June 2017

Genetic Diversity of Mycobacteriophages and the Unique Abilities of Cluster K

Shea A. Morris
Georgia College and State University

Introduction:
Mycobacteriophages make up a large portion of particles in the biosphere reaching 10^{30} in number (Hatfull et al., 2006). These phages are viruses that also infect large, diverse amounts of bacteria, destroying the bacteria that they infect. Mycobacteriophage are so numerous that it is estimated 10^{23} phage infections of their bacterial host occur every second. Their genetics are extremely diverse, although, each individual mycobacteriophage has a specific bacterial host range that is typically restricted to a single genus of bacterium. On rare occasions; however, some phage have a host range outside one genus of bacteria (Pope et al., 2011a). Due to mycobacteriophages large host range, each phage has a specific viral life cycle which allows them go through lytic, lysogenic cycles, or sometimes a combination of the two. This diversity in pathogenic mechanisms allows mycobacteriophage to have a powerful effect on each bacterium it is introduced to and that is why the group is attractive for viral- based antibiotic agents to assist or replace conventional antibiotics. Conventionally, these viral approaches to bacterial control are often referred to as "phage therapy" in medical practice.

Mycobacteriophages are arranged in clusters and subclusters based on similar genetics, and the cluster K is of particular interest as it is rather small compared to other clusters, but also incredibly unique. The phages within the cluster K are temperate phages and many believe have the ability to aid in research on tuberculosis genetics and act as a bactericidal to fight mycobacterial infections (Pope et al., 2011b). Through past and present therapies these phage may be used as an oral cocktail for patients to drink or as an injection that induces sensitivity to antibiotics in the bacterial cell, allowing for it to be killed by either antibiotics or the phage. Through this treatment, less antibiotic resistance would develop and any genetic changes made by bacteria to resist mycobacteriophage would induce a similar genetic mutation in the phage group to maintain its viral abilities against mycobacteria (Flatow & Spellberg, n.d.). In this paper it will be these qualities that are discussed to emphasize the importance of mycobacteriophage and what makes them genetically unique and useful in medical practice.

The Virus:
Mycobacteriophage are bacterial viruses (bacteriophages) that are compromised of nucleic acid surrounded by a protein capsid (Pierce, 2005). Their chromosomes are composed of double stranded DNA and show much morphological and genetic diversity. The vast entity that bacteriophages represent in the biosphere has allowed for around 8,000 bacteriophages to be found (The Actinobacteriophage Database, n.d.). Therein, these phages are divided into clusters and subclusters based on their genetic similarities and host range. However, even phage within the same cluster or subcluster can be extremely diverse. Broad genomic studies are necessary to improve what science knows about mycobacteriophage because there are too many for all to be studied specifically. While only a few phage have been studied

specifically, enough study of their genetic data leads researchers to believe that phage within the same cluster and subcluster can effectively infect the same mycobacteria. If this assumption is consistent with all mycobacteriophage then there may well be a large variety of phage to combat most any mycobacterial disease.

Mycobacteriophage Morphologies:

Based on characteristics noted to date, all mycobacteriophages are classified as either siphoviral or myoviral. The more common, siphoviral phage have long, flexible tails that do not contract, while myoviral phage have shorter contractile tails. Their tail length can be from 135 nm to 350 nm and usually clusters show similar tail length. Each capsid head size is based on the genome length of the phage (Hatfull, 2014b). All mycobacteriophage have an isometric icosahedral capsid head, although for some, the viral head is more prolated. Additionally, all phage that have double stranded DNA and tails are referred to as Claudovirales (Pope et al., 2011a). In Cluster K, most members are siphoviral and their tails range from 185 nm to 200 nm (Hughes & Wolyniak, 2014). Sometimes Cluster K phage have tail fibers at the end of their tails and their capsid heads are usually around 50 to 60 nm in diameter (Pope et al., 2011b). These characteristics have been determined by transmission electron microscopy, typically with a negative stain such as uranyl acetate. Mycobacteriophages that have been examined show fewer morphological differences compared to other types of viruses. Whether that is because differences do not exist, or have yet to be found, is unknown. Nevertheless, many researchers believe this restricted phage morphology has been under positive selection throughout their evolution.

Host Range:

All mycobacteriophage have a specific host range, typically for only one genus of mycobacterium. The most common mycobacterium used to isolate mycobacteriophage is *M. smegmatis*. Almost all known mycobacteriophage were isolated on *M. smegmatis* lawns; however, many other bacteria are also used as hosts, such as, *Pseudomonas, Salmonella, Staphylococcus,* and many others (Pope et al., 2011a). Because of this, the full host range of some phages have yet to be fully examined; therefore, it is possible that many other mycobacteria may be potential hosts that have yet to be experimented with (Sarkis & Hatfull, 1998). Phage are isolated on bacteria by either direct plating or enrichment. While enrichment produces the most amount of phage, some research indicates that it may lower the genetic diversity of the phage by continuing to grow it through enrichment (Pope et al., 2011a). The gene that most determines host range has yet to be found, but many viral researchers have come to believe that it resides on the phage's tail proteins (Sarkis & Hatfull, 1998). The further scientists study the host range of phages, the more information they will have to explain some of the evolutionary patterns of phage development as well as many other aspects (Hatfull

et al., 2010). Through genome analysis, many mycobacteriophage share similar nucleotide sequences in comparison to the mycobacterium they infect which is why the study of bacteriophages offers extensive insight on bacterial genetics and life cycles (Hatfull et al., 2010).

One of the most exciting aspects of phage biology is the fact that many mycobacteriophage infect *M. avium* and *M. tuberculosis*. These two hosts have received a lot of study in recent years due to their resistance to antibiotics and their roles in pathogenesis. Mycobacteriophage have the potential to solve this problem as some mycobacteriophage can infect and kill both bacteria (Broxmeyer et al., 2002). The greatest issue in phage research is the development of a vehicle that can transport phage DNA safely into a plasmid and into the bacteria cell of the host. These vehicles are being created today and are called phasmids (Bardarov et al., 1997).

The integration of phage DNA into a bacteria cell can alter the physiology of the bacteria cell by adding viral genes and altering bacterial genes (Hatfull, 2014a). Phages either enter a host in a lytic or lysogenic cycle and those that enter a lysogenic phase may shift to a lytic cycle if the bacterial environment is no longer favorable in maintaining the prophage. The lytic virus enters the cell and begins to replicate itself within the host destroying the bacteria's DNA. The virus takes over the cell and once the virus has encoded enough new parts it lyses the cell and releases the new phages. In a lysogenic viral cycle, the virus remains dormant and a part of the hosts DNA, this is known as a prophage. Virulent phages only exhibit the lytic cycle and are very dangerous for the bacteria in which they infect. Some mycobacteriophage, such as Cluster K, are temperate phages. These phages can enter and exit from lytic and lysogenic cycles whenever it fits their needs (Sarkis & Hatfull, 1998). Cluster K also exhibit a host range between both slow and fast growing bacteria including *M. tuberculosis, M. avium, M. uclerans and M. smegmatis*. The plaque formation on *M. tuberculosis* is different for each Cluster K phage; however, they infect and kill the bacterium effectively (Pope et al., 2011b).

Grouping:

Phage are grouped using a simple dotplot nucleic acid hybridization analysis that compares genomes. Many dotplots have been performed over various groups of mycobacteriophage and they all show significant diversity throughout the entire phage population. In order to easily identify phage, clusters and subclusters have been formed to show the genomic similarities of certain phage. In order for a phage to be part of a cluster it must share up to 50% of the same nucleotide sequence of at least one other phage within the cluster. If a phage does not show significant similarities in its nucleotide sequence to any group of phage it will be placed into a separate new cluster; these phages are known as singletons and only five are known at this time (Hatfull, 2014b). Along with using dot plots to analyze phage, researchers also use a program called Phamerator which divides the phages into

'phamilies' by sorting their amino acid sequence similarities (The Actinobacteriophage Database, n.d.). Through this program and others, it can be demonstrated that mycobacteriophage have gone through extensive horizontal gene transfer during their evolution, as this gene transfer is not only within a phage cluster, but generally exhibited over the entire mycobacteriophage population. Based on this knowledge it is thought that most phage come from a common, but constantly evolving, gene pool (Pedulla et al., 2003).

Cluster K phage show extensive similarities among subcluster K1 phages, particularly toward the left side of their genome, and there is a more distant relationship between the nucleotide sequences between K1 and, K2 and K3. Their genome architecture is similar with each subcluster sharing many similar genes (Pope et al., 2011b). Recently researchers are working towards developing a new method of phage classification that only involves analyzing the "tapemeasure" gene of phage. This gene is the longest and through this method it would be much easier to place phages into specific clusters (Smith et al., 2013). In all there are a total of 19 clusters from A to S and a few of these are then divided into subclusters. While the use of clusters makes searching for phage with a specific genome easier, it does not explain the full evolutionary history of these phage as evolutionary patterns determined by cluster analysis cannot entirely reveal the pattern by which they evolved (Hatfull, 2014b).

Genome:
Gene analysis of mycobacteriophage further shows the diversity and mosaicism of mycobacteriophage. The genomes vary in length from 25 to 200 kilobase pairs and the use of illegitimate recombination leads to huge amounts of diversity in the phage population (Hatfull et al., 2006). Each segment of phage genome has a unique evolutionary history; some phage lack certain segments compared to their relatives which can indicate the order of when specific phage evolution began. Even so, mapping of this evolutionary history is extremely difficult; however, as the research moves forward, a better understanding of these patterns will be discovered that explain the uniqueness of each phage.

Mycobacteriophage genomes tend to have a precise order for their virion and assembly genes; typical they are on the left side of the genome and can be easily identified (Hatfull et al., 2010). The genes are tightly packed and have little noncoding space, roughly ranging from 400 to 800 base pairs (Hatfull et al., 2010). While it is likely that nearly all the viral genes have been predicted, many are genes that as yet have no known function (Hatfull, 2014b). The average amount of protein coding genes per genome is 114 and the protein domains are commonly small, containing less than 60 residues (Hatfull et al., 2010). Some phage encoding proteins could serve as antibacterial products in medical therapies (Drulis-Kawa, Majkowska-Skrobek, Maciejewska, Delattre & Lavigne, 2012). Mycobacteriphage have proteins that encode for lytic or lysogenic cycles, and therefore is a key gene

determinate for whether a phage is virulent and of use in therapeutic practice. A few mycobacteriophage have a protein that catalyzes recombination; however, this gene cannot be found on all mycobacteriophage, leading researchers to believe there are many variations of enzymes that catalyze recombination in mycobacteriophage (Hatfull et al., 2010). Similarly, certain phage have DNA polymerase to help synthesize their DNA and others do not. The appearance and loss of genes definitely play a significant role in the history of phage, and in some cases, they may indicate one phage being a parent of another. Due to complexities or missing information, they may, in other, cases exist in too many variations of a particular gene to be discoverable (Hatfull et al., 2010). Most phage have a reasonable amount of tRNA genes and a few even have tmRNA genes (Hatfull et al., 2006). tmRNA helps terminate translation of mRNA that becomes blocked due to the lack of a stop codon in the transcript, a common occurrence when the mRNA is degraded from its 3' end. Thus, the encoding of tRNA and tmRNAs by the phage ensures it will have sufficient translational machinery to express proteins needed to complete its lifecycle. These are just a few differences in the phage genomes, and demonstrate why these phage are so genetically diverse.

The most obvious genomic difference in Cluster K phages is the size of the genome of the phage TM4, compared to all other phages within the cluster. TM4 is smaller than all the other phage in Cluster K and it is also the only phage within the cluster that is not temperate. It lacks a section of the genome which codes for different parts of the lytic and lysogenic cycles in other Cluster K phages, leading researchers to believe it has gone through a deletion from its original parent's genome. Cluster K is like most mycobacteriophage in that their virion and assembly genes lie to the left and they have a section of genes with no known function (Pope et al., 2011b).

Phage Therapy:
Phage therapy is not a new topic, it has been discussed and practiced for decades since the first mycobacteriophage was plated on *M. tuberculosis.* However, with the adoption of widespread use of traditional antibiotics therapies (e.g., penicillin) in the western world, phage therapy has had little new experimentation and discussion since the 1930s, and it has remained this way until the past two decades. The focus of phage therapy is introducing phage as the new way to kill bacteria instead of antibiotics. Due to the extensive use of antibiotics, antibiotic resistance, and lack of new discoveries in the field of antibiotic development, effective new antibiotic treatments are waning. Indeed, the population most affected by this detrimental dilemma are patients suffering from chronic bacterial infections. In 1993 the World Health Organization declared multi-drug resistant tuberculosis (MDRTB) as a global emergency due to its high risk of fatalities and its antibiotic resistance (Bardarov et al., 1997). Since this time researchers have begun to intensely study the use of phage in tuberculosis therapy as Cluster K, as well as

other phage clusters, can infect and kill *M. tuberculosis.* The most difficult process has been creating a transport vehicle that would allow mycobacteriophage to enter the human body; however, many therapies are being developed and used to over come this.

The Beginnings of Phage Therapy:
Phage therapy was first practiced in Europe over one hundred years ago and has since been continued to be studied in that area. In France, around 1919, Dr. d'Herrelle began experimenting in intravenous injection of phage to children with dysentery; however, many of his first clinical trials were not printed until he felt that he had done enough research on phage to write a set of books explaining his clinical applications and extensive study of phage (Alavidze & Kutateladze, 2001). Russia has used phage therapy since the 1930s when Georgia began a facility specifically for producing phage and sending them to Russian war camps. Indeed, Russian soldiers in this period were known to carrier canisters of phage which were used to treat diarrhea and wounds. Russia and Georgia both performed phage therapy throughout the past 90 years and continue to produce new phage cocktails for the open market. In the western world, the study of phage began to dissipate when the use of traditional antibiotics began in the medical field. Since traditional antibiotics, prior to the development of antibiotic resistance, worked tremendously well as an effective treatment for bacterial infections, most phage research was set aside (Abedon, Kuhl, Blasdel & Kutter 2011). With the advent of growing antibiotic resistance, research on phage therapy has ramped up in the last twenty years. Clinical trials are now being performed by multiple United States universities and through continued trials it is expected that phage cocktails may reach the pharmaceutical market in potentially less than five years (Flatow, Spellberg & Turner, n.d.).

Therapy Types:
The beauty of phage therapy is the ability for it to be used in many ways. Intravenous therapy was one of the first experimented therapies. In this practice, small amounts of phage were isolated and mixed into a saline solution which was then infused into the blood over a bacterial infection. Patients experienced a decrease in the bacterial infection after 48 hours; however, the trial notes indicate that some patients may have experienced violent reactions at the beginning of their infusion (Abedon, Kuhl, Blasdel & Kutter 2011. Since these early experiments, little intravenous phage therapy has been published. Researchers, working in the Eastern European country of Georgia, have developed two phage cocktails that are used in a myriad of ways. "Intestiphage" and "Pyophage" are the two phage cocktails on the open market in Georgia. "Intestiphage" can be ingested orally and fights against 20 different types of gut bacteria and is often distributed in war zones. Other therapies with "Pyophage" have been created to make "PhagoBioDerm," a bandage or wrap with the "Pyophage" cocktail inside that allows for the phage to

slowly infuse over a wound over days or be placed directly inside the wound (Abedon, Kuhl, Blasdel & Kutter 2011).

In the United States, clinical trials are being performed on another version of phage therapy (Chan et al. 2016). These trials involve injecting phage cocktails into the area of bacterial infection, this allows for the phage to attack the bacterial cell and the cell may either protect itself from the virus or maintain its defense systems against traditional antibiotics; however, it cannot do both. Therefore, either the bacterial cell is destroyed by the phage or traditional antibiotics are used while the bacterial cells defenses are decimated as it protects itself against the phage. This genetic trade off with the bacteria either evolving to increase the resistance against phage or become sensitive to antibiotics has the potential to become a viable bacterial infection treatment throughout the United States (Chan et al. 2016). This combined antibiotic-phage therapy allows for any potential bacteria that are not harmful to be saved while the infectious bacteria are killed. Currently, antibiotics can kill both useful and detrimental bacteria, sometimes leading to more issues for the patients once the natural flora of bacteria are killed (Bacteriophage Therapy, 2015).

Cluster K's role in fighting Tuberculosis:
Cluster K phages infect *M. tuberculosis*, making them prime candidates for the studies discussed above. They are temperate phages which allows them to enter both lytic and lysogenic cycles while in the bacteria cell (Pope et al., 2011b). Potentially they could be inserted into the body to fight the *M. tuberculosis* bacteria that is in the body at that time and then remain in a lysogenic form until more *M. tuberculosis* is found. An oral cocktail could also be developed using these mycobacteriophage. The other potential therapy would be the injection of phage into the infected area and causing the bacteria to evolve and protect itself from the phage while the antibiotics go in and kill the cell. At the moment, only phage TM4 in Cluster K has had extensive research performed on it; however, hopefully in the future more Cluster K phages can be used and put into practice with these therapies.

Our Lab's Role in Cluster K Phage Discovery
At this time, two novel Cluster K phages have been isolated in our lab, in the Department of Biological & Environmental Sciences at Georgia College. Phage Milly belongs to subcluster K2 and has been sequenced and the genome has been annotated. The other phage is Adonis, which is part of subcluster K1 and has recently been isolated and sequenced. Future plans for the lab include the genome annotation of the most Cluster K phage Adonis, as well as lysogeny tests both isolated phage to compare infection cycles and abilities.
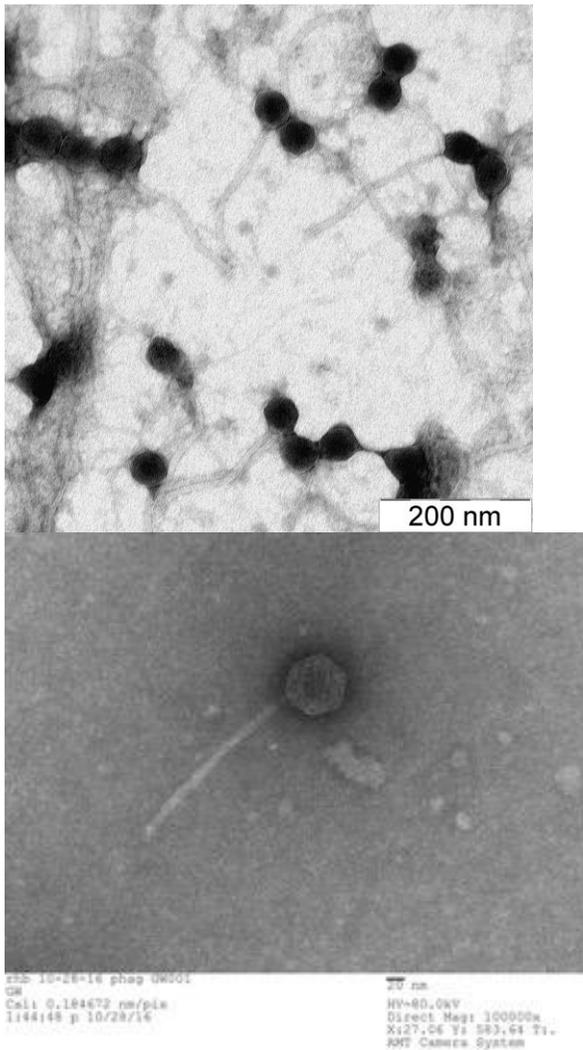
Figure 1: Electron microscopy pictures of Milly (left) and Adonis (right)

Conclusion:

Mycobacteriophages are extremely diverse in their genetics with many diverging evolutionary path ways. They infect many types of mycobacteria and remain extremely specific to their host range. Each phage goes through lytic and/or lysogenic cycles allowing them to destroy the bacteria they infect. Due to their broad host range and genetic differences each phage cluster has the potential to push through antibiotic resistance barriers and transform the world of antibiotics. Specifically, Cluster K phages can infect *M. tuberculosis,* and could serve as a tool to eradicate the growing disease of MDRTB. As research moves forward new and old therapies will be expanded and the use of mycobacteriophage, such as Cluster K, will likely become more common.

Literature Cited:

1. Abedon, S. T., Kuhl, S. J., Blasdel, B. G., & Kutter, E. M. (2011). Phage treatment of human infections. *Bacteriophage, 1*(2), 66-85. doi:10.4161/bact.1.2.15845

2. Alavidze, Z., & Kutateladze, M. (2001). Bacteriophage Therapy. *Encyclopedia of Genetics,* 175-179. doi:10.1006/rwgn.2001.0981

3. Bacteriophage therapy: An alternative to antibiotics? An interview Professor Clokie. (2015). from http://www.news-medical.net/news/20151202/Bacteriophage-therapy-an-alternative-to-antibiotics-An-interview-Professor-Clokie.aspx

4. Bardarov, S., Kriakov, J., Carriere, C., Yu, S., Vaamonde, C., Mcadam, R. A., . . . Jacobs, W. R. (1997). Conditionally replicating mycobacteriophages: A system for transposon delivery to Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences, 94*(20), 10961-10966. doi:10.1073/pnas.94.20.10961

5. Broxmeyer, L., Sosnowska, D., Miltner, E., Chacón, O., Wagner, D., Mcgarvey, J., . . . Bermudez, L. (2002). Killing of Mycobacterium avium and Mycobacterium tuberculosis by a Mycobacteriophage Delivered by a Nonvirulent Mycobacterium: A Model for Phage Therapy of Intracellular Bacterial Pathogens. *The Journal of Infectious Diseases J INFECT DIS, 186*(8), 1155-1160. doi:10.1086/343812

6. Chan, B. K., Sistrom, M., Wertz, J. E., Kortright, K. E., Narayan, D., & Turner, P. E. (2016). Phage selection restores antibiotic sensitivity in MDR Pseudomonas aeruginosa. *Sci. Rep. Scientific Reports, 6*, 26717. doi:10.1038/srep26717

7. Drulis-Kawa, Z., Majkowska-Skrobek, G., Maciejewska, B., Delattre, A., & Lavigne, R. (2012). Learning from Bacteriophages - Advantages and Limitations of Phage and Phage-Encoded Protein Applications. *Current Protein and Peptide Science, 13*(8), 699-722. doi:10.2174/138920312804871193

8. Flatow, I., Spellberg, B., & Turner, P. (n.d.). Old Ideas May Help Us Fight New Superbugs - Science Friday. from http://www.sciencefriday.com/segments/old-ideas-may-help-us-fight-new-superbugs/

9. Hatfull, G. F., Jacobs-Sera, D., Lawrence, J. G., Pope, W. H., Russell, D. A., Ko, C., . . . Hendrix, R. W. (2010). Comparative Genomic Analysis of 60 Mycobacteriophage Genomes: Genome Clustering, Gene Acquisition, and Gene Size. *Journal of Molecular Biology, 397*(1), 119-143. doi:10.1016/j.jmb.2010.01.011

10. Hatfull, G. F., Pedulla, M. L., Jacobs-Sera, D., Cichon, P. M., Foley, A., Ford, M. E., . . . Hendrix, R. W. (2006). Exploring the Mycobacteriophage

Metaproteome: Phage Genomics as an Educational Platform. *PLoS Genetics PLoS Genet, 2*(6). doi:10.1371/journal.pgen.0020092

11. Hatfull, G. F. (2014a). Mycobacteriophages: Windows into Tuberculosis. *PLoS Pathog PLoS Pathogens, 10*(3). doi:10.1371/journal.ppat.1003953

12. Hatfull, G. F. (2014b). Molecular Genetics of Mycobacteriophages. *Microbiology Spectrum, 2*(2). doi:10.1128/microbiolspec.mgm2-0032-2013

13. Hughes, J. B., & Wolyniak, M. J. (2014). 'Bioinformatics: The Study and Analysis of Cluster K Mycobacteriophage Genomic Data. *H-SC Journal of Sciences, 3*.

14. Pedulla, M. L., Ford, M. E., Houtz, J. M., Karthikeyan, T., Wadsworth, C., Lewis, J. A., . . . Hatfull, G. F. (2003). Origins of Highly Mosaic Mycobacteriophage Genomes. *Cell, 113*(2), 171-182. doi:10.1016/s0092-8674(03)00233-2

15. Pierce, B. A. (2005). *Genetics: A conceptual approach*. New York: W.H. Freeman.

16. Pope, W. H., Jacobs-Sera, D., Russell, D. A., Peebles, C. L., Al-Atrache, Z., Alcoser, T. A., . . . Hatfull, G. F. (2011a). Expanding the Diversity of Mycobacteriophages: Insights into Genome Architecture and Evolution. *PLoS ONE, 6*(1). doi:10.1371/journal.pone.0016329

17. Pope, W. H., Ferreira, C. M., Jacobs-Sera, D., Benjamin, R. C., Davis, A. J., Dejong, R. J., . . . Hatfull, G. F. (2011b). Cluster K Mycobacteriophages: Insights into the Evolutionary Origins of Mycobacteriophage TM4. *PLoS ONE, 6*(10). doi:10.1371/journal.pone.0026750

18. Sarkis, G. J., & Hatfull, G. F. (1998). Mycobacteriophages. *Mycobacteria Protocols,* 145-174. doi:10.1385/0-89603-471-2:145

19. Smith, K. C., Castro-Nallar, E., Fisher, J. N., Breakwell, D. P., Grose, J. H., & Burnett, S. H. (2013). Phage cluster relationships identified through single gene analysis. *BMC Genomics, 14*(1), 410. doi:10.1186/1471-2164-14-410

20. The Actinobacteriophage Database | Phamerator Information and Requirements. (n.d.). from http://phagesdb.org/Phamerator/faq/