes but graphical display of many genomes at once is challenging. The most numerous cluster of bacteriophage isolated from Mycobacterium smegmatis, Cluster A, offers a wide variety of genomes to evaluate new comparative genomic tools. Using the multivariate analysis method known as hierarchical clustering, a two-way cluster analysis was produced that simultaneously displays associations between Cluster A bacteriophages based on phams and grouping phams based on their occurrence in different phages. The two-way cluster dendrogram correctly recovered the each phage’s subcluster membership. The graphic also distinctly displays highly conserved phams amongst phages in Cluster A, including nine phams found in every subcluster A phage and seven phams in nearly all subclusters. These highly conserved phams code for essential structural and functional components of phages such as portal, tape measure, and head-to-tail interface proteins. However, these highly conserved phams were not found in continuous stretches of the genome. In contrast, 558 phams occurred only once and 126 phams occurred only twice across all cluster A phages. Two-way clustering provides a different way to visualize the mosaic character of phage genomes at the level of pham conservation within subclusters.