



# Isolation, Characterization, and Annotation: The Search for Novel Bacteriophage Genomes

Samantha Hawtrey<sup>1</sup>, Lori Lovell<sup>1</sup>, and Rodney King<sup>2</sup>

Student<sup>1</sup>, Mentor<sup>2</sup>: The Gatton Academy of Mathematics and Science

at Western Kentucky University, Department of Biology  
Bowling Green, KY 42101-1080

## Abstract

**Bacteriophages are the most common DNA-containing entities on earth, yet very few have been characterized. The purpose of our research was to increase knowledge of phage biodiversity by isolating and characterizing two novel mycobacteriophages from the environment and analyzing the genome of two other novel bacteriophages. Mycobacterium smegmatis was used as the host strain. This microbe is a common inhabitant of soil and is harmless to humans. Soil samples from our hometowns were used to enrich for mycobacteriophages. The newly isolated phages were named Luke117 and Novus. To complete phage analysis, large numbers of phage particles were prepared. The morphology of the phage was determined by electron microscopy, and the capsid and tail were measured. Genomic DNA was isolated from the purified phage particles and analyzed by DNA restriction analysis, spectrophotometry, and gel electrophoresis. In the spring of 2011, we annotated portions of two recently sequenced phages. Using a variety of annotation programs, we examined each potential gene to decide whether or not it was real. Both of these genomes will be published on GenBank's database, further expanding our knowledge of bacteriophage genomics and gene function.**

## Introduction

The names of bacteria-based illnesses, such as leprosy, tuberculosis, and cholera, are familiar to everyone. Harmful and sometimes even lethal, bacteria are known to cause a wide variety of sicknesses. A significant challenge is finding ways to cure these diseases. Although antibiotics have been used to control bacterial diseases, bacteria readily develop resistance to common antibiotics<sup>1</sup>. However, there is another potential way to control harmful bacteria: bacteriophages (Fig. 1). Named after the Greek term “phagein” (meaning “to eat”), bacteriophages inject

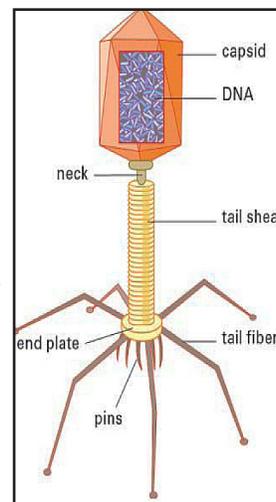
their DNA into bacteria and cause lysis. As viruses, phages are not living, but they contain genetic material. Upon encountering a host bacterium, a phage attaches to the host, penetrates the cell membrane and injects its DNA. The genetic information can then follow two potential paths. In the lysogenic cycle, the DNA is incorporated into the host genome as a prophage and remains a part of the host's genome as long as conditions remain stable for the prophage. When conditions that cause damage to DNA or inhibit DNA replication are present, the cellular SOS response is activated by regulatory proteins. Upon activation of the SOS regulatory system, the phage enters into the lytic cycle<sup>2</sup>. In contrast to the lysogenic cycle, in the lytic cycle, the phage's DNA remains separate from the host genome and is replicated by the bacterium, resulting in the production of many phage particles. Once mature, a phage-coded lysozyme breaks down the cell wall, effectively destroying the cell through osmotic lysis and releasing mature phages from their previous host<sup>3</sup>. There are an estimated  $10^{31}$  different bacteriophages on earth, making them one of the most numerous DNA-containing entities in existence<sup>4</sup>. Our research focuses on the isolation and characterization of mycobacteriophages, dsDNA tailed phages that infect mycobacterial hosts<sup>5</sup>. Because mycobacteriophage genomes are usually between 50,000 to 100,000 base pairs long, they are relatively easy to sequence and analyze, making them prime candidates for genome annotation. The phages isolated in this study infect *M. smegmatis* cells. While most phages infect a single type of host bacteria, a few have been found with a larger host range within a genus<sup>6</sup>. The ability for phages to infect multiple hosts would enable scientists to use harmless, phage-containing bacteria to destroy harmful bacteria through lysis<sup>8</sup>. Because *M. smegmatis* is closely related to *M. tuberculosis*, a phage that infects *M. smegmatis* cells could potentially infect and lyse the bacterial cells that cause Tuberculosis, a disease that killed approximately 1.7 million people in 2009. While a few phages have been found with a multi-host capability<sup>7</sup>, there are presumably many more to be discovered.



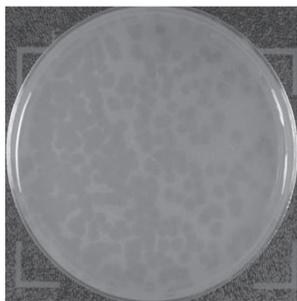
Of the  $10^{31}$  phages on earth, less than 2,000 mycobacteriophages have been sequenced. A larger database of novel bacteriophages is necessary to provide the scientific community insight into the variety of genes, structures, and functions of phages, including multiple host capabilities. As it is dangerous to work with pathogenic bacteria and the incubation period for studying bacteria like *M. tuberculosis* is much longer than *M. smegmatis*, it is beneficial to use a harmless host that replicates quickly (such as *M. smegmatis*) to isolate phage, and then test the resulting novel phages for the capability to infect more dangerous bacteria.

Scientists from the former Soviet Union and Eastern Europe proposed the use of bacteriophage to fight bacterial infections prior to the invention of antibiotics. In 1919, D'Herelle, a French scientist, used phage to treat dysentery. Within a matter of years D'Herelle had a commercial laboratory in Paris that produced phage preparations, and research spread throughout Europe, to the Soviet Union, and even in the US<sup>10</sup>. With antibiotic resistance a growing issue worldwide, phage therapy has been once again proposed as an alternate option for use against resistant bacteria. In fact, phages have been reported to be more effective in the treatment of certain infections in humans including *Staphylococcus aureus*<sup>10</sup>.

The first step to effective practice of phage therapy is to better understand phage genomics. This understanding can only come about through the addition of phages and their genomes to international databases. The purpose of our research was to expand phage databases by isolating and annotating novel bacteriophages. By adding to the scientific community's understanding of phage biodiversity, our research broadens genome sequence databases and may generate new insights into gene function.



**Figure 1.** The anatomy of bacteriophages varies, but a few characteristics hold true for most phage. All bacteriophage have a protein capsid to hold and protect genetic material. Many also possess a tail which serves as a channel for DNA during infection of a bacterium. This tail has an end plate and pins which allows the phage to attach to specific bacteria.



**Figure 2.** Plaques are clear areas found in confluent bacterial growth. Plaque morphology greatly differs between phage and can vary due to growth conditions including incubation time. Differing in size and turbidity, plaques form when a phage infects and lyses a bacterium, leaving a portion of the bacterial lawn clear. Lytic phages form clear plaques, while lysogenic phages form turbid ones. Turbidity signifies the presence of lysogenic phage. The number of phage in a solution can be determined by calculating the number of plaque forming units per milliliter of phage. These are plaques formed when *M. smegmatis* cells were infected by BarrelRoll, one of the sequenced phages.

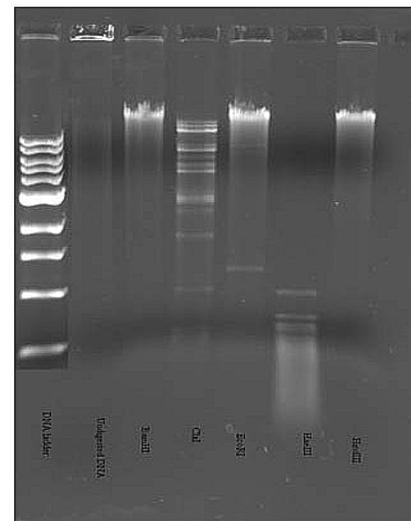
## Materials and Methods

Students in the Genome Discovery and Research course began by collecting individual soil samples from a location of their choice. The GPS coordinates, time, date, and site characteristics were noted for each soil sample. The phages isolated by the entrants were named Luke117 and Novus. Luke117 was taken from moist soil near a pond, while Novus was taken from dry dirt beside a creek bed. To enrich for mycobacteriophages, a culture was created as follows: 40 ml sterile H<sub>2</sub>O, 5ml of 10X glycerol broth, 5 ml of AD supplement (provided by the sponsoring institution), 0.5 ml of 100 mM CaCl<sub>2</sub>, and 5ml of *M. smegmatis* culture and 1 g of the soil sample. This mixture was then incubated for 24 hours at 37°C with shaking at 270 rpm. The next day, the culture was transferred into a 50 ml conical tube and centrifuged at 2000rpm for 10 minutes to precipitate material and bacterial cells into a pellet. The supernatant from the centrifuged sample was transferred to a 50 ml conical tube. A 1ml aliquot from the conical tube was then filter-sterilized using a syringe and .22 μm filter unit. To determine if viable phage were present, serial 10-fold dilutions of the filter-sterilized sample were made using Phage Buffer (a saline solution containing CaCl<sub>2</sub>). The negative control consisted solely of phage buffer. A sample of bacteriophage D29 (a lytic mycobacteriophage) was used as the positive control. Aliquots of the dilutions were mixed with *M. smegmatis*, and incubated for 20 min to allow for phage attachment. The infected cells were then plated and placed in the incubator. After 24 hours of incubation at 37°C, the plates were examined for plaque formation (Fig. 2). If plaques had not yet formed, the plates were returned to the incubator for an additional 24 hours of 37°C incubation. The same day, direct plating was performed by extracting phage from the soil. The sample was flooded with phage



buffer and mixed well with a vortex. After settling for 20 minutes, a 1ml aliquot from the sample was filtered with a 0.22  $\mu\text{m}$  filter. This filtered sample was combined with *M. smegmatis* culture and allowed to sit for 20 minutes. Similarly, D29 and filter-sterilized PB were each combined with *M. smegmatis* culture to serve as positive and negative controls respectively. After 20 minutes elapsed, each sample and control were combined with 4.5 ml of top agar and plated in agar plates to be incubated at 37°C for 24 hours to allow plaques to form, with an additional 24 hours added if necessary. After noting morphology, a pipette tip was used to sample a single plaque from one of the plates from either the phage dilutions or the direct plating. This sample was then added to 100 $\mu\text{l}$  of PB in a microcentrifuge tube and 10 fold serial dilutions were prepared in PB. Each dilution was then mixed with 0.5 ml of *M. smegmatis* cultures and allowed to sit for 30 minutes giving the phages time to infect the bacteria. Top Agar (4.5ml) was added to each sample, mixed and then pipetted immediately onto an agar plate. After 24 hours of incubation at 37°C, the researchers noted the morphology of the plaques produced and incubated the plates an additional 24 hours if needed. This isolation procedure was repeated at least 3 times to ensure the purity of the phage isolate. Purity was established by observing consistent plaque morphology after each round of purification. Novus consistently formed clear plaques of approximately 1 mm in diameter, while Luke117 characteristically produced turbid plaques approx. 2 mm in diameter. After attaining consistent morphology, the titer of the sample (number of plaque forming units per mL of the sample) was determined. The Luke117 stock titer had 4X10<sup>9</sup> pfu/mL, and Novus' was similarly high. To have enough phage particles for electron microscopy and DNA isolation, it was essential that this figure was at least 10<sup>9</sup> pfu/mL. The researchers then selected a titer of the phage particles that produces a webbing pattern of overlapping plaques. This pattern is essential as it produces the largest quantity of viable phage particles. A phage sample was then diluted to the specific titer necessary to produce a webbing pattern and used to produce ten identical plates. Following incubation for 24 hours at 37°C, these plates displaying webbing patterns were each flooded with 8 mL of PB and stored at 4°C overnight. The following day, the

buffer was removed from the plates using a pipet aid and refrigerated in 50ml conical tubes. This stock was used for all future analysis of the phage. DNA isolation began by adding a nuclease mixture to 10ml of filter-sterilized phage stock in order to remove bacterial DNA and RNA from the lysate. The nuclease and lysate were incubated for 30 minutes at 37°C and left to sit at room temperature undisturbed for 1 hour. Afterwards, 4.0 ml of phage precipitant solution was added to the nuclease-treated lysate. After mixing gently, the lysate was incubated overnight at 4°C. The tube was then placed in a high-speed centrifuge and spun for 20 minutes at 10000Xg. The supernatant was removed, leaving a pellet of genetic material at the bottom of the tube which was re-suspended in 1ml of H<sub>2</sub>O. A protein denaturant (present in the Wizard clean up resin) was then added at 37°C to denature the phage protein capsids. This solution was pushed through DNA-containing columns and centrifuged. Salts and proteins were washed away by pushing 2 ml of 80% isopropanol through the column. The phage



**Figure 3. Agarose electrophoresis gel of restriction digests of Luke117, isolated by one of the entrants. Multiple restriction enzymes were used to digest the DNA, cutting wherever the recognition sequence was present. When an electric current passes through the gel, DNA fragments move towards the positive electrode, with the smallest fragments moving most quickly. Lines at the left of the gel mark large segments of DNA, while those close to the right represent smaller segments that moved through the gel more quickly. At the very far left of the second row (HAEIII), a blur marks the presence of cuts too small to form distinct lines. This image was captured by inserting EtBr into the DNA and photographing it under a UV light.**

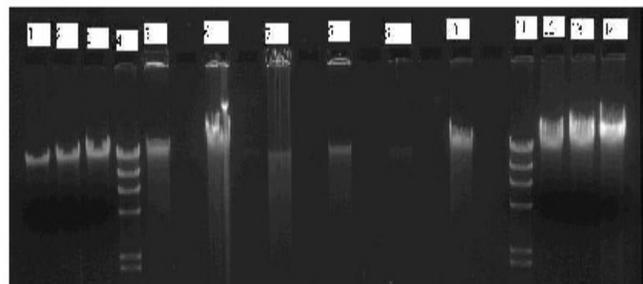


genomic DNA was then eluted from each column by adding pre-warmed 80°C TE to the resin in the column. After 30 seconds, the columns were centrifuged for 1 minute in microcentrifuge tubes. The DNA samples were then transferred into clean tubes and stored at 4°C. To check the quantity of DNA, the researchers used a spectrophotometer to measure concentration and protein content, ensuring that all proteins had been removed from the DNA solution. The next step in the process was to perform a restriction analysis of the purified DNA sample using gel electrophoresis (Fig. 3). This process can help classify phages and determine their novelty. Phage genomic DNA was combined with the enzymes BamHI, ClaI, EcoRI, HaeIII, and HindIII as well as reaction buffer, BSA, and ddH<sub>2</sub>O. Following a 2 hour incubation period in a 37°C water bath, the samples were stored in the freezer overnight. The next day, a 0.8% (weight/volume) agarose gel was prepared by combining agarose, 1X TBE Buffer (a buffer solution containing a mixture of Tris base, boric acid and EDTA) and Ethidium Bromide (a dye that allows visualization of DNA) and heating. While the gel solidified, tracking dye was added to each restriction digest and heated at 65°C for 5 minutes. The samples, as well as a DNA standard, were then pipetted into the gel wells. The samples were electrophoresed until within 1 to 2 cm of the end of the gel and photographed under UV.



Figure 4. This image, captured while using Transmission Electron Microscopy, allowed the researchers to better understand Novus, a phage isolated by one of the entrants. Novus has an average tail length of 264.3 nm, tail width of 14.3 nm, and 50.0 nm capsid diameter. An electron micrograph can be used to determine capsid diameter, tail length and width, and even help classify phage into different subgroups.

After preparing concentrated phage stock, each researcher performed electron microscopy in addition to DNA isolation (Fig. 4). First, a phage sample was prepared by transferring between 0.5 and 1.0 ml of phage lysate into a microcentrifuge tube. The sample was centrifuged for one hour at 4°C at 10000xg. The supernatant was removed using a micropipettor and the pellet was re-suspended in



**Figure 5.** This image is of a gel run for several of the phage isolated by the cohort, with the DNA standards in lanes 1-4 and 11-14. BarrelRoll is in lane ten, and Gemini is in lane six.

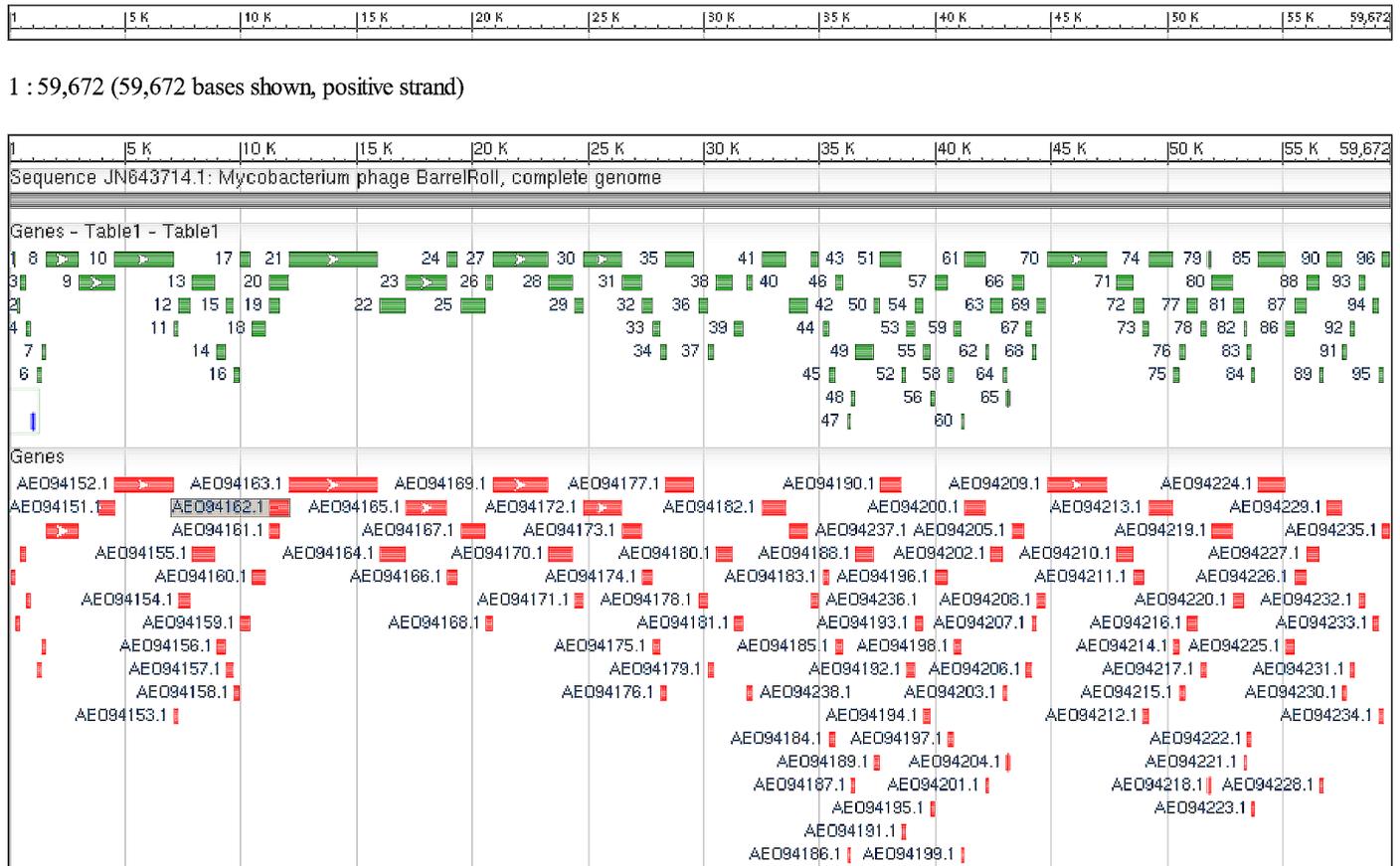
100µl of fresh PB and stored at 4°C until time of electron microscopy. With assistance of a microscopy professor, we mounted and stained the phage using a copper grid. The stained grids were viewed and photographed under the electron microscope. These images display many phenotypic traits of the phage such as tail length and capsid diameter, which are both helpful in determining the phage's cluster. Luke117 had a tail length of 142.5 nm and capsid diameter of 72.5 nm, similar to phages in cluster K1. Novus had an approximate tail length of 264.3 nm, tail width of 14.3 nm and capsid diameter of 50nm, characteristics which are commonly found in cluster B3. The final step of the isolation process was to confirm the quality of each phage genomic DNA sample using electrophoresis (Fig. 5). A 1% agarose gel in 1X TAE buffer (a buffer solution containing a mixture of Tris base, acetic acid and Ethylenediaminetetraacetic acid) containing 0.15-µg/ml ETBR was prepared. One µl of each sample was pipetted into a microcentrifuge tube with 5µl of 1X tracking dye. The samples were electrophoresed for 40 minutes at 120V with Lambda DNA Standard. The gel was then photographed and analyzed. The brightness of each sample signified the amount of DNA present. The two highest quality DNA samples out of the 17 phages isolated were sent to the sequencing center.

The genomic sequences of the two highest quality samples, BarrelRoll and Gemini, were returned us. Several bioinformatic tools were used to annotate the genomes. The first step was to ensure consensus throughout the sequence as well as the direction of possible genes using Consed, a program used for viewing, editing, and finishing DNA sequence assemblies.



**Mycobacterium phage BarrelRoll, complete genome**

gi|347450105|gb|JN643714.1|



**Figure 6.** FASTA image from the Genbank entry of BarrelRoll, showing annotation of potential genes from BarrelRoll’s full genome sequence.

Geneious, a bioinformatics tool for sequence alignment, and Phamerator, a tool for comparative bacteriophage genomics, were then used to compare the DNA sequences of both genomes to similar sequences in GenBank. Genome analysis with these programs gave great insight into the formation and structure of phage genes. Apollo, a gene annotation tool, was then used to analyze potential genes. The research class divided the genomes into 5-6 kbp segments to be analyzed by the students in pairs. Each segment was annotated by multiple groups. Using Apollo, annotations were added to the genomes with length, start codon, gap/overlap, Shine Dalgarno sites, and associated scores recorded for each putative gene. BLAST, or Basic Local Alignment Search Tool, is a tool used to find regions of similarity between DNA sequences. BLAST was used during our annotation to compare potential genes to other mycobacteriophage genomes in the sequence database. E-scores were assigned by this program to show how similar or dissimilar two sequences were found to be. Through comparison, we were better able to understand the putative functions of genes. Through the use of the programs mentioned above, a final annotation was assembled by the research professor using the information gathered and reviewed by each two member student team. These annotations were loaded onto the Gene Workflow annotation editor, to GenBank, the NIH genetic sequence database(Figure 6).

**Results**

Seventeen diverse bacteriophages were isolated by the Genome Discovery and Exploration participants. A concentrated, high-volume solution of phage particles was archived for each of the phages. Phages Novus and Luke117 were isolated by the entrants. While accurate grouping of phages requires genomic sequencing, analysis



of the restriction digests enabled an approximate categorization of the phages into the phage groups found in GenBank. Clusters, or groups of similar phages, range from A to O, while some singleton phages exist which do not fit in any cluster. Novus, which was cut with enzymes HaeIII and BamHI, was grouped into subcluster B3 and Luke117, which was cut with ClaI, HaeIII, and EcoRI, was categorized into subcluster A1. The genomes of two bacteriophages isolated in the research class, BarrelRoll and Gemini, were sequenced. Using data gained through the genomic sequencing, BarrelRoll was placed in group K1 and Gemini in group A2. These genomes were fully annotated by the Genome Discovery and Exploration class participants. The genome of BarrelRoll is 59,672 base pairs long and the genome of Gemini is 52,886 base pairs long. The genomic sequences of these phages were reviewed using Consed to ensure sequence reliability throughout the genome. For each phage, the results of all sequence runs were assembled by Consed to form a complete sequence with a high level of consensus. Because there were many overlapping sequences, we are confident in the completeness and correctness of the genomic sequences to be added to GenBank. While DNA sequence is a beneficial addition to GenBank, annotation adds utility to the database. For BarrelRoll's sequence, we called 95 genes, and 96 genes were called for Gemini. With each putative gene, the start and stop codons, Shine Dalgarno site, gene function, and length were recorded within the annotation. Putative genes were analyzed using both Apollo and BLAST by at least two independent groups of students before submitting the annotations to ensure confidence in their validity. After repeated analysis of putative genes, the annotation of both phages was completed. BarrelRoll has recently been submitted to GenBank and Gemini will be added to the database in the months to come. These two genomes will then be available in BLAST searches, and the annotated genomes can be compared with thousands of other novel phages. We are confident in the reliability of Novus and Luke117 and of the annotation of BarrelRoll and Gemini. During the isolation phase, phages underwent many rounds of purification to ensure only one type of phage was isolated. Plaque morphologies were noted and compared after each round of purification to confirm that the same phage type was isolated. In the electron microscopy, students checked the uniformity

of the phage structure, capsid diameter and tail length. Plaque morphology, restriction digest gels, and electron microscopy photos were compared with those of other phages within the laboratory and at a national level to confirm that the bacteriophages are truly novel. Post sequencing, each genome was annotated at least twice by independent student groups, with extensive discussion over any discrepancies. Finally, the laboratory mentors reviewed all of the annotations, checking one last time for any incongruities. With this level of thoroughness, we have considerable reason to believe that the isolated bacteriophages are pure, novel, and properly annotated, yielding reliable results to add to the diversity of phage found in the sequence databases.

Because of the number of phage isolated, our research contributes to the diversity of phage sequences in the DNA database. The thorough annotation of Gemini and BarrelRoll make their respective genomes useful as comparisons for future gene annotations. The more people that study a gene, the greater the likelihood of its accurate annotation. Therefore the addition of these two fully annotated genomes will prove helpful in the annotation of many other phages to come. In addition, the sequences will soon be available worldwide via GenBank. Such widespread availability will enable these phages to undergo study by other researchers who are interested in phage genomics. There is considerable opportunity for future work involving these phages.

Primarily, our work brings about an increased knowledge of bacteriophages for both the researchers themselves and for the scientific community worldwide. As we worked with phages and learned more about how they function and how their genes work, we were able to refine our own methods of isolation and annotation through gained experience. Understanding what steps were necessary to prevent contamination and knowing what factors mattered in selection of Shine Dalgarno sites was helpful when certain steps had to be repeated. As a result of the feedback given to the creators of the gene annotation programs, these programs will also be improved upon, enabling them to function more efficiently. More important than assisting in the evolution of methods and software, our research adds seventeen novel phages, two with annotated genomic sequences to a relatively small DNA sequence database. By adding to the biodiversity



of isolated phage and annotated genomes found within GenBank, we are offering greater insight into both gene function and genome evolution. Novel genes have been added to the database, and a better understanding of pre-existing genes can be formed through comparison of the phages isolated by our laboratory and those already found in the phage database. While these phages have not been proven to affect bacterial diseases, the benefits yielded through increasing insight into phage genomics has brought us one step closer to more effective use of phages in the medical field. Faster annotation programs, more efficient isolation techniques, and more in-depth genomics knowledge all aid in the search for phage-based cures for bacterial illnesses by providing more data for medically focused phage researchers to work with and more phages available for practical applications. One primary purpose of our bacteriophage research was to increase the breadth of phage genomics knowledge within the scientific community. Although phages are ubiquitous, a mere 1,521 mycobacteriophages have been isolated, 246 sequenced, and only 176 have been added to GenBank.<sup>9</sup> By identifying and analyzing bacteriophages and increasing general understanding of phage structures and functions, scientists can better categorize phage and test more of them for potential phage therapy against bacteria similar to *M. smegmatis*, such as *M. Tuberculosis*. Through this project, we have contributed two isolated bacteriophage DNA and phage particle stock samples, and we contributed to the annotations of two bacteriophage genomes. The more phage isolated and analyzed, the more plausible phage therapy becomes. With the number of annotated phage increasing, the potential that one of them will be useful for phage therapy also increases. The general procedure of bacteriophage isolation and annotation is readily available and described in “Exploring the Mycobacteriophage Metaproteome: Phage Genomics as an Educational Platform”<sup>11</sup>. While this article describes our general research format, the results of our work is entirely unique to our research project. The phages isolated in our laboratory are each unique compared to other phage within our lab, although similar phages may be isolated in other parts of the country. The process of confirming that a phage is most likely novel can be completed without sequencing with considerable accuracy. For Luke117, the electron

microscopy was first compared to those of each cluster in the phage database<sup>9</sup>. Phages with similar tail length and capsid structure were found in clusters A and D. To further narrow the potential cluster, the restriction digest of Luke117 was compared to the virtual digests of each subcluster in A and D. Excessive cuts with Cla1 and HaeIII suggested that Luke117 belongs to cluster A1. Within cluster A1, phages BillKnuckles and SkiPole showed high phenotypic resemblance with tail length to capsid diameter ratios of 2:1 and significant correlation between restriction digest patterns. However, Luke117 lacks several prominent cuts in the digest that are evident in BillKnuckles and SkiPole, showing that the genome is different from those of BillKnuckles and SkiPole. Because of the restriction digest data suggesting that the sequence is similar to those of other phages in the cluster, it is possible that Luke117 is a novel phage that belongs in subcluster A1. Likewise, the same method of analysis was performed for Novus. Through analysis of the EM picture and plaque formations, clusters B, H, and M were selected as potential clusters for Novus. Upon comparison of the restriction digest, Novus’ subcluster was predicted as B3. Athena displayed greatest genomic similarity with the restriction digest. However, there were several gaps in Athena’s digest where there were cuts in Novus. The lack of cuts in these locations supports the likelihood of Novus being a novel phage. The evidence of unique restriction digests for the Luke117 and Novus, along with entire genome comparisons for bacteriophages that were sequenced (BarrelRoll and Gemini), provide substantial evidence supporting our conclusion that these phages are unique.

## Discussion

Contributing the phage genomes and annotations of BarrelRoll and Gemini to the phage DNA database has led to more impacting results by adding to the understanding of phage biodiversity. The entire scientific community now has expanded access to knowledge of bacteriophages and their diverse genomes. Our data has contributed to the pool of available phage information. The full extent of phage biodiversity remains unknown, and will probably never be fully understood. However, the discovery of these seventeen additional phages has brought us one step closer to understanding phage biodiversity.



Furthermore, this information may help with the implementation of phage therapy as an alternative cure to bacterial illnesses. Our results are only a small portion of the overall goal of gaining insight into phage diversity. Because our research is a part of a much larger endeavor, the methods are repeatedly tested by other researchers across the country. Because our phage isolation underwent many rounds of purification and the genome annotation was checked multiple times by fellow researchers and laboratory mentors, we are confident in the validity of our results. Future work would include the following: one of our first projects would be the sequencing and annotation of the phages we isolated, Novus and Luke117, enabling additional phage genomic sequences and annotations to be added to the national databases. Another intriguing prospect would be to isolate phages using different host bacteria such as *M. tuberculosis* or *M. leprae*, which could directly apply to phage therapy. However, this sort of work would require a laboratory with specific biological safety precautions due to the harmful nature of the bacteria. In addition, these experiments would require a significant investment of time since these bacteria replicate at a much slower rate than *M. smegmatis*. A useful endeavor would be the determination of currently unknown gene functions through study of bacteriophage gene expression. This would be done through the comparison of similar bacteriophages, one containing the gene of interest and one without. Phages would be analyzed to see how the absence of that gene affected phenotype and plaque morphology. This observation would aid in the determination of that gene's function. Finally, it would be important to continue isolating additional novel bacteriophages to add to the phage database. This continuation is a reality, as the phage isolation and annotation process is repeated by participating universities each year. As our methods are regularly practiced by universities nationwide, we encountered very few problems with the methods. However, if we were to restart our research today, we would complete further phage analysis, including additional restriction digests. We would also further stress aseptic technique to prevent contamination of materials that led a few researchers to repeat purification steps of the isolation process. More than anything, we would like to extend the amount of research performed throughout the two semesters by increasing the number of phage

isolated and characterized.

The main question our research is seeking to answer still remains: exactly how diverse is the bacteriophage community? The only way this can ever be answered satisfactorily is through further isolations of phage across the globe. The function of many putative genes remains a mystery also. By connecting genes with structure and function, we would be able to better understand phage biology. Lastly, there is the question of whether or not a phage with a host closely related to *M. tuberculosis* or *M. leprae* could be effective in treating TB or Leprosy. As bacteriophage have been known to have a host range of as many as 29 species<sup>12</sup>, finding a phage that infects a safer and quicker replicating host than *M. smegmatis* is a worthy goal. By continually repeating the phage isolation process, the breadth of phage research will expand, and these questions will hopefully be answered while even more intriguing queries unfold.

## References

1. Garcia P, Garcia JL, Lopez R, Garcia E. 2005. Pneumococcal Phages. In: Waldor K, Friedman D, Adhya S, editors. Phages: Their Role in Bacterial Pathogenesis and Biotechnology. Washington (DC): ASM Press. P 335-361.
2. Little JW. 2005. Lysogeny, Prophage Induction, and Lysogenic Conversion. In: Waldor K, Friedman D, Adhya S, editors. Phages: Their Role in Bacterial Pathogenesis and Biotechnology. Washington (DC): ASM Press. P 37-54.
3. Harisha S. 2006. An Introduction to Practical Biotechnology. New Delhi (In) : Laxmi Publications. p. 225.
4. Hendrix RW. Bacteriophage Evolution and the Role of Phages in Host Evolution. In: Waldor K, Friedman D, Adhya S, editors. Phages: Their Role in Bacterial Pathogenesis and Biotechnology. Washington (DC): ASM Press. P 55-65.
5. Pope WH, Jacobs-Sera D, Russell DA, Peebles CL, Al-Atrache Z, et al. 2011
6. Expanding the Diversity of Mycobacteriophages: Insights into Genome Architecture and Evolution. PLoS ONE 6(1): e16329. doi:10.1371/journal.pone.0016329
7. Jensen EC, Schrader HS, Rieland B, Thompson TL, Lee KW, Nickerson KW, Kokjohn TA. 1997.



Prevalence of Broad-Host-Range Lytic Bacteriophages of *Sphaerotilus Natans*, *Escherichia Coli*, and *Pseudomonas Aeruginosa*. *Applied and Environmental Microbiology* 64(2):575-580.

8. Faruque SM, Bin Naser I, Islam MJ, Faruque ASG, Ghosh AN, Nair GB, Sack DA, Mekalanos JJ. 2004. Seasonal Epidemics of Cholera Inversely Correlate With the Prevalence of Environmental Cholera Phages. *PNAS* 102(5):1702-1707.

9. Mycobacteriophage Database [Internet]. [updated 2011 Sep 04]. Pittsburgh (PA):Howard Hughes Medical Institute; [cited 2011 Sep 04]. Available from: <http://phagesdb.org/>

10. Sulakvelidze A, Zemphira A, Morris JG. Bacteriophage Therapy [Internet]. Washington (DC): American Society for Microbiology; 2001 [cited 2011 Sep 4]. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC90351/>

11. Hatfull GF. Exploring the Mycobacteriophage Metaproteome: Phage Genomics as an Educational Platform [Internet]. San Francisco (CA): PLoS Genetics; 2006 [cited 2011 Sep 4]. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1475703/>

12. Kutter K, Sulakvelidze A. 2005. Bacteriophages Biology and Applications. Boca Raton (FL): CRC Press. p 78.

### Acknowledgements

The authors would like to thank Dr. Rhinehart for their help with our experiments. We would also like to thank the Howard Hughes Medical Institute Science Education Alliance for making this project possible.