An Investigation of a Cluster S Mycobacteriophage Genome, Corazon, Genes 4-16: Location and Function

Marina Mehling
Purdue University, mmehling@purdue.edu

Eli Cline
Purdue University

Remington Wilson
Purdue University

McKenna Pineau
Purdue University

An Investigation of A Cluster S Mycobacteriophage Genome, Corazon, Genes 4-16: Location and Function

Eli Cline[1], Marina Mehling[1], McKenna Pineau[1], Remington Wilson[1], Ikenna Okekeogbu[1], Kari L. Clase[1]

[1]Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN 47907

Abstract

The purpose of this investigation is to establish the presence, location, and function of genes within the genome of a novel bacteriophage, Corazon and add to the Actinobacteriophage Database. The Corazon genome was analyzed with the program DNA Master as well as NCBI BLAST, HHPred, and Phamerator to determine the location and function of genes within the auto-annotated range of genes 4-16. Only one of these genes was assigned a function, and the investigation of significant gaps in the genome resulted in an additional gene being added.

Introduction

Mycobacteriophages are a type of virus that specifically attack mycobacteria. This attribute can be exploited to fight antibiotic resistance mycobacteria. As of February 2019, only 14 Cluster S types (a specific group of mycobacteriophage) have been completely sequenced and published in the Actinobacteriophage Database. The Corazon genome was analyzed to determine the accurate locations and functions of 13 auto-annotated genes and one additional manually discovered gene. The main goal of this investigation was to prepare a high-quality annotation of the phage genome. The published work will be used for future investigations.

Contributing to the global understanding of bacteriophages is of interest since the phage-bacteria model has expanded scientists' capabilities of studying evolution and exploring novel medical applications. Publishing these annotations will allow generations of researchers to compare their results to this member of Cluster S and potentially identify a new candidate for phage-mediated transduction, phage therapy, or other application.

Materials and Methods

Procuring A Phage
- A dirt sample was obtained
- A phage was isolated and a high concentration phage lysate was made

Analyzing the Phage
- The phage was viewed through electron microscopy
- DNA analysis was done through Restriction Digest

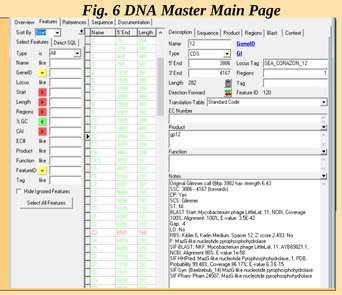
Analyzing Phage Genomes
- Using several programs genes' locations and functions were determined

Genome Analysis
- Auto Annotating the phage genome using the program DNA Master

Confirming gene start site locations and gene functions using
- GeneMark, Starterator, Phamerator Maps, NCBI BLAST, PhagesDB BLAST, HHPred, and Ribosomal Binding Sites scores

Figure 6 shows the program used to evaluate the genes, The Green text shows each gene numerically and its base pair position. The bottom right hand corner shows the notes we found from researching the gene.

Fig. 6 DNA Master Main Page



References and Acknowledgements

[1]LietzkeMehling. (2018). Retrieved from https://phagesdb.org/phages/LietzkeMehling/
[2]LittleLaf. (2018). Retrieved from https://phagesdb.org/phages/VLittleLaf/
[3]Phamerator. (n.d.). Retrieved March 29, 2019, from https://phamerator.org/phages
[4]VasuNzinga. (2018). Retrieved from https://phagesdb.org/phages/VasuNzinga/

Corazon was entered into the Actinobacteriophage Database by Kimberly Mae Manalang at Lafayette College and sequenced at the Pittsburgh Bacteriophage Institute. Genome annotation was performed at Purdue University with guidance from Dr. Kari Clase, Ikenna Okekeogbu, and Emily Kerstiens. Procedures for isolation and analyzation were provided by the Howard Hughes Medical Institute Science Education Alliance. Also recognized are the Biotechnology Innovation and Regulatory Science Center at Purdue University, Bindley Bioscience Center at Purdue University, Polytechnic Institute, and the Department of Agricultural and Biological Engineering at Purdue University.

Results

Procuring a Phage

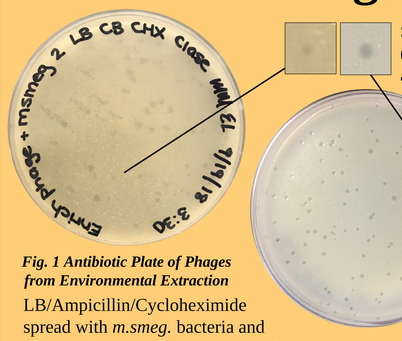


Fig. 1 Antibiotic Plate of Phages from Environmental Extraction
LB/Ampicillin/Cycloheximide spread with *m.smeg.* bacteria and soil sample extraction. Presence of circular clearings in bacterial lawns denote the presence of plaques.

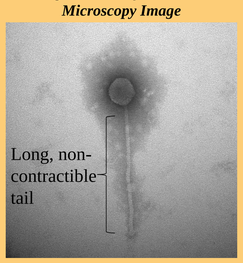
Small, circular plaques (1-3 mm diameter) are associated with lytic phage.

Fig. 2 Beelzebub (cit) Plate with Purified Plaque Morphology
LB/Ampicillin/Cycloheximide spread with *m.smeg.* bacteria and lysate (purified plaque sample). Presence of consistent plaque morphology suggests a single phage type is isolated.

Characterizing the Phage

Fig. 3 LietzkeMehling[1] Electron Microscopy Image

Fig. 4 LittleLaf[2] Electron Microscopy Image



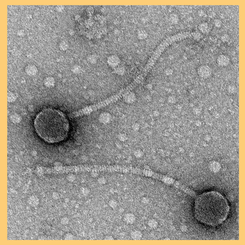


Fig. 5 VasuNzinga[4] Electron Microscopy Image

LietzkeMehling, LittleLaf, and VasuNzinga are all Siphoviridae phages. They are characterized by having long, non-contractible tails and geometric, rounded capsids (heads).

Bioinformatics

Fig. 6 Corazon Genes 4-16 Start Sites

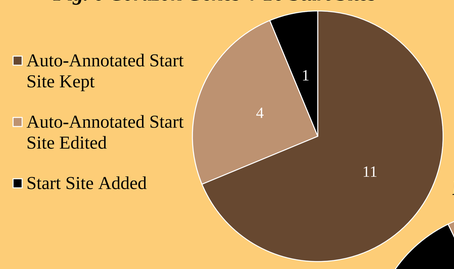


Fig. 6 Denotes how the start sites of genes were modified or added during annotation.

Fig. 7 Summarizes the function calls of the 14 total genes annotated

Fig. 7 Corazon Genes 4-16 Function

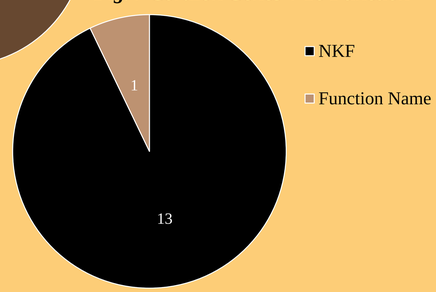


Figure 7 shows the BLAST results show genes from other phages who are very similar to the selected gene. The BLAST results are analyzed for similarity to determine the function and startsite of an unknown gene based on this information from known, similar genes. The other similar genes are shown in the top left corner, and various scores describing similarity are shown on the bottom right.

The Phamerator map (Figure 8) compares the genome of Corazon with several similar phages, including Beelzebub, MosMoris, LittleLaf, and Marvin. The similar colored boxes correspond to similar genes. The purple regions show strong correlations between genomes.

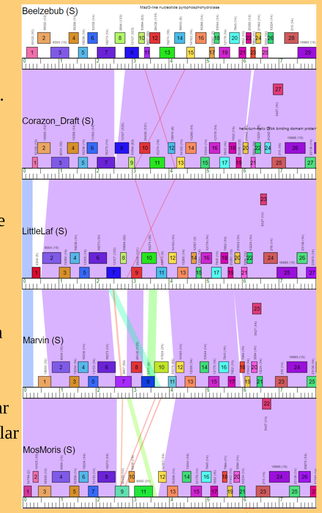


Fig. 8 Phamerator Map[3]

Conclusions and Future Directions

Including *m.smegamitis* bacteria on the initial extraction plates and subsequent attempts to plate an isolated plaque morphology (Figures 1 and 2) ensure that the bacteriophage present are mycobacteria. The presence of small, circular plaques observed on the plates (Figures 1 and 2) suggest that the phage particles quickly lyse their host bacteria, which is characteristic of a lytic life-cycle. The electron microscopy images (Figures 3-5) were used to characterize their respective phages as members of the Siphoviridae family. Siphoviridae have long, non-contractible tails and geometric, round capsids. Cluster S is a family of phage genomes that encompasses Siphoviridae mycobacteria with a lytic life cycle. The wet-lab portion of this investigation successfully isolated a phage type from an environmental sample and characterized it to be a member of Cluster S.

Subsequent DNA sequencing allowed the genome of a Cluster S member, Corazon, to be obtained. Investigating Genes 4-16 of Corazon with bioinformatics tools revealed the phage was highly similar to the phages Beelzebub, LittleLaf, Marvin, and MosMoris in regards to their genes' locations and functions. The investigation uncovered that a gene had to be added to the genome due to a significant gap in genes and coding potential. The study found only gene 12 in this group had a function. There were no other significant discoveries in this study.

The initial hypothesis was rejected since the annotated Corazon genes had no known function (Figure 7). A future investigation could involve protein analysis on bacteriophage proteins to potentially characterize their functions, discover what type of interactions they perform within phage, and relate the activity of phage to its expression level.