CONSIDER FOR TALK

2021 SEA Symposium Abstract

University of California, Los Angeles

Los Angeles CA

Corresponding Faculty Member: Jordan Parker (jordan.p.parker@kp.org)

Utilizing Comparative Genomic Analysis to Define Superclusters as a Method for Understanding Intercluster Phage Diversity Across Different Host Types

Dominic Garza, Vivian Nguyen, Nathan Wobig, Andrew Kapinos, Canela Torres, Amanda C Freise, Jordan Moberg Parker

Phage diversity exists along a “continuum”; as such, current methods of clustering phages obscure similarities that exist between phages of separate clusters. Implementation of superclusters, a higher order phage organization system, could help capture these intercluster relationships by highlighting genomic similarities between different clusters of phages that infect separate host types. While prior literature has employed the concept of superclusters, the definition remains inconsistent. Therefore, this study strived to provide a clearer definition of superclusters by identifying potential genomic characteristics of supercluster relationships. A number of comparative genomic tools such as Phamerator, Gene Content Similarity (GCS) Calculator, Average Nucleotide Identity (ANI) Calculator, and GEPARD-generated nucleotide dotplots were used to identify genomic patterns among supercluster members that could be connotative of supercluster relationships. This analysis identified four superclusters comprising phages from different clusters infecting a variety of hosts such as *Arthrobacter*, *Gordonia*, *Microbacterium*, *Mycobacterium*, *Rhodococcus*, and *Streptomyces*. All these superclusters’ members were revealed to share intercluster GCS values near or surpassing the 35% clustering threshold despite having intercluster nucleotide sequence similarity confined to short, well-conserved genomic regions. The analysis of each supercluster group’s shared genes revealed that most had functions related to either phage structure or DNA/RNA processing. Phages are the ‘dark matter’ of the biological world, so refining systems of phage classification such as superclusters will help increase our understanding of phage diversity and phages’ impact on the world.