CONSIDER FOR TALK

8th Annual SEA-PHAGES Symposium Abstract

Florida International University

Miami FL

Corresponding Faculty Member: Mauricio Rodriguez-Lanetty (rodriguezlanetty@gmail.com)

Bias on gene strand location mirrors Mycobacteriophages genetic clusters

Samuel Salazar, Michael Herrera, Mauricio Rodriguez-Lanetty

Mycobacteriophages infecting the bacterial host Mycobacterium smegmatis comprise an extensive majority of sequenced phage genomes. As part of an ever-expanding roster, newly isolated mycobacteriophages are categorized into groups known as clusters, which are further divided into sub-clusters. This process of clustering phages is performed via comparison of whole genome sequence similarity. Noticing differences in the number of ORFs on the lagging strands in specific mycobacteriophages, we sought to determine the correlation between cluster and the number of ORFs found on the lagging strand in a genome. Our results indicate that strand orientation may serve as a quick and reliable method for cluster prediction. The term “strand orientation” here refers to the preference of ORFs to be located on the leading (forward) strand or lagging (reverse) strand of viral dsDNA, and their relative use compared to the other strand. In order to determine a relationship between strand orientation and cluster, 174 mycobacteriophage genomes were chosen to represent a span of 48 different subclusters. Using Phamerator, the total number of reverse strand ORFs for each selected phage were counted and used to calculate the % of reverse ORFs and Forward/Reverse ratio (F/R) in each genome. The analysis showed that similarity in reverse strand orientation is high correlated to current whole-genome clustering classification. Phages with similar percentage of reverse ORFs and F/R ratio were routinely classified in the same cluster or subcluster. Presence of reverse ORFs in coding regions ranged from a maximum of 73% in cluster A to as low as 0% in multiple, smaller clusters. Regions of the genome coding in the reverse direction were also noticed to be highly conserved among phages of the same cluster. Interestingly, clusters A and B, which together comprise >50% of all M. smegmatis phages, demonstrate the greatest use of their reverse strand. These findings contribute to expanding the overall understanding of mycobacteriophage genomic diversity and cluster relationships. Further research may include insight into the advantages and selective pressures that may be affecting greater lagging strand usage in highly represented clusters.