DO NOT CONSIDER FOR TALK

11th Annual SEA Symposium Abstract

University of Maryland, Baltimore County

Baltimore MD

Corresponding Faculty Member: Steve Caruso (scaruso@umbc.edu)



Michelle L Guldan

Of clusters and mischief: the Streptomyces phageome

Kristina E Atanasoff, Ashley Batista, Courtney L Colson, Gregory Furletti, Stephen Goralski, Michelle L Guldan, Matthew S Kane, Jenni Kelleher, George Kyaw, Sarah V Lance, Dorjan Leka, Shirin Ludwig, Victoria Mahaney, Phoebe Mariano, Tanzeem Naqvi, Thomas O'Brien, Sana Rana, Chih Shiang K Shee, Andrew L Simpson, Shreya Singh, Ryland N Spence, Racheal A Spruill, Lowen E Wachhaus, Claudia L Xie, Ivan Erill, Steven M Caruso

The 2018 UMBC Phage Hunters successfully isolated, characterized and sequenced the genomes of six bacteriophages infecting two hosts: the phytopathogenic *Streptomyces scabiei* and heavy metal-tolerant *Streptomyces mirabilis*. Five of the phages are *Siphoviridae*, including: two BB2 phages, Heather and RemusLoopin; a BG phage, Mischief19; a BI2 phage, PherryCruz; and a singleton, RosaAsantewaa (accession numbers MK686069-MK686072). The sixth phage, Forthebois, is a *Tectiviridae* and is in cluster BO (accession number MK620900). Analysis of the annotated genomes with Phamerator, Geppard plots, and the Gene Content Similarity (GCS) metric revealed that the cluster assignment for some of these phages may need to be revised. For instance, RosaAsantewaa is currently classified as a singleton, but shares 60% GCS with subcluster BI2. Conversely, re-assigning Mischief19, currently annotated as a BG phage, as the first member of subcluster BG2 would significantly decrease the dispersion of GCS values within cluster BG. To gain further insight into the accuracy of these putative reassignments, we explored the use of two highly conserved structural genes (those coding for the tape measure and terminase proteins, often syntenic) for phylogeny-based cluster assignment. The tape measure protein (TMP) generates noisier Gepard plots than the terminase protein due to the presence of internal repeats, limiting its applicability for clustering purposes. Further analysis enabled us to identify the repeat motifs in several clusters, and phylogenetic analysis also confirmed that repeats have an impact on alignment and subsequent tree inference. Using the tree distance for clustering, we observe that the results essentially recapitulate the established clusters, with some notable exceptions. Most notably, the TMP of Mischief19 bears little resemblance to that of its assigned cluster, supporting again that Mischief19 should probably define its own subcluster BG2. This phage displays a large insert between its capsid and tail fiber genes and almost half its genome is composed of orphams. Our bioinformatic analysis also detected the presence of a predicted protein in RosaAsantewaa with a partial match to the coliphage superinfection exclusion protein B (PF14163). This predicted protein appears to be conserved among BI phages, which have been consistently annotated as lytic. Examination of the BB2 subcluster, including phages Heather and RemusLoopin, revealed a high degree of conservation in this subcluster and the likely acquisition of a *thyX* gene in one of the subcluster phages (Sebastisaurus). We discuss the possible role of superinfection exclusion proteins in lytic phages, and of thymidylate synthases in lysogenic ones.