Purification and Analysis of Mycobacteriophage Alice

Coreen M. Manley, Stephanie E. Simon, Dr. Lee E. Hughes, Dr. Robert C. Benjamin

University of North Texas Biological Sciences
1155 Union Circle #305220
Denton, TX 76203
Abstract

The purpose of this research is to expand the knowledge of mycobacteriophage and analyze a single mycobacteriophage genome to be archived for future use. Mycobacteriophages are found in soil almost anywhere with an abundance of host mycobacteria. The analysis of mycobacteriophage may help lead to breakthroughs in DNA recombination and help the scientific community in the search for a cure for genetic diseases as well as diseases caused by mycobacteria including tuberculosis. In this study, a single phage, Alice, was isolated from a soil sample found in a compost pile. The sample was directly plated using *M. smegmatis* as the host organism. After the plaques were verified to be formed by phage, a series of dilutions to isolate and purify the phage were performed. The final product from the dilutions was a high titer lysate that, when prepared, was used to extract the DNA from the phage which was then sequenced and analyzed. Compared to other organisms, mycobacteriophage are relatively unknown. Only about 60 have been sequenced, so the range of comparison for conservation in Alice’s genome is very slim. In this study we hope to increase the scientific community’s knowledge of mycobacteriophages and how they affect the areas in which they are found.

Introduction

Mycobacteriophages infect bacteria that live in soil and are found in almost any climate that will support said bacteria. They are harmless to humans which makes them excellent to study. The host bacteria used in this study was the mycobacterium *M. smegmatis* in a late log/early stationary phase 48 hour culture. A phage’s primary purpose is to infect a host and replicate using the host’s organelles. Eventually the new phage will become too numerous for the cell to hold and it will lyse, killing the cell and releasing the new phage to the environment where they can then find new hosts to infect.
Using the knowledge of how the phage infects a bacterium, it could then be possible to use phages to insert DNA into an organism’s genome to or to study similarities in other organisms as well as many other scientific uses. The phage D29 was used as a control. Mycobacteriophages are found almost anywhere on the planet but there are only a few known mycobacteriophages that have been identified and sequenced. Most mycobacteriophage are relatively small, consisting on average of about 50 kilobase pairs (kbp). Alice is comparatively large for a mycobacteriophage, measuring roughly 150kbp. There are several interesting aspects regarding the size of the phage and the genome that will be discussed later. Though there are a few similar mycobacteriophage similar to the one discussed here, Alice is, at the moment, relatively unique. In preliminary findings using the gene recognition software GeneMark and Glimmer, Alice shows to have an unusually high amount of tRNA, another unique trait of this phage. In purifying and sequencing this phage, we hope to contribute to the scientific community’s knowledge of phage and how we can use them to fight disease and in genetic research.

Methods and Materials

Most of the methods used in this research are referenced from the NGRI Phage Laboratory Manual. There are some modifications which will be noted. The phage D29 is used throughout these procedures as a control.

Purification of the phage. Since mycobacterium and mycobacteriophages live in soil, two separate soil samples were obtained in the hopes that there would be bacteria and phage living in at least one of them. The first sample was taking from a location near some mushrooms, thus the sample’s name, Mushroom. The other, Garden, was taken from a compost pile. In order to determine the existence of phage in Sample Mushroom, an enrichment process was
preformed. Sample Garden was not enriched. It was soaked in a phage buffer and then the supernatant was directly plated. Using sample Garden, a spot test, as stated in the NGRI Phage Laboratory Manual was preformed to ensure the presence of phage. Unless stated otherwise, the phage or plaques from here on will be from sample Garden. Next phage purification is used to isolate a phage. A ten fold serial dilution of the sample is preformed as stated in the NGRI Phage Laboratory Manual. Second and third purifications are preformed as described previously with the exception of the third, where two separate plaques where tested.

**Harvesting of a High Titer Lysate solution.** Using plates from the final purification, flood with phage buffer(PB) and harvest the mixture of PB and phage. Follow procedure as shown in the NGRI Phage Laboratory Manual for the filter sterilization technique and the spot test of the lysate. This is the medium titer lysate. A titer calculation is required for further steps. To make a phage stock from a ten plate lysate, the calculation of the number of plaque forming units (pfu’s) required for a web pattern is needed as well as the average area of plaques and plate sizes. Plate ten identical plates using top agar(TA) and *M. smegmatis* then flood plates with PB. Aspirate the liquid after three hours and transfer to a conical tube. Using steriflips, vacuum sterilize this sample. This is the final titer lysate. This lysate is then used in a final dilution, an infection of *M. smegmatis* with $10^{-6}$ through $10^{-11}$. This final test of the lysate ensures its purity and a high enough quantity of DNA for testing.

**DNA and Phage Analysis.** Using the high titer lysate, preform a DNA preparation as described in the NGRI Phage Laboratory Manual with the following modifications: Using a 1000μL pipette, aspirate excesses supernatant away from the pellet after draining the it. Also use the 1000μL pipette to mix the resin and pellet mixture thoroughly by aspirating and expelling the liquid. When filtering the resin/pellet mixture through the columns, use five columns instead of
two. Keep the purified DNA on ice or in the freezer. Using the purified genomic phage DNA, a restriction digest was performed using 0.8μL of DNA and the following enzymes: BanHI, ClaI, EcoRI, HaeIII and HindIII. Prepare the gels using the procedure in the NGRI Phage Laboratory Research Manual. Electrophorese the gel at 100V for approximately 35 minutes. Also using the phage DNA prepare an electron microscopy grid. Refer to the NGRI Phage Laboratory Research Manual for procedure, with the following modification: allow phage to soak in for two minutes, then begin the washes.

**Genomic Structure Analysis.** Once the genome had been sequenced, several software programs were used to predict genes and other structures in the genome. GeneMark and Glimmer were the two prediction softwares used, and Apollo was used for editing.

**Results**

Since soil is where the host bacteria of mycobacteriophages live, two soil samples were collected over the course of the experiment for study. The first sample, Mushroom was gathered in a grassy lawn where a small colony of mushrooms had proliferated. The soil is watered regularly and in part sun and was damp when collected. This sample was used for the enrichment procedures. Enrichment is used to help boost the amount of phage in a sample thus increasing chances for phage to appear when the enriched filtrate is plated. Sample Mushroom did not yield any plaques. This could be due to human error, contamination of the sample after it was taken, or contamination of the soil with chemicals such as pesticides and herbicides designed to kill unwanted bugs, weeds, and pathogens such as bacteria. If pesticides killed off all the mycobacteria, there would be very few if any mycobacteriophages due to the lack of a host, thus making an enrichment very difficult to do.
The second soil sample, Garden, was collected from a compost pile in a garden where organic practices are used. The pile from which the soil was collected is about six months old and consists of decaying fruits, vegetables, and other miscellaneous plant matter. The soil was damp and warm. Only top soil was collected since most bacteria is aerobic and cannot survive deeper than one or two inches, therefore no phage would be deeper than one or two inches. The filtrate from this sample did not undergo any type of enrichment, and was directly plated. The direct plating of Garden yielded plaques, indicating, but not verifying the presence of phage. The spot test of the direct plate verified the presence of phage.

Sample Mushroom has been discarded and now only sample Garden will be considered. Since the spot test was positive for phage, the next step is to isolate and purify on phage out of that spot test. A total of four plaque purifications were preformed to isolate the phage from sample garden. On the second purification, there were two different sized plaques, one small, about 1mm across, and another larger one, about 1.5mm across. Two separate purifications were preformed, one on each plaque to ensure that only one phage was creating the plaques. The resulting plates had both size plaques present, indicating that the phage creates both size plaques.

Using the $10^{-4}$ plate (Fig. 1) from the final purification, flood the plate with phage buffer. This buffer will suspend the phage, pulling them out of the agar so they can be collected. The objective is to have enough phage stock that can be empirically tested for phage. A spot test
must be performed next to calculate the amount of plaque forming units per milliliter. A spot test (Fig.2) yielded a titer of $4 \times 10^{10}$ pfu/mL.

Next, a phage stock must be prepared from the lysate. Approximately ~1846 pfu are needed to cover a plate in a web pattern and $4.6 \times 10^{-8}$ mL of phage stock was needed to accomplish the desired pattern. To get to that amount, another dilution is necessary which resulted in $500\mu$L of $10^{-6}$ dilution. Ten identical plates were infected with this dilution. All ten plates resulted in web patterns. All ten plates were flooded with PB for 2 to 3 hours. This is done to allow the maximum amount of phage to be soaked up out of the agar by the PB. After the PB/phage mixture is harvested, it has to be centrifuged and vacuum sterilized to filter out any contaminants such as dead bacterial cells or agar. Once the phage is purified, it must go through one more dilution to find the final titer. The final lysate’s titer was $11.9 \times 10^{10}$ pfu/mL.
The DNA preparation was done to discard debris in the lysate that was not phage and break down the capsid to get to the phage DNA. The first attempt at filtering the resin/phage mixture through the columns did not work. There was most likely still too much debris for it to be pushed through the filters. The second time five columns were used to reduce the amount of filtrate pushed through each column. The product of the DNA prep was about 53.0 μL of phage DNA. This was kept on ice to prevent denaturation and degradation of the DNA. The total yield of DNA was approximately 33.3 μg and the concentration was 628.4 μg/μL. This DNA is used in the restriction digests. The restriction digests use digestion enzymes from bacteria to break down the phage DNA and allow for an estimate at the genome size by calculating the size of the fragments (Fig. 4a and 4b). The calculated genome size of phage Alice was approximately 82,890
bases long. This was calculated by graphing the log fragment length of the ladder in base pairs (bp) against the distance of the band’s migration (cm). The equation of the line of best fit was then used to calculate the fragment length of each band on the gel. The fragment sizes were then added up to find an approximate estimate of the phage’s genome.

Electron microscopy is used to determine the size and appearance of the phage (Fig. 5).
<table>
<thead>
<tr>
<th>Tube #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rxn Buffer 10x</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
</tr>
<tr>
<td>DNA</td>
<td>0.8 µL</td>
<td>0.8 µL</td>
<td>0.8 µL</td>
<td>0.8 µL</td>
<td>0.8 µL</td>
<td>0.8 µL</td>
</tr>
<tr>
<td>10 x BSA</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
<td>2 µL</td>
</tr>
<tr>
<td>Restriction enzyme</td>
<td>none</td>
<td>Bam HI 0.5 µL</td>
<td>Cla I 2 µL</td>
<td>Eco RI 0.5 µL</td>
<td>Hae III 1 µL</td>
<td>Hind III 0.5 µL</td>
</tr>
<tr>
<td>dd H₂O</td>
<td>15.2 µL</td>
<td>14.7 µL</td>
<td>13.2 µL</td>
<td>14.7 µL</td>
<td>14.2 µL</td>
<td>14.7 µL</td>
</tr>
<tr>
<td>Total</td>
<td>20 µL</td>
<td>20 µL</td>
<td>20 µL</td>
<td>20 µL</td>
<td>20 µL</td>
<td>20 µL</td>
</tr>
</tbody>
</table>
Figure 5
Electron microscopy of Alice. The capsid is approximately 75nm wide and the tail is only 50 nm long.
After the restriction digest, the remaining DNA was split, half sent for archiving for future study and use, and the other half was sent to be sequenced. From the sequencing we learned that Alice has an unusually large genome (for a mycobacteriophage), more than 150 kilobase pairs long. Thus far, it has been structurally analyzed, but no functional analysis has been attempted. Using Glimmer and GeneMark to predict the genes, Apollo for gene editing and viewing software, BLAST for nucleotide and amino acid comparison as well as other Howard Hughes Medical Institute softwares, the preliminary structural decisions on Alice have been made. There are over 120 putative genes and twenty-four tRNAs that have been found.

**Discussion**

In order to isolate a single type of phage, soil was collected because the host organism of mycobacteriophage are mycobacteria. Plaque purifications were performed to isolate one phage out of many possible phages in the soil sample used. The plaques formed from phage Alice were very small though the phage itself is large. This indicates that the burst size of phage Alice is much smaller than that of the control phage D29, whose plaques were much larger than Alice’s. Due to the clarity of the plaques, it is most likely that Alice is a lytic phage. The burst size is one of many unique properties of Alice. This phage also produces two slightly different sized plaques as shown during the plaque purification series, though the reason for this is unknown.

The high DNA concentration probably indicates that there is a relatively large genome and large amount genomic material in one phage. The restriction digest confirms this theory as does the EM photos. Alice is a large phage with a capsid diameter of approximately 75-100nm across and with a short tail of 25 to 50 nm. The large head correlates with the small burst size. Not many fully constructed Alice phage would be able to fit into a *M. smegmatis* bacterium,
causing it to lyse much faster than a bacterium infected with D29 phage. Also, most of the energy of phage production most likely goes into the formation of the capsid and genetic information. The length of the tail is probably due to this fact. A large capsid indicates a large amount of genetic information. The initial estimation Alice’s genome size made after a restriction digest was performed. Each band on the gel represents two cuts made on the DNA at specific sequences. Once the bands are measured and put into the equation of the line of best fit for the ladder, the genome size can be calculated. Alice’s estimated genome size was approximately 82,890 base pairs long, relatively large for the type of phage. After the sequencing, it was found that that estimate was off by quite a bit. The final size of the genome was almost double of the original estimate.

The size of the genome and the phage itself is very intriguing. The purpose of a phage is simple: infect the host to produce more phage. Most phage DNA consists of a blueprint which is inserted into the host’s genome and then the host produces clones of the original phage which eventually burst out of the cell and infect the bacteria around it. Due to the size of Alice’s genome, it is very likely that there is more than just a simple blueprint hidden in this phage’s DNA. A large phage genome can mean many things. This particular phage may have genes for resisting certain enzymes designed to destroy it, or the ability to infect more types of bacteria. The length of the genome could also make it more easily manipulated by researchers in order to insert or delete multiple genes or to discover how the genes interact with each other. The large amount of tRNA present indicates that Alice may be able to infect a wider range of bacteria, increasing its survivability in an environment.

Acknowledgements
I would like to thank the Howard Hughes Medical Institute for putting on this research project and the University of North Texas for hosting and providing the lab. I would also like to thank Stephanie Simon, Dr. Lee Hughes, and Dr. Benjamin for all their work in making this lab happen and for the stimulating intellectual discussions held in and outside the class and for teaching me the techniques necessary in laboratory work. I would also like to thank my classmates for the help and work they put in during the structural analysis of Alice.

**Funding**

Funding provided by University of North Texas and the Howard Hughes Medical Institute.

**References**

Howard Hughes Medical Institute. 2009. NGRI Phage Laboratory Manual.