CONSIDER FOR TALK

2023 SEA Symposium Abstract

University of South Florida

Tampa FL

Corresponding Faculty Member: Richard Pollenz (pollenz@usf.edu)

Annotation of Azira a CT cluster phage that infects G. rubripertincta

Elijah Algarin-Martinez, Victoria Correa, Don Danielson, Wynter Dean, Maria Denipoti Cavasini,, Joseph French, Naleena Gaskin, Urvi Jain, Jeanette Janvier, Breanna Macumber, Francesca Martini, Sydney Mazzei, Coen McGarrah, Jennifer Mujica, Oslow Odegaard, Corissa Quarterman, Taylor Rand, Nina Seidensticker, Tyler Serrano, Anna Soltys, Max Ungrey

Bacteriophages require host metabolic pathways to replicate and survive. Azira is a CT cluster phage and infects *Gordonia rubripertincta*. The Azira genome was annotated using DNA Master, NCBI BLAST, HHpred, PhagesDB, and Phamerator. Azira is a lytic phage that has a genome of 45,336bp and contains 67 protein coding genes. During the annotation, several thymine metabolic enzymes were identified including: deoxycytidylate deaminase (gp44), thymidylate synthase (gp46), dUTPase (gp49) and thymidylate kinase (gp50). These genes are organized in an “operonic” structure suggesting they function together. Interestingly, there are two genes of unknown function (gp47 and 48) between these different enzymes. Within CT cluster, there are 38 phages with a similar genome structure. Thymidylate synthase is found in all of the cluster CT phages. However, three phages do not have dUTAPase, but have an annotated Maz G nucleotide pyrophosphohydrolase that may be able to function like a dUTAPase. Polynucleotide kinase was also annotated instead of thymidylate kinase in 19 out of 38 CT phages. All CT phages have several NKF genes within the identified operon. Since phages do not carry out metabolism and utilize nucleotides produced by the host bacteria, it is intriguing to find a series of metabolic genes in a phage genome. These enzymes have important functions in humans. The enzyme dUTPase has been shown to help to maintain low intracellular dUTP/dTTP ratios that prevents extensive mutations resulting in double strand breaks. It is also interesting that the four genes identified in Azira do not constitute the full thymine synthesis operon and that there are two genes of unknown function within the identified “operon” that may encode proteins involved in the pathway. Perhaps these enzymes play a role in modifying the thymine used in the synthesis of the phage genome as a defense to bacterial enzymes. Lysis cassettes are composed of endolysins such lysin A, lysin B, and transmembrane proteins (TM) called holins. Azira’s lysis cassette is composed of a split lysin A, containing genes that encode proteins with an L-ala-D-glu peptidase domain and another containing a glycosyl hydrolase domain. These genes are followed by three that encode a 4TM protein, a 2TM protein, and a 1TM protein that appears to have an N-terminal signal sequence and suggests it may be secreted. It is unknown how these TM proteins interact to lyse the bacteria, especially if one protein is secreted, but the finding of a distal protein with a single TM domain is consist with other cassettes containing multiple TM genes. Azira also features a downstream lysin B gene located in the middle of a metabolic operon described above. The future analysis of this cassette may lead to greater understanding of the lysis pathway in Gram positive bacteriophages. Collectively, this research will further the understanding of phage evolution and fitness. The SEA PHAGES consortium is acknowledged for their support.