Characterizing Mab Cluster R Prophage of Pathogen
Mycobacterium abscessus (Mab)

Madeline Kimble
CHARACTERIZING MAB CLUSTER R PROPHAGE OF PATHOGEN

MYCOBACTERIUM ABSCESSUS (MAB)

by

Madeline Kimble

A Thesis Submitted in Partial Fulfillment
of the Requirements for a Degree with Honors
(Biochemistry)

The Honors College
University of Maine
May 2021

Advisory Committee:
Sally D. Molloy, Assistant Professor of Genomics, Advisor
Benjamin King, Assistant Professor of Bioinformatics
Keith Hutchison, Professor Emeritus of Molecular and Biomedical Sciences
Melissa Ladenheim, Associate Dean of Honors College
Melody Neely, Associate Professor of Molecular and Biomedical Sciences
ABSTRACT

*Mycobacterium abscessus* (Mab) is an emerging pathogen that can cause pulmonary, skin and disseminating infections. It is one of the most drug-resistant pathogens and infections typically result in high morbidity and mortality. Understanding mechanisms of antibiotic resistance is critical for developing more effective treatments. Prophage, integrated viral genomes, are known to contribute to bacterial virulence and antibiotic resistance, yet Mab prophages remain largely uncharacterized. My research aims to characterize the diversity of the novel cluster MabR prophage genomes. The Molloy lab has demonstrated that the prophage McProf increases mycobacterial resistance to antibiotics. Using the McProf prophage genome sequence, we probed the PATRIC *M. abscessus* database to identify bacterial strains that carry prophage genomes related to McProf. We identified 25 related genomes, 8 of which were unique. This group of prophages are genetically distinct from prophages already described and we assigned them to a new cluster, MabR. Prophage genome ends were defined, and prophage sequences were extracted from bacterial genomes. MabR genomes are highly conserved, particularly across the structural genes in the right arm and the immunity cassette in the left arm. All nine genomes share a tyrosine-integrase and nearly identical attachment sites. All nine members share one of two types of Type VII secretion system polymorphic toxin systems, adjacent to the right attachment site that potentially improve fitness of the bacterial host. In future research we will investigate the role of MabR polymorphic toxins in drug resistance and bacterial fitness.
ACKNOWLEDGMENTS

I would like to thank Dr. Sally Molloy for her support and guidance throughout my past four years as an undergraduate. I would like to thank the members of my advisory committee: Dr. Benjamin King, Dr. Keith Hutchison, Dr. Melissa Ladenheim, and Dr. Melody Neely for their support, encouragement, and flexibility throughout my thesis project. The following members of the Molloy Lab have all contributed to this project: Colin Welch, Dakota Archambault, JC Ross, Anna Schumann, Jacob Cote, and the Genomic Research Collaborative group. The following departments have supported our research: The University of Maine Honors College, the Department of Molecular and Biomedical Sciences, and Dr. Graham Hatfull and the SEA-PHAGES program.

Lastly, Research reported in this project was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the national Institutes of Health under grant number P20GM103423.
TABLE OF CONTENTS

INTRODUCTION 1

LITERATURE REVIEW 4

*Mycobacterium abscessus* 4

Antibiotic resistance 4

Mycobacteriophage 6

Impact of prophage on bacterial fitness 8

McProf and its known role in *whiB7* expression and drug resistance 10

Type VII secretion polymorphic toxin systems 10

MATERIALS & METHODS 12

DNA collection and prophage identification/extraction 12

Mab cluster R genome analysis 12

Promoter site analysis 13

RESULTS 14

Identifying McProf-like sequences in *M. abscessus* genome sequences 14

Extraction and identification of MabR cluster genomes. 15

Cluster MabR prophage genome comparative analysis 16

Attachment sites and location of integration. 20

MabR prophage integrase analysis 20

MabR integration locations in *M. abscessus* genome 22

Repressor and Cro promoter analysis 23

DISCUSSION 25

RECOMMENDATIONS FOR FUTURE WORK 32
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>McProf genome Map</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>Pairwise comparison of the McProf genome sequence with the GD collection of <em>M. abscessus</em> prophage genomes described by Hatfull et al. in the phagesdb.org database (Dedrick et al., 2021; Russell et al., 2017)</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>Method of Identification and extraction of cluster MabR prophage sequences in <em>M. abscessus</em> genome sequences</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>Comparison of cluster MabR prophage genome organization and sequence similarity</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>MabR prophage prophiFSAT01-1 genome map</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>Comparative analysis of MabR attachment sites and integration location sites.</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>McProf and Lambda integrase core-binding (CB) domain alignment</td>
<td>21</td>
</tr>
<tr>
<td>8</td>
<td>Cro and repressor promoter sequences of MabR prophage</td>
<td>23</td>
</tr>
</tbody>
</table>
INTRODUCTION

*Mycobacterium abscessus* is an emerging opportunistic lung pathogen that is among the most resistant organisms to antibiotics (Johansen, 2020; Shaw, 2020). Clinical treatments of *M. abscessus* infections are less than 50% effective, resulting in high morbidity and mortality (Hurst-Hess, 2017). *M. abscessus* is found in contaminated soil and drinking water, infecting susceptible individuals with pulmonary health issues (Johansen, 2020; Shaw, 2020). Little is known about the mechanisms of increased antibiotic resistance of mycobacteria, specifically those of *M. abscessus*. Therefore, further investigation into the characterization of *M. abscessus* genome would allow for better drug target opportunities in the future.

The genomes of *M. abscessus* strains typically carry one or more prophage that are suspected to play a role in the alteration of host physiology, but there is little to no characterization of the viral portion of the genome (Dedrick, 2021). Nearly all pathogenic bacteria carry prophage, or integrated viral genomes, within their genomes (Brüssow et al., 2004). Viruses infect bacterial hosts and form lysogens by integration of their viral DNA into the host genome, through the use of a phage encoded enzymes, integrase, that targets phage and bacterial sequences called attachment sites (Hoess et al., 1978). Many well-known pathogenic bacteria carry prophage, including *E. coli* and *Vibrio cholerae*, that encode toxins that result in harmful effects on the host (Nanda, 2014). Without the presence of the prophage, however, these bacteria are harmless to the host (Nanda, 2014). The Molloy lab has identified a mycobacteria prophage, McProf, within the genome of pathogenic mycobacteria, *M. chelonae* (Molloy, unpublished). Dr. Molloy has recently demonstrated an increased antibiotic resistance response in *M.
chelonae when McProf is present (Molloy, unpublished). The presence of the McProf genome also increases expression of intrinsic mycobacterial antibiotic resistance genes when exposed to stresses such as sub-lethal concentrations of antibiotics or superinfection by a second phage (Molloy, unpublished). Given the role that McProf plays in altered drug resistance in M. chelonae, we aimed to identify prophage like McProf in M. abscessus, which could play a similar role in increased virulence and drug resistance.

To date there has only been one major report describing M. abscessus prophages. The Hatfull laboratory sequenced the genomes of 82 clinical isolates of M. abscessus and identified 67 novel M. abscessus prophage genomes (Dedrick, 2021). These prophages were sorted into 17 distinct clusters, MabA-Q (Dedrick, 2021). The gene content and organization of prophage genomes for each of these Mab clusters were characterized, defining the similarities and differences between groups and subgroups (Dedrick, 2021). Cluster MabR prophages, including McProf, have not yet been described.

This study aimed to characterize the genome features of eight novel prophage genomes that belong to cluster MabR. Sequence databases were probed using the McProf genome sequence to identify M. abscessus strains that carry MabR genome sequences. The McProf-like prophage genomes were identified as novel MabR prophage, then extracted and annotated. We performed comparative analysis to determine similarities and distinctions between integration locations and overall genome structure. The attachment sites, integrase core-binding domains, and integration locations were compared throughout the MabR genomes. The integration locations of mycobacteria are sought to play a role in the alteration of host physiology and perform
efficient site-specific recombination for stable lysogenic infection. Our research provides insight into the 9 MabR prophage genome structures, along with comparative analysis into the attachment sites, integration locations, and integrase enzymes.
Mycobacterium abscessus

Mycobacteria are diverse organisms that can be categorized as tuberculosis causing mycobacteria and non-tuberculosis causing mycobacteria (NTM). NTM are abundantly found in environmental sources, such as soil and drinking water, which can lead to human-pathogen interaction (Johansen, 2020). An emerging opportunistic NTM pathogen is *Mycobacterium abscessus* (*M. abscessus*). *M. abscessus* is among the most drug-resistant organisms and treatments are less than 50% effective (Hurst-Hess, 2017). Susceptible individuals are those with pulmonary diseases, such as Cystic Fibrosis, causing soft-tissue, pulmonary, and disseminating infections (Johansen, 2020; Shaw, 2020). *M. abscessus* is able to escape the human innate immune response through the production of inflammatory cytokines and the formation of granulomas (Johansen, 2020). By transitioning from a smooth to rough colony morphology, *M. abscessus* escapes the adaptive immune response. This variant morphology allows degradation of the granuloma and the formation of extracellular bacterial cords that inhibit B and T cell recruitment (Johansen, 2020). The cell wall of *M. abscessus* also is rich in mycolic acids making it impermeable to many external factors, aiding in the intrinsic drug-resistance of the mycobacteria (Johansen, 2020). The triggers or mechanisms of the escape from host immune responses is unclear and the *M. abscessus* genome is not well characterized in the literature.

Antibiotic Resistance

Mycobacterial antibiotic resistance is due to acquired and intrinsic resistance mechanisms. Acquired resistance occurs with mycobacterial genotypic changes (Nessar,
Mutations may occur in genes that are targeted by antibiotics, making the gene product unrecognizable by the antibiotic and increasing resistance (Nessar, 2012). Aminoglycosides and macrolides are cornerstone drugs used to treat NTM infections, by targeting the rRNA operon (Nessar, 2012). Specifically, aminoglycosides target the 16s rRNA thereby inhibiting protein synthesis (Nessar, 2012). Spontaneous mutations within the mycobacterial 16S rRNA, that change or remove the aminoglycoside binding site make the antibiotics no longer effective (Nessar, 2012). Macrolides interfere with the peptidyl transferase, inhibiting ribosomal translocation (Nessar, 2012). Acquired resistance to macrolides occurs by mutations in the 23s rRNA, or in a gene that results in increased methylation of the 23s rRNA, both of which prevent macrolides from binding and inhibiting translocation (Nessar, 2012).

Intrinsic resistance results from expression of genes already encoded in the bacterial genome. These genes are often upregulated upon exposure to environmental stresses such as sub-lethal concentrations of antibiotics (Nasiri et al., 2017). The mycobacteria have waxy cell walls rich in mycolic acids that make the outer envelope more impermeable to diffusion of small hydrophilic antibiotics across the cell wall (Hurst-Hess, 2017). The diffusion of antibiotics into the cell causes transcriptional reprogramming of the cell and induction of genes within the antibiotic resistance response (Geiman, 2006). Other genes that contribute to intrinsic resistance encode efflux pumps and antibiotic modifying enzymes that either increase the flux of antibiotics out of the cell or modify the antibiotic or its target within the bacteria (Hurst-Hess, 2017). A key player in intrinsic resistance is \textit{whiB7}, a conserved mycobacterial transcriptional activator (Geiman, 2006). \textit{whiB7} can turn on hundreds of genes in \textit{M.}
*abscessus*, but some well-known genes within its regulon are *erm*, a methyl transferase, *tap*, a multi-drug efflux pump, and *eis2*, an enhanced intracellular survival protein, which are induced by the presence of macrolides (Nessar, 2012; Rominski et al., 2017). This regulon of genes is known to increase antibiotic resistance in *M. abscessus* (Hurst-Hess, 2017). The macrophage environment and sub-lethal concentrations of antibiotics induce *whiB7* and in turn the *whiB7* target genes. It is not clear how these conditions lead to *whiB7* induction. The Molloy lab has recently demonstrated bacteriophage infection also contributes to *whiB7* expression and intrinsic drug resistance.

**Mycobacteriophage**

Viruses that infect the *Mycobacterium* species are known as mycobacteriophage. Mycobacteriophage are a diverse group of biological entities and to understand their relationships to each other, they have been sorted into 29 clusters, A-Z and AA-AC (Jacobs-Sera, 2012). The majority of mycobacteriophage studied to date were isolated in a single host, *Mycobacterium smegmatis* (Jacobs-Sera, 2012; Dedrick et al., 2017). The diversity of mycobacteriophage allows for a variety of host preferences of differing *Mycobacterium* species. A few of the bacteriophage strains are able to infect *M. abscessus* (Dedrick, 2021). Some bacteriophages are found integrated into bacterial genomes, aiding in the microbial success of many pathogens (Dedrick et al., 2017)

Mycobacteriophage have two modes of infection: lytic and lysogenic. The viral DNA is injected into the host and directed by other elements into either lytic or lysogenic lifestyle. Much of what we understand about lysogeny comes from studies on the *E. coli* phage, Lambda. The elements of the Lambda genetic switch are Cro, CI, and CII (Ptashne, 2004). The expression of CI promotes the entrance into the lysogenic infection
cycle by repressing early lytic genes and promoting CI and integrase expression, respectively (Ptashne, 2004). This results in the integration of the viral DNA into the host genome, forming a prophage and the silencing of the lytic genes (Broussard, 2014). The expression of Cro inhibits CI, in turn inhibiting lysogenic gene expression, and allowing the phage to enter the lytic life cycle (Broussard, 2014). Lytic infection entails the production of new viral phage and the death of the host cell. Lysogenic infection, unlike lytic, allows the bacterium to survive and replicate, and with it, the prophage genome replicates. The success of the prophage is now tied to that of the bacterium and therefore prophage have evolved to encode genes that improve the fitness and survival of the bacterial genome.

Viral integration occurs between the bacteriophage genome and the host bacterial genome. The viral genome contains an integrase enzyme that is crucial for the recombination reaction (Broussard, 2014). The integrase uses attachment sites from the bacterial genome and viral genome as substrates, known as attB and attP, respectively. Integrase enzymes contain a core-binding (CB) domain that defines the specificity for the attachment sites (Biswa et al., 2005). The attachment sites contain a common core that also play a role in specificity, allowing for the integrase to target the correct attP and attB for site-specific recombination, creating two new sites, attL and attR, that flank the prophage genome (Hoess et al., 1978). Two types of integrases can be utilized in the lysogenic mechanism, either a serine integrase or a tyrosine integrase. A serine integrase utilizes a serine residue in site-specific recombination for DNA substrate cleavage displacing the DNA’s 3’ bridging O, forming a 5’ phosphoserine linkage (Grindley et al., 2006). It also contains a carboxyterminal domain that allows for regulatory reactions to
Tyrosine integrase, however, use a tyrosine residue to displace the 5’ DNA bridging O forming a bond to the broken DNA 3’ strand end (Grindley et al., 2006). Simply, the serine integrases create a double-strand break before recombination, whereas tyrosine integrases break one strand at a time for each duplex, and forming a Holliday junction (Grindley et al., 2006). The majority of prophage encode a tyrosine integrase, with lambda integrase being one of the most well studied integrases (Grindley et al. 2006). The N-terminus of the Lambda integrase includes the CB domain and C-terminus contains the recombinase domain. The tyrosine recombination begins with the cleavage by a tyrosine nucleophile on one strand of each duplex of DNA, forming the previously described phosphotyrosine linkages, along with 5’ free hydroxyl groups (Grindley et al. 2006). A Holliday junction is then formed by the 5’ end and the opposing DNA 3’ phosphotyrosine, further isomerizing and repeating the process a second time and resolving the Holliday junction (Grindley et al. 2006). The strand exchange is then completed and two new attachment sites for the integrated prophage are formed (Hoess et al., 1978).

**Impact of Prophage on Bacterial Fitness**

Prophages are prevalent in pathogenic bacteria because they alter the fitness of their bacterial hosts in multiple ways. One such way is encoding genes such as toxin antitoxins that increase bacteria's ability to tolerate stress such as exposure to antibiotics or immune response during infection of eukaryotic hosts (Nanda, 2015). Prophage also encode toxins that increase the virulence of bacterial pathogens. Prophage present in *E. coli* encode the Shiga toxin, which leads to life-threatening diarrheal diseases (Nanda, 2014). The CTX prophage in *Vibrio cholerae* encodes for the cholera toxin leading to
food poisoning and other gastrointestinal issues (Nanda, 2014). Both *E. coli* and *V. cholerae*, in the absence of prophage, are harmless to the host. The role of mycobacterial prophage in host fitness and virulence is not well characterized in the literature.

The Molloy lab has previously demonstrated the impact of prophage on *M. chelonae*, a closely related mycobacteria to *M. abscessus* and *Mycobacterium tuberculosis*. *M. chelonae* contains a natural prophage called McProf (Figure 1). The Molloy lab demonstrated that the McProf prophage enhances the bacterium’s drug resistance in response to stress in the form of sub-lethal concentrations of antibiotics or superinfection by a second phage (Molloy, unpublished). The superinfecting phage interacts with McProf in an unknown way that leads to increased expression of antibiotic resistance gene *whiB7* and increased resistance to aminoglycosides (Molloy, unpublished).

**Figure 1. McProf genome map.** Genome organization of prophage McProf. The coordinates of the McProf genome are represented by the ruler. Genes are shown as colored boxes above (transcribed rightwards) or below (transcribed leftward) the ruler. The map was generated using Phamerator (Molloy, unpublished).

McProf and its known role in *whiB7* expression and drug resistance.
The McProf genome has a genome length of 67,657 bp and encodes 98 genes (Molloy, unpublished). McProf contains the two attachment sites, attL and attR, made up of 45 bp, with the attR overlapping a leftward tRNA-Lys (Molloy, unpublished). Adjacent to the attL site is a rightward tyrosine integrase, gp1, and a leftward putative immunity repressor, gp3 (Figure 1) (Molloy, unpublished). Following the integrase and repressor, are the early lytic genes, late lytic genes/structural genes, the lysin cassette, and the accessory genes, which contain the polymorphic toxin cassette (Figure 1). Although McProf has a functional integration system, there is no evidence of lytic gene expression in *M. chelonae*. The expression profile of McProf has been determined for lysogenic infection expression (Molloy, unpublished). McProf genes that may drive changes in whiB7 expression are currently being investigated. The most promising McProf gene candidates are three reverse genes located adjacent to the left attachment site, known as the Type VII secretion system polymorphic toxin cassette (Molloy, unpublished). The role of the polymorphic toxin is still unknown but is thought to play a role in the increased antibiotic resistance after exposure to stress (Molloy, unpublished). Only one other paper has described the role of polymorphic toxin systems in *M. abscessus* strains, with still very little known about the impact of the prophage genomes on the antibiotic resistance of *M. abscessus* (Dedrick, 2020).

**Type VII secretion polymorphic toxin systems**

The McProf genome encodes a type VII secretion system (T7SS) polymorphic toxin immediately adjacent to the right attachment site (Figure 1). The cassette includes a ~100 amino acid EsxA-like protein with a WXG-100 motif, a 732 amino acid polymorphic toxin with an N-terminus WXG-100 motif and C-terminus Tde-like DNase
toxin motif. The third gene in the operon is a cognate immunity protein that likely binds to and neutralizes the toxin domain of the polymorphic toxin (Figure 1). In other bacterial systems, polymorphic toxins allow kin bacteria within a population to communicate with each other, discriminate non-kin population members, and respond to stress or antagonism in the environment including escape from macrophage phagosomes, tolerating antibiotic exposure, or defense against phage superinfection (Kumar, 2019). Polymorphic toxins are different than toxin/antitoxin systems in that they are secreted into the environment or directly into bacterial or eukaryotic cells. They therefore depend on bacterial secretion that allow transport of the toxin across the cell envelope (Kumar, 2019). The transport allows for stress signals to be given to neighboring kin bacteria or to defend against antagonistic non-kin community members. Studies on other bacterial species have described the role of similar PT systems in pathogenic Agrobacterium and Pseudomonas in communicating danger and stress conditions to kin cells (Ma et al., 2014). The McProf polymorphic toxin is likely secreted by one of the host’s T7 secretion systems, either the Esx-3 or Esx-4 secretion systems. Although T7SS substrates have been studied in M. tuberculosis, their precise function in virulence is still unknown, but they are often involved in virulence. For example, the Esx-1 system in M. tuberculosis involves secretion of dimerized substrates EsxA and EsxB and these interact with the macrophage phagosome membrane, facilitating mycobacterial escape from the phagosome (Houben et al., 2014). The role of T7SS substrates in M. abscessus is also largely unknown (Gröschel, 2016). The specific type VII ESX system defined in M. abscessus are ESX-4 and ESX-3, but very little is known about the role these systems play in other mycobacterial strains (Johansen, 2020).
MATERIALS AND METHODS

DNA collection and prophage identification/extraction

Genome sequences of clinical isolates of *M. abscessus* that carry MabR prophage genomes sequences were identified through BLAST analysis of the McProf prophage genomes sequence at phagesdb.org (Russell et al., 2017) using the PATRIC *M. abscessus* database. Bacterial genomes that carried regions with high sequence similarity to the McProf genome were downloaded from NCBI (the National Center for Biotechnology Information). Phaster (Arndt et al., 2016) was used to identify candidate mycobacteriophage genome coordinates. Regions of the bacterial genome sequence that contained the potential prophage sequence were examined in Benchling (Benchling [Biology Software], (2021). The left and right attachment sites were identified using McProf attachment sites as models to identify repeat sequences among the contig that included the prophage genome. The trimmed genomes were extracted sequences were trimmed to these sequences which determine the prophage genome ends. Trimmed genomes were auto annotated in PECAAN and DNA Master.

Mab cluster R genome analysis

Prophage sequences were auto annotated using Glimmer (Delcher,et al., 1998), Genemark (Besemer et al., 2005), in DNA Master (Pope et al., 2018) and PECAAN. Gene starts were analyzed and determined by manual inspection. HHPRed (Söding, 2005), NCBI BLASTP (Altschul et al., 1997), and the Conserved Domain Database (Lu et al., 2020) were used to predict gene functions. Membrane proteins were identified through programs TmHmm (Krogh et al., 2001) and SOSUI (Hirokawa et al., 1998). Comparative genomic analysis was performed using Clustl Omega (Sievers, et al., 2011)
and Phamerator (Cresawn et al., 2011). A database of MabR cluster prophage genomes was constructed using Phamerator, which includes 9 prophage genomes.

**Promoter site analysis**

Promoter sequences of *Mab* cluster R prophages were identified by analyzing intergenic regions between the Cro and repressor genes in PhiSite using the Promoter Hunter tool (Klucar et al., 2010; Stano et al., 2011). RNAseq data from McProf lysogens of *M. chelonae* was used in conjunction with PhiSite to predict promoters in the McProf genome. Comparative analysis of intergenic regions and promoter sequences was performed using Clustal Omega.

**RESULTS**
Identifying McProf-like sequences in *M. abscessus* genome sequences

The Hatfull laboratory has characterized 82 clinical *M. abscessus* strains, GD01-GD111 isolated from 78 different patients and sequenced their genomes (Dedrick et al., 2021). From this collection of GD strains, 67 unique prophage sequences were extracted and annotated and sorted into 17 clusters, MabA - Q (Dedrick et al., 2021). To determine if the McProf prophage genome is related to any of these prophages, the McProf genome sequence was analyzed by BLASTN against this collection of prophage genome sequences in the phagesdb.org database (Russell et al., 2017)(Figure 2a). The McProf genome shared little nucleotide identity with the GD prophages. Through the pairwise comparison, the McProf genome shared 0-10% genome content with the GD strain database (Figure 2A). Prophages with <35% shared gene content are sorted into distinct Mab clusters (Dedrick, 2021) and therefore McProf was assigned to a novel cluster, MabR. To determine if other MabR prophage genomes exist in published *M. abscessus*
genome sequences, the McProf genome was BLASTed against a wider range *M. abscessus* sequences in the PATRIC database (Wattam et al., 2014) (Figure 2B). BLAST analysis identified 25 *M. abscessus* strains with genome regions that showed high similarity to the McProf genome (Figure 2b). Of the 25 MabR genomes identified, only eight of these were unique.

**Extraction and identification of MabR cluster genomes**

The genome sequence of each bacterial strain carrying a MabR prophage was analyzed by the program PHASTER (Arndt et al., 2016) in order to identify all the prophage sequences in the bacterial genome and the coordinates of those prophage sequences (Figure 3). Extraction of the genomes began with inputting the contig of the Mab bacterial strain the similar sequence was identified within (Figure 3). The contig was input into PHASTER program, with output determining the estimated coordinates of the prophage sequence within the contig (Figure 3). Using these coordinates, we extracted a sequence from the bacterial genome including an extra 5 - 10 thousand base pairs of flanking sequence, to ensure that the entire prophage genome was included in the sequence. This sequence was uploaded into the web-based program PECAAN which provides an auto-annotation and BLASTP and HHpred data for each predicted gene. We used this data to identify key prophage genes such as the integrase and structural genes. The genomes' ends were determined as repeat sequences adjacent to the integrase gene and the accessory genes. The trimmed fastA file was then auto annotated in both PECAAN and DNA Master and genes were manually inspected for gene starts and functions (Figure 2). Extraction and identification demonstrated all MabR genomes to
have a left attachment site overlapping the 3' end of a tRNA-Lys, along with encoding a tyrosine integrase.

Cluster MabR prophage genome comparative analysis

*M. abscessus* prophage genomes are diverse, for there are 18 different clusters, including MabR. The MabR genomes are conserved within the MabR cluster, but some are more closely related to each other than others. A phylogenetic analysis of the Mab cluster R prophage genomes presented subgroupings of prophages with more similarities in
relation to others (Figure 4a). The MabR cluster prophage are not dissimilar enough to form subclusters however, they can be sorted into three groups: 1) prophiFVLQ01-1; 2) prophiFSMS01-1, prophiFSQJ01-1, and prophiFSOD01-1; and 3) McProf, prophiFSAT01-1, prophiFSIL01-1, prophiFVMH01-1, and prophiFSIG01-1 (Figure 4a). McProf is most similar to prophiFSAT01-1 (Figure 4a). Group 2 appears to be the most distinct. A Phamator map with all MabR prophages aligned by group was created to present the differences between groups 1 and 3 and the more distinct group 2 MabR
prophages (Figure 4b). All the MabR prophages have a similar genome organization. We will use prophiFSAT01-1 as our prototype MabR genomes (Figure 5). All MabR genomes begin with a tyrosine integrase gene, gp1, adjacent to the left attachment site (Figure 4b). This is followed by the immunity cassette, which includes the integrase, immunity repressor, and Cro (Dedrick, 2017). The Cro and repressor genes are highly conserved throughout the genomes (Figure 4b). The early lytic genes are more diverse across the MabR genomes however the structural genes are highly conserved. The lysis cassette of the MabR genomes all contain a lysin A, lysin B, and holin (Figure 4b, 5).

**Figure 5. MabR prophage prophiFSAT01-1 genome map.** Genome organization of prophage prophiFSAT01-1. The coordinates of the prophiFSAT01-1 genome are represented by the ruler. Genes are shown as colored boxes above (transcribed rightwards) or below (transcribed leftward) the ruler. The map was generated using Phamerator.
There is some variation in the accessory gene region, including two different types of polymorphic toxin systems located adjacent to the attR (Figure 4b, 5).

![Image of Clustal Omega alignment and NCBI BLAST analysis]

**Figure 6. Comparative analysis of MabR attachment sites and integration location sites.** (A) Clustal Omega alignment of left and right attachment sites for all 9 MabR cluster prophage genomes. Three groups of attachment sites are illustrated, Group 1 (yellow), Group 2 (green), and Group 3 (blue). Nucleotides conserved across all attachment site sequences are indicated by an asterisk (red). (B) Integration locations for each prophage were determined through NCBI BLAST analysis. There were four different examples of integration locations. All prophage integrated into the 3’ end of a tRNA-Lys however the flanking bacterial genes differed between some prophage. The bacterial genes adjacent to attL and attR are described in “Results” and their transcriptional orientation are indicated by red (reverse) and green (forward) arrows.
Attachment sites and location of integration

The attachment sites of a prophage (attL/attR) are made up of the attachment sites of the bacterial and bacteriophage genome, which are attB and attP, respectively (Hoess et al., 1978). Therefore, attL and attR are suspected to have a common core between the two that aligns with the same common core of attB and attP. Thus, the alignment of attL and attR of all MabR prophages was performed in order to find a similar sequence between the two attachment sites throughout all prophages. The alignment output demonstrated a conserved region of nucleotides among the att sites that reads “GGGxTCGAAxC” (Figure 6a). It’s possible these nucleotides are conserved because they are part of the common core that the integrase requires in order to carry out integration or excision. Based on sequence and length the attachment sequences from each of the MabR prophage can be sorted into three groups. Group 1 includes prophage prophiFSMS01-1, prophiFSOD01-1, and prophiFSQJ01-1. Group two includes prophage McProf, prophiFSAT01-1, prophiFSIL01-1, and prophiFVMH01-1. Group three includes the single MabR prophage, prophiFSIG01-1.

MabR prophage integrase analysis

All the MabR genomes share a tyrosine integrase, gp1, immediately adjacent to the left attachment site. Tyrosine integrases typically have a core binding site in the N-terminus and recombination domain in the C-terminus (Biswas et al., 2005). To identify which residues of the MabR integrases may be involved in binding the attachment sites, we aligned the MabR integrase protein sequence with that of Lambda. Further analysis with HHpred demonstrated a 100% probability alignment to the lambda tyrosine integrase. The lambda integrase has a defined core-binding domain (CB domain),
catalytic core, and active site conserved residues (Biswas et al., 2005). Alignments of the MabR prophage integrases to the CB domain of lambda proved notable similarity between residues (Figure 7). The CB domain is not exactly identical to the integrases of MabR but demonstrated the coordinates of residues that can be suspected as the CB domain of the MabR prophages. The coordinates within the integrase gene estimated for the CB domain of MabR cluster prophage are residues 119-199 (prophiFSQJ01-1, prophiFSOD01-1, prophiFSMS01-1), 114-194 (prophiFSIL01-1, prophiFVMH01-1, and prophiFSIG01-1), and 115-194 (FVLQ01-1) (Figure 7). The CB domains of prophages prophiFSQJ01-1, prophiFSOD01-1 and prophiFSMS01 are distinct from the CB domains of the other MabR integrases. Comparative analysis was performed on the catalytic domain of lambda integrase in relation to the MabR cluster prophages integrases. Similar

![Figure 7. McProf and Lambda integrase core-binding (CB) domain alignment.](image)

Clustl Omega comparative analysis was performed on lambda integrase core binding domain and MabR prophage integrases. A) The CB domain of lambda integrase and the entire integrase sequence of MabR prophage was investigated. Two groups of integrase CB domains were identified, Group 1 (yellow) and Group 2 (green). Conserved (red) and similar (purple) residues were identified between prophage genomes and the lambda CB domain by asterisks and dots, respectively. Numbers on the right side indicate residue number in the sequence.
to the CB domain, the alignment was not 100% identical, but showed very high similarity
including across the catalytic core of MabR cluster prophage integrases (Figure 7).

MabR integration locations in \textit{M. abscessus} genome

MabR prophage have similar attachment sites all of which overlap the 3' end of a tRNA-Lys. However, the landscape of the attachment sites differs between the prophages. There are four distinct location types. The integration of the prophages involves integration through a tRNA, which overlaps the attB site. All MabR cluster prophage integration involves a tyrosine integrase, whereas other clusters of Mab prophage have been found to use a serine integrase. All Mab cluster R prophage genomes are integrated using tRNA-Lys (Figure 6b). The location within the bacterial genome, however, differs when identifying the adjacent bacterial genes to attL and attR. Type 1 integration location is found within prophiFSAT01-1 and prophiFVLQ01-1, containing a reverse transcribed (3R)-hydroxyacyl-ACP dehydratase subunit HadA adjacent to the attL site, and a forward transcribed lipoprotein LppS adjacent to the attR site (Figure 6b). Prophages prophiFSIG01-1, prophiFSIL01-1, and prophiFVMH01-1 all share Type 2 with a forward transposase-like protein at the attL site and an uncharacterized protein near the attR site (Figure 6b). Type 3 was identified in prophiFSMS01-1 and prophiFSQJ01-1, containing a reverse niacin/nicotinamide transporter NaiP gene adjacent to the attL site and a major facilitator transporter at the attR site (Figure 6b). FSOD, Type 4, is almost identical, in integration location sites, to prophiFSMS01-1 and prophiFSQJ01-1, but the attR bacterial gene is uncharacterized protein, rather than the major facilitator transporter (Figure 6b). In relation to the phylogenetic representation, the types of integration locations of MabR genomes are found within already determined similar genomes.
Repressor and Cro promoter analysis

Proper encoding of the immunity repressor allows for stable lysogeny to take place. The immunity repressor works in conjunction with the Cro protein to determine the life cycle of the phage. If the repressor is expressed, then the lysogenic life cycle takes place as a result of the Cro protein being blocked from initiating late lytic genes and entering the lytic life cycle. The McProf promoter for the repressor was also determined using RNAseq data from a McProf lysogen. This allowed us to determine where reads begin to map the genome region upstream of the immunity repressor. We used this data in conjunction with the PhiSite promoter analysis to predict the most likely promoter for the McProf promoter. Based on this data, we examined these genome regions in the other MabR genomes to determine if they carried similar or different promoters in this

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>prophiFSMS01-1</td>
<td>AATCCAGTGGTTGACAGTCATAGTGTGCTAGTAAATTGAGTCACCGCATCC</td>
</tr>
<tr>
<td>prophiFSOS01-1</td>
<td>AATCCAGTGGTTGACAGTCATAGTGTGCTAGTAAATTGAGTCACCGCATCC</td>
</tr>
<tr>
<td>prophiFSQJ01-1</td>
<td>AATCCAGTGGTTGACAGTCATAGTGTGCTAGTAAATTGAGTCACCGCATCC</td>
</tr>
<tr>
<td>prophiFVL01-1</td>
<td>ATATCCGCTGTTGACAGACGACCTGCTAGTTGAGTCTATGGCATCC</td>
</tr>
<tr>
<td>prophiFSAT01-1</td>
<td>ATATCCGCTGTTGACAGACGACCTGCTAGTTGAGTCTATGGCATCC</td>
</tr>
<tr>
<td>prophiFSIL01-1</td>
<td>ATATCCGCTGTTGACAGACGACCTGCTAGTTGAGTCTATGGCATCC</td>
</tr>
<tr>
<td>Mcprof</td>
<td>ATATCCGCTGTTGACAGACGACCTGCTAGTTGAGTCTATGGCATCC</td>
</tr>
</tbody>
</table>

* * *********** ********* ********* * *

<table>
<thead>
<tr>
<th>B</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>prophiFSMS01-1</td>
<td>CAGC---------------AACCACGTAGCTAGCTACAGCTAGCTCC</td>
</tr>
<tr>
<td>prophiFSOS01-1</td>
<td>CAGC---------------AACCACGTAGCTAGCTACAGCTAGCTCC</td>
</tr>
<tr>
<td>prophiFSQJ01-1</td>
<td>CAGC---------------AACCACGTAGCTACAGCTAGCTACAGCTCC</td>
</tr>
<tr>
<td>prophiFVL01-1</td>
<td>---------------TTACTGGTACGACGACACCAGATAGCTGTAACAGCCGTAA</td>
</tr>
<tr>
<td>prophiFSIL01-1</td>
<td>GAGGTACGTGGGATTGGGTTACAGAGAACGAAATAGCTGTTAGG---</td>
</tr>
<tr>
<td>prophiFSAT01-1</td>
<td>GAGGTACGTGGGATTGGGTTACAGAGAACGAAATAGCTGTTAGG---</td>
</tr>
<tr>
<td>Mcprof</td>
<td>GAGGTACGTGGGATTGGGTTACAGAGAACGAAATAGCTGTTAGG---</td>
</tr>
</tbody>
</table>

* * * * * * * * * *

Figure 8. Cro and repressor promoter sequences of MabR prophage. A) A Clustl Omega alignment of Cro promoter sequences of MabR prophage cluster. The leftward yellow box is the -35 box, and the rightward yellow box is the -10 box of the promoters. B) Clustl Omega alignment of the repressor promoter of MabR prophage cluster. There are three types of promoters identified, type 1 (yellow), type 2 (red), and type 3 (green). The asterisks indicate conserved nucleotides.
region. The promoter sequences between the repressor and Cro genes were analyzed within Mab cluster R prophages. All of the Mab cluster R prophages have the same Cro promoter sequences (Figure 8a). The repressor promoter differed throughout the cluster. There were three different types of repressor promoters identified (Figure 8b). MabR prophage, prophiFSAT01-1 and prophiFVLQ01-1, each had distinct promoter sequences from the rest of the cluster R prophages. The prophages prophiFSMS01-1, prophiFSOD01-1, and prophiFSQJ01-1 all have the same repressor promoter sequence (Figure 8b), which is to be expected since they share the same relation in the phylogenetic analysis. Through further PhiSite analysis, and the start of the RNAseq repeats, a repressor promoter sequence was determined for McProf (Figure 8b).
DISCUSSION

*Mycobacterium abscessus* is an opportunistic lung pathogen posing a real threat in immunocompromised individuals and cystic fibrosis patients (Johansen, 2020; Shaw, 2020). The high incidence of multi-drug resistance in *M. abscessus* results in ineffective treatments and mortality, with less than 50% of treatments being effective (Hurst-Hess et al., 2017). There is little to no understanding of the role prophage play in mycobacterial virulence. Characterization of prophage and their role in virulence and pathogenicity of *M. abscessus* could provide opportunities to develop better therapies, including drug treatments and phage therapy (Dedrick, 2019). The Molloy laboratory has demonstrated that mycobacteria, in the presence of prophage, have increased expression of antibiotic resistance genes, which warranted the investigation of similar prophage present in the genomes of clinical *M. abscessus* strains. McProf, a prophage discovered in the Molloy laboratory that is present in the genome of the *M. chelonae* strain, is distinct from the 67 Mab prophages characterized by the Hatfull laboratory (Dedrick, 2021). Therefore, McProf-like genomes were identified through the PATRIC database and grouped into a novel cluster, MabR. The characterization of cluster MabR prophage is crucial for further understanding of differing *M. abscessus* prophage genomes and their roles in antibiotic resistance. Understanding how prophage genome content impacts host physiology is currently being studied by the Molloy lab. The location of prophage integration into the bacterial genome can also impact host physiology. This study describes the integration site of the novel MabR prophages, and regulatory sequences that impact the lysogenic decision.
McProf is a novel MabR prophage and is distinct from prophages characterized by the Hatfull group. Given the Molloy lab has demonstrated that McProf impacts drug resistance of the pathogen, *M. chelonae*, it is important to understand the prevalence of this type of prophage in other mycobacterial genomes. We used the McProf genome sequence to probe published *M. abscessus* genome sequences to identify related prophage genomes and identified 25 additional MabR genomes. Eight of these prophage sequences were unique. We have annotated and characterized these novel MabR prophage genomes and location in the *M. abscessus* genomes. MabR genomes are fairly conserved, however, there are some differences that allow them to be sorted into two groups (Figure 4a). There are mainly differences in the early lytic genes and accessory genes that distinguish the two groups, illustrated in the phamerator map between the MabR prophages prophiFSOD01-1 and FVLQ01-1 (Figure 4b). The phylogenetic analysis clusters the prophages into three distinct groups: 1) prophiFVLQ01-1; 2) prophiFSMS01-1, prophiFSQJ01-1, prophiFSOD01-1; and 3) McProf, prophiFSAT01-1, prophiFSIL01-1, prophiFVM01-1 and prophiFSIG01-1 (Figure 4a). These differences don’t warrant sub clustering, but the more distantly related prophage of group 2 does correlate with differences observed in attachment site sequence, integration location and integrase core binding (CB) domain.

The organization of MabR prophage genomes are similar to that of other *M. abscessus* genomes. Prophages from the Mab clusters A1, A3, and R encode a forward oriented tyrosine integrase (gp1) immediately adjacent to the left attachment site (Dedrick, 2021). MabA1 integrates into the 5' end of a tRNA-Lys, similar to that of MabA3 prophage (Dedrick, 2021), while MabR integrates into the 3' end of a tRNA-
Lys. Like other Mab prophage clusters, the integrases of MabR are followed by an immunity cassette, early lytic genes, structural genes, a lysis cassette, and finally a large set of accessory genes. Other *M. abscessus* prophage, including MabA1, have reported encoded genes that may contribute to antibiotic resistance, stress tolerance, and virulence, including toxin/antitoxin (TA) systems and polymorphic toxins (PT) (Dedrick, 2021). There are no apparent TA systems, but MabR genomes do all encode type 7 secretion system polymorphic toxins that are hypothesized to contribute to drug resistance (Molloy, unpublished). Along with PT systems, the location of prophage integration within the bacterial host has proven to influence bacterial physiology (Dedrick, 2021).

The MabR prophages all integrate into the 3’ end of tRNA-Lys genes, however regions that flank the tRNA gene can be sorted into four types (Figure 6b). The 4 types of integration locations differ in the bacterial genes adjacent to the attL and attR of the prophage genomes (Figure 6b). The location of integration is important to consider due to the effects the transcribed bacterial and viral genes can have on the host physiology (Dedrick, 2021). Integration types 1, 2, and 3 have a reverse bacterial gene adjacent to the attL, and a forward bacterial gene adjacent to the attR site (Figure 6b). The difference is found within the type of bacterial genes that are present, but the direction of transcription is of importance in relation to impacting host physiology. Type 2 integration location contains a forward bacterial gene adjacent to the attL and attR sites (Figure 6b). Prophages have demonstrated an impact on bacterial gene expression by interrupting genes or by read through from phage genes that flank bacteria genes (Dedrick, 2021). The Hatfull group demonstrated within MabI and MabJ disrupt the host
gene expression due to integration location overlapping an opening reading frame (Dedrick, 2021). MabR prophage, however, integrate into the 3’ end of tRNA-Lys within an intergenic region, suspecting no alteration in flanking gene expression of the host. It is possible there are more than one tRNA-Lys genes within the strain genomes that the prophage are using for integration. The *M. abscessus* ATCC(199977) strain encodes 3 tRNA-Lys genes. We only explored tRNA-Lys genes in the prophiFSIL01-1 strain, which carries 2 tRNA-Lys genes. The second unoccupied tRNA-Lys corresponds to the Type 1 integration location described for the MabR prophages, prophiFSAT01-1 and prophiFVLQ01-1 (Figure 6b). It is unclear whether prophage that integrate into different integration locations prefer one tRNA-Lys over another, or whether each prophage can integrate into both tRNA-Lys genes. The tRNA-Lys genes should be identified in all strains of MabR, but was not investigated in this study. Further investigation into the integration location attachment sites would provide better understanding of how the differing locations may impact host physiology and site-specific recombination.

The MabR genomes can also be sorted into two major groups based on differences in their attachment site sequences. The prophage genome coordinates are defined by the attL and attR sites on each end of the genome. The attachment sites, attP and attB, act as substrates for the integration reaction between the integrase, viral genome, and bacterial genome (Hoess et al., 1978). The integration mechanism, as previously described, combines the attB and attP sites to divide into attL and attR. The result of recombination is a common core between attL, attR, attB, and attP, that has demonstrated specificity properties for the integrases enzyme and site-specific recombination (Hoess et al., 1978). A Clustal Omega alignment of attL and attR for all 9
MabR prophage demonstrated a conserved region of nucleotides. The conserved region is suspected to contain the common core of the attachment sites (Figure 4a). Without experimental analysis we cannot know which nucleotides are part of the common core; however, we can predict that the nucleotides that are conserved across all the MabR attachment sites are likely part of the common core. The attachment sites are crucial for the integrase to bind specifically and perform genetic recombination (Hoess et al., 1978). Analysis on L5 integration in mycobacterial species has demonstrated only a need for integrase and the attP site to be present for efficient recombination (Lee et al., 1993). Investigation into the MabR prophage integrase is warranted to determine if there are similar patterns between the differing attachment sites and integrases that may impact binding specificity and efficient lysogenic recombination.

We suspected if there were three different groups of attachment sites within MabR prophage, then there would be two different examples of core-binding domains of the integrases to facilitate binding specificity. MabR cluster prophage contain a tyrosine integrase with estimated core-binding domains similar to lambda, but overall have two different examples of integrases. After characterization, all of the MabR prophage genomes contained a tyrosine integrase within the same “pham”. The tyrosine integrase CB domain, as discussed, is crucial for binding of the integrase for lysogeny. Although the protein sequences differ there is some conservation of sequence between the integrases. HHPred and Clustl Omega alignments identified a conserved region of residues that aligned the lambda integrase CB domain and a region within the MabR prophage integrases that is likely involved in DNA binding. Among the MabR integrase sequences, the C-terminus recombinase sequence was highly conserved but there was
there was more variability within the N-terminus CB domain. The CB domain was
distinct for the same group of prophages that had distinct attachment sites, including
prophiFSMS01-1, prophiFSOD01-1, and prophiFSQJ01-1 (Figure 6a, 7). The second
example of attachment sites and CB domains were shared by the other 6 MabR prophage
genomes. The integrase is crucial for efficient lysogeny, but the maintenance of lysogeny
can also impact host physiology, and therefore the promoter sequences for the immunity
repressor and Cro genes was investigated.

The genetic switch between lysogenic and lytic life cycle is determined by the
interactions of the immunity repressor and cro proteins with the regulatory region
between these two genes (Broussard et al., 2014). The Cro promoters were highly
conserved across all nine MabR genomes (Figure 8a). Although all nine MabR genomes
share an immunity repressor from the same pham, there were differences in promoter
sequence across the nine genomes (Figure 8b). All 9 MabR Cro promoter sequences were
identical (Figure 8a). The repressor promoters, however, demonstrated differences
among the MabR prophage. There were three different groups identified of repressor
promoters (Figure 8b). The MabR prophage prophiFSMS01-1, prophiFSOD01-1, and
prophiFSQJ01-1 all had the same repressor promoter, while prophiFSIL01-1,
prophiFSAT01-1, and McProf shared the same repressor promoter sequence (Figure 8b).
The MabR prophage prophiFVLQ01-1 was different from all other MabR repressor
promoters and contained its own unique repressor promoter sequence (Figure 8b). Its
possible that the repressor promoter differences could impact the efficiency of
transcription of the immunity repressor and therefore stability of the MabR prophage.
Promoters for immunity repressors are often weaker promoters and rely on positive
autoregulation by the immunity repressor to stabilize the RNA complex (Dodd et al., 2001), while Cro requires a strong promoter because it solely relies on being expressed for activation.

MabR cluster prophage are not distinct enough to be subclustered, but they do demonstrate differences in attachment sites, integration locations, and integrase sequences that suggest they could have preferred integration sites. The second phylogenetic grouping of prophages prophiFSMS01-1, prophiFSQJ01-1, and prophiFSOD01-1 share the same attachment site sequence and integrase protein sequence and integrate into the type 3 and 4 integration locations, along with sharing the same repressor promoter sequence. The remaining prophages from phylogenetic group 1 (prophiFVLQ01-1) and group 3 (McProf, prophiFSAT01-1, prophiFSIL01-1, prophiFVM01-1 and prophiFSIG01-1) share more similar attachment sites, integrase CB sequences and integrate into integration location types 1 and 2. Phylogenetic group 1 is clearly distinct in multiple ways. They are phylogenetically distant from the other prophage and share the same attachment site sequences, core-binding domains of integrases, and repressor promoter sequences. This distinction doesn’t warrant sorting into subclusters but may indicate differences in integration location/site preference and lysogeny maintenance. It’s possible that these prophages do not distinguish between tRNA-Lys genes and a closer analysis of the tRNA-Lys genes in all nine bacterial strains should be completed. To determine the true specificity of MabR prophages, we would need carry out studies on the MabR integrase genes and perform integration assays with varying attB DNA targets.
RECOMMENDATIONS FOR FUTURE WORK

Cluster MabR prophages are novel and distinct from any other *M. abscessus* prophage genome identified. Therefore characterization was important for further understanding of the impact that this type of prophage could have on the fitness of *M. abscessus* strains. However, there is still more characterization that needs to be performed, some of which is being carried out by other members of the Molloy laboratory. This includes analysis of the co-habitating prophage of the MabR prophage carrying bacterial strains and a comparison of the accessory genes and polymorphic toxin systems within the MabR genomes. The role of McProf gene expression is being carried out to determine how prophage gene expression impacts antibiotic resistance and gene expression. An overall characterization of Mab prophage genomes will help further the understanding of the role prophages play in virulence and fitness of *M. abscessus*, allowing for more beneficial and efficient treatment options in the future.
REFERENCES


doi:10.1016/j.ygeno.2011.07.004

AUTHOR’S BIOGRAPHY

Madeline K. Kimble was born in Cleveland, Ohio on July 14, 1999. She was raised in Avon, Ohio and graduated from Saint Joseph Academy High School in 2017. Maddie is majoring in Biochemistry. She is a member of the University of Maine softball team and is a two-time captain. Maddie is also involved in student-athlete mentoring and a Student-Athlete Advisory Committee Representative for softball. She has received an INBRE fellowship. Upon graduation, Maddie plans to study Arts and Humanities in Medicine at the University of Maine before pursuing a career in the medical field.