Isolation and Genomic Analysis of Mycobacteriophages TopGun and ScoobyBlue, as well as Gordoniaphage TinaLin

Abigail T. Peterson, William Lara, Oscar Gonzalez, Nick Hillar, Tina Hailer, Jarelin Huliganga, Angelica Villa-Reyes, Kanoe Saffery, Tina Slowan-Pomeroy, Laura A. Briggs

Truckee Meadows Community College

Science Education Alliance · HHMI

Phage Hunters Advancing Genomics and Evolutionary Science

ABSTRACT

The goal of this research was to isolate, purify, and characterize bacteriophages found in Northern Nevada soils. In conjunction with the Howard Hughes Medical Institute's Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program, this research expands our understanding of the diversity of bacteriophages in this region.

This study focused on three different bacteriophages; two Mycobacteriophages, Topgun and ScoobyBlue, and one Gordoniaphage, TinaLin. Topgun and ScoobyBlue were isolated on the host *Mycobacterium smegmatis mc2155*. Phage TinaLin was isolated on the host *Gordonia terrae NRRL B-16283*. Topgun and TinaLin were found in 2019 and ScoobyBlue was discovered in 2017. Each phage was isolated and purified by the plaque purification method until uniform plaques were obtained. Phage DNA was extracted from each lysate and sent for sequencing at the Pittsburg Bacteriophage Institute, followed by annotation using PECAAN and Phamerator.

TopGun is in the mycobacteriophage A1 subcluster with a 63.9% GC content and a genome length of 50,977 base pairs with 91 genes and no tRNAs. ScoobyBlue is a subcluster A3 mycobacteriophage with a GC content of 63.9% and a genome length of 49,638 base pairs with 88 genes and two tRNAs. TinaLin is a subcluster CU1 Gordoniaphage with a 65.4% GC content and a genome length of 43,366 base pairs with 75 genes and one tRNA. TinaLin shows 82.71% and 82.71% sequence similarity to CU1 phages Splinter and Vendetta, respectively. TopGun shows 99.74% sequence similarities to A1 phages Wilkins and CactusRose, respectively. ScoobyBlue shows 99.11% and 98.87% sequence similarity to A3 phages HelDan and Fred313.

The genome arrangement of ScoobyBlue is unlike any other A3 mycobacteriophage, making it unique among its genomic cluster which is an area of interest for future research. ScoobyBlue can cross infect a non-pathogenic strain of *Mycobacterium tuberculosis* H37Ra. TinaLin also is a point of interest due to it being one of the few Gordoniaphage that TMCC has discovered. Future research could include host range analysis on these three phages to explore the differences between the phage captured at TMCC to other phages in the same cluster to find genomic differences.

REFERENCES

1. Bittel, Jason. "Phage Hunters." *Pitt Magazine*, 2019, Fall, www.pittmag.pitt.edu/news/phage-hunters.
2. Pope, Welkin H., et al. "Expanding the Diversity of Mycobacteriophages: Insights into Genome Architecture and Evolution." Ed. Ramy Aziz. *PLoS ONE* 6.1 (2011): E16329. Web. 22 Apr. 2015.
3. A.B. Monk. *et. al.*; Bacteriophage applications: where are we now?. Lett appl Microbiol. 2010 Oct; 51(4): 363-369. Published online 2010 Aug 26. doi: 10.1111/j. 1472-765X.201002916.x
4. Li, Q. *et. al.*; Mycobacteriophage SWU1 gp39 can potentiate multiple antibiotics against Mycobacterium via altering the cell wall permeability. *Sci. Rep.* 6, 28701; doi: 10.1038/srep28701 (2016)

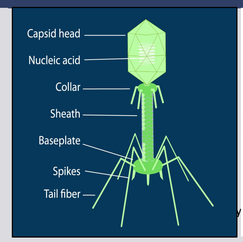
ACKNOWLEDGEMENTS

Thank you to NV INBRE for funding this research and to the Howard Hughes Medical Institute for inclusion in the SEA-PHAGES program. A very special thank you to our mentors Tina Slowan–Pomeroy and Dr. Laura Briggs for their continual patience and guidance throughout our research.
Funding Source – NSF #1118679
Nevada IDeA Network of Biomedical Research Excellence (INBRE)
HHMI Science Education Alliance - Phage Hunters Advancing Genomics & Evolutionary Science (SEA-PHAGES)

Nevada INBRE | IDeA Network of Biomedical Research Excellence

INTRODUCTION

Bacteriophages (phages) are viruses that infect and kill bacteria. Bacteriophages are the most abundant entities in the biosphere (1). It is estimated that there are approximately 10^{31}, which is roughly about 10 million more than there are stars in the universe (1). Phages are host specific and a single phage can only infect a small range of bacterial hosts. Bacteriophages consist of a capsid ("head") that contains genetic material (usually dsDNA) and a tail, which allows the phage to attach and inject its genetic material into the host. (Figure 1). There are two different life cycles a phage may take, lytic or lysogenic (Figure 2).After the phage injects its nucleic acid into the host bacterial cell, phages in the lytic life cycle will replicate using the host's replication machinery and lyse the bacterial cell, releasing hundreds of phage particles that can go on to infect new bacterial hosts.The lytic cycle presents as clear areas, called plaques, on a bacterial lawn. In the lysogenic cycle, the phage genome will integrate into the host chromosome and lie dormant while replicating along with the bacterial host cell - this presents as turbid, or cloudy, plaques. Mycobacteriophages are bacteriophages that infect and kill bacteria from the genus *Mycobacterium*, to find TopGun and ScoobyBlue a popular mycobacterial host, *Mycobacterium smegmatis (M. Smegmatis) mc^{2}155* was used. Gordoniaphage are bacteria in genus *Gordonia terrae*, to find TinaLin the host *Gordonia terrae NRRL B-16283* was used. In this study, the two Mycobacteriophages TopGun and ScoobyBlue were analyzed along with the Gordoniaphage TinaLin.These phages were found in soil in northern Nevada and annotated at Truckee Meadows Community College. TopGun is a sub-cluster A1 with 63.9% GC content. ScoobyBlue is a sub-cluster A3 with 63.9% GC content. TinaLin is a sub-cluster CU1 with a 65.4% GC content. Genome annotation was done at TMCC. Annotation is important for figuring out which genes are in the genome and using that information we can begin to understand which genes cause certain phages to behave in the way they do.



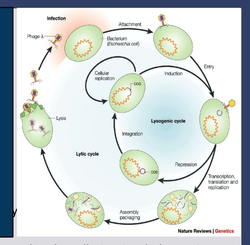


Figure 1) Illustration of bacteriophage showing all structural elements.
Figure 2) Illustration of lytic vs lysogenic cycles

METHODS AND MATERIALS

In Situ - Isolation, Purification, DNA isolation and Restriction, Transmission Electron Microscopy:
Mycobacteriophage ScoobyBlue was isolated from an enriched soil sample collected at 39.54621 N, 119.72121 W, Mycobacteriophage Topgun was isolated from an enriched soil sample collected at 39.5727 N, 119.797 W, and Gordoniaphage TinaLin was isolated from an enriched soil sample collected at 39.571297 N, 119.79599 W; all three were found in Reno, NV. Purification of the two mycobacteriophages was completed by serial dilution onto Luria Agar (LA) plates overlaid with *M. smegmatis mc^{2}155* and Top Agar (TA). Plates were incubated for 24 hours at 37°C, the same steps were done with TinaLin with the exception that the host *Gordonia terrae NRRL B-16283* was used. These steps were repeated until pure phage was acquired. A lysate of sufficiently high titer was obtained for all three phages: 2.37E7 PFU/mL for TopGun, 2.0E8 PFU/mL for TinaLin, and 2.7E10 PFU/mL for ScoobyBlue. DNA was isolated from each phage at a sufficient concentration for sequencing. Restriction analysis of phage genomic DNA and agarose gel electrophoresis was performed to analyze integrity.

In Silico - Genome Sequencing and Bioinformatics Analysis:
Annotations of TopGun, ScoobyBlue, and TinaLin were performed using PECAAN and phamerator. Putative ORFs were blasted using PhagesDB, NCBI, and HHPRED. Once start sites and gene function calls were assessed and annotations were completed, Phage DNA for TinaLin and Top Gun were then sent for sequencing at Pittsburgh Bacteriophage Institute and ScoobyBlue was sent to North Carolina State Genomic Sciences Laboratory using Illumina Sequencing.

CONCLUSIONS

Using PECAAN and phamerator, undergraduate research students were able to completely annotate mycobacteriophages ScoobyBlue and TopGun as well as Gordoniaphage TinaLin. Topgun and TinaLin have been sent for final consideration by Pittsburgh University for submission to GenBank. ScoobyBlue is an A3 mycobacteriophage, one of 98. The genome of ScoobyBlue is 49,638bp in length with a GC content of 63.9%, 88 genes, and 2 tRNAs. TinaLin is one of 7 CU1 Gordoniaphage. The TinaLin genome length is 43,366bp with 65.4% GC content, 75 genes, and 1 tRNA. Mycobacteriophage Topgun is an A1 subcluster, one of 163; the genome length of Topgun is 50,977bp with a GC content of 63.9% and no known tRNA. TinaLin has one known orpham, an unique gene that hasn't been found in another phage's genome yet, and TopGun has 2. No orphams were found in ScoobyBlue. Many of the usual genes are present in these phages, however the function and purpose of many genes remain unknown and further research is needed.

ScoobyBlue is a point of interest due to its unusual genome arrangement compared to other A3 mycobacteriophages. The lysis cassette (Lysin A, Lysin B, and Holin genes) controls the timing of the lysis whereby the replicated phage particles can burst out of the bacterial host cell. ScoobyBlue's lysis cassette is roughly 7 genes earlier than other A3 phage. The effect of having a lysis cassette this early in a A3 phages genome is undiscovered and research will need to be done to see what this difference means in terms of functionality This could also imply that the 7 genes that ScoobyBlue is "missing" have no effect on how the phage functions as a whole. Also ScoobyBlue can cross infect *Mycobacterium tuberculosis* H37Ra which seems to be a common trait among the A3 phages. Host range analysis would be the next step to see how well ScoobyBlue can infect *M. tuberculosis* as well as if it can infect other hosts.

A total of 47 functional genes were discovered in Topgun and a majority of the functions were of a structural nature. For example, many genes were found to either be minor tail proteins or assembly chaperones. Gene #36 was found to be a serine integrase which was a unique find amongst the 93 genes annotated. Serine integrase is a bacteriophage enzyme that is responsible for the integration/insertion of viral DNA into the host's genome to begin a lysogenic life cycle for the virus.

A total of 12 bacteriophage genomes have been characterized by TMCC undergraduate students and 10 of these 12 have been published in GenBank. We anticipate many more novel phages to be discovered, annotated, and published in GenBank from Nevada.

RESULTS

	ScoobyBlue	TinaLin	Topgun
Bacterial Host	*Mycobacterium smegmatis*	*Gordonia terrae*	*Mycobacterium smegmatis*
Sub-Cluster	A3	CU1	A1
Length (bp)	49,638	43,366	50,977
GC Content (%)	63.9	65.4	63.9
Total Genes	88	75	92
tRNA	2	1	-

Table 1: Shows the standard data for bacteriophage. Details on the host, cluster, length, GC content, number of genes present and tRNAs help the annotation process.

Table 2 (on the right): Displays the commonly found genes in bacteriophage and their functions

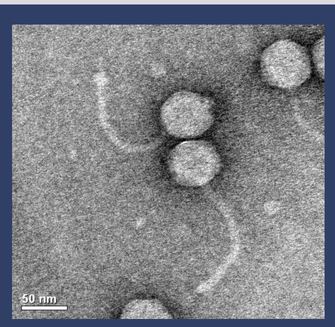


Figure 3) A EM picture of a mycobacteriophage
Firgure 4) Plating of phage TopGun. The plaques indicate areas where the bacteria were lysed.

Gene Product	Gene Function
Terminase, Small Subunit	ATP-driven motor that packages genetic material into the capsid of the phage
Terminase, Large Subunit	ATP-driven motor that packages genetic material into the capsid of the phage
Portal Protein	Aids in phage DNA entering host cell
Capsid Maturation Protease	Required for DNA replication
Scaffolding Protein	Required for phage capsid assembly
Major Capsid Protease	Part of phage head capsid assembly
Head to Tail Connector	Assembly of phage head capsid to phage tail
Major Tail Subunit	Assembly of phage tail
Tail Assembly Chaperone	Aids the assembly of phage head capsid to phage tail
Tapemeasure	Determines phage tail length
Minor Tail Protein	Part of phage tail assembly
Holin	Aids in phage DNA exiting the host cell
Lysin A	Degrades host cell wall during final stages of lytic cycle
Lysin B	Degrades host cell wall during final stages of lytic cycle
Nucleiotidyl transferase	Adds 3' terminial CCA to tRNA; require for protein synthesis
Nrdh glutradoxin	Catalyzes the reduction of ribonucleotide reductase
Lsr2 bridging protein	Cell wall biosynthesis; antibiotic resistance
Integrase	Integration of phage genome into host genome & excision out
Immunity Repressor	Prevents transcription of lytic genes; maintain lysogeny
CRO protein	Prevents transcription of lytic genes; maintain lysogeny
Excisionase	Aids integrase in integration of phage genome into host genome & excision out
DNA Primase	Invloved in DNA replication; type of RNA polymerase
DNA Helicase	Separate ddDNA into single strands; required for DNA replication
Whi B Transcription Factor	Transcriptional regulator; DNA-binding properties
DNA Methylase	Adds methyl groups to DNA, represses gene transcription
SS DNA Binding Protein	Protects SS DNA from being digested by nucleases
RNA Ligase	Enzyme that catalyzes ATP into AMP, releasing energy
ClpP Protease	Enzyme that cleaves proteins
DNA Binding Protein	Protects DNA from being digested by nucleases
Endonuclease	Aids terminase in packaging DNA in phage capsid
CMP Hydroxymethyl Hydrolase	Nucleotide hydrolase
DNA Polymerase	Required for DNA replication
RecA DNA Recombinase	Essential for the repair and maintenance of DNA
Ribonucleotide Reductase	Enzyme required for DNA synthesis and repair
Exonuclease	Catalyzes the excision of nucleotides from DNA strands
ATPase	Enzyme that catalyzes ATP into ADP, releasing energy
Transposase	Enzyme required for transposon excision and integration