DO NOT CONSIDER FOR TALK

2021 SEA Symposium Abstract

Worcester Polytechnic Institute

Worcester MA

Corresponding Faculty Member: JoAnn Whitefleet-Smith (jaws@wpi.edu)



Erin Bryan



Sophia Strano

Genes of Unknown Function

Erin Bryan, Sophia Strano, Robert Belmonte, Ulises Cardano, Domenica Ferrero, Monet Norales, Ilaria Wernick

At the beginning of this laboratory course, we analyzed the draft genomes of two hypothetical bacteriophages, ManAs and Abigail. After doing a thorough and complete analysis of the genomes as they were presented on PECAAN, the team found many genes of unknown function to still be present, particularly within the Abigail phage. The goal of this project was to identify the functions of genes that were previously labeled as having “No Known Function”, or NKF. The genes that were picked for this project were genes that were not only labeled NKF but were also adjacent to other genes that did have functions.   
We hypothesized that, based on data from PECAAN and information about the synteny of the phage genome, inferences can be made about the function of these undetermined genes. Specifically, we hypothesized that the hypothetical function of these genes is strongly influenced by the functions of genes neighboring their start and stop within the genome. We chose to specifically examine genes of unknown function that were located between two genes of a known function for this purpose, such as gene 7 in the Abigail phage genome, which is located between a minor capsid protein and a major tail protein.   
Tools that were used for this investigation were mainly independent runs of programs that were collectively run in PECAAN, the site in which the genes in question were first labeled NKF based on the PECAAN-collected evidence. These independently run programs included NCBI BLAST, NCBI SmartBLAST, NCBI COBALT, Vista, HHPred, Phamerator, PhagesDB, SOSUI, and Conserved Domain Database.   
 Research that could be explored afterward was based on early observations from the project. Independent runs of NCBI SmartBLAST and NCBI COBALT turned up alignments for Abigail gene 57 that were not shown in the evidence presented by PECAAN’s bioinformatics programs. This could be an interesting line of research to pursue, as independent program runs were identifying potential function calls that PECAAN was not. A way to research these relatively inconclusive calls in the future would be benchwork involving the Abigail gene, which is not something that the team has the resources to do in the timeframe of this class. An additional question to be explored is how the functions of the NKF genes could be investigated if independent runs of the programs encompassed by PECAAN did not result in any new information or potential function calls to pursue.