Corazon Genome Annotation Project

Eung Baek kim  
*Purdue University*, kim2353@purdue.edu

Eunhui Yoo  
*Purdue University*, yoo95@purdue.edu

Mai Liu  
*Purdue University*, liu1938@purdue.edu

Chao Fu  
*Purdue University*, fu186@purdue.edu

Yejin Jang  
*Purdue University*, jang103@purdue.edu

Follow this and additional works at: https://docs.lib.purdue.edu/purc

**Recommended Citation**

kim, Eung Baek; Yoo, Eunhui; Liu, Mai; Fu, Chao; and Jang, Yejin, "Corazon Genome Annotation Project" (2019). *Purdue Undergraduate Research Conference*. 66.  
https://docs.lib.purdue.edu/purc/2019/Posters/66

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Abstract
Corazon genomes were annotated by establishing relationship of Corazon and other phages to understand the overall genomic architectures, running automated gene prediction proteins and functional data on the predicted gene, and reviewing the prediction and making necessary changes to delete or identify any missing genes. The group annotated the mycobacterium Corazon genes 33-48. Based on the evidence supported by the programs, most of the Corazon genes from 33 to 48 were structural genes, responsible for the formation of phage structure. Gene 47 is a frameshift gene which a nucleotide reads more than once or omitted to have two genes in the same space.

Introduction/Background
Mycobacterium phage, Corazon, was found at Lafayette College in Easton, Pennsylvania in 2017. Its plaque were small, round and clear with siphoviridae morphology type. Research questions regarding this research include the function of the gene section, and methods to analyze the gene.

Processes to obtain our result
Start Sites
1. RBS
2. Blast Result
3. Phamrator
4. Starterator

Function
1. Phamrator
2. NCBI Blast
3. HHPred
4. Phagesdb

Flow chart of process

Result
1. Gene Functions
   - Portal Protein
   - Capsid Maturation Protease
   - Scaffolding Protein
   - Head-to-tail Connector Complex Protein
   - major capsid protein
1. Frameshift was present in Gene 47
   - The same start site between Gene 46 & 47
   - Gene 47 overlaps Gene 46
   - Function : tail assembly chaperone.

Applications of our findings
- Gene therapy
- Phage-based technology

Summary
Corazon genome was annotated by DNA master, NCBI, HHpred, Phamrator. The results from those analytical softwares gave evidence to identify the type of function of the gene, and an interesting result was found throughout the analysis of the genome. The analysis showed that Corazon has a protein related function. Also, there was a frameshift on gene 46 and 47 of Corazon showing the overlap between those two genes.

Conclusion
The assigned regions of genes were found related to the phage capsid formation. Additionally, Gene 47 is a frameshift which has a function of tail assembly chaperone. Since Gene 47 and 46 start at the same base-pair, Gene 47 overlaps the part of Gene 46.

Acknowledgement
Howard Hughes Medical Institute, Science Education Alliance Biotechnology Innovation and Regulatory Science Center, Purdue University Hindley Bioscience Center, Purdue University Polytechnic Institute, Purdue University Department of Agricultural and Biological Engineering, Purdue University
Emily Kerstiens, Gillian Smith, Mikael Reuhs, Joey Krampen, Jacob Riedel, Sarah Bell, Janice Chan, Christina Sanchez, Austin Larson