5-2021

Mysteries and Uncertainties in Tracing Cryptic Viral Infections

Melaina Jacoby

Mysteries and Uncertainties in Tracing Cryptic Viral Infections

A thesis submitted in partial fulfillment of the requirement

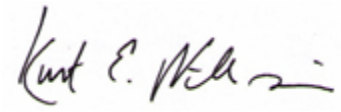
for the degree of Bachelor of Science in Biology from

William & Mary

by

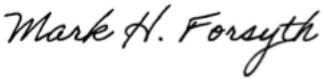
Melaina Jacoby

Accepted for Honors



Dr. Kurt Williamson, Director

Dr. Lizabeth Allison



Dr. Mark Forsyth

Dr. Doug Young

Williamsburg, VA

Defended May 5, 2021

ACKNOWLEDGEMENTS

I would first like to thank Kurt Williamson for his guidance and support throughout the years, as this project would not be possible without his patience, knowledge, and refreshing sense of humor. I would also like to thank Lizabeth Allison, Mark Forsyth, and Doug Young for serving on my Honors committee. Most importantly, I would like to thank my friends and family for listening to me and offering constant encouragement during this process, and I am especially grateful for my parents' unwavering support of my endeavors in science.

ABSTRACT

Viruses are the most abundant biological entities on Earth. Viral impacts are evident from the level of individual cells and population all the way up to ecosystems and global elemental cycles. Since bacteriophages (viruses that infect bacteria) were first identified in the early twentieth century, the study of these fascinating entities has shown how viral dynamics within ecosystems can influence microbially-mediated processes at a large scale. Viral infections can impact hosts and host-mediated processes in in multiple ways, one of which is through cryptic infections. This state, in which a bacterium may harbor a cryptic phage infection, is known as lysogeny. Such infections provide an evolutionary advantage to the phage, to survive times when host cells are scarce or few progeny can be made. Further, such infections may provide advantages to the host cell, as alleles carried by the phage genome may impact host phenotype. These cryptic infections can be activated (induced) chemically through laboratory assays, enabling us to determine the extent of lysogeny within bacterial assemblages. Most of the information we possess regarding lysogeny has come from research in aquatic ecosystems; however, the few studies that have investigated lysogeny in soil bacteria suggest that this is an important mechanism for phage replication and survival in soil environments. This study aimed to elucidate potential temporal trends in lysogeny in soil bacterial communities, as well as to compare the efficacy of inducing agents commonly used in induction assays. After collection and analysis of soil samples over a six-month period, results suggested that samples from only two out of six months showed evidence of prophage induction. Subsequent experiments compared induction responses of soil and aquatic bacteria, and the lysogenic bacterium E. coli W3104, to six different inducing agents. Results were highly variable; while most inducing agents, particularly mitomycin C, did appear to generate an increase in extracellular phage particles, few

of these increases were accompanied by bacterial lysis. This is important because determinations of lysogeny within bacterial communities depends on both increases in phage abundance and decreases in bacterial abundance in order to score the response. These results suggest that a reevaluation of the ways in which lysogeny is measured may be necessary, as calculations based on induction assays frequently return biologically nonsensical results. This study offers data supporting the prevalence of lysogeny within soil bacterial communities, and proposes potential directions for future research in this field.

TABLE OF CONTENTS

Page

INTRODUCTION

"There are few studies more fascinating, and at the same time more neglected, than those of the teeming populations that exist in the dark realms of the soil". These words were written by Rachel Carson in her 1962 classic novel "Silent Spring". Almost sixty years later, the mysteries of the soil have yet to be revealed by modern science, especially when it comes to microbial dynamics within these environments. Although Carson describes the roles of bacteria, fungi, algae, insects, and even earthworms in essential soil processes, she omits a potentially critical piece of the puzzle: viruses.

Viruses are among the most abundant and influential biological entities on the planet. Current estimations suggest that there are 10^{31} viruses on Earth, which is twelve times the total number of prokaryotes (Cobián Güemes et al., 2016). In fact, there are more viruses in a liter of coastal seawater than there are people on the planet (Suttle, 2013). There is a broad range of viral diversity, partially owing to the fact that viruses are able to flexibly evolve to changing environmental circumstances and fill many ecological niches (Wasik and Turner, 2013). It is even believed that all cellular organisms can be infected by at least one type of virus (Fuhrman, 1999). Bacteriophages, or viruses that infect bacteria, are more abundant than viruses that infect eukaryotic organisms, and seem to have important roles in regulating microbial community composition (Weitz and Wilhelm, 2012), carbon and nutrient cycling, biogeochemical cycles (Fuhrman, 1999; Weitz and Wilhelm, 2012), and even the shaping of global climate (Fuhrman, 1999). The study of viral impacts on these phenomena has been mainly conducted in aquatic ecosystems, where bacteriophage abundance is known to be high – around 2.5×10^8 per mL (Ashelford et al., 2002) -- and on average, 10 times larger than bacterial abundance. However, up to 97% of viruses in the world are located in soils and sediments (Cobián Güemes et al., 2016),

and more research is necessary to determine whether these models of viral impacts developed predominantly in marine systems can be applied to soils.

One important way in which the ecology and impacts of bacteriophages may differ between marine and soil systems lies in viral replication pathways. After infecting a host, a phage can follow one of two main pathways. The lytic pathway involves the phage inserting its genome into the bacterium and subsequently hijacking the host's replication machinery to immediately begin the process of copying phage genomes and generating progeny phage particles (Court et al., 2007). This eventually leads to lysis of the host cell and release of progeny phage particles. Lytic replication is believed to provide an evolutionary advantage when host abundance and productivity are high, as there are plenty of resources for the phage to replicate successfully and maximize production of progeny (Payet and Suttle, 2013). On the other hand, there may be instances in which these resources are much more limited, therefore making it evolutionarily disadvantageous to immediately reproduce via the lytic cycle. Certain phages, called temperate phages, are able to sense such resource limitations and reproduce via the lysogenic pathway, which occurs when the phage genome is incorporated into the host's chromosomal DNA instead of immediately replicating. This incorporated phage genome is called a prophage, and can exist indefinitely until reverted into the lytic cycle via a process termed induction (Court et al., 2007). A diagram depicting the relationship between these two cycles can be seen in Figure 1.

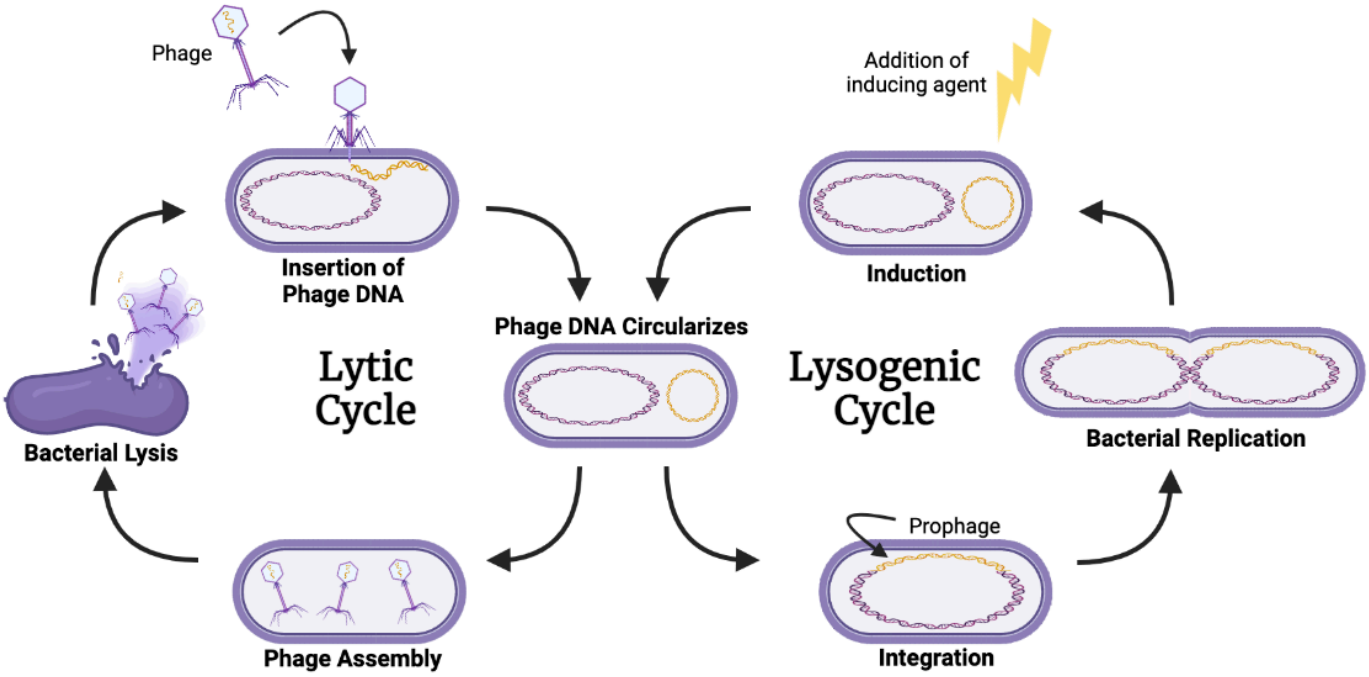


Figure 1. Comparing the lytic versus lysogenic lifecycles of a phage.

Image made on biorender.com

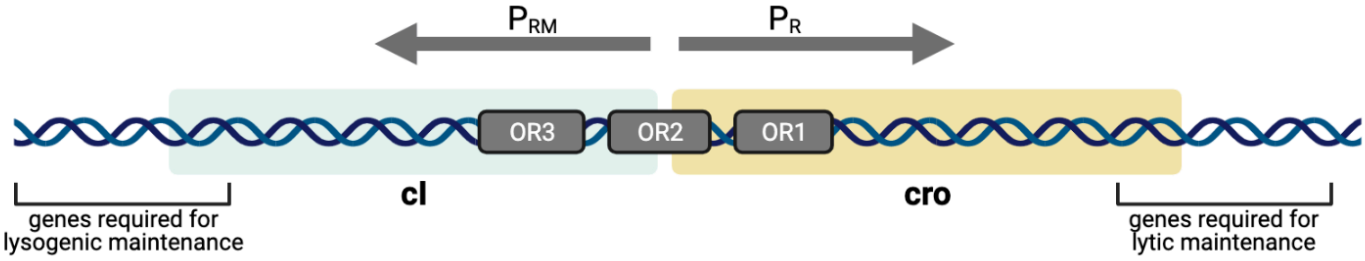


Figure 2. The regulatory circuit of phage lambda (Ptashne, 1986).

Image made on biorender.com.

The switch between the lytic and lysogenic cycle is controlled by a complex regulatory genetic network, the details of which have been identified in bacteriophage λ, a model temperate phage that infects *E. coli* (Ptashne, 1986; Echols, 1972; Lwoff, 1953). Contained in the λ chromosome are the genes *cI* and *cro*, which encode two different repressor proteins that control this regulatory network (Figure 2). Between these genes are three operator sites – O_R1, O_R2, and O_R3. The promotor P_R points RNA polymerase rightward towards transcription of *cro,* and P_{RM} points RNA polymerase leftward towards transcription of *cI.* In lysogeny, the λ prophage is stably integrated into the bacterial chromosome and the repressor protein cI will be at high abundance within the cell. The cI repressor has the highest affinity for O_R1, which overlaps with the *cro* gene. When the cI repressor is bound to O_R1, RNA polymerase cannot transcribe *cro* or any of the phage genes associated with lytic replication. The main promoter available under lysogeny is P_{RM}, which regulates expression of *cI* and the generation of more repressor protein.

This state of lysogeny is stable and can last many generations of cell division, producing daughter cells that are all lysogens, themselves. While stable, however, lysogeny is not indefinite or interminable. One well characterized signal that can terminate lysogeny is DNA damage to the host cell. Through DNA damage to the cell, whether through UV light exposure, chemical mutagens, or toxic compounds, the cellular protein RecA is activated. RecA is an important protein in the bacterial SOS response, a set of bacterial genes that are activated in response to DNA damage and help repair damage. RecA has proteolytic activity, and specifically targets and cleaves the host protein LexA during the SOS response. Cleavage of this host repressor protein allows transcription of the genes involved in the SOS response for DNA repair. The phage cI repressor is structurally similar to the cellular repressor LexA, and therefore gets cleaved during the SOS response in a way that prohibits its dimerization. Since dimerization of the repressor

protein is required to bind to the operator sequences, eventually enough cI repressor is cleaved that the prophage operator sequences remain unbound. This allows RNA polymerase to bind to P_R and begin transcribing the *cro* gene, as well as other downstream genes associated with lytic replication. The protein product of *cro* is, itself, another repressor that preferentially binds to O_R3, blocking transcription of the *cI* repressor. This prevents RNA polymerase from binding to P_{RM} and allows the switch to be irreversibly flipped towards lytic growth.

Of course, not all infections by phage λ will result in a lysogenic infection. This decision is made by the phage when first infecting the host, and is influenced by environmental and host conditions at that point. It would be evolutionarily detrimental to the phage to attempt lytic replication in a host that is starved or has limited resources necessary for viral growth, since fitness is determined by the number of progeny created. For example, a starved host would likely lack the cellular resources needed to maximize phage production, and low host density would decrease the probability of subsequent infections of nearby hosts by progeny viruses (Ghosh et al., 2009). The working hypothesis is that temperate phages like λ will enter into lysogeny when hosts are starved or host density is low, and essentially "wait" until chances of maximizing reproduction are greater. When conditions finally improve, prophage induction can occur. In λ, the decision of whether to engage the lytic or lysogenic replication pathway is primarily controlled by the cII protein (Ptashne, 1986). cII is susceptible to a variety of generic cellular proteases, and cells that are metabolically active can be expected to have high enough concentrations of these proteases to degrade the cII protein before it can act as a promoter enhancer. In nutrient replete conditions, cells will be actively growing and cII will be degraded by bacterial proteases; therefore, no cI repressor will be synthesized, ultimately leading to lytic replication. However, in nutrient deplete conditions, bacterial growth will be slow or possibly

stopped, and low levels of bacterial proteases will lead to the persistence of cII; therefore, cII is able to activate transcription of genes necessary for lysogeny.

Although these mechanisms of control are well-elucidated in bacteriophage λ, this model and these specific genes and proteins most likely cannot be applied to every temperate phage. However, the important take-away is that temperate phages have complex genetic circuits that allow them to sense their environment and switch between lytic or lysogenic replication, based on evolutionary selection: either increasing the phage's chance at survival or maximizing production of progeny. It therefore follows that varying conditions of nutrient availability, weather conditions, microbial community composition, etc., may require temperate phages to adopt a different replication strategy to maximize survival versus making progeny. In aquatic environments, attempts at determining the prevalence of lysogeny through induction assays have been met with variable results. A study of ice-covered Antarctic lakes reported very high levels of lysogeny, up to 89.5% of the bacterial community contained inducible lysogens (Lisle and Priscu, 2004). In contrast, one study of coastal seawaters found no evidence of lysogeny (Wilcox and Fuhrman, 1994). There have also been conflicting seasonal patterns reported. Williamson et al. (2002) and Cochran et al. (1998) both described trends of increased lysogeny between the months of February and October in Tampa Bay, FL, while a study by Laybourn-Parry et al. (2006) in saline Antarctic lakes reported a trend of highest lysogeny in winter and spring and a decline in summer. It is possible that changes in climate and location may greatly affect these trends, making it even more prudent to study this phenomenon in soils.

Soil environments in particular are chemically, physically, and biologically diverse around the globe, and nutrient availability seems to be strongly correlated to both geographic location and plant characteristics in the area (Jobbagy and Jackson, 2001). It has also been shown

that increased nitrogen and phosphorus input can change the composition of microbial communities (Leff et al., 2015; Koorem et al., 2014), and that soil depth also plays a role in bacterial community composition (Liang et al., 2019). Based on the heterogeneity of soil environments around the globe, it would follow that such diverse environments would select for viral communities and replication strategies best suited to those environments.

Stewart and Levin (1984) and Marsh and Wellington (1994) originally proposed that the lysogenic lifestyle should offer an evolutionary benefit to phages in soil environments, as typical soil characteristics such as long periods of host inactivity and limited opportunities for movement may not allow lytic growth to be optimal for the phage. This hypothesis has held up over the years as more studies on lysogeny in soils have been conducted. Estimates of inducible fractions of bacteria have ranged from 30% in Delaware soils (Williamson et al., 2008) to 4.6 - 21.1% in Antarctic soils (Williamson et al., 2007), and Liang et al. (2019) suggest that the fraction of lysogenic bacteria increases with soil depth. However, there are still very few studies that have focused on lysogeny in soils, and therefore a lack of information exists on possible seasonal trends or variations in the fraction of inducible bacteria in different soil types.

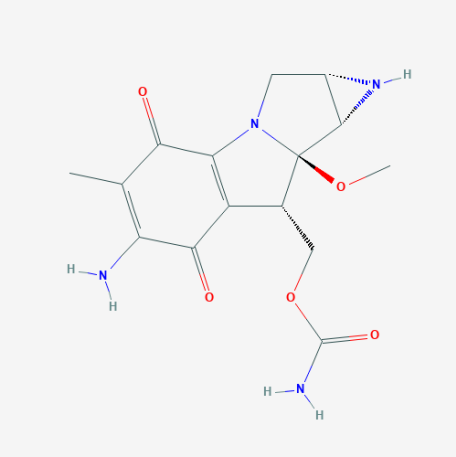


Figure 3. 2D-Structure of mitomycin C (CID 5746, PubChem)

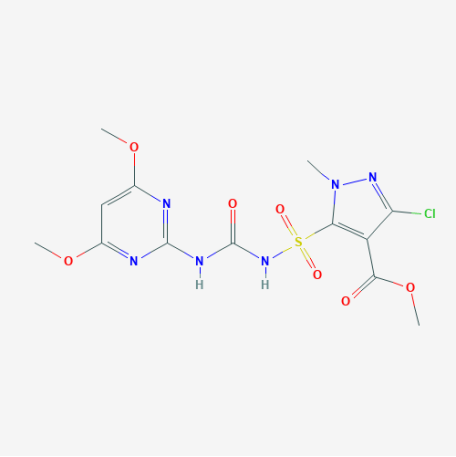


Figure 4. 2D-Structure of halosulfuron-methyl (CID 91763, PubChem)

In most studies to date, an antibiotic called mitomycin C has been used as the gold standard inducing agent (Figure 3). Mitomycin C is an alkylating agent that cross-links complementary strands of DNA and thus inhibits DNA synthesis (Verweij and Pinedo, 1990). Beyond its use in prophage induction assays, mitomycin C has a wide spectrum of antitumor ability, and is used to treat certain types of cancer. However, its usefulness in prophage induction assays has been questioned because mitomycin C is not usually found in natural environments. Another commonly used inducing agent is UV light, which offers a cheaper method than mitomycin C, but unfortunately has been shown to be much less effective of an inducing agent (Loessner et al., 1991).

Alternatively, a previous William & Mary honors thesis project investigated the herbicide SedgeHammer (Gowan Company, Yuma, AZ), which is used around campus to eliminate nutsedge, a nuisance weed (Hart, 2010). The active ingredient in SedgeHammer is halosulfuron-methyl, shown in Figure 4. SedgeHammer inhibits the acetolactate synthase enzyme, which produces three amino acids without which DNA replication cannot continue. In Hart's Honors Thesis work, SedgeHammer caused the largest increase in viral direct counts in aquatic environmental samples tested, showing its potential use as an inducing agent for environmental samples. Finally, bacterial quorum sensing molecules called acyl-homoserine lactones (AHLs) were successfully used (Ghosh et al. 2009) to induce both *E. coli* lambda lysogens and bacterial communities from environmental samples. The results of this study with AHLs supports the hypothesis that host density may play a role in whether a phage replicates through the lytic or lysogenic cycle, as quorum sensing itself is a cell-density dependent phenomenon. More studies to determine the most effective inducing agents will be helpful to optimize future work studying lysogeny, particularly in soils.

The main goals of my project were to: 1) collect monthly soil samples and analyze them for evidence of prophage induction; 2) determine potential seasonal trends in lysogeny in soil microbial communities; and 3) compare the efficacy of different inducing agents across soil, aquatic, and *E. coli* samples. The results of this work will provide insight into the influence of lysogeny in these understudied soil environments, as well as recommend the most effective methods for performing induction assays on environmental samples in future studies.

MATERIALS AND METHODS

Induction of Monthly Soil Samples

Sample Collection:

Soil samples were collected from a field site in the College Woods (Helsley et al., 2014) near Lake Matoaka in Williamsburg, VA, USA (37.269 N, 76.721 W). Approximately 500g of soil was collected with a small trowel via the random walk method and placed into a quart-sized Ziploc plastic bag. The soil sample was then transported to the lab and sieved to 4mm.

Bacterial Extraction:

Ten grams of sieved soil was added to each of 2 pre-chilled blender cups on ice, and 100 ml of chilled (4°C) 1% potassium citrate buffer (per liter:10 g potassium citrate, 1.44 g $Na_2HPO_4 \cdot 7H_2O$, 0.24 g KH_2PO_4, pH 7) (Williamson et al., 2003) was added to each blender cup. Samples were blended on high for 3 minutes and the resulting slurry allowed to settle on ice for approx. 1 min. Slurries were then processed as follows.

For the months of November 2018, December 2018, January 2019, and February 2019: six polyallomer SW 41 Ti tubes were prepared per replicate blender extraction. Two mL of Nycodenz (1.3 g ml^{-1}; stock prepared by dissolving 24g Nycodenz in 30mL 1% potassium citrate buffer) was added to each tube. Then 9 mL of slurry was carefully layered on top using a 10mL serological pipette. Care was taken not to disturb the interface.

For the months of April 2019 and June 2019, three sterile polypropylene Oak Ridge tubes were prepared per replicate blender extraction. Five milliliters of Nycodenz (1.3 g ml^{-1}) were added to each tube. Then 15 ml of slurry were carefully layered on top using a 10 ml serological

pipette. Care was taken not to disturb the interface. Tubes were then centrifuged at 8,000 x g at 4°C for 20 minutes to sediment soil particles. The resulting supernatant was decanted from each tube and pooled in two sterile 50 ml centrifuge tubes. The supernatant was homogenized by pouring back and forth between the two tubes. The homogenized supernatant was then divided between 4 sterile 50 ml tubes, resulting in 16mL supernatant (bacterial extract) in each tube.

Induction Procedure:

For the months of November 2018, December 2018, January 2019, and February 2019, two control and two treatment replicates were prepared; for the months of April 2019 and June 2019, three control and three treatment replicates were prepared. In all cases, treatment tubes received mitomycin C at a final concentration of 0.5 μg ml^{-1}, and control tubes received an equal volume of sterile water (8 μl). Samples were then incubated in a rotary shaker at 140 rpm and 28°C for 24 hours.

Sample Storage:

After 24 hours, samples were removed from the rotary shaker and 1 ml aliquots were dispensed into sterile cryovials. Aliquots were frozen with liquid nitrogen and stored in a -80°C freezer until slide preparation occurred (nominally, 1 – 4 weeks).

Slide Preparation:

Samples were thawed in a warm water bath (~50°C) and then stored on ice. Samples (20 μl) were suspended in sterile deionized water to make up 100 μl total. Suspended samples were immobilized on Whatman Anodisc filter membranes (13mm diameter, 0.02μm pore size, Whatman, Maidstone, England) held in 13mm polypropylene Swinnex filter holders. Using a vacuum, sample was pulled through filter for approximately one minute. Filters were stained

using 100µL of a 2.5 X SYBR Gold solution (ThermoFisher, Waltham, MA), and incubated in the dark for 15 minutes. After incubation, the SYBR Gold solution was vacuumed through the filter holder for approximately one minute, and the filter was removed and allowed to air dry in the dark. Dry Anodisc filters were mounted on glass slides with 10 µl of Antifade (20 mM Tris-HCl, pH 8; 0.5% p-phenylenediamine; 90% glycerol), and covered with a cover slip. Slides were either analyzed immediately or stored at -20°C until analysis (nominally, 2-3 days)

Epifluorescence Microscopy:

An Olympus BX51 microscope (Olympus, Center Valley, PA) fitted with an Olympus U-RFL-T mercury lamp, FITC excitation filter, and 100x/1.30 oil lens was used to image the prepared slides. Fifteen fields per replicate were digitally photographed at ×1,000 magnification with a Hamamatsu C8484 CCD camera. Efforts were made to select fields of view randomly as to properly sample the potential variation within each slide. Photos were captured and analyzed using MetaMorph software (MetaMorph, Nashville, TN).

Inducing Agent Tests

Sample Collection:

Soil samples for both March 2020 and February 2021 were collected and sieved as previously described.

Water samples for the March 2020 collection date were collected from a dock near the Keck Environmental Laboratory on Lake Matoaka in Williamsburg, VA, USA (37.271 N, 76.723 W). Surface water was collected with a 250mL polycarbonate Nalgene bottle by triple-rinsing and then filling the bottle completely.

Water samples for the February 2021 collection date were collected from the Crim Dell Pond on the campus of William & Mary in Williamsburg, VA (37.2706 N, 76.7135 W). Surface water was collected with a 250mL polycarbonate Nalgene bottle by triple-rinsing and then filling the bottle completely.

Bacterial Extractions:

Bacterial extractions were performed on soil samples as described previously. During the March 2020 induction procedure, a mixture of polyallomer SW 41 Ti tubes and Oak Ridge polypropylene tubes were used during the centrifugation process. During the February 2021 induction procedure, multiple rounds of centrifugations were conducted using the SW 41 Ti tubes and the supernatant was refrigerated between rounds.

Escherichia coli W3104 Culture Preparation

E. coli W3104 was purchased from Carolina Biological Supply Co. (Burlington, NC). This strain of *E. coli* is a lambda lysogen and should produce phage particles upon induction. In

a 500 ml flask containing 160 ml of sterile tryptic soy broth, 150 µl of overnight *E. coli* W3104 culture was added. The inoculated media was incubated for 4 hours at 37°C and 200 rpm. Induction assays were conducted immediately after the incubation period had completed and the bacterial culture was in early exponential phase.

Inducing Agents:

Mitomycin C was purchased from three different vendors: Fisher Scientific (Pittsburgh, PA), ApexBio (Houston, TX), and Research Products International (RPI, Mount Prospect, IL); all were added to induced samples at a final concentration of 0.5 µg ml^{-1}. Ultraviolet light inductions were accomplished by decanting samples into sterile petri dishes and exposing them to the germicidal lamp (UV-C) in a Labconco Class II biosafety cabinet (Labconco, Kansas City, MO) for 2 minutes. SedgeHammer (Gowan Company, Yuma, AZ) is a selective herbicide for control of nutsedge and was prepared by dissolving 0.5 g of SedgeHammer in 10 ml of sterile deionized water. SedgeHammer was added to induced samples at a final concentration of 0.05 µg ml^{-1} (Hart, 2010). A mixture of acyl-homoserine lactones (AHLs: *N*-Hexanoyl-L-homoserine lactone and *N*-Tetradecanoyl-DL-homoserine lactone, Sigma-Aldrich, St. Louis, MO) was used, and each AHL was added to induced samples to achieve a final concentration of 1 µM (Ghosh et al., 2008).

Induction Procedure:

For the March 2020 tests, thirty sterile 50 mL tubes were prepared. Fifteen of those tubes received 12 ml of soil bacterial extraction, and the remaining fifteen tubes received 12 ml of whole water from Lake Matoaka. All treatments were done in triplicate; Fisher mitomycin C,

ApexBio mitomycin C, SedgeHammer, and the AHL mix were used as inducing agents. The remaining three tubes were controls and received an equivalent volume (12 µl) of sterile water.

For the February 2021 tests, fifty-four sterile 50 mL tubes were prepared. Eighteen of those tubes received 8 mL of soil extraction, eighteen tubes received 8 mL of whole water from the Crim Dell Pond, and the remaining eighteen tubes received 8 mL of prepared *E. coli* W3104 culture. Treatments were done in triplicate, with each inducing agent type being added to three tubes from each sample type. Fisher mitomycin C, ApexBio mitomycin C, and RPI mitomycin C SedgeHammer, and UV light were used as inducing agents. The remaining three tubes were controls and received 4 µL of sterile water.

In both the March 2020 and February 2021 trials, all samples were blinded (coded) to prevent identification of samples and reduce bias in results. All samples were placed in a rotary shaker at 140 rpm and 28°C for 24 hours. Following this incubation period, 1 ml aliquots of each sample were transferred to sterile cryovials, frozen in liquid nitrogen, and stored at -80°C until slide preparation (nominally, 1 – 4 weeks).

Slide Preparation:

Slide preparation was performed using the same protocol as described previously, with slight variations to the volume of sample added to the filter depending on each sample type: for soil samples, 20 µl sample and 80 µl sterile water was added to the Anodisc filter; for water samples, 100 µl of sample were added to the filter; for *E. coli* samples, 2 µl sample and 98 µl sterile water was added to the filter. The remainder of slide preparation was carried out as described above.

Epifluorescence Microscopy:

Microscopy was conducted using the same microscope and software as described previously.

Data Collection:

Photos were taken using the same protocol as previously described. 15 pictures were taken for each sample.

Each photo for each sample was then analyzed for viral and bacterial abundance data. For March 2020 samples and February 2021 *E. coli* samples, viral and bacterial counts were conducted via the manual counting method described previously. For February 2021 soil and aquatic samples, viruses and bacteria were discriminated from each other (and cell debris) based on pixel dimensions and counted using MetaMorph software (MetaMorph, Nashville, TN).

Data Analysis

Abundance Calculations:

Viruses per ml were calculated using Formula 1:

$$VDC\ ml^{-1} = \frac{VDC_{avg} * A_{filter}}{A_{image} * V}$$

Formula 1

VDC ml^{-1} is the viral direct count per mL of sample. VDC_{avg} is the average direct count of viruses per image. A_{filter} is the total area of the Anodisc filter. A_{image} is the total area of the viewing field. V is the total volume that was filtered through the Anodisc (for soil samples, V = 0.02 ml; for aquatic samples, V = 0.100 ml; for *E. coli* samples, V = 0.002 ml).

Bacteria per ml were calculated using Formula 2:

$$BDC\ ml^{-1} = \frac{BDC_{avg} * A_{filter}}{A_{image} * V}$$

Formula 2

BDC ml^{-1} is the bacterial direct count per mL of sample. BDC_{avg} is the average direct count of bacteria per image. A_{filter} is the total area of the Anodisc filter. A_{image} is the total area of the viewing field. V is the total volume that was filtered through the Anodisc (for soil samples, V = 0.02 ml; for aquatic samples, V = 0.100 ml; for *E. coli* samples, V = 0.002 ml).

Burst size and inducible fraction were calculated using the following formulas:

$$B_Z = \frac{P_I}{B_L}$$

Formula 3

B_Z is the burst size, or the average number of viruses produced per bacterial cell upon lysis. P_I is the number of prophage induced, which is found by subtracting VDC of controls from VDC of induced samples. B_L is the number of bacteria lysed, which is found by subtracting BDC of induced samples from BDC of controls.

$$IF = \frac{P_I / B_Z}{BDC_{control}} \times 100$$

Formula 4

Inducible fraction (IF) is the percentage of cells in the sample that are capable of being chemically induced. P_I and B_Z values were found during burst size calculations. $BDC_{control}$ represents the calculated BDC ml^{-1} of the control samples. Previous studies have used the terms "lysogenic fraction" and "fraction of chemically inducible cells" (Knowles et al., 2017) to describe this phenomenon, but we will use the term "inducible fraction" since this is a more accurate description of the data being collected during this study (Williamson et al., 2007).

Statistical Analysis:

Analyses were conducted using GraphPad Prism software (GraphPad, San Diego, CA). For monthly samples, a two-tailed unpaired t-test with a 95% CI was conducted on controls and treatments to determine statistically significant differences between the two groups. Significance was defined as $p \leq 0.05$. For induction comparison experiments, a one-way ANOVA was run with Tukey's multiple comparison post-test across all sample types to determine significance.

RESULTS

Monthly Sampling and Inductions

For each month, inductions were performed using ApexBio mitomycin C as the inducing agent. Viral direct counts (VDC) per milliliter of soil extraction were determined for each sample using epifluorescence microscopy. Mean VDC ranged from a minimum of 4.39×10^5 ml^{-1} in February control samples to a maximum of 1.57×10^6 ml^{-1} in December control samples. An unpaired t-test was conducted for each month, showing no significant increase in VDC between viral controls and treatments for the months of November, December, January, February, or June (November: $p = 0.3784$; December: $p = 0.3062$; January: $p = 0.2759$; February: $p = 0.2865$; June: $p = 0.6249$). However, a significant increase in VDC occurred in the April samples ($p = 0.0035$) (Figure 5A-F).

Monthly VDC Comparisons

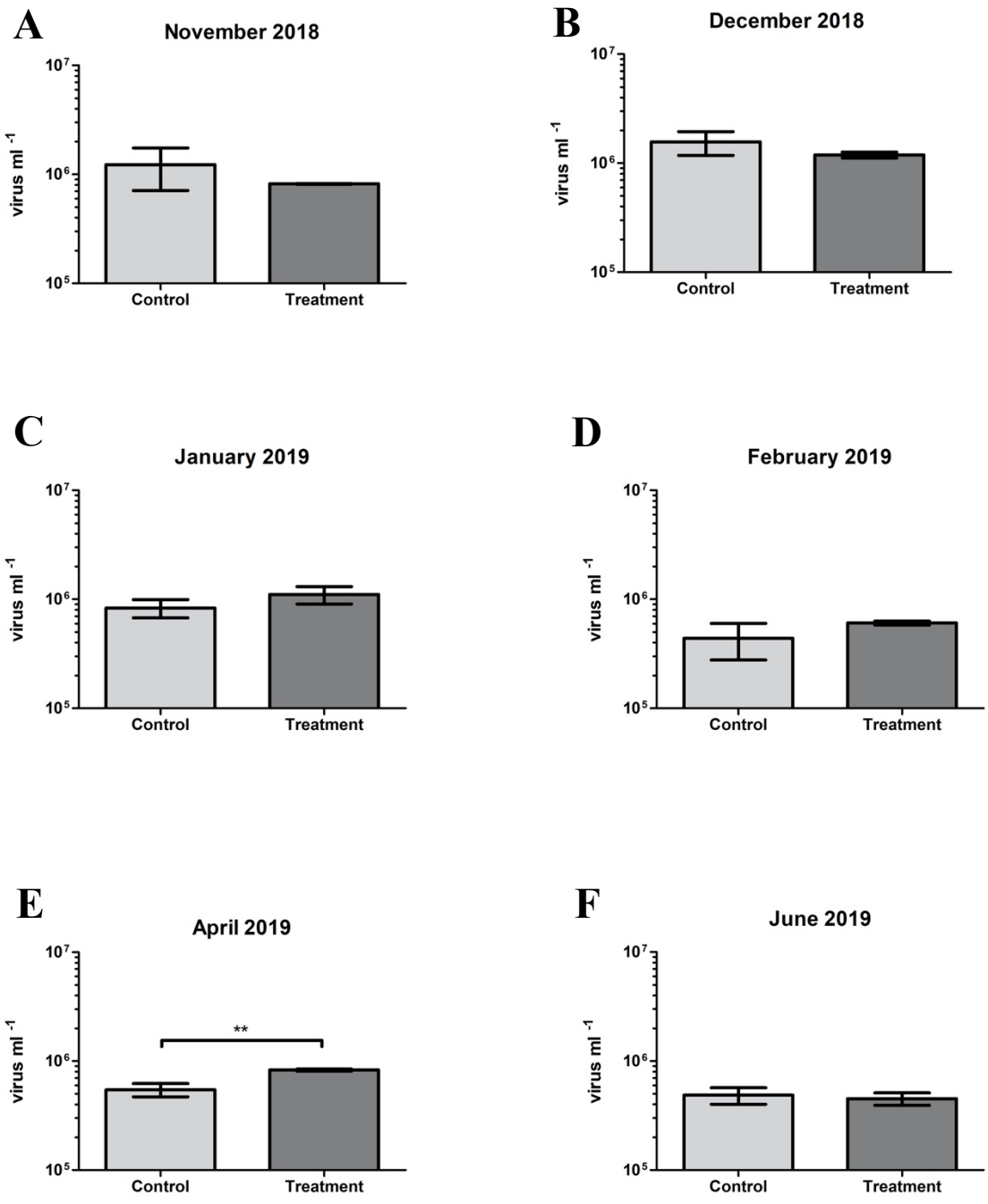


Figure 5. Comparing VDC in control and treatment samples for each time point. Error bars represent standard deviation (SD). N = 2 for the months of November, December, January, and February; N = 3 for the months of April and June. Asterisks indicate level of significance (signifies p ≤0.01).**

The number of prophage induced (P_I) by mitomycin C was determined for each monthly sample by subtracting VDC of control samples from VDC of treatment samples (Figure 6). A positive value of P_I was observed in January, February, and April samples, with no significant differences between these three data points. November, December, and June samples exhibited a decrease in VDC in treatments compared to controls. This led to negative values which do not make biological sense for later burst size and inducible fraction calculations, so these values were excluded from this graph.

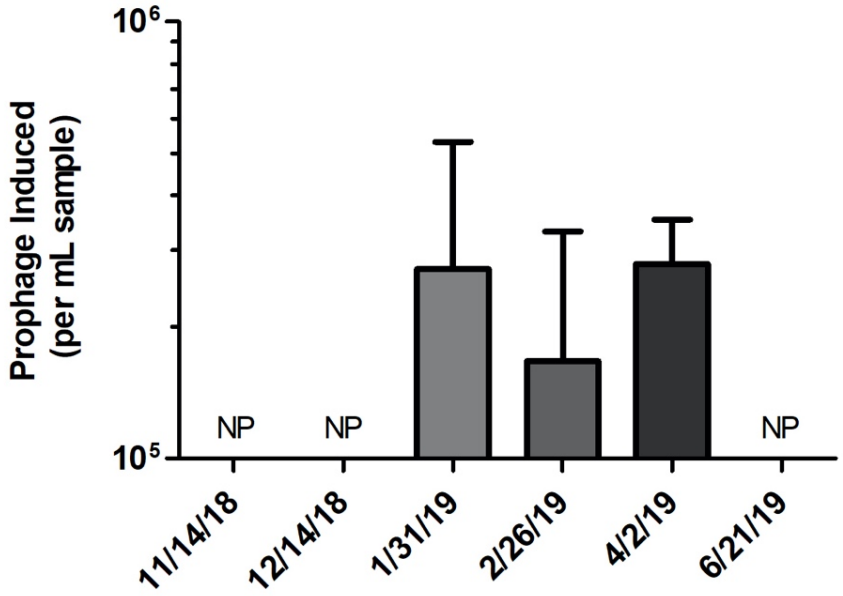


Figure 6. The number of prophage induced per mL of monthly soil extraction. Error bars represent standard deviation (SD). N = 2 for the months of November, December, January, and February; N = 3 for the months of April and June. NP = not presented due to negative value of prophage induction.

Bacterial direct counts (BDC) were determined for each sample using epifluorescence microscopy. Mean BDC ranged from a minimum of 2.40×10^6 ml^{-1} in February treatment samples to a maximum of 6.95×10^6 ml^{-1} in April control samples. The number of bacteria lysed (B_L) assumed due to prophage induction was calculated by subtracting BDC of treatments from BDC of controls (Figure 7). Positive values for B_L were observed in December, February, April, and June samples; whereas a slight increase in BDC in treatments compared to controls occurred in the months of November and January. This resulted in a negative value for B_L, which was excluded from the graph. A significant difference ($p \leq 0.05$) was only observed between the February and April samples.

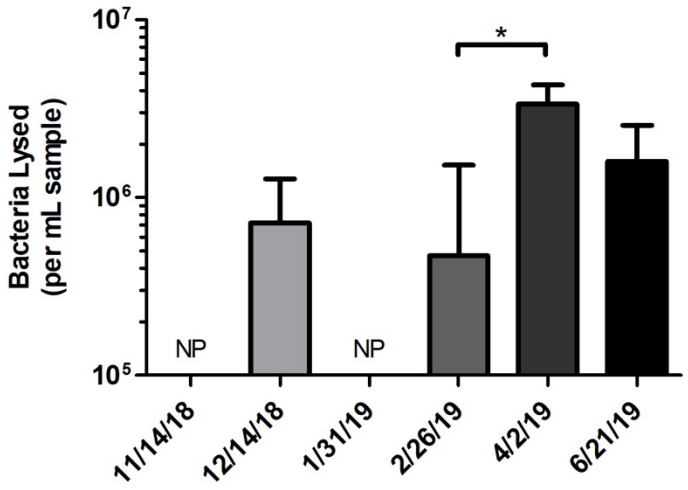


Figure 7. The number of bacteria lysed per mL of monthly soil extraction. Error bars represent SD. N = 2 for the months of November, December, January, and February; N = 3 for the months of April and June. NP = not presented due to negative value of bacterial lysis. Asterisks indicate level of significance (* represents $p \leq 0.05$).

Sample	*Prophage Induced (ml^{-1} extraction)*	*Bacteria Lysed (ml^{-1} extraction)*	*Calculated Burst Size (B_Z)*	*Inducible Fraction (%) based on calculated B_Z*	*Inducible Fraction (%) based on B_Z = 20*
11/14/18	---	---	---	---	---
12/14/18	---	$7.20 \times 10^5 \pm 5.51 \times 10^5$	---	---	---
1/31/19	$2.72 \times 10^5 \pm 2.59 \times 10^5$	---	---	---	---
2/26/19	$1.67 \times 10^5 \pm 1.64 \times 10^5$	$4.71 \times 10^5 \pm 1.05 \times 10^6$	0.354 ± 0.864	16.421 ± 43.306	0.201 ± 0.099
4/2/19	$2.79 \times 10^5 \pm 7.84 \times 10^4$	$3.36 \times 10^6 \pm 9.52 \times 10^5$	0.083 ± 0.033	48.362 ± 23.783	0.291 ± 0.768
6/21/19	---	$1.59 \times 10^6 \pm 9.56 \times 10^5$	---	---	---

Table 1. Calculations of inducible fraction. Values are represented as mean between replicates ± standard deviations.

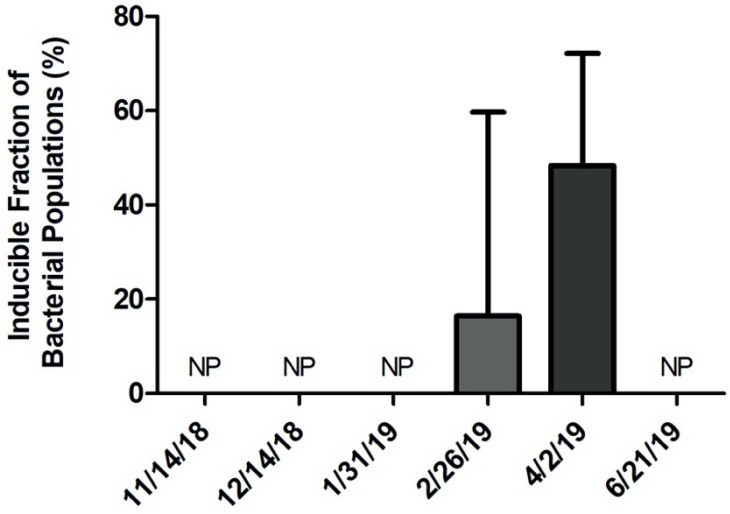


Figure 8. Inducible fraction of the bacterial population using calculated burst size. Error bars represent SD. N = 2 for the months of November, December, January, and February; N = 3 for the months of April and June. NP = not presented due to negative value of prophage induction or bacterial lysis.

For samples that exhibited positive induction (positive values of both P_I and B_L), burst size and inducible fraction were calculated (Table 1). Inducible fraction was determined using both the calculated burst size for each sample (Figure 8) and an assumed burst size of 20 commonly used in the literature (Williamson et al., 2007). The April 2019 samples seemed to have the highest inducible fraction using both methods of calculation, suggesting that lysogeny was the most prevalent in these samples.

March 2020 Inducing Agent Tests

Soil Sample Comparisons

Bacteria were extracted from the College Woods field site. Inductions were performed using Fisher mitomycin C, ApexBio mitomycin C, SedgeHammer, and AHLs as the variable inducing agents. VDC for treatment and control samples were determined via epifluorescence microscopy, and mean VDC ranged from a minimum of 2.77×10^5 ml^{-1} in controls to a maximum of 1.34×10^6 ml^{-1} in SedgeHammer treatments. A one-way ANOVA was run with a Tukey's multiple comparison post test to determine significant increases in VDC between controls and treatments. A significant increase in VDC was observed relative to controls for all inducing agents tested, with the highest level of significance correlating to the ApexBio, SedgeHammer, and AHL treatments (Figure 9).

BDC were determined for each sample using epifluorescence microscopy. Mean BDC ranged from a minimum of 1.69×10^6 ml^{-1} in control samples to a maximum of 4.98×10^6 ml^{-1} in SedgeHammer samples. All treatment samples exhibited an increase in BDC compared to controls.

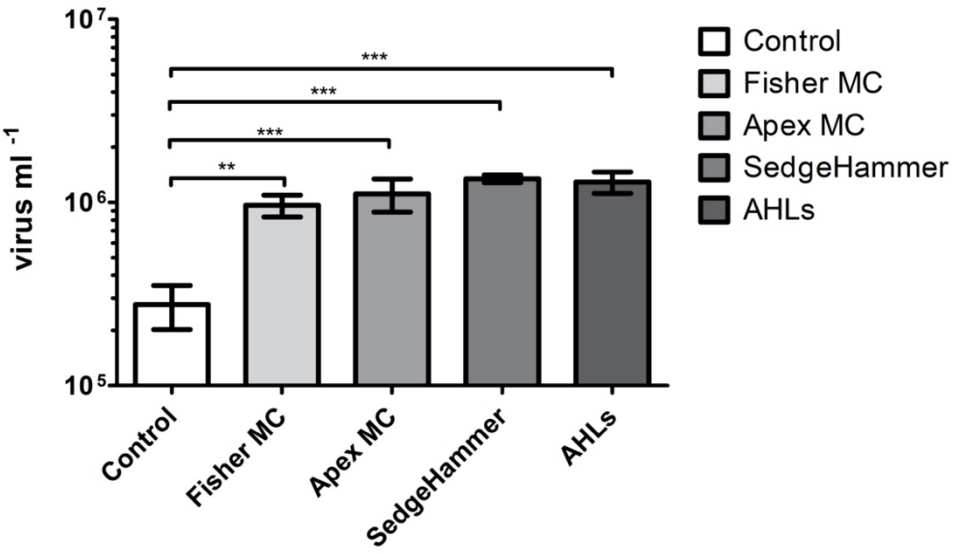


Figure 9. Comparison of VDC for March 2020 soil inducing agent tests. Error bars represent SD (N = 3). Asterisks indicate level of significance (represents $p \leq 0.01$; *** represents $p \leq 0.001$).**

Aquatic Sample Comparisons

Surface water samples were obtained from Lake Matoaka at the Keck Lab dock. Inductions were performed using Fisher mitomycin C, ApexBio mitomycin C, SedgeHammer, and AHLs as the variable inducing agents. VDC for treatment and control samples were determined via epifluorescence microscopy, and mean VDC ranged from a minimum of 1.09×10^5 ml^{-1} in controls to a maximum of 4.72×10^5 ml^{-1} in AHL treatments. A one-way ANOVA was run with a Tukey's multiple comparison post test to determine significant increases in VDC between controls and treatments. A significant increase was only observed in AHL treatments, however, all inducing agents still exhibited an increase in VDC compared to controls (Figure 10).

BDC were determined for each sample using epifluorescence microscopy. Mean BDC ranged from a minimum of 2.07×10^5 ml^{-1} in Fisher mitC samples to a maximum of 5.40×10^5 ml^{-1} in SedgeHammer samples. The only induced sample that exhibited a decrease in BDC compared to controls was Fisher MC.

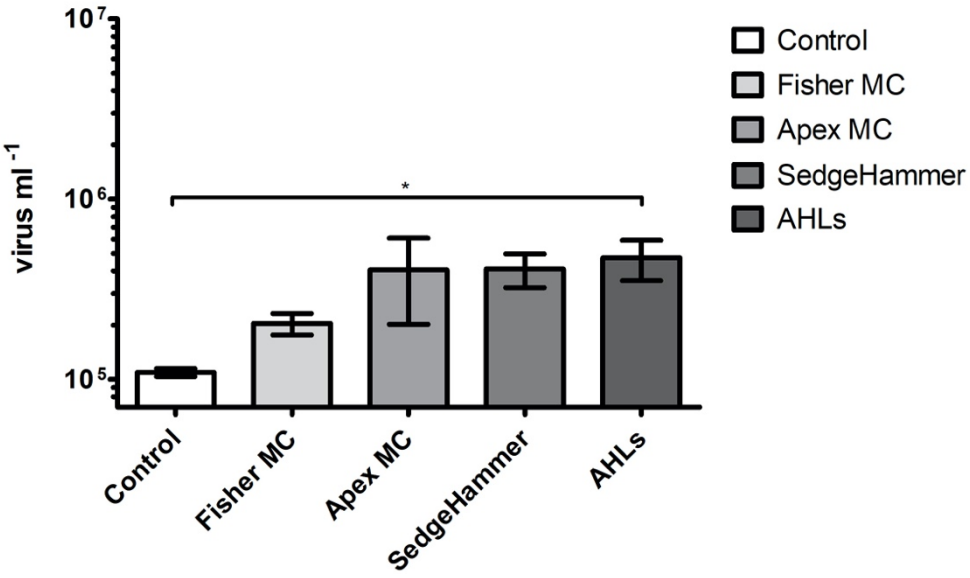


Figure 10. Comparison of VDC for March 2020 aquatic sample (Lake Matoaka) inducing agent tests. Error bars represent SD (N = 3). Asterisks indicate level of significance (* represents $p \leq 0.05$).

Inducible Fraction and Other Characteristics

The number of prophage induced by each inducing agent for both soil and aquatic inducing agent comparisons was determined by subtracting VDC of control samples from VDC of treatment samples (Figure 11). A positive value of prophage induced was found for all samples, and ranged from a minimum of 9.75×10^4 ml^{-1} for aquatic Fisher MC to a maximum of 1.07×10^6 ml^{-1} for soil SedgeHammer samples. No significant differences were found between inducing agent types within soil and aquatic sample types.

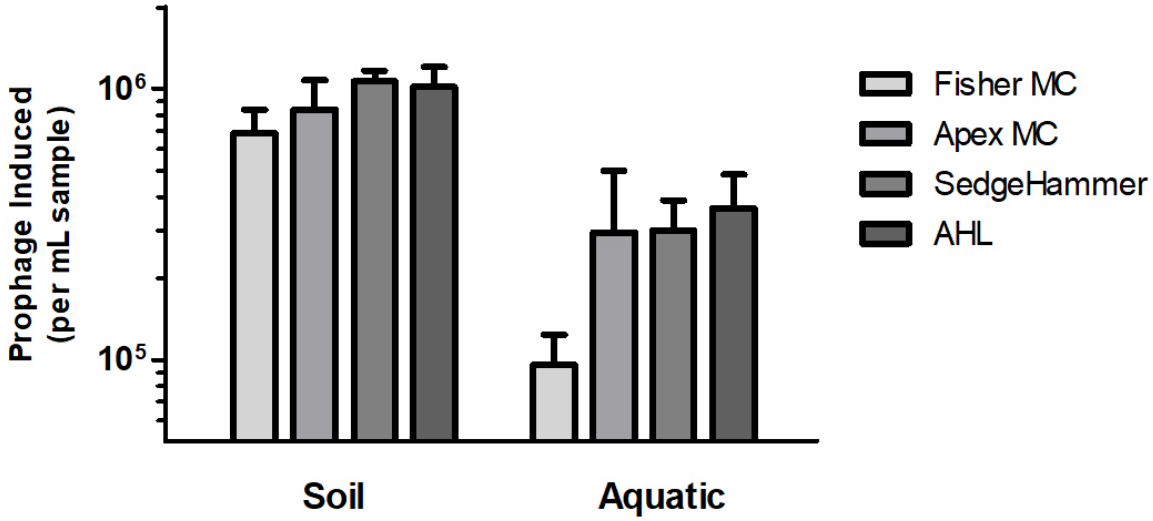


Figure 11. The number of prophage induced per mL of sample for both soil and aquatic inducing agent tests. Error bars represent SD (N = 3).

The number of bacteria lysed by prophage induction for both soil and aquatic inducing agent comparisons was calculated by subtracting BDC of treatments from BDC of controls. No soil treatments were found to have a positive value of B_L. The only aquatic treatment with a positive value of B_L was Fisher MC, with $B_L = 2.09 \times 10^4$ ml^{-1}. Therefore, no calculations of burst size or lysogenic fraction could be conducted on the other samples, as a negative B_L value would lead to meaningless values for these measures. The calculated burst size of Lake Matoaka bacteria induced with Fisher MC was 4.571 ± 8.561, and the inducible fraction was 9.170 ± 17.389.

February 2021 Inducing Agent Tests

Soil Sample Comparisons

Inductions were performed using Fisher mitomycin C, ApexBio mitomycin C, RPI mitomycin C, SedgeHammer, and UV light as the inducing agents. VDC for treatment and control samples were determined via epifluorescence microscopy, and mean VDC ranged from a minimum of 5.79×10^6 ml^{-1} in UV treatments to a maximum of 2.05×10^7 ml^{-1} in Apex MC treatments. A one-way ANOVA was run with a Tukey's multiple comparison post test to determine significant increases in VDC between controls and treatments. No significant increases in VDC were observed for any of the inducing agents, although there were slight increases in the means of Fisher MC, Apex MC, and RPI MC samples when compared to controls, suggesting that mitomycin C may be the best inducing agent in this case (Figure 12).

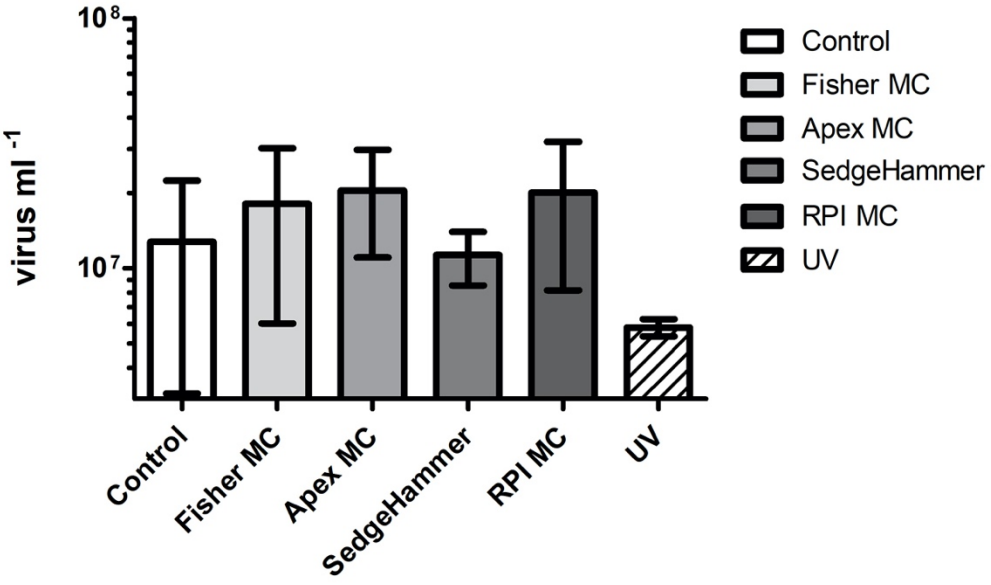


Figure 12. Comparison of VDC for March 2020 soil inducing agent tests. Error bars represent SD (N = 3).

BDC were determined for each sample using epifluorescence microscopy. Mean BDC ranged from a minimum of 2.41×10^6 ml^{-1} in RPI mitC samples to a maximum of 4.68×10^6 ml^{-1} in UV samples. Fisher MC, Apex MC, and RPI MC induced samples exhibited a decrease in BDC compared to controls.

Aquatic Sample Comparisons

Inductions were performed using Fisher mitomycin C, ApexBio mitomycin C, RPI mitomycin C, SedgeHammer, and UV light as the inducing agents. VDC for treatment and control samples were determined via epifluorescence microscopy, and mean VDC ranged from a minimum of 5.79×10^6 ml^{-1} in UV treatments to a maximum of 2.05×10^7 ml^{-1} in Apex MC treatments. A one-way ANOVA was run with a Tukey's multiple comparison post test to determine significant increases in VDC between controls and treatments. None of the groups exhibited a statistically significant increase in VDC, although all treatment groups except for UV light increased slightly in comparison to controls (Figure 13).

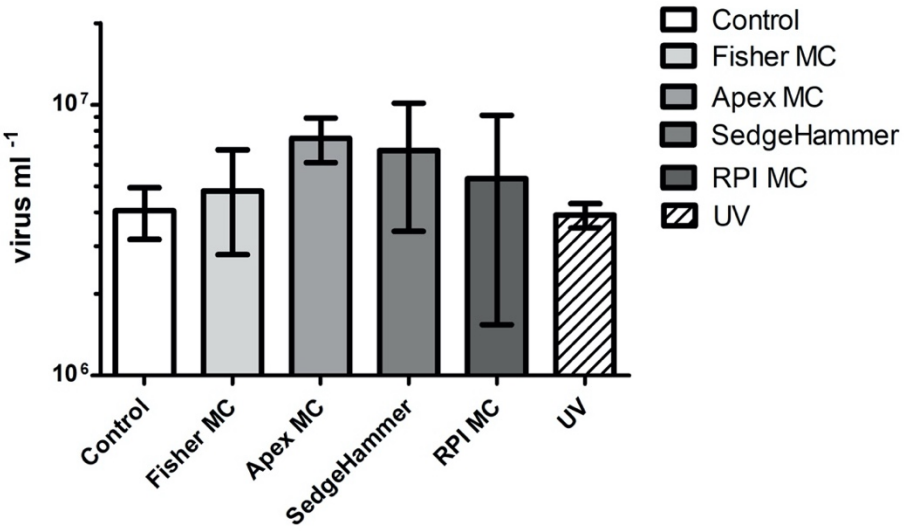


Figure 13. Comparison of VDC for March 2020 aquatic inducing agent tests. Error bars represent SD (N = 3).

BDC were determined for each sample using epifluorescence microscopy. Mean BDC ranged from a minimum of 4.53×10^5 ml^{-1} in RPI MC treatments to a maximum of 9.01×10^5 ml^{-1} in SedgeHammer treatments. Apex MC, RPI MC, and UV light exhibited a decrease in BDC compared to controls, while the other two treatments did not.

E. coli Sample Comparisons

Inductions were performed using Fisher mitomycin C, ApexBio mitomycin C, RPI mitomycin C, SedgeHammer, and UV light as the variable inducing agents. VDC for treatment and control samples were determined via epifluorescence microscopy, and mean VDC ranged from a minimum of 5.15×10^6 ml^{-1} in UV treatments to a maximum of 1.16×10^8 ml^{-1} in Apex MC treatments. A one-way ANOVA was run with a Tukey's multiple comparison post test to determine significant increases in VDC between controls and treatments. None of the groups exhibited a statistically significant increase in VDC, although treatment with both Apex and RPI MC caused an increase in VDC (Figure 14).

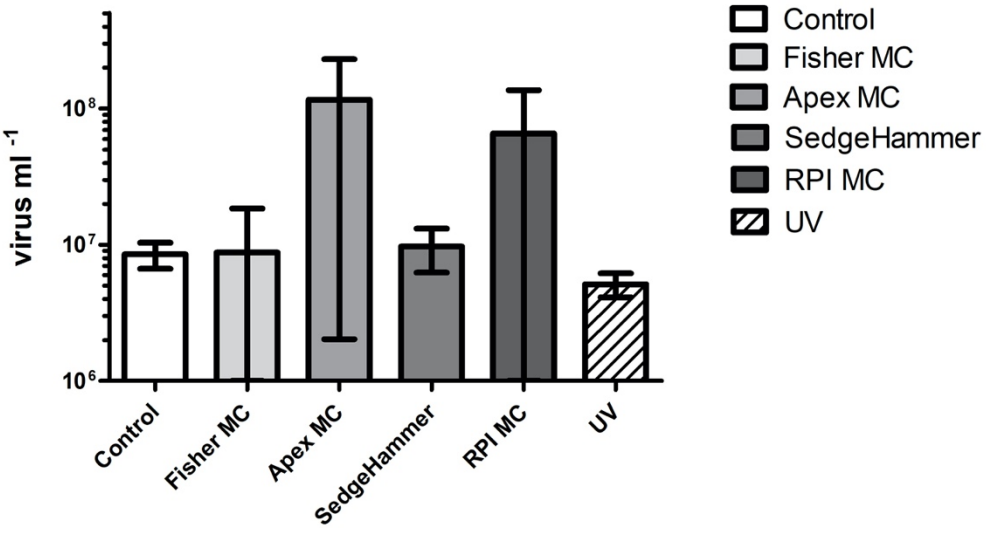


Figure 14. Comparison of VDC for March 2020 *E. coli* inducing agent tests. Error bars represent SD (N = 3).

Inducible Fraction and Other Characteristics

The number of prophage induced by each inducing agent for both soil and aquatic inducing agent comparisons was determined by subtracting VDC of control samples from VDC of treatment samples. A positive value of P_I was found for every sample except soil bacteria induced with SedgeHammer, soil bacteria induced with UV, and aquatic bacteria induced with UV; values ranged from a minimum of 7.22×10^5 ml^{-1} in aquatic Fisher MC treatments to a maximum of 7.71×10^6 ml^{-1} in soil Apex MC treatments. No significant differences were found between inducing agent types within soil and aquatic sample types (Figure 15).

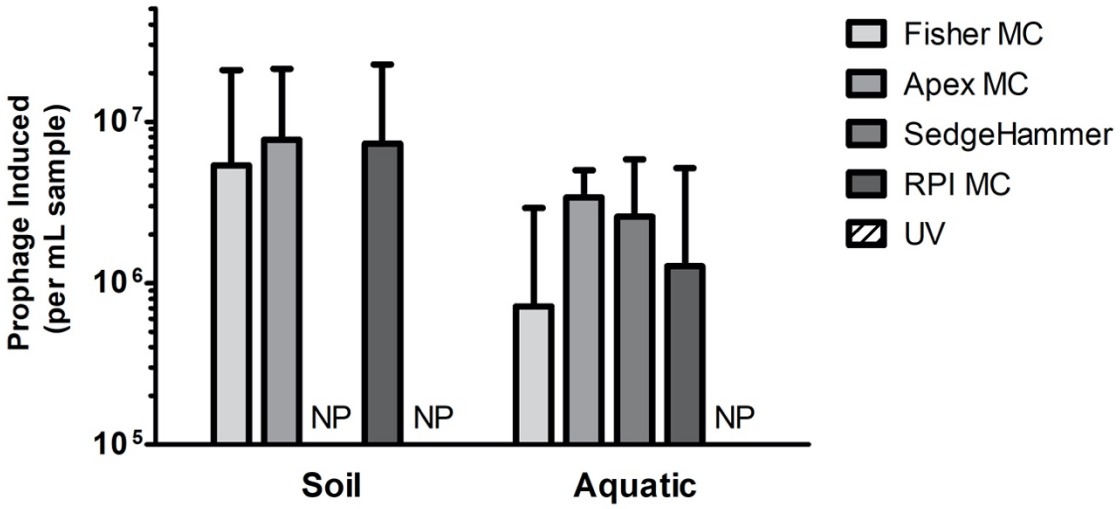


Figure 15. The number of prophage induced per mL of sample for both soil and aquatic inducing agent tests. Error bars represent SD (N = 3). NP = not presented due to negative value of prophage induction.

The number of bacteria lysed by prophage induction for both soil and aquatic inducing agent comparisons was calculated by subtracting BDC of treatments from BDC of controls. A positive value of B_L was found for every sample except soil SedgeHammer, soil UV, aquatic Fisher MC, and aquatic SedgeHammer treatments. B_L values ranged from a minimum of 8.73×10^4 ml^{-1} in aquatic Apex MC treatments to a maximum of 1.39×10^6 ml^{-1} in soil RPI MC treatments. No significant differences were found between inducing agent types within soil or aquatic sample types (Figure 16).

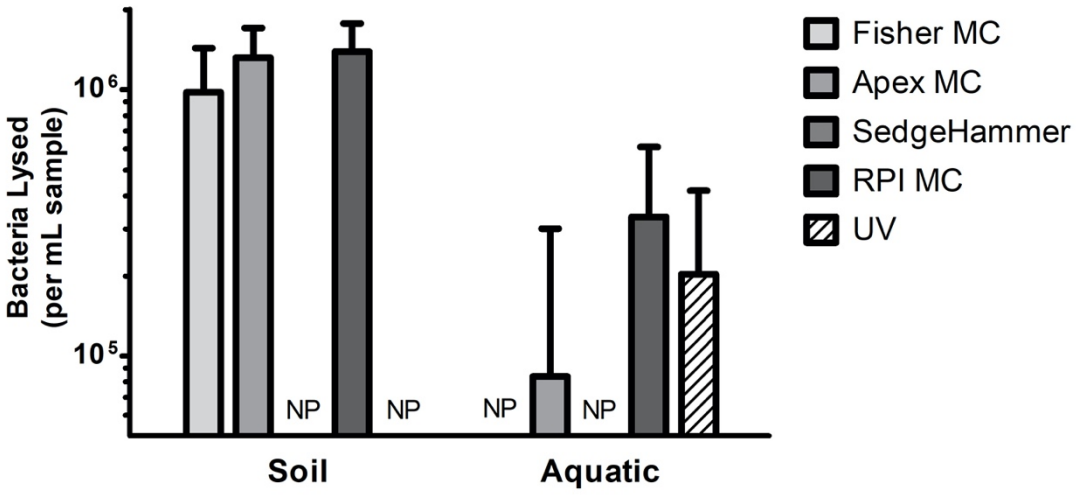


Figure 16. The number of bacteria lysed per mL of sample for both soil and aquatic inducing agent tests. Error bars represent SD (N = 3). NP = not presented due to negative value of bacterial lysis.

For samples that exhibited positive induction (positive values of both P_I and B_L), burst size and inducible fraction were calculated. Inducible fraction was determined using the burst size calculated for each sample and ranged from a minimum of 10.664 ± 28.694 in aquatic Apex MC treatments to a maximum of 42.361 ± 186.549 in aquatic RPI MC treatments (Figure 17).

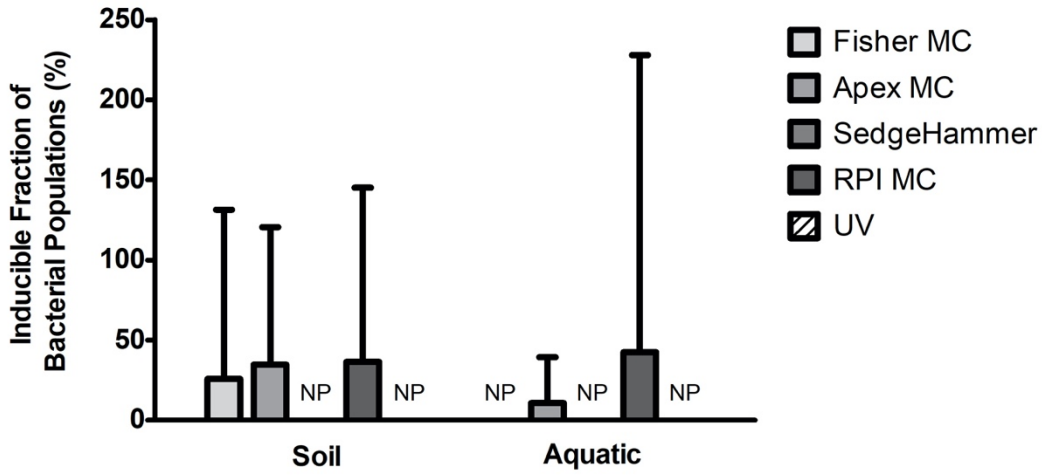


Figure 17. Inducible fraction of the bacterial population in soil and aquatic samples using calculated burst size. Error bars represent SD (N = 3). NP = not presented due to negative value of prophage induction or bacterial lysis.

DISCUSSION

This project began with the goal of elucidating potential seasonal trends in lysogeny in soil bacterial communities. As time went on, it became clear that achieving this goal was more difficult than originally thought, as evidence of lysogeny was scarce in the results of the monthly induction assays. In an attempt to determine whether the inducing agent used in the monthly experiments was the cause of these puzzling results, follow-up experiments were developed to compare the efficacy of different inducing agents in causing prophage induction. Six inducing agents were tested throughout two separate experiments, but gave rise to highly variable results that will be discussed below.

Limitations of this Study

The first limitation of this study was the small sample size. As this was an individual project, both time and resources limited how many replicates could be examined for each monthly time point and inducing agent comparison. Although the induction protocol is fairly simple, the process of collecting quantitative data via epifluorescence microscopy is time-consuming and limited the number of samples that could be collected and processed. For the first four months of sampling, I conducted the experiment in duplicate, and then increased to conducting the experiment in triplicate for the last two months. All inducing agent comparison tests were conducted in triplicate as well. Even though a sample size of three is statistically stronger than a sample size of two, additional replicates would capture the natural variability of

these environmental samples, as well as increase the strength of any statistical testing in picking out meaningful differences amidst a noisy background.

Secondly, because of this natural variability, our monthly or single time point samples may have been too infrequent to represent the complexity of these ecosystems. Both soil and aquatic ecosystems can be affected by many outside factors such as temperature, nutrient availability, soil water content (Williamson et al., 2017), etc., and it is known that such factors influence microbial community composition. In future studies, it would be beneficial to collect samples on a more frequent basis in order to capture the effects of these potential changes in microbial community composition.

Lastly, there may be some inherent limitations in the methods used to enumerate viral and bacterial direct counts via epifluorescence microscopy. It is assumed in these experiments that any increase in viral direct counts (VDC) observed in treatment samples is due to release of virus particles upon bacterial lysis. However, background lytic infections could also contribute to differences in VDC between treatment and control samples. There may be additional background induction occurring in response to natural mechanisms, or even bacteria that are metabolically inactive and unable to induce. All of these scenarios could potentially affect estimates of burst sizes and of lysogenic fraction. It is also assumed that changes in bacterial direct counts (BDC) are solely due to induction-mediated cell lysis, but this may exclude bacterial mortality due to other factors, potentially decreasing estimates of burst sizes and of lysogenic fraction. As will be discussed in more detail below, it is also possible for bacterial cells to continue dividing even as prophage induction progresses, paradoxically leading to increases in BDC concomitant with increases in VDC. This possibility poses eminent challenges to interpreting inducible fraction in various bacterial communities based on the present mathematical equations. Finally, it is

important to note that microscopy-specific issues such as staining artifacts, appearance of virus particles in different focal planes, or even misclassification of virus particles vs. small bacterial cells during manual counts may introduce random errors with unknown effects on the results of this study.

Monthly Sampling and Inductions

Between the months of November 2018 and June 2019, six monthly samples were collected from soils in the College Woods near the campus of William and Mary in Williamsburg, VA. The samples were then analyzed for evidence of prophage induction. Based on the observed results for these samples, it is difficult to elucidate a temporal trend in lysogeny in the soil bacterial community (Figures 5 - 8). In a purely biological sense, one would expect that induction would be characterized by an increase in viral counts and a decrease in bacterial counts in treatment samples compared to controls, since as the switch from lysogenic to lytic replication occurs, extracellular phage particles are released, and bacterial cells are lysed following induction. However, this is difficult to reliably observe in environmental samples, as there is often a lot of background noise in the sample, or the changes in abundance are too small to observe. Examples of background noise may include a high level of virus particles that crowd the slide and make increases difficult to detect, or even fluorescence in the images that may reduce accuracy in counting. In soil samples in particular, the bacterial extraction method yields an extraction that is still very concentrated with soil microorganisms and needs to be diluted during slide preparation in order to clearly distinguish viral particles during microscopy and counting. Combined with the sometimes-subjective nature of the manual counting protocol,

small increases or decreases in viral or bacterial counts in treatments relative to controls may be obscured.

Therefore, I originally analyzed the monthly samples by defining induction as a significant increase in VDC as the most accurate determinant of induction. Using this definition, only the April samples induced, as that was the only month with a significant increase in VDC in treatments compared to controls (Figure 5). I then incorporated the bacterial direct counts using the inducible fraction formula, which is the most commonly used method in the literature to determine prevalence of lysogeny in a sample. First, burst size (B_Z), or the average number of viruses produced per cell upon lysis, is calculated by dividing the number of prophage induced by the number of bacteria lysed. The inducible fraction (IF) formula then uses this burst size to determine the percentage of cells in the sample that induced. The calculated values for prophage induced, bacteria lysed, and inducible fraction need to be positive in order for the formula to be valid.

When applied to our monthly data, only the February and April samples met this criterion (Table 1). These samples yielded an inducible fraction of ~16% and ~48%, respectively. As this is the first study examining the possibility of seasonal lysogeny in soils, it is impossible to directly compare these calculated values to existing literature. In similar studies in aquatic environments, there seemed to be a trend of higher prevalence of lysogeny occurring from the months of February to October (Williamson et al., 2002; Cochran and Paul, 1998). Our results seem to exhibit a similar trend, as a positive lysogenic fraction was observed in the months of February and April, but not in the months of November, December, or January. Our June samples exhibited a negative value of prophage induced, and I was not able to collect samples during the summer months due to summer break, but it would be interesting to continue monthly

sampling in future work and determine if this trend is reproducible. Even so, the error value associated with both calculated IF was very large (Table 1). These calculations also yielded burst sizes < 1, which doesn't make biological sense, as a fractional number of viruses could not be produced upon bacterial lysis. This discrepancy raises doubts on the ability of these methods to accurately and/or reproducibly measure lysogeny in soil bacterial communities.

Indeed, in previous studies, attempts to estimate the prevalence of lysogeny of bacteria in natural soil environments have led to mixed results. These variable results are evidenced by an IF of 4 - 20% found in Antarctic soils, an IF of 22 - 68% found in temperate Delaware soils (Williamson et al., 2007), and an IF of 80 - 89% found in a study that utilized Bio-Sep beads to sample active soil microbiota (Ghosh et al., 2008). This variability is exacerbated by the fact that there are still very few studies focusing on lysogeny in soils and IF estimates in aquatic environments cannot necessarily be applied to soils. It is therefore difficult to know whether the findings of our study are consistent with the estimates of lysogeny found in previous work in soil environments because of the lack of consistency across these previous studies.

Inducing Agent Comparisons

After obtaining these results for the monthly samples, I began to wonder if the use of a different inducing agent could potentially affect the outcome of these experiments. The majority of induction studies use mitomycin C. In fact, as of 2017 there have been approximately 40 independent studies that used mitomycin C to study rates of lysogeny in natural environments (Knowles et al., 2017). Its mechanism of action involves causing DNA damage to induce prophage, which is very effective but also can easily be over- or under-dosed, leading to false

negatives or inhibition of samples, respectively. Another commonly-used inducing agent is UV light, which has a similar mechanism of damaging bacterial DNA as mitomycin C. However, studies have shown UV light to be significantly less effective in inducing prophage compared to mitomycin C, as well as less consistent (Williamson et al., 2007; Loessner et al., 1991).

Over the years, some novel inducing agents have shown some promising effects. One of these novel approaches is to use environmental pollutants such as herbicides (Hart, 2010), aliphatic and aromatic hydrocarbons (Jiang and Paul, 1996), and even sunscreen (Danovaro and Corinaldesi, 2003) to induce prophage. These pollutants may represent a more realistic mode of induction in natural environments, as a main criticism of mitomycin C is that it doesn't represent realistic responses to inducing agents that would naturally be found in these environments. Among the environmental pollutants that may act as a prophage inducing agent is SedgeHammer, an herbicide commonly used on the William and Mary campus to treat nutsedge and other weeds. SedgeHammer inhibits the bacterial acetolactate synthase enzyme, eventually causing a halt in DNA replication and, potentially, subsequent prophage induction. It has been shown to induce prophage in ambient lake water (Hart, 2010), but had not yet been tested in soils. Finally, there is a long-standing hypothesis that bacterial density may play an important role in prophage induction, which seems to be supported by a 2009 study that showed the ability of essential gram-negative quorum sensing molecules called acyl-homoserine lactones (AHLs) to induce phage lambda in *E. coli* systems, as well as show positive induction in environmental bacterial assemblages (Ghosh et al., 2009).

I believed it would be beneficial to determine the efficacy of these novel inducing agents as compared to the more traditional mitomycin C and UV approaches, as well as to test them in more diverse environmental samples. I also wanted to compare mitomycin C purchased from

different vendors: my monthly experiments used mitomycin C purchased from ApexBio, whereas the majority of previously published induction assays used mitomycin C purchased from Fisher Scientific. I hypothesized that perhaps prophage induction efficacy differs between brands.

The first inducing agent comparison was conducted in March of 2020, during which I compared Fisher MC, ApexBio MC, SedgeHammer, and AHLs on both soil and aquatic bacterial communities. In soil samples, all the inducing agents tested elicited a strong induction response, as measured by significant increases in VDC relative to controls (Figure 9). Out of all inducing agents tested, SedgeHammer induced the largest number of prophage, but differences between inducing agents were not statistically significant. Aquatic samples did not have as strong of an induction response, as only AHLs exhibited a significant increase in VDC relative to controls (Figure 10). However, all other inducing agents still had positive values of prophage induction. Bacterial counts for this experiment were less promising, and all but one sample (aquatic Fisher MC) exhibited negative values of bacterial lysis. Therefore, the inducible fraction could only be calculated for this one sample. Unlike the monthly samples, however, the burst size and inducible fraction calculations made biological sense and agreed with literature values for other soil and aquatic induction studies.

In order to confirm these findings, I conducted a second round of inducing agent comparisons. This experiment was conducted in February of 2021, since the COVID-19 pandemic delayed research capacities almost immediately after the first inducing agent comparisons were performed. This time, I used Fisher MC, ApexBio MC, Research Products International (RPI) MC, SedgeHammer, and UV light as the inducing agents, and I conducted this experiment using bacterial communities extracted from soil samples, aquatic samples, and

Escherichia coli W3104, a lysogen carrying phage lambda. This experiment produced even more variable results than the previous one, which raised questions as to the reasons behind these confusing findings.

The *E. coli* W3104 samples were originally included in the experimental design to act as a control to the environmental samples, since W3104 is a known lysogen that should have reliably induced. Before conducting the full inducing agent comparisons, multiple preliminary tests were conducted that confirmed the strain's ability to be induced with UV light (data not shown), thus I was confident that induction would be observed with the UV light treatment in the larger trial. However, no inducing agents exhibited a significant increase in VDC compared to controls, although both Apex and RPI MC showed positive values for prophage induction. Most surprisingly, UV-treated samples actually showed a decrease in VDC compared to controls, and Fisher MC and SedgeHammer VDC increased negligibly (Figure 14).

In soil and aquatic samples, results were equally as variable. None of the inducing agents elicited a significant increase in VDC compared to controls, although values of prophage induction were positive across both sample types for all three mitomycin C treatments. SedgeHammer treatments only showed positive values of prophage induction in aquatic samples, and UV-treated samples showed negative values of prophage induction across the board (Figure 15). However, as opposed to the March 2020 inducing agent comparison, most of the samples exhibited positive values for bacteria lysed, which made it possible to perform calculations of inducible fraction for five out of ten samples. Although the calculated burst size and inducible fractions for these samples agreed with literature values, their associated errors were very large.

These two experiments gave us some valuable insight into which inducing agents may work best in these environmental samples. I believe it is safe to say that UV light was consistently the least effective, as it did not consistently elicit prophage induction. Next, there is a definitive middle field of inducing agents, including SedgeHammer, AHLs, Fisher MC, and RPI MC. Although AHLs and RPI MC consistently exhibited higher levels of prophage induction as compared to UV, they were only tested in one of the two experiments, and it is therefore difficult to definitively compare their efficacy. AHLs are also expensive to purchase, need to be prepared in cocktails to maximize response, and not all bacteria in the sample can be expected to respond to AHLs. These factors limit the standard use of AHLs as an inducing agent in future assays. Fisher MC and SedgeHammer showed perhaps the most variable results, as they exhibited higher levels of prophage induction in some experiments but showed negative values in others. Finally, the inducing agent that generated the most reliable results across all experiments was ApexBio MC. It was the only inducing agent that exhibited positive prophage induction in all trials, and was almost always among the highest levels of VDC, bacteria lysed, and inducible fraction. In future induction experiments in both soil and aquatic samples, it seems that using ApexBio MC as the inducing agent would produce the most consistent results.

In light of this conclusion, my original concerns about the ability of ApexBio MC to induce the monthly samples may have been unfounded. It also seems that the consistency of ApexBio MC in the inducing agent tests can give us more confidence that the results of the monthly inductions may be legitimate, and not erroneous due to an ineffective inducing agent. Nonetheless, this selection of ApexBio MC as the most consistent inducing agent is still somewhat of a subjective choice given the inconsistency of the data, and more research will be needed to definitively choose the best inducing agent for these studies.

Conclusions and Future Directions

This project illuminated many of the complications often encountered with induction experiments: it can be difficult to get consistent results, and limitations on time and resources frequently result in small numbers of replicates, which can compound problems with picking out statistically significant signals against considerable background noise in the data. These problems call into question the validity of chemical induction assays for assessing lysogeny in environmental samples. In a 2017 paper by Knowles et al., the accuracy of inducible fraction calculations across published environmental induction studies was analyzed. The authors found that across the literature, negative IF values made up one-third of total reported IF values. Furthermore, reported IF values were inconsistent even within technical replicates in the same study. Because biologically nonsensical values are found so often, we may need to reevaluate the ways in which we calculate the inducible fraction of bacteria within these environmental samples. A potential way forward would be to focus on changes in viral counts as opposed to incorporating both viral and bacterial counts, as changes in VDC and measurements of induced prophage seem to be the most reliable way to determine whether or not a sample has induced. This approach has its obvious shortcomings, as the number of cells that produced phage particles is still a valuable part of any calculation determining induction. Thus, future research will be necessary to evaluate these approaches and determine the best way forward.

There are many ways in which future studies could build off of this work. First, it may be helpful to incorporate viral reduction approaches into sample processing prior to induction (e.g., Williamson et al 2002). This method uses centrifugation or tangential flow filtration to concentrate cells and resuspend them in virus-free medium, reducing the number of free virus particles in the sample before induction occurs. This may generate a higher signal-to-noise ratio,

in that the difference in VDC between controls and treatments is more obvious when performing direct counts on the samples. Virus reduction approaches may also limit the amount of background noise that clouds the samples, and may lead to cleaner, more consistent results. However, one potential drawback to this approach is that some bacterial cells will be lost in the process. Nevertheless, this approach has not yet been evaluated for induction of soil bacterial communities and is therefore worth investigating.

It may also be helpful to determine the concentrations at which these inducing agents work best, particularly for soil bacteria, as there is a chance that administration at suboptimal concentrations is causing induction to be unsuccessful or even inhibited. While some experiments have been performed testing different mitomycin C concentrations for aquatic bacteria (Cochran et al., 1998), no one has yet attempted to titrate inducing agent dosages for soil bacteria. Finally, if sufficient time and resources are available in the future, it would be helpful to repeat the soil bacterial induction experiments with larger sample sizes, more frequent sampling, and more diverse sample sites to truly elucidate any trends in lysogeny that may exist among soil bacteria.

Temperate phages can protect bacterial hosts from further phage infections, change the phenotype of their host in a process called lysogenic conversion, and even affect nutrient cycling at the global level (Williamson et al., 2017). Continuing this work would be another important step towards determining the importance of lysogeny in these natural environments, allowing us to further understand these phenomena and perhaps even uncover some mysteries of the "dark realms of the soil".

REFERENCES

Ashelford, K. E., Day, M. J., & Fry, J. C. (2003). Elevated Abundance of Bacteriophage Infecting Bacteria in Soil. *Applied and Environmental Microbiology*, *69*(1), 285–289. https://doi.org/10.1128/AEM.69.1.285-289.2003

Carson, Rachel. *Silent Spring*. Penguin Books, in Association with Hamish Hamilton, 2015.

Cobián Güemes, A. G., Youle, M., Cantú, V. A., Felts, B., Nulton, J., & Rohwer, F. (2016). Viruses as Winners in the Game of Life. *Annual Review of Virology*, *3*(1), 197–214. https://doi.org/10.1146/annurev-virology-100114-054952

Cochran, Pk, Kellogg, C., & Paul, J. (1998). Prophage induction of indigenous marine lysogenic bacteria by environmental pollutants. *Marine Ecology Progress Series*, *164*, 125–133. https://doi.org/10.3354/meps164125

Cochran, PK, & Paul, J. (1998). Seasonal abundance of lysogenic bacteria in a subtropical estuary. *Applied and Environmental Microbiology*, *64*(6), 2308–2312. PubMed. https://doi.org/10.1128/AEM.64.6.2308-2312.1998

Court, D. L., Oppenheim, A. B., & Adhya, S. L. (2007). A new look at bacteriophage lambda genetic networks. *Journal of Bacteriology*, *189*(2), 298–304. https://doi.org/10.1128/JB.01215-06

Echols, H. (1972). Developmental pathways for the temperate phage: Lysis vs lysogeny, *Annual Review of Genetics*, *6*(0), 157–190. https://doi.org/10.1146/annurev.ge.06.120172.001105

Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature*, *399*(6736), 541–548. https://doi.org/10.1038/21119

Ghosh, D., Roy, K., Williamson, K. E., Srinivasiah, S., Wommack, K. E., & Radosevich, M. (2009). Acyl-homoserine lactones can induce virus production in lysogenic bacteria: An alternative paradigm for prophage induction. *Applied and Environmental Microbiology*, *75*(22), 7142–7152. https://doi.org/10.1128/AEM.00950-09

Ghosh, D., Roy, K., Williamson, K. E., White, D. C., Wommack, K. E., Sublette, K. L., & Radosevich, M. (2008). Prevalence of Lysogeny among Soil Bacteria and Presence of 16S rRNA and trzN Genes in Viral-Community DNA. *Applied and Environmental Microbiology*, *74*(2), 495–502. https://doi.org/10.1128/AEM.01435-07

Helsley, K. R., Brown, T. M., Furlong, K., & Williamson, K. E. (2014). Applications and limitations of tea extract as a virucidal agent to assess the role of phage predation in soils. *Biology and Fertility of Soils*, *50*(2), 263–274. https://doi.org/10.1007/s00374-013-0855-x

Jobbágy, E. G., & Jackson, R. B. (2001). The distribution of soil nutrients with depth: Global patterns and the imprint of plants. *Biogeochemistry*, *53*(1), 51–77. https://doi.org/10.1023/A:1010760720215

Knowles, B., Bailey, B., Boling, L., Breitbart, M., Cobián-Güemes, A., Del Campo, J., Edwards, R., Felts, B., Grasis, J., Haas, A. F., Katira, P., Kelly, L. W., Luque, A., Nulton, J., Paul, L., Peters, G., Robinett, N., Sandin, S., Segall, A., … Rohwer, F. (2017). Variability and host density independence in inductions-based estimates of environmental lysogeny. *Nature Microbiology*, *2*, 17064. https://doi.org/10.1038/nmicrobiol.2017.64

Koorem, K., Gazol, A., Öpik, M., Moora, M., Saks, Ü., Uibopuu, A., Sõber, V., & Zobel, M. (2014). Soil Nutrient Content Influences the Abundance of Soil Microbes but Not Plant Biomass at the Small-Scale. *PLOS ONE*, *9*(3), e91998. https://doi.org/10.1371/journal.pone.0091998

Laybourn-Parry, J., Marshall, W. A., & Madan, N. J. (2007). Viral dynamics and patterns of lysogeny in saline Antarctic lakes. *Polar Biology*, *30*(3), 351–358. https://doi.org/10.1007/s00300-006-0191-9

Leff, J. W., Jones, S. E., Prober, S. M., Barberán, A., Borer, E. T., Firn, J. L., Harpole, W. S., Hobbie, S. E., Hofmockel, K. S., Knops, J. M. H., McCulley, R. L., Pierre, K. L., Risch, A. C., Seabloom, E. W., Schütz, M., Steenbock, C., Stevens, C. J., & Fierer, N. (2015). Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academy of Sciences*, *112*(35), 10967–10972. https://doi.org/10.1073/pnas.1508382112

Liang, X., Wagner, R. E., Zhuang, J., DeBruyn, J. M., Wilhelm, S. W., Liu, F., Yang, L., Staton, M. E., Sherfy, A. C., & Radosevich, M. (2019). Viral abundance and diversity vary with depth in a southeastern United States agricultural ultisol. *Soil Biology and Biochemistry*, *137*, 107546. https://doi.org/10.1016/j.soilbio.2019.107546

Lisle, J. T., & Priscu, J. C. (2004). The occurrence of lysogenic bacteria and microbial aggregates in the lakes of the McMurdo Dry Valleys, Antarctica. *Microbial Ecology*, *47*(4), 427–439. https://doi.org/10.1007/s00248-003-1007-x

Loessner, M. J., Goeppl, S., & Busse, M. (1991). Comparative inducibility of bacteriophage in naturally lysogenic and lysogenized strains of Listeria spp. By u.v. Light and Mitomycin C. *Letters in Applied Microbiology*, *12*(5), 196–199. https://doi.org/10.1111/j.1472-765X.1991.tb00538.x

Lwoff, A. (1953). Lysogeny. *Bacteriological Reviews*, *17*(4), 269–337. PubMed.

National Center for Biotechnology Information (2021). PubChem Compound Summary for CID 5746, Mitomycin. Retrieved May 14, 2021 from https://pubchem.ncbi.nlm.nih.gov/compound/Mitomycin.

National Center for Biotechnology Information (2021). PubChem Compound Summary for CID 91763, Halosulfuron-methyl. Retrieved May 14, 2021 from https://pubchem.ncbi.nlm.nih.gov/compound/Halosulfuron-methyl.

Payet, J. P., & Suttle, C. A. (2013). To kill or not to kill: The balance between lytic and lysogenic viral infection is driven by trophic status. *Limnology and Oceanography*, *58*(2), 465–474. https://doi.org/10.4319/lo.2013.58.2.0465

Ptashne, Mark. *A Genetic Switch: Phage Lambda Revisited*. CSHL Press, 2004.

Stewart, F. M., & Levin, B. R. (1984). The population biology of bacterial viruses: Why be temperate. *Theoretical Population Biology*, *26*(1), 93–117. https://doi.org/10.1016/0040-5809(84)90026-1

Suttle, C. A. (2013). Viruses: Unlocking the greatest biodiversity on Earth. *Genome*, *56*(10), 542–544. https://doi.org/10.1139/gen-2013-0152

Verweij, J., & Pinedo, H. (1990). Mitomycin C: Mechanism of action, usefulness and limitations. *Anti-Cancer Drugs*, *1*, 5–13.

Wasik, B. R., & Turner, P. E. (2013). On the biological success of viruses. *Annual Review of Microbiology*, *67*, 519–541. https://doi.org/10.1146/annurev-micro-090110-102833

Weitz, J. S., & Wilhelm, S. W. (2012). Ocean viruses and their effects on microbial communities and biogeochemical cycles. *F1000 Biology Reports*, *4*. https://doi.org/10.3410/B4-17

Wilcox, R. M., & Fuhrman, J. A. (1994). Bacterial viruses in coastal seawater: Lytic rather than lysogenic production. *Marine Ecology Progress Series*, *114*(1/2), 35–45. JSTOR.

Williamson, K. E., Corzo, K. A., Drissi, C. L., Buckingham, J. M., Thompson, C. P., & Helton, R. R. (2013). Estimates of viral abundance in soils are strongly influenced by extraction and enumeration methods. *Biology and Fertility of Soils*, *49*(7), 857–869. https://doi.org/10.1007/s00374-013-0780-z

Williamson, K. E., Radosevich, M., Smith, D. W., & Wommack, K. E. (2007). Incidence of lysogeny within temperate and extreme soil environments. *Environmental Microbiology*, *9*(10), 2563–2574. https://doi.org/10.1111/j.1462-2920.2007.01374.x

Williamson, K. E., Radosevich, M., & Wommack, K. E. (2005). Abundance and Diversity of Viruses in Six Delaware Soils. *Applied and Environmental Microbiology*, *71*(6), 3119–3125. https://doi.org/10.1128/AEM.71.6.3119-3125.2005

Williamson, K. E., Schnitker, J. B., Radosevich, M., Smith, D. W., & Wommack, K. E. (2008). Cultivation-based assessment of lysogeny among soil bacteria. *Microbial Ecology*, *56*(3), 437–447. https://doi.org/10.1007/s00248-008-9362-2

Williamson, K. E., Wommack, K. E., & Radosevich, M. (2003). Sampling Natural Viral Communities from Soil for Culture-Independent Analyses. *Applied and Environmental Microbiology*, *69*(11), 6628. https://doi.org/10.1128/AEM.69.11.6628-6633.2003

Williamson, K., Fuhrmann, J., Wommack, K. E., & Radosevich, M. (2017). Viruses in Soil Ecosystems: An Unknown Quantity Within an Unexplored Territory. *Annual Review of Virology*, *4*, 201–219. https://doi.org/10.1146/annurev-virology-101416-041639

Williamson, S. J., Houchin, L. A., McDaniel, L., & Paul, J. H. (2002). Seasonal Variation in Lysogeny as Depicted by Prophage Induction in Tampa Bay, Florida. *Applied and Environmental Microbiology*, *68*(9), 4307. https://doi.org/10.1128/AEM.68.9.4307-4314.2002