DO NOT CONSIDER FOR TALK

2022 SEA Symposium Abstract

University of Ottawa

Ottawa

Corresponding Faculty Member: Adam Rudner (arudner@uottawa.ca)

Characterization of Arthrobacter globiformis Bacteriophages Crewmate, Lizalica, Cassia, ObiToo, Warda and JohnDoe using Phylogeny.fr and Gene Content Similarity (GCS) programs.

Celine Fawagreh, Danya Alomar, Danyaal Ansari, Jaina Patel, Jinane Elhage, Laila Alfadhli, Naveed Shahzada, Omar Ferri, Tanvi Ahluwalia, Sarah Lapolice, Katia Koziel Ly, Ciara Heffernan, Astha Chandra, Yasmine Elmi, David Jung, Gapisha Karunakaran, Tala Salaheddin, Rama Sarakbi, Erika Znamenski, Julia Douglas Freitas, Kanchi Baijal, Liam McCarthy, Elizabeth C Williams, Lisa D'Ambrosio, Keith Wheaton, Kin Chan, Adam D Rudner

Six AZ cluster bacteriophages, Crewmate, Lizalica, Cassia, ObiToo, Warda and JohnDoe were isolated by students at the University of Ottawa. The phages were isolated using *Arthrobacter globiformis* B-2979, a bacterium found in soil that serves as the host these phages. All six bacteriophages were obtained from a variety of soil and organic material samples in Ottawa, Canada.

The selected phages belong to the Siphoviridae family. Siphoviridae is a family of double-stranded DNA viruses characterized by long, non-contractile, and flexible tails. They differ in their mechanisms of DNA packaging and genome organization. The head is typically 60 nm in diameter and the tail, which is made up of six subunits, is 150 nm in length.

Upon successfully purifying and isolating the phages, Illumina sequencing at the University of Pittsburgh was utilized to sequence the DNA of all six phages. Student groups at uOttawa used PECAAN, GeneMark, Phamerator, HHpred, NCBI databases, PhagesDB, and Starterator to complete the functional annotations of all the genomes.

This project aims to characterize the bacteriophages using a robust phylogenetic analysis software, Phylogeny.fr, and Gene Content Similarity (GCS). These are comparative genomics techniques that provide more insight on key genomic similarities and differences.

Gene Content Similarity (GCS) is a measure of near and distant phage relationships. The calculated value refers to the number of phams shared between the genomes being compared, divided by the total number of phams in the genome. This value is then determined for each of the two genomes in the pair being compared, and is averaged to produce a final value. Based on initial gene content similarity percentages, Lizalica and Warda are the most similar to one another as they have the highest gene content similarity value, and the most shared phams. Lizalica and ObiToo are the least similar as they have the fewest shared phams and the lowest gene content similarity values.

Phylogeny.fr is a program that will be used to reconstruct a phylogenetic tree based on the protein sequences of the genes of interest. The phylogenetic trees will reveal inter-relationships between the six selected phages in regards to the genes of interest.