2016-06-01

Characterization of Five Brevibacillus Bacteriophages and Their Genomes

Michael Allen Sheflo
Brigham Young University

Characterization of Five *Brevibacillus* Bacteriophages and Their Genomes

Michael Allen Sheflo

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Donald P. Breakwell, Chair
Julianne H. Grose
Sandra Hope

Department of Microbiology and Molecular Biology

Brigham Young University

June 2016

ABSTRACT

Characterization of Five *Brevibacillus*
Bacteriophages and Their Genomes

Michael Allen Sheflo
Department of Microbiology and Molecular Biology, BYU
Master of Science

Brevibacillus laterosporus (*B. laterosporus*) is a pathogen difficult to distinguish from *Paenibacillus larvae* (*P. larvae*), and contributes to Colony Collapse Disorder (CCD) of honeybees. To develop a biocontrol agent to limit its presence, bacteriophages were isolated from Utah County soil samples and used to infect *B. laterosporus* isolated from Utah County honey and larvae samples. Since CCD is prevalent in Utah beehives, bacteriophage that infect and lyse *B. laterosporus* may be isolated and characterized.

Pathogens were isolated from soil samples, and 16S rRNA gene tests initially identified the strains as *P. larvae*. Bacteriophages were isolated, purified, and amplified sufficiently to obtain images by electron microscope and genome sequencing by 454 pyrosequencing. Genomes were annotated with DNA Master, a Multiple Document Interface (MDI) program. Open reading frames (ORF's) were compared to the National Center for Biotechnology Information's (NCBI) database of primary biological sequence information via the Basic Local Alignment Search Tool (BLAST) algorithm.

Later testing determined the pathogen to actually be *B. laterosporus.* Plaques demonstrated lytic activity, and electron microscopy revealed bacteriophages of the myoviridae family. The five sequenced genomes were composed of linear dsDNA ranging from 45,552 to 58,572 base pairs in length, 92 to 100 genes per genome, and a 38.10% to 41.44% range of G + C content.

Discovering and describing new bacteriophages is a reasonably reproducible process and contributes to appreciating the diverse relationships between bacteriophage, bacteria, and eukaryota. Scientific facilitation of the bacteriophages role in limiting detrimental bacteria may contribute as an adjunctive therapy for CCD.

Keywords: American Foulbrood, bacteriophage, *Brevibacillus laterosporus*, colony collapse disorder, European Foulbrood, genome, *Paenibacillus larvae*, Utah

ACKNOWLEDGEMENTS

I offer my sincere gratitude to Dr. Don Breakwell for his enthusiasm, optimism, and academic support throughout my time at BYU. I extend this gratitude to my graduate committee, Dr. Sandra Hope and Dr. Julianne Grose, for their ever-present wealth of knowledge and resources. I thank the research assistants from the Breakwell Laboratory from 2011 to 2013 for their assistance and creativity: Alicia Brighton, Josh Fisher, Adam Gardner, Jordan Jensen, Bryce Lunt, and Bryan Merrill. I thank my special research team for their support: Nic Ferguson, Ashley Kennedy, Carson Russell, and Kurt Williams.

I thank the entire Utah County Beekeepers Association (UCBA) for supplying this endeavor with all necessary samples from their bee farms. I thank the BYU Department of Microbiology and Molecular Biology and the BYU College of Life Sciences for funding this research. I thank the Research Instrumentation Center (RIC) facility for access to their useful equipment, the BYU Sequencing Center (DNASC) for providing genomic information, and the BYU Microscopy Lab for all the amazing images. I thank the American Society of Microbiology for publishing the annotated genomes described here.

I thank all the faithful tithe paying members of The Church of Jesus Christ of Latter-Day Saints for their sacrifice that makes possible the unique laboratory of learning that is Brigham Young University. Finally, I thank my parents, Larry and Mary, and my wife, Chelsie, for all their support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AFB – American Foulbrood

BLAST – Basic Local Alignment Tool

BYU – Brigham Young University

dsDNA – Double Stranded DNA

LB – Lysis Broth

ORF – Open Reading Frame

PCR – Polymerase Chain Reaction

Phage – Bacteriophage

PL – *Paenibacillus larvae*

TAE – tris acetate EDTA

TE – tris EDTA

TEM – transmission electron microscopy

CHAPTER 1 – Introduction and Background

The health and survival of honeybees is vital to sustaining our current world economy and ecosystem. Besides being the only insect that directly produces food consumed by people, honeybees also participate in nearly one third of all plant pollination worldwide (Johnson, 2010). Wild honeybees or commercially managed hives are used to pollinate fruit and nut trees not only throughout the United States, but globally (Morse, 2002). Truly, our ecosystem has developed to rely on the beneficial activities of honeybees, and our massive agricultural production has also benefited.

Recently, honeybees have suffered an unusually high incidence of colony collapse, with an unprecedented decline of nearly half of all North American colonies (Johnson, 2010). Experts are still trying to determine if the decline is a novel phenomenon, but many have concluded that it is the result of a variety of causes. Such causes under investigation include weather fluctuations, human activities, and many well-studied bacterial, fungal, and viral diseases (Runckel, 2011). One cause is the bacterial disease, American Foulbrood (AFB) (Genersch, 2008; Eischen, 2005). Although AFB has been studied and treated extensively since 1906, it continues to cause more worldwide beehive destruction than any other known factor (Antunez, 2012; Di Pinto, 2011). European Foulbrood (EFB) is a similar bacterial infection, although it is less severe. Together, these bacterial diseases are referred to as Foulbrood Disease.

American Foulbrood (AFB) is a devastating disease that kills honeybee larvae, contributes to colony collapse, and limits agricultural yields (de Graaf, 2006; Genersch, 2010; Johnson, 2010). Regardless of the best efforts, the endospore-forming bacterium that causes AFB, *Paenibacillus larvae* (*P. larvae*), continues to build resistance to the most effective antibiotic treatments that combat and contain its proliferation (Martinez, 2009). Also, the

presence of endospores in honey contribute to their propagation (Piccini, 2002; Lindstrom, 2008; Lauro, 2003). Honeybee larvae of less than a week old consume the endospores contained in the food fed to them by worker bees. Inside the young larval gut, the endospores germinate and the bacteria rapidly divide. Eventually the gut epithelial layer of larvae rupture and the larva die. This is a relatively quick process that can produce nearly 100 million new bacteria and endospores in a single larva (Genersch, 2006; Gillard, 2008). Although only young bee larvae are killed, highly active and social adult honeybees carry the latent endospores to other parts of the hive resulting in rapid spread of the disease. In the first decade of the 20th century, the disease was described, named, and attributed to a bacterium called *Bacillus larvae*, which was later reclassified as *Paenibacillus larvae* (Antunez, 2007; de Graaf, 2006; Genersch, 2005; Rauch 2009).

Paenibacillus larvae has been extensively studied for decades primarily because of its influence on agriculture (Eischen, 2005). Less is known about *M. plutonius* and *B. laterosporus.* Although *P. larvae* is specifically associated with the honeybee, it belongs to a diverse genus. The *Paenibacillus* genus contains numerous species that live in all types of environments including water, soil, larvae, and in the laboratory (Genersch, 2008). Some *Paenibacillus* species are considered to directly benefit the economies of mankind and have been recently used as a source of chemical agents for biotechnology (Morse, 2002). Even so, other species are admired for their complex colony formations as well as playing a role in the global nitrogen cycle through biological nitrogen fixation.

Although paenibacilli live in many environments, there are only a few microbes that reside within the honeybee gut (Runckel, 2011). It is difficult to successfully isolate *P. larvae* from an adult honeybee or a larva. The most successful attempts come from isolating endospores

from honey of infected hives (Gillard, 2008 and Pohorecka, 2008). Endospores are able to withstand harsh treatments used to isolate them. Most hives are considered to be latent carriers of the endospores and any effective eradication of the pathogen must address the presence of endospores (Genersch, 2010). Hives infected with Foulbrood Disease are commonly burned in order to prevent it from spreading. This means that the functional problem of Foulbrood Disease in beehives is not the spread of active bacteria, but rather the presence of endospores.

As new information was obtained about this genus, *Paenibacillus* went through some taxonomic changes including reclassification, and a significant expansion of the known species (Genersch, 2006). *P. larvae* was originally named *Bacillus larvae.* Initially, *P. larvae* was the combination of *Bacillus larvae* and *Bacillus pulvifaciens*. The distinctions between these two were eventually believed to fall short of a species distinction and were later reclassified to *Paenibacillus larvae* subspecies *larvae* and *Paenibacillus larvae* subspecies *pulvifaciens* (Genersch, 2006). The differences were observed to be virulence factors and colony morphology (De Graaf, 2006). *P. larvae* subspecies *larvae* was considered the more virulent of the two subspecies and *P. larvae* subspecies *pulvifaciens* produced an orange-pigmented colony. Eventually, analysis of the 16S rRNA gene sequence was performed with many isolated strains of both subspecies which revealed that there was a negligible difference between the subspecies at a genetic level (Antunez, 2007).

As a result, the subspecies classification was abolished and today the only causative agent of AFB is *P. larvae*. Insights derived from better understanding the bacteriophage-host relationship between *P. larvae* and their phages may provide additional understanding. The official classification of *P. larvae* is Kingdom Monera, Phylum Firmicutes, Class Bacilli, Order

Bacillales, Family Paenibacillaceae, and Genus *Paenibacillus*. Nearly 70 species have been identified within genus *Paenibacillus*, and P. *larvae* is among the best studied.

These bacteria have been isolated from many environments including soil, water, vegetation, and insects. Bacteria of this genus stain gram-positive, are facultative anaerobes, and can form endospores (Schuch, 2001). *P. larvae* was originally cultivated on sheeps blood agar containing naladixic acid, and now a semi-selective medium has been used to isolate theses bacteria, one of which is called *Paenibacillus Larvae* Agar (PLA) (Hornitzky, 1991). PLA is not without its shortcomings, as a variety of closely related bacteria grow on it. However, *P. larvae* colonies are generally unique enough to distinguish them from other bacteria such as its close relative, *Brevibacillus laterosporus*. Other tests are vital to maintain accuracy of the isolation such as the catalase test. Typically, if a colony contains gram-positive bacteria and is catalase positive, sequencing the 16S rRNA gene can confirm the isolate's identity. (Dingman, 2012; Ryba, 2009; and Schuch, 2001).

To control and prevent AFB, oxytetracycline hydrochloride and tylosin tartrate are used, but the bacteria show increasing signs of antibiotic resistance (Lipsitch, 2002). The only significant treatment is to simply burn, bleach, or gas infected hives. Although a variety of responses to this disease have been used, bacteriophage therapy is yet to be tested. Phages, are a potential therapy because they can target and destroy their bacterial hosts (Gouchnauer, 1970; Valerianov, 1976; Haq, 2012; Jonczyk, 2011; and Stahly, 1999).

Historically, the scientific interest of phages has risen, declined and risen again. The simple reason for a reignited interest in the area of bacteriophages is to study them for their own sakes. Although in some sense, the study of bacterial viruses, or phages, has never left the forefront of biology. They were present at each significant step deeper and deeper into the

molecular understanding of life. Though not studied for their own sakes, phages were used as models to unravel the mysteries of higher life forms mainly because of their extremely simple biology. It was phages that were first used to combat pathogens before the discovery and mass production of penicillin. The British scientist Twort was the first to observe round clearings on bacterial lawns called plaques that are caused by bacteriophage lysis of host bacteria. Shortly thereafter, the French Canadian d'Herelle not only observed the same phenomenon, but began an effort known as bacteriophage therapy to harness the natural bacteriocidal attribute of phages to eliminate unwanted bacteria. From its inception bacteriophage therapy developed roots in the country of Georgia, where the Soviet leader Joseph Stalin allocated funds to finance the therapeutic research.

For decades, the use of phages to combat bacterial diseases was common in Eastern Europe, but for a variety of reasons some of which may be political, the therapy never quite caught on in the Americas. Penicillin seemed to be an effective and frugal alternative; however, bacterial resistance has more recently demonstrated that the continued use of a variety of antibiotics ranging from penicillin to vancomycin is unsustainable. It seems possible that the bacteria that have been historically treated with antibiotics are evolving resistances to such treatments. Primarily for this reason, there has been a renewed interest in bacteriophage therapy as an alternative to combating unwanted bacterial growth.

This reinvigorated interest in bacteriophage therapy seems a reasonable scientific and economic pursuit. The characteristics of phages may in some respects allow them to be an even better bacteriocidal treatment. Antibiotics harness the naturally zero-sum environmental factors of fungus and bacteria; meaning, their competitive relationship can be characterized as fairly evenly matched. Phages, on the other hand, are exponentially more abundant than bacteria, grow

faster and in more abundance, and are naturally capable of avoiding bacterial efforts to resist bacteriophage propagation. Simply put, there are more phages than bacteria, and phages have stronger adaptive qualities than bacteria whereas traditional antibiotics do not have these characteristics.

While such logic may become more accepted in the western scientific communities, realistic use of phages still has many hurdles to overcome. Already, research efforts are beginning to explore phages at an unprecedented level with advancements to genome sequencing and mass spectrometry on a larger scale. The number of phages studied in depth in the last few years far out numbers the number phages not only studied in the last few decades, but also in the depth in which they are studied. Technology has increased the breadth and depth of our understanding of phages.

Phages are particularly abundant in the biosphere. Some estimates have put the global bacteriophage population at 10^{31} virions (Weinbauer, 2002). They are extremely diverse with thousands of different phages infecting a single bacterial strain (Hatfull, 2014). Because they are great reservoirs of genetic material and are mediators of horizontal gene transfer, bacteriophages play a significant role in evolution (Pope, 2011, Farrar, 2007 and Hatfull, 2010). One way this occurs is through their method of reproduction.

The bacteriophage reproductive mechanism primarily works by hijacking a host's machinery to produce viral progeny. This reproductive mechanism has a high efficiency of energy to progeny ratio, meaning that with very little effort on the part of the bacteriophage, it can quickly set in motion a series of events that exponentially produce progeny. Once a bacteriophage adheres to the surface of a host cell, it inserts its own DNA into the interstitial fluid of the bacterium. From there, proteins and RNA involved in transcription assist the

promotion of the viral DNA into the host genome. At this point, the viral DNA could trigger a lytic phase or a lysogenic phase to proliferate the viral DNA (Xu, 2004).

Along with the high diversity of bacteriophages, comes a high specificity or preference of hosts. This is central to the bacteriophage therapeutic strategy because it suggests that bacteriophages can be used to safely treat patients. We can be confident that the only cells infected and lysed are unwanted bacterial pathogens (Haq, 2012). Host specificity provides many great tools to this holistic approach. Because of this specificity, bacteriophages can successfully target pathogens in a large and complex environment, kill only the pathogens, and limit their own proliferation. Once a bacteriophage kills all of its hosts, eventually, it will also be destroyed. Bacteriophages are self-limiting tools of destruction. There is no concern about a bacteriophage going rogue on neighboring cells once their specific hosts are nowhere to be found.

In addition to host specificity, bacteriophages also vary greatly in morphology. Nearly 95% of all characterized bacteriophages are classified as siphoviridae, but the remaining 5% are composed of nearly a dozen other classifications (Ackermann, 1998 and Veesler, 2011). Siphoviridae are a subset of caudovirdeae which are tailed bacteriophages. Myoviridae are the other subset of tailed bacteriophages, but are differentiated by their ability to retract and extend their tails. Conversely, Siphoviridae typically have long, non-retractile tails. Both classifications have icosahedral heads where the double stranded DNA is stored until infection (Xu, 2004). Regardless of morphology, phages have also been studied for their therapeutic potential.

Phage therapy harnesses the bacteria-bacteriophage relationship to allow nature to combat the disease using tools that have been created from millions of years of evolution (Cairns, 2009; Deresinski, 2009; Vessler, 2011). Bacteriophages have been at the forefront of scientific advancement for decades (Jonczyk, 2011). They have proven to be useful because of

their relatively simple biology and ease of compliance in the laboratory. But, their use as a therapy for bacterial disease has been controversial.

Bacteriophage therapy was initially a Soviet concept, and as a result did not catch much attention in the United States for decades. Today, Europe has discovered many effective uses of bacteriophages, and the Soviet stigma of bacteriophage therapy is a problem of the past. The remaining stigma comes from a general fear among the American populace of being intentionally infected with a virus. Although it would be a more accurate portrayal of bacteriophages to be seen as friendly, viral infections remain an area of fear (Farrar, 2007).

The circumstances surrounding the discovery of bacteriophage therapy partially explains the difficulty of the scientific community to embrace its significance (Jonczyk, 2011). The idea of bacteriophage therapy began in the 1920's in Eastern Europe. Around the same time, the discovery of penicillin was embraced in The West. Because of decades of political differences between the Soviet Union and the United States bacteriophage therapy was not well embraced by western science. Even after these political differences calmed down, bureaucratic resistance in the United States to bacteriophage therapy proved to be overbearing (Jonczyk, 2011). For phages to be therapeutically useful, multiple strains must be isolated periodically from the environment to avoid bacterial resistance (Jones, 2007 and Kysela, 2007). As a result, FDA approval and any sanctioned use of bacteriophage "cocktails" is seemingly out of reach for their actual use. For now, this obstacle resides primarily with therapeutic use of treating humans with phage. Luckily, fewer regulations on honeybees will make them a useful model for testing bacteriophage therapy in the United States. In recent years, there has been a revival of interest in bacteriophages (Deresinski, 2009 and Jones, 2007) even though they have consistently been at the forefront of

scientific discovery (Hershey, 1952; Luria, 1943; Sanger, 1977). In the United States they have only been studied as a model organism and in some cases for genetic manipulation.

For example, phages may also be used as a molecular tool in a process called recombineering (Murphy, 1998 and Zhang, 1998). Phages are used in genetic research to deliver specific DNA sequences into bacteria. This has provided new opportunities to develop new gene therapy techniques. The idea is that we can use bacteriophages to insert genes into microbes that will either kill them or produce progeny that will not be pathogenic. It is possible for a bacteriophage to infect a microbe without killing the host for many generations (van Kessel, 2008).

Recombineering is just one of many tools of bacteriophage therapy. Bacteriophage therapy is an area of interest primarily due to the decreasing capabilities of antibiotics. The basic idea behind the decreasing effectiveness of antibiotics is that bacteria evolve faster than the antibiotic treatments that are developed. Bacteriophages, on the other hand, are more abundant by an order of magnitude than bacteria, and are therefore able to evolve faster than the bacteria. Also, bacteriophages have much faster generation times than bacteria (Pope, 2011).

As a result, it is unlikely that bacteria will be able to develop significant resistance to bacteriophages (Haq, 2012). The most likely relationship between bacteriophage and host is ancient and has developed a balance of power between the two entities. This means that scientists are choosing to harness the already biological balance within soil and water to treat diseases (Jones, 2007). Instead of using a man-made tool as a cure, let us use a biological battle that has been going on for millions of years. Bacteriophage therapy recognizes that we can let biology work for us rather than trying to create biological cures ourselves. Such an approach

harnesses the evolutionary powers that have been refined through the crucible of surviving the harshest and most unforgiving environment in the world for millennia.

Because of new genomic technologies, genome annotation has proven to be great tool in understanding the evolutionary impact of bacteriophages. Although bacteriophages are not classified by the normal phylogenic trees, they are too diverse to be classified as species or sub species. In other words, even within morphologically identical bacteriophages genomic analysis often reveals no more than 70% homology. Most species usually vary at the genomic level no more than 5% to 10%. As a result, bacteriophages have been further sub classified into clusters and sub clusters based on their genomic relationships (Henry, 2010).

Within the siphoviridae classification, genomes have taken on a few identifiable patterns. First, domains that coded for structural proteins are nearly always located somewhere in the first half of the genome. Next, domains encoding for functional proteins that assist in structural protein assembly, hijacking of host resources, or preparation of lyses follow the structural domains. Finally, the remaining open reading frames (ORF's) tend to code for unknown functional domains (Henry, 2010).

These domains are significantly shorter than the domains in the first half of the genomes, and most have been shown to be non-essential domains. Some hypothesize that these unknown domains may provide some evolutionary context of the bacteriophage. Furthermore, these regions may have coded for important genes earlier on in the phage's life, but are no longer necessary and are constituted as artifacts (Pope, 2011).

Another important genomic discovery is the mosaic relationship of bacteriophage domains. Mosaicism describes how genes can be shuffled around to form new genes. Most results of the reshuffling cause an inactivation of function, some cause bacteriophage mortality,

and rarely does it result in an increased fitness of the progeny. Since the scale of bacteriophage turnover and population is so large, the rare event of increased fitness becomes a frequent occurrence in some sense. Because there are so many bacteriophages and such a high fluctuation of genes, these entities evolve faster than all other organisms on earth (Casjens, 2011).

Although there are many variables that must be considered when trying to identify and characterize a previously unknown bacteriophage, ideas generated from known bacteriophages from other orders of phylogeny have proven to be helpful. For example, the high level of understanding produced about mycobacteriophages seems to provide the largest body of reference even though the comparison is across more fundamental phylogenic relationships (Hatful, 2010). On the other hand, studies focused on bacteriophage infection of *Bacillus subtilis* have proven to also be helpful.

Bacillus subtilis is a phylogenic cousin of *Paenibacillus*. Since bacteriophage infection is primarily concerned with the structure of bacterial cell wall, *B. subtilis* is extremely useful to consider for our purposes. Although the large body of mycobacterial infection does suggest using calcium ions as a facilitator of infection, knowledge of calcium levels and infection time for *B. subtilis* are extremely detailed. The concentration of calcium that optimized infection was 75mM. This level of calcium is nearly twenty times more concentrated than standardized protocols for mycobacterium infection (Steensma, 1979).

Furthermore, infection time for *B. subtilis* is around thirty to forty minutes, while infection time for *Mycobacterium* is about twenty minutes. The higher dose and longer infection time can help pinpoint the best concentrations and burst time for *Paenibacillus larvae* infection. Although the actual mechanistic role of calcium continues to be elusive, it is hypothesized that it facilities binding of the siphovirdic tail to the host cell wall as well as play a role in incorporation

of bacteriophage DNA into the host genome (Steensma, 1979). Even though the contributions of *Bacillus* and *Mycobacterium* are informative, previous research directly involved with *P. larvae* are also helpful.

Because of the popularity of phages in Eastern Europe, it should not be a surprise to known that the first bacteriophage of *P. larvae* was isolated in Russia by Smirnova in 1953. Until now, there are no known studies that describe bacteriophages for *B. laterosporus.* Although AFB had been studied in America for nearly five decades, Smirnova was the first to discover an entity in the environment that only lysed *P. larvae*. His initial intent was to find an effective way to diagnose AFB, and shortly thereafter in 1954 he described his attempts to use it therapeutically. His work further investigated the use of phages as a prophylactic rather than as a treatment. Later on in 1961, he developed a phagovaccine that could help protect bee larva.

When Smirnova started working on *P. larvae* bacteriophage therapy in Leningrad, only a few years later in 1955 in the Midwest of the United States, Gouchnauer also isolated a bacteriophage of *P. larvae*. Although the intent of his research was not always clear, he provides much information regarding the properties of these phages (Gouchnauer, 1954). From his writings, he seemed hesitant that these phages could be used as a treatment, but optimistic that they could be used for AFB diagnosis. His main concern regarding therapeutic use was that an effective treatment could only come from a bacteriophage cocktail containing many different strains (Gouchnauer, 1958). In other words, it was not clear at that point in time how specific phages were to their hosts. There seemed to be some debate regarding the infectivity of phages across an entire species. His findings demonstrate his skepticism by showing that some phages that infect a specific host strain from one region have no ability to infect another host strain of a separate region. Regardless of his intent, his work provided the most extensive contribution to

Paenibacillus bacteriophage research. He was able to determine the effects of temperature and pH on bacteriophage growth, infection times, a burst time curve, the number of progeny produce per infected cell, and cross-infection frequency (Gouchnauer, 1970).

There was one more brief contribution made in 1976 by Valerianov in which another bacteriophage was isolated in Bulgaria, but the next significant contribution came in 1984 and 1985 from Dingman, Bakhiet, Field, and Stahly. Two more phages, PBL1 and PBL0.5, were isolated in Iowa and were observed under an electron microscope. Also, a physical map of the PBL1 genome was described to have cohesive ends, similar to *B. subtilis*, and was hypothesized to form concatemers in the host.

Most recently in 1999, the same Stahly that contributed to the 1984 work, discovered a virulent mutant that could potential have the therapeutic properties long ago described by Smirnova. Such a discovery could potentially alleviate the concerns once held by Gouchnauer about virulence of phages beyond diagnostic uses. While working with PPLc1 and PBL1c, a mutant of the 1984 strain PBL1, he was able to determine that the genome was composed of dsDNA and about 40 kbp in length.

Finally, if the efforts of this work demonstrate a useful contribution to the current body of knowledge regarding bacteriophages, it will do so by providing at least a partial description of nearly a dozen new strains, and some of the first fully sequenced and annotated genomes. This work also describes the preliminary steps that could someday lead to an alternative treatment of Foulbrood Disease. Bacterial infections are becoming more and more difficult to eliminate from bee hives, but understanding and using the bacteriophages that infect and lyse them may prove to be another treatment. Utah, a state struggling with the persistence of Foulbrood Disease, acts as a reasonable environment to extract, study, and use bacteriophages of pathogens in the

environment and in the laboratory. This work hopes to be another example of the general importance of studying phages.

It is for these reasons that this research has invested curiosity in the realm of bacteriophages. The use and study of phages is so vast that a variety of scientific designs could potential be advantageous. Such designs could attempt to superficially look at a large number or variety of phages. This is the situation for many advantageous metagenomic efforts. Other designs could attempt to focus on a single bacteriophage or cluster of bacteriophages that infect the same bacteria. This is the situation for many phage hunter programs studying mycobacteriophages (Hatfull, 2010). It is the latter research model that this thesis attempts to apply on a small scale in Utah where we have observed a local problem. As an initial step towards a bacteriophage therapy for Foulbrood Disease, *B. laterosporus* strains were isolated and characterized, as well as bacteriophages that infect them. We established a working relationship with the Utah County Beekeepers Association and obtained strains of both the bacteria and the phage. We also obtained PPLC1 from Alippo, a well-studied bacteriophage of *Paenibacillus* Bacteriophage from outside of Utah.

The new strains of *B. laterosporus* were isolated from honey and larva from local bee keepers that do and do not show signs of the disease. Characterization of *B. laterosporus* included a monitored growth on selective media, Gram reaction test (including a KOH quick test), catalase test, and 16S rRNA gene sequencing. The new strains of the bacteriophage were isolated from soil near local bee hives that do and do not show signs of the disease. Bacteriophages were characterized with an electron microscope, a restriction endonuclease digest of their DNA, and genome sequencing. The novel contribution of this study lies in the annotation of the bacteriophage genomes.

Although beyond the scope of this work, these goals are intended to provide a foundation for future therapeutic use of bacteriophages. With an archive of host bacteria and bacteriophages, the next step would be to treat bee hives with a cocktail spray of bacteriophages in the environment. To see this work to completion, the bacteriophages should be tested on bee hives already infected with Foulbrood and either shows signs of the disease or not. Although some useful information would be gathered from treating new hives that have yet to pick up the bacteria, the treatment of a bacteriophage cocktail spray would provide the largest benefit by testing its usefulness to already infected hives. Most well established hives carry some level of *M. Plutonius, B. laterosporus,* and *P. larvae* whether she exhibits symptoms of Foulbrood or not.

Finally, it is imperative to understand that at the core of this work are three actors: Honeybees (*Apis mellifera*), a honey bee pathogen (*B. laterosporus*), and bacteriophages of that pathogen (*Brevibacillus* bacteriophages). The relationship between these three entities is complex, and although there has been extensive research performed to understand the honeybee and the pathogen, little is known about the bacteriophage. These bacteriophages of interest play an essential role not only in the environment in general, but specifically can have a profound impact on this pathogen and the honeybee.

CHAPTER 2 – Materials and Methods

Samples were collected and donated by members of the Utah County Beekeepers Association. Each beekeeper was carefully instructed on how to gather samples following standardized protocols. Specifically, honey samples were placed in sterile 15 mL conical tubes. Collections of larvae used a toothpick or probe to remove either an infected or healthy sample and placed in a sterile 1.5 mL Eppendorf tube. Soil samples were taken using a probe, pocketknife or spoon to scrape the soil surface near a hive that did or did not include dead bee carcasses and were placed in a sterile 50 mL conical tube. With each sample, the sampling date, health of bees, approximate ambient temperature, location of soil sample in relation to hive, hive GPS coordinates, and a brief description of the sample location were recorded. Sample containers were closed tightly and stored in a dark, dry, room-temperature or cooler location until the samples were delivered to the laboratory.

Methods to isolate *P. larvae* and *B. laterosporus* spores from honey are well established, and we followed the published protocols with exactness (Hornitzky 1991, Schuch 2001, de Graaf 2006). Spore suspensions that were extracted from honey samples were streaked onto *Paenibacillus larvae* agar (PLA) plates and incubated at 37°C (Dingman 2012). Plates were checked every 24 hours for colony growth. Colonies were streaked to purity and their morphology described. Since *P. larvae* and *B. laterosporus* strains are Gram positive and produce catalase, each isolate was examined for Gram reaction and catalase reaction. A loopful of each strain was combined with 800µL of Lysis Broth and 800µL of 40% glycerol in a cryovial and stored at -80°C.

Although we did not use a significant portion of the larvae samples, we felt the samples might be beneficial in the future. Nevertheless, we developed protocols to isolate *B. laterosporus*

from larvae carcasses. To isolate *B. laterosporus* from a larva, we placed a larva in a 15mL tube with 6mL of Phosphate Buffer Saline and homogenized it until most of the tissue was separated. We then centrifuged at 4000 RPM for 40 minutes and heated the tube at 80°C for 40 minutes to better isolate the endospores. Finally, the isolated spores were plated on PLA and incubated at 37°C for two to three days.

Identity of isolates was confirmed using polymerase chain reaction (PCR) to amplify the 16S rRNA gene (Martinez, 2010; Ryba, 2009; Dingman, 2012; Bakonyi, 2003). A colony from each strain was boiled in 60µl of ddH20 in a PCR tube for 5 minutes. This provided the template for the PCR cocktail containing 5µL 10x REDtaq Buffer, 1µL nucleotides (200µM of each dNTP), 1µM Forward Primer (ACTCCTACGGGAGGCAGCAGT), 1µM Reverse Primer (CGATTACTAGCGATTCCGACTTCA) (Liu et al., 2005), 2.5µL REDtaq DNA polymerase (0.05 unit/µL), 1µL template (200pg/µL), and 38.5µL ddH2O. This 50µL cocktail was made for each *P. larvae* strain. The reaction was performed as follows: 95°C for 4 min then 30 cycles of 95°C for 1 min, 37°C for 30 sec, and 55°C for 2 min. After the cycles, the PCR stayed at 55°C for 5 minutes and held at 4°C. Ten microliters of each strain result was electrophoresed on a 1% TAE gel at 150 Volts for 30 minutes. Products were sent to the BYU DNA sequencing center for *454 GS-FLX Titanium pyrosequencing*. A loopful of each strain was combined with 800µL of Lysis Broth and 800µL of 40% glycerol in a cryovial and stored at -80°C.

Next, soil samples were used to obtain strains of *Brevibacillus* bacteriophage. We tried a variety of methods to isolate robust bacteriophage from the soil samples, and we were successful with multiple protocols (Gouchnauer, 1955; Hatfull, 2010; Henry 2010; Jakutyte, 2011). Twenty-five milliliters of Lysis Broth was placed in a sterile 250mL Erylenmyer flask. A loopful of either a colony from a plate or defrosted freezer archive sample was used to inoculate the

flask. The flask was placed in a 37°C incubator shaker. Once the broth culture of a single *B. laterosporus* strain exhibited sufficient turbidity, a spoonful of soil was placed in the flask and returned to the same incubator shaker for 48 hours.

After 48 hours the enrichment culture was transferred to a 50mL conical tube and centrifuged at 4000 RPM for 10 minutes. The supernatant was then filtered through a 0.2µm filter. Each soil enrichment filtrate was stored and the strain that was used to initially enrich the bacteriophage was recorded. The isolated bacteriophage was then tested for infectivity of *B. laterosporus* strains.

Bacteriophages were then subjected to multiple passages for purification and preparation of lysates. Five hundred microliters of *B. laterosporus* broth was transferred to a test tube with 100µL of calcium (1M $CaCl_2 * 2H_2O$), and 100µL of bacteriophage lysate. After 45 minutes incubation, it was mixed with 4.5mL of 1x Lysis Top Agar and transferred to a Lysis Agar plate. The contents were allowed to solidify at room temperature and placed in a 37°C incubator for 24 hours. After 24 hours, the plates typically showed a yellow opaque lawn with many circular plaques. If only a few plaques were present, the plaque would be picked with a sterile, 200µl sterile pipette tip and transferred to 100µL of Lysis Broth in a 1.5mL tube. This bacteriophage suspension was used to repeat the same plating process multiple times.

Once a high titer of bacteriophage lysate was obtained, ten web plates would be soaked in 5mL of Lysis Broth for three hours and liquid would be harvested and stored into a new 50 ml conical tube. This high titer bacteriophage lysate would be used to test cross infection of *B. laterosporus* strains, electron microscopy, and DNA analysis. Each bacteriophage lysate was archived by combining 800µL of the lysate and 800µL of 40% glycerol and stored at -80°C.

The first step in characterizing the bacteriophage was to view them with a transmission electron microscope (Dingman, et al 1984). Ten microliters of bacteriophage lysate was placed on Parafilm and a new EM cooper grid was placed (dark side down) on top on the droplet. After 20 minutes, the grid was transferred (dark side down) to a 10μL droplet of 2% phosphotungstic acid for 2 minutes. Finally, filter paper was used to remove any remaining liquid and the dry copper grid, dried, and stored in a protected case and sent to the BYU electron microscopy center for imaging. Images were obtained by Michael Standing from the BYU Electron Microscopy Center using a FEI Tecnai T-12 microscope.

Once the bacteriophage structure was identified with electron microscopy, we anticipated the need for a rigorous extraction method of DNA. The bacteriophage head appeared to be assembled with strong links between the proteins, and we experimented with a variety of extraction techniques until we found one method that produced clean results. Twenty milliliters of bacteriophage high titer lysate was placed in a clean, autoclaved 50mL polycarbonate oak ridge tube and 10μL of nuclease mix was added. After 2 minutes of mixing by inversion, the tube was placed at 37°C for 60 minutes. Next, the tube was left undisturbed at room temperature for 30 minutes. Twenty milliliters of Phage Precipitant Solution was added, mixed by inversion for 2 minutes, and placed on ice for 30 minutes. The tube was centrifuged in a Sorvall RC 5C Plus with a SS-45 rotor at 20000RPM for 10 minutes. The supernatant was discarded and the tube was inverted and dried for 5 minutes. The pellet was resuspended in 300μL of TE buffer. Slow pipetting loosened the pellet from the tube to produce a homogenous suspension.

The suspension was transferred to a 1.5mL tube along with 20μL of Proteinase K and incubated for 120 minutes at 55°C. The sample was mixed with 600mL of equilibrated phenol for 5 minutes until a milky white homogenous solution appeared. The sample was then

centrifuged at 14000RPM for 15 minutes. The aqueous layer was transferred to a new 1.5mL tube and the process was repeated except with chloroform.

The sample was washed with phenol and chloroform, 1mL of 100% ethanol was thoroughly mixed with the sample and allowed to sit overnight in a -20°C freezer. After 24 hours, the ethanol-precipitate sample was centrifuged at 14000RPM and the supernatant was removed. The pellet was then rinsed with 95% ethanol and allowed to sit on ice for 20 minutes. The sample was centrifuged again at 14000 RPM and the supernatant was removed and the pellet was dried on a heating block at 55°C for 5 min. Once all the liquid had evaporated, the pellet was resuspended in 300µL of TE buffer. If the pellet did not resuspend completely, then a little more TE buffer was added and the sample was placed on the heating block again at 55°C for 5 minutes. The resulting DNA solution stored at -20°C.

Two methods were used to assess the quality of the extracted DNA: spectrophotometric determination using a NanoDrop ND-1000 Spectrophotometer performed at the BYU Research Instrumentation Core Facility (RIC), and fluorometry as performed by the BYU DNA Sequencing Center prior to genome sequencing. Each sample that was sequenced passed quality control test at the BYU DNA Sequencing Center.

For gel electrophoresis, a 1% agarose gel was prepared with 1x TAE and high grade agarose. A DNA sample and 1µl of 10x loading dye was combined and loaded into each loading well. Once the wells were loaded, the samples were electrophoresed at 120 volts for 40 minutes. The sample was oriented to flow in the gel toward the cathode (+). The results were recorded using a DS-34 GelCam.

Raw genomic information obtained from the BYU DNA Sequencing Center was analyzed using DNA Master, a Multiple Document Interface (MDI) program. Open reading

frames (ORF's) were identified and compared to the National Center for Biotechnology Information's (NCBI) database of primary biological sequence information via the Basic Local Alignment Search Tool (BLAST) algorithm. Identifying Shine-Dalgarno (SD) sequences also helped identify which ORF's had the highest potential for producing putative gene products. Putative gene products were then determined among a team of genomic analysts to prepare the annotated genomes for publication.

CHAPTER 3 – Results of Isolation and Characterization

B. laterosporus isolated from Utah County honey samples were consistent in morphology with other known *P. larvae* strains described in previous studies (De Graaf, 2006 and Hornitxky, 1991). Table 01 briefly summarizes the findings of fifteen isolated colonies. Each colony was grown on PLA and characterized based on standardized colony protocols established by the American Society for Microbiology at http://www.microbelibrary.org/component/resource/laboratory-test/3136-colony-morphology-protocol.

Although each strain was inoculated multiple times with the phage, the most useful strains were BL 02 and BL 06 (see Table 01). Typically, broth cultures inoculated from the refrigerator colony (3°C) showed sufficient growth after 24 hours while the freezer stock (-80°C) showed sufficient growth after 48 hours. Each strain, except for BB 08, was able to be infected by each of the five phages at least to some extent. BL 02 and BL 06 were the most susceptible and were able to provide many web plates for high titers. BB 08 passed the initial morphology, KOH, and H2O2 tests, but failed the 16s test. The BB 08 16s test revealed that the strain was actually a close cousin named *Brevibacillus brevis*. We used BB08 as a negative control as well as performed some minor cross-infection tests that resulted in no significant outcomes. Although not described in this work, we also received a strain of *P. larvae* from Alippi in Argentina (1995). Initial test showed this strain to be highly susceptible to our phage, but additional tests will be described at a later time.

Among the many bacteriophages isolated from numerous soil samples from Utah County, five were virulent enough to produce high quantities of high titers for additional characterization. The bacteriophage were named after basketball players from the 2011 BYU Basketball starting lineup (see Table 02). Of those that were fine-tuned, the most significant variables included

calcium levels and infection time (Steensma, 1979). Plating bacteriophage on bacterial lawns is considered by some to be an arduous task to produce titers on a large scale, yet it was the most conservative approach to insuring results. In other words, liquid culture methods were attempted and some results were produced, but because no significant protocol of liquid culture is known and plating techniques were used as the alternative.

Plaque morphologies are presented in Figure 01. These morphologies were most often well-defined, clear circles of 2-4 mm in diameter similar to the images for Jimmer1 and Jimmer2. Plaques with larger diameters from 3-4 mm (Abouo) and 9-10 mm (Davies) were seldom. The most rare plaque formations occurred early on in the isolation process and were similar to the "s-shaped" clearing of Emery. This morphology occurred only during the first or second pass of Jimmer1, Jimmer2, and Emery. Morphologies similar to this were not observed once the bacteriophage was considered to be isolated from other strains. The widths of these clearings were 9-10 mm with an additional 3-4 mm more opaque halo. Although halos such as this were observed, multiple retests lead to the conclusion that all bacteriophage are lytic and produce uniformly clear plaques.

Initial TEM images revealed bacteriophage particles that were similar to the siphoviridae classification. There was a tendency to see bacteriophage tails and no bacteriophage capsid. However, additional TEM revealed fully intact phages and are shown in Table 02. These newer images included icosahedral bacteriophage particles leading to the conclusion that the bacteriophages are myoviruses.

Prior to the 454 pyrosequencing of the each phage, a sample of each DNA was observed on gel electrophoresis. Figure 03 provides the results of restriction endonuclease digest. Four samples were tested, but sequencing revealed five unique genomes. The top left RE Digest

(Jimmer2) and top right RE Digest (Jimmer1) were nearly identical. The similarities in fragment lengths were later supported by the high similarity in genome sequencing. These results led to naming of these bacteriophage as Jimmer1 and Jimmer2. Their distinctions are indicated with the numeral 1 and 2 at the end of each name.

The bottom left RE Digest was unusually abundant in DNA (See Figure 03). Although at the time it was believed that this sample was pure, subsequent genome sequencing revealed two unique bacteriophage genomes. This would account for the substantial brightness of the DNA fragments. These bacteriophage were named Emery and Abouo and are consistent with slight differences in bacteriophage particles obtained by TEM from the same titer. Unfortunately, only after further attempts to isolate these two bacteriophage will a clearer RE Digest reveal two unique banding patterns. Finally, the bottom right RE Digest is for Davies. It confirms that the extraction successfully isolated DNA, but is inconclusive in establishing a unique banding pattern. Later, the genome sequencing revealed Davies to be a unique phage.

Whole genome sequencing revealed five unique genomes of *Brevibacillus* bacteriophage. Table 02 provides a brief summary of each genome. In addition to the data provided in the table, all base one calls were made between the putative terminase gene and the previous gene and each genome was determined to be circular by Newbler. The base one call for Emery was made at an Integrase gene (See Table 05). This decision was made because a terminase gene could not be identified, and the genome alignment at this location showed the greatest homology to the other four genomes. Jimmer1 had an average fold coverage of 250.16. Jimmer2 had an average fold coverage of 271.55, Emery had an average fold coverage of 143.2, Abouo had an average for coverage of 116.9, and Davies had an average fold coverage of 130.78. Jimmer1, Jimmer2, and Davies were each assembled by multiple contigs while Emery and Abouo were sequenced in

their entirety. No tRNAs were found.

Genome annotations were submitted to Genbank and assigned accession numbers (See Table 02). Tables 03, 04, 05, 06, and 07 provide information about each gene from the five genomes. There are a total of 485 genes published on GenBank as a result of this research. Further analysis of these genomes revealed sixty structural proteins and 133 non-structural proteins. 90% ± 3% of the genes were located on the forward strands with an average G + C content of 39.48% ± 1.41% (Merrill, 2014).

As recorded in Tables 03 and 04, Jimmer 1 and Jimmer 2 appear to be nearly identical with the only differences at gp13 where the Jimmer 1 protein is a helix-turn-helix domain-containing protein and the Jimmer 2 protein is similar in function to hypothetical protein DesyoDRAFT_1114. Table 05 lists the gene products for Emery and demonstrates a genome that is most unlike the other four genomes. Among the 100 gene products, only twenty-three have functions similar to that of known proteins. Of these twenty-three known proteins, ten have no similar protein found in the other four genomes. These ten unique gene products include a glycoside hydrolase (gp39), ricin B lectin (gp45), tyrosine recombinase (gp46), rare lipoprotein A (gp59), an activator of middle period transcription (gp61), prophage lamdaBa04 DNA-binding protein (gp62), AbrB family transcriptional regulator (gp70), carbonic anhydrase (gp80), virulence-associated E family protein (gp94), VRR-NUC domain-containing protein (gp96), and SNF2-like protein (gp97).

Jimmer 1 and 2, Abouo, and Davies (Tables 03, 04, 06, and 07) share and interesting relationship where all four share many structural and non-structural proteins with islands of proteins similar in various pairs of the four genomes. All four genomes share forty gene products including similar terminase small and large subunits, bacteriophage portal protein, head

morphogenesis protein, bacteriophage tail sheath and core tail proteins, XRE family transcriptional regulator, single-stranded DNA-binding protein, recombination protein U, dUTPase, MarR family regulatory protein, and RNA polymerase sigma-24 subunit. Abouo and Davies (Tables 06 and 07, respectively) share twenty-nine similar proteins including a prophage protein, putative DNA packaging protein, extracellular solute-binding protein, polyprotein, Kelch repeat type 1-containing protein, exo-glucosaminidase LytG, prophage antirepressor, CRISPR-associated helicase Cas3, integrase, putative prophage LambdaCh01 replication protein O, DNA replication protein, and DNA repair protein RecN.

Jimmer 1, Jimmer 2, and Davies (Tables 03, 04, and 07, respectively) share six similar gene products and only one has a known function as a site-specific DNA methylase (gp81 and gp80, respectively). Jimmer 1 and Jimmer 2 (Tables 03 and 04, respectively) share only one uniquely similar gene product with Abouo which is GumA (gp54 and gp48, respectively). Jimmer 1 and Jimmer 2 have fifty-one gene products that have no known similar gene products found in the other three genomes including thaxtomin Synthase B (gp42), histidine kinase (gp45), chromosome segregation ATPase (gp69), and a pyrophosphokinase (gp94). Abouo (Table 06) has twenty-two gene products not among the other four genomes, which includes a cell division protein (gp49), a putative NAD dependent epimerase/dehydratase (gp57), a plasmodium membrane protein (gp58), a protein similar to zinc metalloprotease (gp68), phage N-6-adenine methyl transferase (gp77), and a putative catechol dioxygenase (gp78). Davies (Table 07) has only eighteen gene products unique among the five genomes and with known functions including a second phage terminase small subunit (gp49), homoserine kinase (gp58), and a tonb-dependent receptor protein (gp77).

Among the gene products of all five genomes (Tables 03, 04, 05, 06, 07), additional studies showed five similar assembly or structural proteins and seven regulatory or non-structural proteins. The five assembly or structural proteins include a terminase large subunit (gp2, gp2, gp2, gp3 respectively with Jimmer 1 and 2 sharing the same gene product numbering), SPP1 Gp7 family head morphogenesis protein (gp4, gp4, gp5, gp4 and gp5), tail length tape measure protein (gp20, gp20, gp20, gp16), baseplate J family protein (gp27, gp27, gp29, gp21), and a tail protein (gp15, gp15, gp16, gp12 and gp18). The seven regulatory or non-structural proteins include LysM domain-containing protein (peptidoglycan binding) (gp22, gp22, gp25, gp17 respectively with Jimmer 1 and 2 sharing the same product numbering), peptidoglycan hydrolase (gp36, gp36, gp38, 3p31), phage-like element PBSX protein (gp26 and gp28, gp26 and gp28, gp28 and gp30, gp22), bhlA/Bacteriocin (gp34, gp34, gp36, gp29), DNA replication protein (gp66, gp70, gp77, gp99), site-specific DNA methylase (gp79, gp81, gp88, gp33), and RNA polymerase sigma-70 factor (gp91, gp91, gp99, gp94) (Merrill, 2014).

Most of the significant differences in gene products are contained in the Emery genome (Table 05). This genome includes six proteins that are not contained in the other four genomes. These proteins include an integrase family protein (gp1), a phage virion morphogenesis family protein (gp10), a prohead core scaffolding/protease (gp6), AbrB family transcriptional regulator (gp72), a DNA-dependent DNA polymerase family A (gp81), and a virulence-associated E family protein (gp96). Jimmer 1 and 2 (Tables 03 and 04, respectively) have four missing proteins that are in the other three genomes as well as three proteins that are only found in their genome. The four missing include a tail fiber protein, a membrane protein, tyrosine recombinase XerC, and an accessory gene regulator. The three unique Jimmer 1 and 2 proteins include a

subtilisin-like serine protease (gp42), serine recombinase (gp49), and phage replication protein O (gp76) (Merrill, 2014).

CHAPTER 4 – Discussion

P. larvae is a well-studied bacterium and this research demonstrated that the already well-established protocols for growing it are reproducible and verifiable. The results we achieved from growing the bacteria in 24 hours from refrigerated stock and 48 hours from frozen stock in the broth recipe provided from the literature were consistent with what was anticipated.

Although each strain of *P. larvae* that was isolated and grown in the laboratory, it is not clear why only BL 02 and BL 06 produced high enough titers to produce a sufficient amount of web plates in order to adequately amplify the bacteriophage. For now, only suspicion regarding the infectious susceptibility of these two strains compared to their more resistant relatives will be reasonable. A full genome analysis of strains BL 02 and BL 06 would likely reveal which combination of absent protective genes or present susceptible genes allowed the bacteriophages to infect to a significant extent to provide high titers. The only genetic test on all *B. laterosporus* strains used was the 16s test to confirm their identity. This test is insufficient to provide insight into the variances between strains of the same species.

During the process confirming the *B. laterosporus* strains, 16s test revealed that BB 08 was *Brevibacillus brevis*. This result allowed us to used BB 08 as a negative control when infecting the other *B. laterosporus* with the bacteriophage. Surprisingly, BB 08 was able to inconsistently produce plaques when infected with *Brevibacillus* bacteriophage. This result may provide insight into the ability of bacteriophages to bacteria across genus and species. Other studies have shown that bacteriophages have been known to sometimes infect across species but not across genus. For example, mycobacteriophages have been known to infect both *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* (Hatfull, 2010). This cross infection

phenomenon is the justification for using *M. smegmatis* for bacteriophage research because it is a safer and faster growing bacteria than its cousin *M. tuberculosis*.

If this same idea holds true for *B. laterosporus* and *B. brevis*, then this could provide additional justification to either expand the scope of bacteriophage cross infection or to again reevaluate the genus and species classification of *B. laterosporus*. Another possibility is that the 16S test inaccurately demonstrated that the host bacteria is *B. laterosporus*. This is especially possible considering the most recent author correction demonstrating that the bacterial host was changed from *Paenibacillus larvae* to be *Brevibacillus laterosporus*. This new information would then show consistency with the already established notion that bacteriophages can cross infect between species. The bacteriophage used is able to infect both *B. laterosporus* and *B. brevis*.

It is interesting that the early use of bacteriophages were to identify *Bacillus larvae* and *Bacillus pulvifaciens* (Genersch, 2006). The division was based on qualitative standards of bacterial virulence. The use of phages for identification purposes supported a division between B. *larvae* and B. *pulvifaciens*, but not necessarily at the species level. Later on when 16S gene sequence were analyzed (Bakonyi, 2003), the distinction between the two bacteria was eliminated without any additional bacteriophage studies supporting the change. Although 16S gene sequences is the current standard, bacteriophage-host specificity which is primarily based on attachment proteins is another reliable test that focuses identification on other variables (Brussow, 2013). During this process, the bacterial strain was believed to be *P. larvae*. At the time, this conclusion was the result of data the National Center for Biotechnology Information website and confirmatory tests we performed. Later on, it was discovered that the bacterium were *Brevibacillus laterosporus* (Sheflo, 2015).

Once the host pathogen was confirmed and plaques were discovered, the next step was to amplify the amount of bacteriophages. The two approaches were to use either bacterial lawns or liquid cultures. Using bacterial lawns is a slower and more stable approach while liquid cultures is faster but at greater risk for contamination. Although the literature provided clear protocols regarding each approach, ultimately using bacterial lawns was the chosen. Although laborious, this approach was successful at producing bacteriophage titers high enough for genome sequencing.

During the first two passages of bacteriophage isolation, there was variation of plaque morphology that included an "s-shaped" clearing by Emery and plaque diameters ranging from 2-4mm by Jimmer 1 and 2 to 9-10mm by Davies (See Figure 01). After a third passage of isolation, all plaques produced were uniformly clear circles about 9-10mm in diameter. The morphology is consistent with lytic bacteriophages. This process confirms the established need for multiple passages of infection to better purify the bacteriophage.

It is possible that during the early passages for isolation that there are other bacteriophages competing to infect the host. Since it can be reasonable assumed that the host has already been adequately isolated, the other possibility is that there are other bacteriophages that either have weaker infectious properties or for any number of reasons do not survive through multiple passages. The protocol developed for this determined that early on in isolation, plaques of different morphologies even when cohabitating on the same plates should be further isolated on subsequent plates.

Electron microscopy confirming that all five bacteriophages are myoviridae provides plenty of information regarding its nature (See Figure 02). Typically, myoviridae are lytic rather than lysogenic and contain linear, dsDNA with a G + C content about 35% (Capparelli, 2007).

Brevibacillus Bacteriophage are consistent with a more lytic nature and contain linear, dsDNA. The 39.48% average G + C content of the genomes demonstrates a deviation from the commonly accepted 35% for the myoviridae family. This difference may be the result of the host G + C content being higher at 44%. Future research could explore this discrepancy within the myoviridae family or for bacteriophages in general.

Bacteria are known to have wide ranges of G + C content which are often correlated with coding regions. This fluctuation may also occur among bacteriophages which could demonstrate a G + C content variation that correlates more with the host content rather than with the bacteriophage morphology. In other words, G + C content may be determined more by the content of the host rather than the type of bacteriophage, or determined by both factors.

At the time when these five genomes were sequenced, they were believed to be the first sequenced genomes of their kind. Soon after, a *Paenibacillus larvae* bacteriophage genome was published in Portugal named philBB_Pl23. Later on, it was discovered that the five Utah genomes belong to a bacteriophage of a different host, *Brevibacillus laterosporus*. This helps support the result (See Tables 03, 04, 05, 06, 07) that a significant portion of the gene products have close homologs to *B. laterosporus* (ten homologs in Jimmer 1 and Jimmer 2, thirty-seven in Emery, thirteen in Abouo, and fourteen in Davies) while there are less homologs closely associate with *P. larvae* (nine homologs in Jimmer 1 and Jimmer 2, three in Emery, seven in Abouo, and seven in Davies). Although this data helped lead to the conclusion that the host was actually *B. laterosporus* the homologs also demonstrate a large array of gene product homologs with other frequent matches with *Desulfitobacterium*, *Clostridium*, and *Bacillus*. The majority of gene product homologs suggest that there is similar proteins found among bacteriophages that infect the firmicute phylum in general.

Fold coverage played a helpful role in sorting out the genomes for sequencing and reassuring that the products for sequencing were sufficient to produce credible results. Having fold coverage well over 100 and even over 200 for some cases helps reassure that the genomes are authentic. The high fold coverage made it possible to distinguish between Emery and Abouo who were isolated together. In other words, these two unique bacteriophages were only distinguishable from one another at the genomic level. Plaque morphology (see Figure 01) and electron microscopy (see Figure 02) was not definitive enough to distinguish between the two bacteriophages. Also, this means that the two bacteriophages were not able to be isolated from each other even after multiple passages of infection.

High fold coverage also made it possible to observe subtle differences between the genomes of Jimmer 1 and Jimmer 2 which were 99.8% similar. Although no noticeable difference between the two bacteriophages regarding infectiousness or morphology, a 0.2% difference in genomes most likely occurred later on in the evolution of these two bacteriophages. It is likely that at the time of isolation in the laboratory, these two bacteriophages had a common ancestor. Through multiple passages of isolation and stages of amplification, one isolate became two genetically unique bacteriophages. It is possible with the number of generations produced in the short time of this research that a 0.2% difference in genomes can occur. This rapid change in genome supports the idea that bacteriophages play a central role in genetic variation.

Next, Table 02 provides interesting numbers about the range of G + C content from bacteriophage to bacteriophage. The host genome G + C content is about 44%, and while G + C contents of bacteriophage closely resemble that of their hosts, this data suggests otherwise. The G + C content of the bacteriophage is significantly lower than the G + C content of the hosts. Additional research may pinpoint the reason for this discrepancy. Some possible causes could

include the addition of pathogenicity islands in the hosts that increase the G + C content. To date, no other *B. laterosporus* genome and G + C content is known. On average, a more consist G + C content may be identified later on as the number of genomes of both host and bacteriophage are characterized. Differences in G + C content between themselves and with their host.

A thorough discussion about the gene products on these bacteriophages is beyond this scope, and has already been investigated adequately in another work (Merrill, 2014). However, a cursory look at the genes list can offer some insight. This myoviridae has a genome of 45,000 bases to 55,000 bases and number of genes between 92 and 100. Although many structural and non-structural proteins are known, the majority of gene products have no known function and are similar to other hypothetical proteins found in bacteriophages and their host within the firmicutes phylum. The most common homologues of *Brevibacillus*, *Paenibacillus*, *Desulfitobacterium*, *clostridium*, and *bacillus* share a similar difficulty of containing a majority of proteins of unknown function. This highlights a point in research where genomic databases have more gene products than known structure or function.

This leads to the direction of future research to explore the role these gene products play. These hypothetical proteins can be anything from old defunct proteins that no longer play a role in the life cycle of these bacteriophages to a subtle regulatory or virulence protein that can mean the difference between survival and extinction. The effort to exploring these possibilities in the relatively simple bacteriophage can serve as a foundation for understanding the role of unknown gene products in other more complex entities.

With the exception on Emery, of which hardly any gene product showed a known function, most structural proteins that make up the actual myoviridae appear to be grouped together at the beginning of the genome (See Table 05). Convention suggests that the terminase

gene product be the determining protein for the base one call, and this convention helps sort the general use of the gene products for these genomes. The most significant order of the genomes is grouped around the base one call. For Jimmer 1 and 2, Abouo and Davies (Tables 03, 04, 06, and 07, respectively), twenty-one of the first thirty-six gene products (58%) are functionally identical while the remaining sixty gene products contain twenty functionally identical matches (33%).

The exception of the base one call of Emery at an Integrase gene product may play a significant role in the activities of the bacteriophage. For example, the integrase gene is used to help integrate the bacteriophage genome into the host genome. If this is occurring, it may suggest that Emery has a more lysogenic activity than lytic activity. This bacteriophage may want to integrate rather than immediately reproduce numerous progeny until the bacterial host bursts. This is another common strategy of a bacteriophage to ensure the propagation of its genomic information. At the moment, there is no clear correlation with plaque morphology of Emery to support a possible lysogenic lifestyle.

Although prior research has provided no information regarding the character of these bacteriophage, this research has provided the first intimate look at the genomic details. High throughput sequencing technology has allowed for lesser studied bacteriophage to be reconsidered for avenues of research. Even now, this research has provided support for follow-up studies that include taking these bacteriophage to a therapeutic level of treatment. This research could in part support the validity needed to seriously consider bacteriophage therapy as a reasonable treatment for CCD. Even beyond the scope of treating this single disease of a single insect, this research hopes to inspire curiosity into the use of bacteriophage therapy for other agriculture diseases and possible treatment for humans.

ADDENDUM

The laboratory effort for this thesis was conducted between 2011 and 2012. During that time, *P. larvae* was the intended pathogen of research because of its direct role in causing American Foulbrood (AFB). The method and materials of isolating *P. larvae* from local bee hives in Utah County was consistent with well-established protocols from published literature. It is now known that this same process was used instead to unintentionally isolate a close cousin, *B. laterosporus*. At the time, the catalase and 16s rRNA tests incorrectly confirmed the bacteria identity as *P. larvae*. These confirmatory tests were considered the gold standard for identification. Laboratory work continued, and bacteriophages that infected *B. laterosporus* were unknowingly isolated and characterized.

In 2015, after additional bacteria and bacteriophage strains were isolated in similar fashion, multiple evidences brought into question the identity of the first group of isolated strains. De Graaf et al. published in 2013 a new method of isolating *P. larvae*. This was the basis for Merrill et al. publishing in 2015 additional *Brevibacillus* and *Paenibacillus* bacteriophage genomes. As suspicion arose, the original 16s rRNA confirmatory tests of the original strains were re-evaluated. Although the top BLAST matches on the NCBI database were *Paenibacillus larvae* subspecies pulvifacens DSM 8442 and *Paenibacillus larvae* subspecies pulvifacens DSM 8443, the majority of the other matches were for *Brevibacillus laterosporus*. It was concluded that the two *P. larvae* matches were misidentified and are currently labeled as “unverified” on NCBI as a result. It followed that the five bacteriophages characterized here are not *Paenibacillus* bacteriophages, but rather *Brevibacillus* bacteriophages.

This new information brings into question the scope of this hypothesis which was to find bacteriophages that can be used to combat against AFB, a contributing factor to CCD. While *P.*

larvae is the causative pathogen of AFB, *B. laterosporus* is a secondary invader associated with *Melissococcus plutonius*, the causative agent of European Foulbrood (EFB). EFB is also a contributing factor of CCD, but to a lesser extent. A more appropriate scope would reflect that *Brevibacillus* bacteriophages can be used to assist in combating EFB. The findings should still be considered relevant to furthering the cause against CCD

The secondary aims of this research remain appropriate considering this new information. *Brevibacillus* bacteriophages still add to the library of genomic information and add relevance to the idea that bacteriophages are useful as potential adjunctive treatments for bacterial infections. These five bacteriophages represent the first fully annotated genomes of its kind, and this research has spawned practical application of bacteriophage therapy on beehives.

This situation has produced an unforeseen avenue of research consisting of establishing more confidence in bacterial identification. The problems of distinguishing between *P. larvae* and *B. laterosporus* have already been the subject of additional endeavors. Topics include the reevaluation of *Paenibacillaceae* taxonomy, reassessing the catalase and 16s rRNA tests, and developing new methods to identify bacteria.

It is significant to appreciate that when Smirnova initially discovered *Paenibacillus* bacteriophages, he was interested in using them to identify the presence of AFB. In a similar way the genomic results of Emery could potentially support this possibility. Thirty-seven of the one hundred gene products were most homologous to *B. laterosporus* genes. No other homologue had nearly as many matches to the gene products of Emery. With the exception of the "unverified" *P. larvae* ssp. Pulvifacens 8442 and 8443, the genome annotation of Emery would have been the first evidence suggesting a misidentification of the bacterial host had occurred.

As new methods and technology come forth, our ability to provide an accurate taxonomy will possibly need to change. From describing colony and plaque morphology to classifications based on microscopy, organizing these entities was effective for its time. Now with the capabilities of gene sequencing, some of the current taxonomic hierarchy may need to be reevaluated. Entities compared at the genomic level will demonstrate a more subtle variation between all levels of taxonomic hierarchy with many crossovers in genes between entities of different classifications. New ways of describing and sorting all entities will provide seemingly endless ways of making sense of the biological world.

Initially, the new information that required this research to submit for a change in the published data from *Paenibacillus* Bacteriophages to *Brevibacillus* Bacteriophages came with concerns. In the end, the integrity of this data required that the change be made, and it highlights the ever-present need to ensure an accurate report of results. This change adds to an already long list of similar changes made in the history of taxonomy, including the bacterial classification changes described earlier from the *Bacillus* genus to the *Paenibacillus* genus and the elimination of the *Paenibacillus* subspecies classification. This situation further highlights the scientific shortcomings of using taxonomy in general to describe groups that actually do not exist in nature. In other words, we use these groups to make better sense of the world, but it will never perfectly fit what actually is going on in these biological entities.

Genomic data now offers the possibility to reach the taxonomic limits of grouping biological entities. In other words, a biological entity can now be placed in a group as specific as a change in a single base pair. For now, there is no realistic way of further classifying beyond a single base pair difference. Perhaps in the future, if warranted, epigenic differences could provide even further distinction between entities that have no differences in base pair sequences.

Genomic studies will continue to provide a pool of opportunities to explore the power of biological information contained in endless combinations of nucleic acids.

TABLES AND FIGURES

TABLE 01 – *Brevibacillus laterosporus* isolated from Utah County honey

Strain	Location in Utah	Color	Form	Margin	Elevation	KOH	H_2O_2	16s
BL 01	Orem	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BL 02	Orem	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BL 03	Farmington	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BL 04	Orem	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BL 05	Orem	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BL 06	Orem	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BL 07	Farmington	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BB 08	Provo	Yellow	Circular	Entire	Convex	+/-	-	*B. brevis*
BL 09	Midway	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BL 10	Midway	Yellow	Irregular	Entire	Convex	+/-	-	*B. laterosporus*
BL 11	Highland	Bright Yellow	Circular	Entire	Pulvinate			
BL 12	Highland	Yellow	Irregular	Entire	Convex	+/-	-	
BL 13	Salt Lake City	Yellow	Irregular	Entire	Convex	+/-	-	
BL 14		Yellow	Irregular	Entire	Convex			
BL 15	Orem	Pale	Punctiform	Entire	Flat			
BL 16	Orem	Pale	Punctiform	Entire	Flat			

TABLE 02 – Genome Summary of Five *Brevibacillus* Bacteriophage

Phage	GenBank Accession Number	Length (base pairs)	Number of Genes	G/C Content
Jimmer1	KC595515	54,312	100	38.11%
Jimmer2	KC595514	54,312	100	38.10%
Emery	KC595516	58,572	100	41.44%
Abouo	KC595517	45,552	92	39.16%
Davies	KC595518	45,798	93	39.16%

TABLE 03 – Detailed List of Jimmer1 Genes

Gene	Start Site	Stop Site	Molecular mass of protein (kDa)	Function	Homologue	E - Value
1	26	463	16.11	Terminase small subunit G1p	Bacillus clausii KSM-K16	0.0e0
2	450	1706	48.41	Pbsx family phage terminase large subunit	Paenibacillus mucilaginosus 3016	0.0e0
3	2523	1786	28.53			
4	2612	4060	55.32	Phage portal protein, SPP1 family	Clostridium botulinum C str. Eklund	0.0e0
5	4057	5100	40.59	Phage putative head morphogenesis protein, SPP1 gp7 family	Paenibacillus larvae subsp. larvae BRL-230010	0.0e0
6	5175	5810	23.40	Phage minor structural protein GP20	Paenibacillus larvae subsp. larvae BRL-230010	0.0e0
6	5175	5810	23.40	Phage minor structural protein GP20	Paenibacillus larvae subsp. larvae BRL-230010	0.0e0
7	5827	6198	12.88	Hypothetical protein Plarl_06935	Paenibacillus larvae subsp. larvae BRL-230010	0.0e0
8	6215	7255	38.69	Phage protein	Paenibacillus larvae subsp. larvae BRL-230010	0.0e0
9	7309	7473	6.10	hypothetical protein Plarl_06945	Paenibacillus larvae subsp. larvae BRL-230010	1.4E-4
10	7473	7832	13.21	putative phage protein	Paenibacillus alvei DSM 29	7.4E-43
11	7826	8188	13.46	hypothetical protein Desde_1086	Desulfitobacterium dehalogenans ATCC 51507	2.2E-43
12	8188	8694	19.56	hypothetical protein PAV_11c00660	Paenibacillus alvei DSM 29	0.0E0
13	8681	9124	7.59	helix-turn-helix domain-containing protein	Desulfotomaculum acetoxidans DSM 771	3.0E-9
14	9108	9284	6.77	hypothetical protein PAV_11c00640	Paenibacillus alvei DSM 29	3.4E-8

15	9286	10599	47.65	phage tail sheath protein	Paenibacillus alvei DSM 29	0.0E0
16	10600	11064	17.70	core tail protein	Clostridium botulinum F str. Langeland	0.0E0
17	12177	11653	19.49	Cro/CI family transcriptional regulator	Streptococcus pyogenes MGAS10394	4.8E-17
18	12334	12567	8.13	hypothetical protein WG8_0646	Paenibacillus sp. Aloe-11	2.2E-12
19	12580	13392	30.55	Prophage antirepressor	Eubacterium siraeum 70/3	0.0E0
20	13519	13848	10.92	hypothetical protein Dred_2594	Desulfotomaculum reducens MI-1	0.24
21	13932	14387	15.80	Phage XkdN-like protein	Desulfosporosinus youngiae DSM 17734	0.0E0
22	14614	15174	20.24	hypothetical protein CLD_2458	Clostridium botulinum B1 str. Okra	0.06
23	15229	17262	76.09	hypothetical protein Plarl_07000	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
24	17255	17932	25.37	uncharacterized protein PPOP_1629	Paenibacillus popilliae ATCC 14706	0.0E0
25	17947	18915	36.83	hypothetical protein Plarl_13404	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
26	18920	19279	13.26	Protein of unknown function (DUF2577)	Desulfosporosinus youngiae DSM 17734	3.9E-31
27	19276	19674	15.11	Protein of unknown function (DUF2634)	Desulfosporosinus youngiae DSM 17734	2.4E-35
28	19671	20741	39.35	putative phage protein	Paenibacillus alvei DSM 29	0.0E0
29	20734	21408	26.29	phage-like element PBSX protein	Paenibacillus sp. JC66	0.0E0
30	21395	21673	10.12	hypothetical protein Desde_1343	Desulfitobacterium dehalogenans ATCC 51507	1.4E-4
31	21677	21967	10.62	hypothetical protein PDENDC454_03740, partial	Paenibacillus dendritiformis C454	2.3E-7

32	21977	23077	41.68	flagellin domain-containing protein	Halanaerobium hydrogeniformans	0.32
33	23092	23439	13.26	WXG repeat protein	Saccharomonospora azurea NA-128	3.77
34	23439	23552	4.26	hypothetical protein HMPREF1025_01333	Lachnospiraceae bacterium 3_1_46FAA	4.5E-3
35	23645	23911	10.28	hypothetical protein BRLA_c28520	Brevibacillus laterosporus LMG 15441	1.4E-39
36	23911	24186	9.87	hypothetical protein BRLA_c21440	Brevibacillus laterosporus LMG 15441	4.5E-19
37	24158	24784	22.93	mannosyl-glycoendo-beta-N-acetylglucosaminidase family protein	Brevibacillus laterosporus GI-9	0.0E0
38	24759	24884	5.12			
39	24881	25852	37.55	hypothetical protein ABC3128	Bacillus clausii KSM-K16	0.0E0
40	25872	26030				
41	26234	27376	38.32	secreted peptidase	Streptomyces hygroscopicus subsp. jinggangensis 5008	2.1E-22
42	27333	27527	7.49	thaxtomin synthetase B	Streptomyces turgidiscabies Car8	9.84
43	27764	27886		hypothetical protein CC1G_04490	Coprinopsis cinerea okayama7#130	13.23
44	29184	27931	47.17	kelch repeat protein	Brevibacillus laterosporus LMG 15441	0.0E0
45	30333	29668	25.25	multi-sensor signal transduction histidine kinase	Oscillatoria sp. PCC 6506	0.60
46	30444	30626	7.47	hypothetical protein BLGI_842	Brevibacillus laterosporus GI-9	3.5E-16
47	30752	31093	13.50	yolD-like family protein	Brevibacillus laterosporus GI-9	0.0E0
48	32774	31203	61.13	site-specific recombinase	Geobacillus kaustophilus HTA426	0.0E0
49	33299	33883	22.16	hypothetical protein BRLA_c05180	Brevibacillus laterosporus LMG 15441	8.4E-45

50	33946	34593	24.55	hypothetical protein BRLA_c05170	Brevibacillus laterosporus LMG 15441	3.1E-36
51	34813	35184	13.43	phage element (ICEBs1)transcriptional regulator (Xre family) protein	Bacillus amyloliquefaciens subsp. amyloliquefaciens DC-12	1.5E-15
52	35465	35244	4.98	ZYRO0G00484p	Zygosaccharomyces rouxii	0.10
53	35460	35573		putative secreted protein	Serratia odorifera 4Rx13	7.24
54	35699	35911	5.62	GumA	uncultured bacterium	4.31
55	35933	36298	13.73	hypothetical protein HMPREF0987_01484	Lachnospiraceae bacterium 9_1_43BFAA	2.0E-15
56	36977	36336	23.93	putative phage repressor	Clostridium difficile ATCC 43255	4.9E-20
57	37125	37367	9.12	transcriptional regulator, XRE family	Desulfotomaculum nigrificans DSM 574	3.9E-13
58	37645	37421	8.49	hypothetical protein SEEM42N_00690	Salmonella enterica subsp. enterica serovar Montevideo str. 42N	0.11
59	37958	37638	11.76	helix-turn-helix domain-containing protein	Desulfotomaculum kuznetsovii DSM 6115	9.9E-21
60	38115	38330	7.59	helix-turn-helix domain-containing protein	Desulfotomaculum acetoxidans DSM 771	3.0E-9
61	38534	38352	6.88	hypothetical protein CKL_2011	Clostridium kluyveri DSM 555	4.4E-21
62	38738	38568	6.07	hypothetical protein Plarl_22353	Paenibacillus larvae subsp. larvae BRL-230010	1.7E-11
63	38852	39004	5.68	hypothetical protein CKL_2010	Clostridium kluyveri DSM 555]	4.1E-6
64	39045	39539	16.56	transcriptional repressor	Bacillus vallismortis DV1-F-3	1.2E-8
65	39523	39708	4.71	hypothetical protein	Trichomonas vaginalis G3	4.01

66	39705	39553	9.02	XRE family transcriptional regulator	Paenibacillus polymyxa SC2	2.5E-27
67	39967	40179	8.22	hypothetical protein BBR47_29000	Brevibacillus brevis NBRC 100599	0.05
68	40169	40342	6.87	C2H2 transcription factor	Beauveria bassiana ARSEF 2860	4.98
69	40391	40639	9.82	chromosome segregation ATPase	Thermosphaera aggregans DSM 11486	0.03
70	40623	40880	10.04	hypothetical protein	Plasmodium berghei strain ANKA	0.02
71	40877	41362	18.10	gp157-like protein	Deep-sea thermophilic phage D6E	1.5E-35
72	41373	41981	22.75	hypothetical protein PAV_5c00050	Paenibacillus alvei DSM 29	0.0E0
73	41974	42387	15.35	single-stranded DNA-binding protein	Brevibacillus laterosporus LMG 15441	0.0E0
74	42368	48467	13.17	hypothetical protein BBR47_35660	Brevibacillus brevis NBRC 100599	5.6E-45
75	42762	43787	39.39	putative prophage replication protein O	Paenibacillus polymyxa M1	6.2E-41
76	43791	44717	35.02	primosomal protein DnaI	Paenibacillus dendritiformis C454	0.0E0
77	44701	44979	11.16	CCR4-Not complex component	Coprinopsis cinerea okayama7#130	0.44
78	44992	45726	28.44	hypothetical protein CBCST_07962	Clostridium botulinum C str. Stockholm	0.0E0
79	45730	45939	8.17	unnamed protein product	Tetraodon nigroviridis	0.14
80	45908	46402	19.50	hypothetical protein IGO_05662	Bacillus cereus HuB5-5	1.6E-18
81	46390	46800	16.08	hypothetical protein CmalA3_01914	Carnobacterium maltaromaticum ATCC 35586	1.3E-19
82	46797	47348	20.23	recombination protein U	Bacillus cereus BAG3X2-2	0.0E0
83	47349	47624	10.35	hypothetical protein PDENDC454_04229	Paenibacillus dendritiformis C454	1.8E-14

84	47728	48384	24.90	dUTPase	Geobacillus thermoglucosidasius C56-YS93	0.0E0
85	48381	48467	4.73			
86	48467	49057	23.51	Site-specific DNA methylase	Bacillus subtilis BSn5	0.0E0
87	49106	49327	7.01	hypothetical protein NCAS_0F02210	Naumovozyma castellii CBS 4309	10.37
88	49368	49991	24.30	hypothetical protein BCAH820_4401	Bacillus cereus AH820	0.0E0
89	50016	50309	11.45	hypothetical protein PMI05_01596	Brevibacillus sp. BC25	3.0E-10
90	50316	50498	7.12	hypothetical protein CC1G_09777	Coprinopsis cinerea okayama7#130	1.00
91	50533	50814	11.02	hypothetical protein BATR1942_07635	Bacillus atrophaeus 1942	2.2E-41
92	50798	51097	9.78	hypothetical protein BCQ_PT51	Bacillus cereus Q1	0.25
93	51129	51782	26.58	hypothetical protein PAV_1c09130	Paenibacillus alvei DSM 29	3.1E-26
94	51779	52102	12.46	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	Phaeobacter gallaeciensis DSM 17395	0.02
95	52114	52356	9.34	hypothetical protein BBR47_35560	Brevibacillus brevis NBRC 100599	1.1E-9
96	52335	52604	10.46	putative MarR family regulatory protein	Pseudomonas fluorescens SBW25	3.92
97	52695	53204	19.76	RNA polymerase, sigma-24 subunit, ECF subfamily protein	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
98	53277	53615	12.67	hypothetical protein Spirs_2785	Spirochaeta smaragdinae DSM 11293	8.9E-37
99	53688	54008	12.68	hypothetical protein BRLA_c22590	Brevibacillus laterosporus LMG 15441	2.1E-25
100	54042	54296	9.18	hypothetical protein bcere0002_54360	Bacillus cereus ATCC 10876	1.95

TABLE 04 – Detailed List of Jimmer2 Genes

Gene	Start Site	Stop Site	Molecular mass of protein (kDa)	Function	Homologue	E -Value
1	26	463	16.11	terminase small subunit G1p	Bacillus clausii KSM-K16	0.0E0
2	450	1706	48.41	pbsx family phage terminase large subunit	Paenibacillus mucilaginosus 3016	0.0E0
3	2523	1786	28.53			
4	2612	4060	55.32	phage portal protein, SPP1 family	Clostridium botulinum C str. Eklund	0.0E0
5	4057	5100	40.59	phage putative head morphogenesis protein, SPP1 gp7 family	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
6	5175	5810	23.40	phage minor structural GP20	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
7	5827	6198	12.88	hypothetical protein Plarl_06935	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
8	6215	7255	38.69	phage protein	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
9	7309	7473	6.10	hypothetical protein Plarl_06945	Paenibacillus larvae subsp. larvae BRL-230010	1.4E-4
10	7473	7832	13.21	putative phage protein	Paenibacillus alvei DSM 29	7.3E-43
11	7826	8188	13.46	hypothetical protein Desde_1086	Desulfitobacterium dehalogenans ATCC 51507	2.2E-43
12	8188	8694	19.56	hypothetical protein PAV_11c00660	Paenibacillus alvei DSM 29	0.0E0
13	8681	9124	15.83	hypothetical protein DesyoDRAFT_1114	Desulfosporosinus youngiae DSM 17734	5.3 E-40
14	9108	9284	6.77	hypothetical protein PAV_11c00640	Paenibacillus alvei DSM 29	3.4E-8
15	9286	10599	47.65	phage tail sheath protein	Paenibacillus alvei DSM 29	0.0E0

16	10600	11064	17.70	core tail protein	Clostridium botulinum F str. Langeland	0.0E0
17	12177	11653	19.49	Cro/CI family transcriptional regulator	Streptococcus pyogenes MGAS10394	4.8E-17
18	12334	12567	8.13	hypothetical protein WG8_0646	Paenibacillus sp. Aloe-11	2.2E-12
19	12580	13392	30.55	Prophage antirepressor	Eubacterium siraeum 70/3	0.0E0
20	13519	13848	10.92			
21	13932	13848	15.80	Phage XkdN-like protein	Desulfosporosinus youngiae DSM 17734	0.0E0
22	14614	15174	20.24	hypothetical protein CLD_2458	Clostridium botulinum B1 str. Okra	0.06
23	15229	17262	76.09	hypothetical protein Plarl_07000	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
24	17255	17932	25.37	LysM domain-containing protein	Desulfitobacterium dehalogenans ATCC 51507	0.0E0
25	17947	18915	36.92	hypothetical protein Plarl_13404	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
26	18920	19279	13.26	Protein of unknown function (DUF2577)	Desulfosporosinus youngiae DSM 17734	3.8E-31
27	19247	19674	15.11	Protein of unknown function (DUF2634)	Desulfosporosinus youngiae DSM 17734	2.3E-35
28	19671	20741	39.52	putative phage protein	Paenibacillus alvei DSM 29	0.0E0
29	20734	21408	26.29	phage-like element PBSX protein	Paenibacillus sp. JC66	0.0E0
30	21395	21673	10.12	hypothetical protein Desde_1343	Desulfitobacterium dehalogenans ATCC 51507	1.4E-4
31	21677	21967	10.62	hypothetical protein PDENDC454_03740, partial	Paenibacillus dendritiformis C454	2.3E-7
32	21977	23077	41.69	flagellin domain-containing protein	Halanaerobium hydrogeniformans	0.36
33	23092	23439	13.26	WXG repeat protein	Saccharomonospora azurea NA-128	3.72

34	23439	23552	4.26	hypothetical protein HMPREF1025_01333	Lachnospiraceae bacterium 3_1_46FAA	4.4E-3
35	23645	23911	10.28	hypothetical protein BRLA_c28520	Brevibacillus laterosporus LMG 15441	1.4 E-39
36	23911	24186	9.87	hypothetical protein BRLA_c21440	Brevibacillus laterosporus LMG 15441	4.4E-19
37	24158	24784	22.93	mannosyl-glycoendo-beta-N-acetylglucosaminidase family protein	Brevibacillus laterosporus GI-9	0.0E0
38	24759	24884	5.12			
39	24881	25852	37.55	hypothetical protein ABC3128	Bacillus clausii KSM-K16	0.0E0
40	25872	26030				
41	26234	27376	38.32	secreted peptidase	Streptomyces hygroscopicus subsp. jinggangensis 5008	2.1 E-22
42	27333	27527	7.49	thaxtomin synthetase B	Streptomyces turgidiscabies Car8	9.73
43	27764	27886				
44	29184	27931	47.17	kelch repeat protein	Brevibacillus laterosporus LMG 15441	0.0E0
45	30333	29668	25.25	multi-sensor signal transduction histidine kinase	Oscillatoria sp. PCC 6506	0.59
46	30444	30626	7.47	hypothetical protein BLGI_842	Brevibacillus laterosporus GI-9	3.5E-16
47	30752	31093	13.50	yolD-like family protein	Brevibacillus laterosporus GI-9	0.0E0
48	32774	31203	61.13	site-specific recombinase	Geobacillus kaustophilus HTA426	0.0E0
49	33299	33883	22.16	hypothetical protein BRLA_c05180	Brevibacillus laterosporus LMG 15441	8.4E-45
50	33946	34593	24.55	hypothetical protein BRLA_c05170	Brevibacillus laterosporus LMG 15441	3.1 E-36
51	34813	35184	13.43	XRE family transcriptional regulator	Sporosarcina newyorkensis 2681	6.4 E-15

52	35465	35244	4.98			
53	35460	35573				
54	35699	35911	5.62	GumA	uncultured bacterium	4.29
55	35699	36298	13.73	hypothetical protein HMPREF0987_01484	Lachnospiraceae bacterium 9_1_43BFAA	1.9E-15
56	36977	36336	23.93	transcriptional regulator, XRE family	Alicyclobacillus acidocaldarius LAA1	4.0E-20
57	37125	37367	9.12	transcriptional regulator, XRE family	Desulfotomaculum nigrificans DSM 574	3.8E-13
58	37645	37421	8.49	hypothetical protein SEEM42N_00690	Salmonella enterica subsp. enterica serovar Montevideo str. 42N	0.11
59	37958	37638	11.76	helix-turn-helix domain-containing protein	Desulfotomaculum kuznetsovii DSM 6115	9.7 E-21
60	38115	38330	7.59	helix-turn-helix domain-containing protein	Desulfotomaculum acetoxidans DSM 771	2.7E-9
61	38534	38352	6.88	hypothetical protein CKL_2011	Clostridium kluyveri DSM 555	4.9E-21
62	38738	38568	6.07	hypothetical protein Plarl_22353	Paenibacillus larvae subsp. larvae BRL-230010	1.7E-11
63	38852	39004	5.68	hypothetical protein CKL_2010	Clostridium kluyveri DSM 555	4.0E-6
64	39045	39539	16.56	transcriptional repressor	Bacillus vallismortis DV1-F-3	1.2E-8
65	39523	39708	4.71	hypothetical protein	Trichomonas vaginalis G3	3.99
66	39705	39953	9.02	XRE family transcriptional regulator	Paenibacillus polymyxa SC2	2.4E-27
67	39967	40179	8.22	hypothetical protein BBR47_29000	Brevibacillus brevis NBRC 100599	0.05
68	40169	40342	6.87	C2H2 transcription factor	Beauveria bassiana ARSEF 2860	4.92
69	40391	40639	9.82	chromosome segregation ATPase	Thermosphaera aggregans DSM 11486	0.03

70	40623	40880	10.04	hypothetical protein	Plasmodium berghei strain ANKA	0.02
71	40877	41362	18.10	gp157-like protein	Deep-sea thermophilic phage D6E	1.5 E-35
72	41373	41981	22.75	hypothetical protein PAV_5c00050	Paenibacillus alvei DSM 29	0.0E0
73	41974	42387	15.35	single-stranded DNA-binding protein	Brevibacillus laterosporus LMG 15441	0.0E0
74	42368	42748	13.17	hypothetical protein BBR47_35660	Brevibacillus brevis NBRC 100599	7.0E-45
75	42762	43787	39.39	putative prophage replication protein O	Paenibacillus polymyxa M1	6.1E-41
76	43791	44717	35.02	primosomal protein DnaI	Paenibacillus dendritiformis C454	0.0E0
77	44701	44979	11.16	CCR4-Not complex component	Coprinopsis cinerea okayama7#130	0.43
78	44992	45726	28.44	hypothetical protein CBCST_07962	Clostridium botulinum C str. Stockholm	0.0E0
79	45730	45939	8.17	unnamed protein product	Tetraodon nigroviridis	0.14
80	45908	46402	19.50	hypothetical protein IGO_05662	Bacillus cereus HuB5-5	1.6 E-18
81	46390	46800	16.08	hypothetical protein CmalA3_01914	Carnobacterium maltaromaticum ATCC 35586	1.3 E-19
82	46797	47348	20.23	recombination protein U	Bacillus cereus BAG3X2-2	0.0E0
83	47349	47624	10.35	hypothetical protein PDENDC454_04229	Paenibacillus dendritiformis C454	1.8E-14
84	47728	48384	24.90	dUTPase	Geobacillus thermoglucosidasius C56-YS93	0.0E0
85	48381	48467	4.73			
86	48467	49057	23.51	Site-specific DNA methylase	Bacillus subtilis BSn5	0.0E0
87	49106	49327	7.01			
88	49368	49991	24.30	hypothetical protein BCAH820_4401	Bacillus cereus AH820	0.0E0

89	50016	50309	11.45	hypothetical protein PMI05_01596	Brevibacillus sp. BC25	3.0E-10
90	50316	50498	7.12	hypothetical protein CC1G_09777	Coprinopsis cinerea okayama7#130	0.99
91	50533	50814	11.02	hypothetical protein BATR1942_07635	Bacillus atrophaeus 1942	2.2E-41
92	50798	51097	9.78			
93	51129	51782	26.58	hypothetical protein PAV_1c09130	Paenibacillus alvei DSM 29	3.1E-26
94	51779	52102	12.46	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	Phaeobacter gallaeciensis DSM 17395	0.02
95	52114	52356	9.34	hypothetical protein BBR47_35560	Brevibacillus brevis NBRC 100599	1.1E-9
96	52335	52604	10.46	MarR family transcriptional regulator	Tistrella mobilis KA081020-065	5.09
97	52695	53204	19.76	RNA polymerase, sigma-24 subunit, ECF subfamily protein	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
98	53277	53615	12.67	hypothetical protein Spirs_2785	Spirochaeta smaragdinae DSM 11293	8.8E-37
99	53688	54008	12.25	hypothetical protein BRLA_c22590	Brevibacillus laterosporus LMG 15441	2.1E-25
100	54042	54296	9.18	hypothetical protein bcere0002_54360	Bacillus cereus ATCC 10876	1.93

TABLE 05 – Detailed List of Emery Genes

Gene	Start Site	Stop Site	Molecular mass of protein (kDa)	Function	Homologue	E -Value
1	36	908	32.67	integrase family protein	Paenibacillus elgii B69	0.0E0
2	886	1617	24.10	hypothetical protein PelgB_38212	Paenibacillus elgii B69	0.0E0
3	1614	3158	58.68	hypothetical protein PelgB_38207	Paenibacillus elgii B69	0.0E0
4	3171	4724	58.50	hypothetical protein BRLA_c33960	Brevibacillus laterosporus LMG 15441	0.0E0
5	4717	5349	17.30	hypothetical protein BRLA_c33960	Brevibacillus laterosporus LMG 15441	0.0E0
6	5357	8377	112.18	hypothetical protein BRLA_c33980	Brevibacillus laterosporus LMG 15441	0.0E0
7	8381	8752	13.54	hypothetical protein BRLA_c33990	Brevibacillus laterosporus LMG 15441	6.6 E-35
8	8721	9248	20.54	hypothetical protein BRLA_c34000	Brevibacillus laterosporus LMG 15441	0.0E0
9	9248	9652	14.83	hypothetical protein BRLA_c34010	Brevibacillus laterosporus LMG 15441	0.0E0
10	9652	10209	21.14	hypothetical protein BRLA_c34020	Brevibacillus laterosporus LMG 15441	0.0E0
11	10212	10742	19.63	hypothetical protein PelgB_37222	Paenibacillus elgii B69	0.0E0
12	10748	12283	56.97	hypothetical protein BRLA_c34040	Brevibacillus laterosporus LMG 15441	0.0E0
13	12283	12717	15.94	hypothetical protein BRLA_c34050	Brevibacillus laterosporus LMG 15441	0.0E0
14	12730	13077	12.76	hypothetical protein BRLA_c34060	Brevibacillus laterosporus LMG 15441	0.0E0
15	13086	13202	4.28	hypothetical protein BRLA_c34070	Brevibacillus laterosporus LMG 15441	7.6E-15

16	13220	16219	108.04	hypothetical protein BRLA_c34080	Brevibacillus laterosporus LMG 15441	0.0E0
17	16219	16815	22.26	hypothetical protein BRLA_c34090	Brevibacillus laterosporus LMG 15441	0.0E0
18	42.84	16808	17950	hypothetical protein BRLA_c34100	Brevibacillus laterosporus LMG 15441	0.0E0
19	17947	18297	13.06	hypothetical protein BRLA_c34110	Brevibacillus laterosporus LMG 15441	0.0E0
20	18294	18680	14.34	hypothetical protein BRLA_c34120	Brevibacillus laterosporus LMG 15441	0.0E0
21	18697	19818	41.55	hypothetical protein BRLA_c34130	Brevibacillus laterosporus LMG 15441	0.0E0
22	19828	20406	21.50	hypothetical protein BRLA_c34140	Brevibacillus laterosporus LMG 15441	0.0E0
23	20391	20714	11.82	hypothetical protein BRLA_c34150	Brevibacillus laterosporus LMG 15441	0.0E0
24	20711	21109	15.10	hypothetical protein BRLA_c34160	Brevibacillus laterosporus LMG 15441	0.0E0
25	21126	23075	72.54	hypothetical protein BRLA_c34170	Brevibacillus laterosporus LMG 15441	0.0E0
26	23098	23343	9.46	hypothetical protein BRLA_c34190	Brevibacillus laterosporus LMG 15441	2.6E-22
27	23343	23465	4.46	hypothetical protein EAT1b_0052	Exiguobacterium sp. AT1b	1.4E-4
28	23452	23655	4.86	hypothetical protein BLGI_5021	Brevibacillus laterosporus GI-9	2.2 E-28
29	23749	24015	10.26	hypothetical protein BRLA_c28520	Brevibacillus laterosporus LMG 15441	0.0E0
30	24018	24272	9.09	hypothetical protein BRLA_c21440	Brevibacillus laterosporus LMG 15441	1.4 E-19
31	24269	25447	43.40	hypothetical protein BRLA_c28530	Brevibacillus laterosporus LMG 15441	0.0E0

32	25573	25451	4.13			
33	25551	26234	26.63	adenine-specific methyltransferase	Paenibacillus alvei DSM 29	0.0E0
34	26306	26656	13.59	hypothetical protein HMPREF1013_05350	Bacillus sp. 2_A_57_CT2	0.0E0
35	26978	26742	9.10	unnamed protein product	Blastocystis hominis	2.52
36	27700	27011	25.16	hypothetical protein PpeoK3_15141	Paenibacillus peoriae KCTC 3763	2.8E-22
37	28070	27876		hypothetical protein SHJG_1456	Streptomyces hygroscopicus subsp. jinggangensis 5008	4.19
38	28345	28345		hypothetical protein PTD2_21262	Pseudoalteromonas tunicata D2	4.26
39	28546	28671	4.54	glycoside hydrolase family 3 protein	Petrotoga mobilis SJ95	2.19
40	29020	28796	8.62	hypothetical protein BRLA_c34310	Brevibacillus laterosporus LMG 15441	6.7E-31
41	29239	29114				
42	29457	29329	4.88			
43	29493	30197	26.10	hypothetical protein BRLA_c34320	Brevibacillus laterosporus LMG 15441	0.0E0
44	30203	30340	5.26	predicted protein	Naegleria gruberi	1.11
45	30625	30344	8.97	Ricin B lectin	Streptomyces griseus XylebKG-1	1.00
46	31828	31858	45.15	tyrosine recombinase XerC	Paenibacillus mucilaginosus 3016	0.0E0
47	32082	31858	8.17	putative transcriptional regulator	Paenibacillus larvae subsp. larvae BRL-230010	9.1E-20
48	32149	33042	30.01	hypothetical protein BRLA_c33480	Brevibacillus laterosporus LMG 15441	0.0E0
49	33160	33243		hypothetical protein WG8_4645	Paenibacillus sp. Aloe-11	1.1E-3
50	33666	33583				

51	33811	34476	25.41	hypothetical protein BRLA_c33170	Brevibacillus laterosporus LMG 15441	1.4 E-18
52	34544	35212	25.40	hypothetical protein BRLA_c33170	Brevibacillus laterosporus LMG 15441	4.3 E-36
53	35205	35723	19.23	accessory gene regulator B family protein	Brevibacillus laterosporus GI-9	2.2 E-39
54	35731	35847	4.11	hypothetical protein BRLA_c33190	Brevibacillus laterosporus LMG 15441	7.88
55	35857	36222	13.70	hypothetical protein BRLA_c33200	Brevibacillus laterosporus LMG 15441	2.1 E-13
56	36524	36378	5.85			
57	36798	36568	8.66	XRE family transcriptional regulator	Acetonema longum DSM 6540	1.5E-15
58	36901	38178	50.08	hypothetical protein Plarl_11826	Paenibacillus larvae subsp. larvae BRL-230010	3.0 E-24
59	38253	38384	4.55	rare lipoprotein A	Thiorhodococcus drewsii AZ1	3.11
60	38571	38419	5.21			
61	38549	8731	7.02	activator of middle period transcription	Enterobacteria phage Bp7	6.31
62	39110	38748	14.06	Prophage LambdaBa04, DNA-binding protein	Bacillus cereus BDRD-ST24	5.3E-18
63	39290	39547	9.61	hypothetical protein MUY_01529	Bacillus licheniformis WX-02	5.4 E-11
64	39546	39815	11.18	hypothetical protein BRLA_c33570	Brevibacillus laterosporus LMG 15441	1.5 E-35
65	39892	40197	8.17	hypothetical protein BRLA_c33580	Brevibacillus laterosporus LMG 15441	6.0 E-35
66	40181	40264		hypothetical protein SEVCU121_1963	Staphylococcus warneri VCU121	6.05
67	40335	40538	7.92	hypothetical protein BRLA_c33630	Brevibacillus laterosporus LMG 15441	5.5E-4

68	40551	40862	12.16	hypothetical protein PDENDC454_04189	Paenibacillus dendritiformis C454	7.9 E-14
69	40995	41219	8.77	hypothetical protein MPER_08472	Moniliophthora perniciosa FA553	1.56
70	41262	41504	8.25	AbrB family transcriptional regulator	Caldicellulosiruptor saccharolyticus DSM 8903	2.5E-22
71	41525	41716	7.63	hypothetical protein PAV_4c00490	Paenibacillus alvei DSM 29	8.4E-5
72	41883	42095	8.30	conserved hypothetical protein	Ixodes scapularis	1.42
73	42139	42675	20.33	Phage protein	Bacillus azotoformans LMG 9581	1.2E-7
74	42139	43080	7.53	hypothetical protein Clocel_0758	Clostridium cellulovorans 743B	1.1 E-13
75	43077	44261	43.90	hypothetical protein	Desulfotomaculum ruminis DSM 2154	0.0E0
76	44524	44493	9.02	hypothetical protein Dalk_2302	Desulfatibacillum alkenivorans AK-01	9.15
77	44490	45044	20.17	hypothetical protein Ccel_3065	Clostridium cellulolyticum H10	0.0E0
78	45103	45615	20.28	conserved hypothetical protein	Listeria monocytogenes FSL F2-208	0.01
79	45608	47680	78.71	DNA-directed DNA polymerase	Desulfotomaculum ruminis DSM 2154	0.0E0
80	47696	47968	10.13	carbonic anhydrase	Vibrio sinaloensis DSM 21326	0.56
81	47958	48302	13.63	conserved protein of DIM6/NTAB family	Pantoea sp. YR343	1.06
82	48299	48478	6.77	hypothetical protein bmyco0002_56490	Bacillus mycoides Rock1-4	1.1 E-16
83	48594	48695				
84	48705	48995	10.12	pyridoxamine 5'-phosphate oxidase-related protein	Mycobacterium ulcerans Agy99	1.25
85	49019	49411	14.76	hypothetical protein IC1_01171	Bacillus cereus VD022	0.04
86	49408	49662	9.47	hypothetical protein PlarlB_06280	Paenibacillus larvae subsp. larvae B-3650	0.32

87	49659	49970	11.63	helix-turn-helix protein	Faecalibacterium cf. prausnitzii KLE1255	2.8 E-32
88	49967	50203	9.61	hypothetical protein bthur0004_55620	Bacillus thuringiensis serovar sotto str. T04001	3.7 E-16
89	50200	50523	11.90	PREDICTED: laminin subunit beta-1	Otolemur garnettii	0.12
90	50520	50684	6.29	hypothetical protein HMPREF1068_03926	Bacteroides nordii CL02T12C05	1.72
91	50681	50941	10.11	hypothetical protein Tsp_01148	Trichinella spiralis	1.23
92	50952	52169	41.48	RNA polymerase sigma factor, sigma-70 family	Paenibacillus alvei DSM 29	2.5E-44
93	52223	53185	36.82	ATPase AAA	Acetohalobium arabaticum DSM 5501	0.0E0
94	53182	55554	91.46	virulence-associated E family protein	Clostridium cellulolyticum H10	0.0E0
95	55623	55109				
96	55829	56104	10.39	VRR-NUC domain-containing protein	Desulfotomaculum ruminis DSM 2154	1.9E-26
97	56101	57486	52.51	SNF2-like protein	Clostridium cellulolyticum H10	0.0E0
98	57487	57723	8.89			
99	57751	58287	20.94	hypothetical protein PaelaDRAFT_2391	Paenibacillus lactis 154	1.3 E-38
100	58400	58504				

TABLE 06 – Detailed List of Abouo Genes

Gene	Start Site	Stop Site	Molecular mass of protein (kDa)	Function	Homologue	E -Value
1	26	463	16.17	terminase small subunit G1p	Bacillus clausii KSM-K16	0.0E0
2	450	1706	48.41	pbsx family phage terminase large subunit	Paenibacillus mucilaginosus 3016	0.0E0
3	1719	3176	56.35	phage portal protein, SPP1 family	Desulfosporosinus youngiae DSM 17734	0.0E0
4	3173	4213	40.28	phage putative head morphogenesis protein, SPP1 gp7 family	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
5	4292	4918	22.64	phage minor structural GP20	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
6	4909	5247	11.84	prophage protein	Lactobacillus pentosus MP-10	5.1E-27
7	5263	6288	38.25	phage protein	Enterococcus sp. C1	0.0E0
8	6303	6605	9.29	hypothetical protein nfa15290	Nocardia farcinica IFM 10152	1.9E-5
9	6586	6963	13.57	Phage QLRG family, putative DNA packaging protein	Desulfosporosinus youngiae DSM 17734	1.9E-34
10	6957	7319	13.52	hypothetical protein Desde_1086	Desulfitobacterium dehalogenans ATCC 51507	2.4E-42
11	7319	7750	16.90	hypothetical protein PaelaDRAFT_2404	Paenibacillus lactis 154	0.0E0
12	7737	8171	17.89	hypothetical protein DesyoDRAFT_1114	Desulfosporosinus youngiae DSM 17734	5.6E-45
13	8164	8340	6.77	hypothetical protein PAV_11c00640	Paenibacillus alvei DSM 29	3.4E-8
14	8342	9655	47.61	phage tail sheath protein	Paenibacillus alvei DSM 29	0.0E0
15	9656	10117	17.13	core tail protein	Clostridium botulinum A2 str. Kyoto	0.0E0
16	10303	10404				

17	10650	11060	15.37	Phage XkdN-like protein	Desulfitobacterium dehalogenans ATCC 51507	0.0E0
18	11120	11230				
19	11272	11991	25.63	extracellular solute-binding protein	Nitrosococcus watsonii C-113	0.11
20	12032	14086	75.49	hypothetical protein Plarl_07000	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
21	14099	14380	10.58	GK18909	Drosophila willistoni	0.30
22	14373	15050	25.45	uncharacterized protein PPOP_1629	Paenibacillus popilliae ATCC 14706	0.0E0
23	15065	16054	37.67	hypothetical protein Plarl_13404	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
24	16038	16397	13.23	Protein of unknown function (DUF2577)	Desulfosporosinus youngiae DSM 17734	2.4E-30
25	16394	16576	7.20	Polyprotein	Hepatitis C virus	0.86
26	16573	16971	15.10	Protein of unknown function (DUF2634)	Desulfosporosinus youngiae DSM 17734	2.1E-35
27	16968	18038	39.46	putative phage protein	Paenibacillus alvei DSM 29	0.0E0
28	18031	18705	26.20	phage-like element PBSX protein	Paenibacillus sp. JC66	0.0E0
29	18692	18970	10.16	hypothetical protein Desde_1343	Desulfitobacterium dehalogenans ATCC 51507	8.8E-5
30	18974	19789	29.62	Kelch repeat type 1-containing protein	Paenibacillus larvae subsp. larvae B-3650	1.0E-12
31	19804	20394	20.93	hypothetical protein GY4MC1_0642	Geobacillus sp. Y4.1MC1	2.6E-8
32	20412	20801	14.82	hypothetical protein	Geobacillus thermoleovorans CCB_US3_UF5	2.5E-26
33	20805	20939	5.32	hypothetical protein HMPREF9469_05014	Clostridium citroniae WAL-17108	0.16
34	21032	21298	10.28	hypothetical protein BRLA_c28520	Brevibacillus laterosporus LMG 15441	5.1E-40

35	21303	21545	8.21	hypothetical protein BRLA_c34210	Brevibacillus laterosporus LMG 15441	8.0E-31
36	21542	22165	22.64	Exo-glucosaminidase LytG	Brevibacillus laterosporus LMG 15441	0.0E0
37	22590	22378				
38	23360	22674	24.10	hypothetical protein BLGI_3418	Brevibacillus laterosporus GI-9	3.4E-42
39	23376	23507				
40	23629	23970	13.56	yolD-like family protein	Brevibacillus laterosporus GI-9	0.0E0
41	24216	24052	6.33	prophage antirepressor	Bacillus sp. M 2-6	4.6 E-15
42	24404	24228	4.55	CRISPR-associated helicase Cas3	Leptospira noguchii str. 2006001870	12.30
43	25935	25935	45.99	Integrase	Geobacillus sp. Y4.1MC1	0.0E0
44	26243	26911	25.70	putative membrane protein	Brevibacillus laterosporus GI-9	7.5E-41
45	26904	27413	18.89	hypothetical protein BRLA_c33180	Brevibacillus laterosporus LMG 15441	1.0E-37
46	27410	27550	4.67	hypothetical protein BLGI_1765	Brevibacillus laterosporus GI-9	3.4E-6
47	27554	27922	14.03	hypothetical protein BRLA_c33200	Brevibacillus laterosporus LMG 15441	3.5 E-30
48	28181	28363	5.62	GumA	uncultured bacterium	4.08
49	28377	28751	14.01	cell division protein FtsQ	Bacillus thuringiensis MC28	1.1E-16
50	29423	28791	24.04	putative phage repressor	Clostridium difficile ATCC 43255	1.4E-21
51	29580	29822	9.08	hypothetical protein BCAH187_A0631	Bacillus cereus AH187	5.7E-13
52	30017	29844	6.89	hypothetical protein CKL_2011	Clostridium kluyveri DSM 555	3.5 E-21
53	30143	30304	5.83	predicted protein	Lactobacillus crispatus MV-1A-US	0.02
54	30362	30640	10.86	protein of unknown function DUF1156	Desulfitobacterium dichloroeliminans LMG P-21439	4.98

55	30637	30885	9.05	XRE family transcriptional regulator	Paenibacillus polymyxa SC2	1.8E-27
56	30899	31111	7.81	hypothetical protein BBR47_29000	Brevibacillus brevis NBRC 100599	2.1E-5
57	31101	31271	6.80	putative NAD dependent epimerase/dehydratase family protein	Streptomyces hygroscopicus ATCC 53653	6.31
58	31331	31579	9.79	conserved Plasmodium membrane protein, unknown function	Plasmodium falciparum 3D7	0.21
59	31563	31826	10.31	hypothetical protein BRLA_c33620	Brevibacillus laterosporus LMG 15441	0.19
60	31823	32308	18.14	gp157-like protein	Deep-sea thermophilic phage D6E	6.8E-35
61	32319	32909	21.99	hypothetical protein PAV_5c00050	Paenibacillus alvei DSM 29	0.0E0
62	32902	33315	15.09	single-stranded DNA-binding protein	Brevibacillus laterosporus LMG 15441	0.0E0
63	33329	33676	13.12	hypothetical protein BBR47_35660	Brevibacillus brevis NBRC 100599	1.4E-45
64	33696	34757	39.67	putative prophage LambdaCh01, replication protein O	Bacillus methanolicus PB1	7.7E-17
65	34747	35445	27.35	DNA replication protein	Paenibacillus popilliae ATCC 14706	0.0E0
66	35438	35707	10.65	DNA repair protein RecN	Nodularia spumigena CCY9414	0.27
67	35720	36454	28.45	hypothetical protein CBCST_07962	Clostridium botulinum C str. Stockholm	0.0E0
68	36473	36799	12.07	similar to zinc metalloprotease	Leptosphaeria maculans JN3	0.05
69	36877	37287	15.91	hypothetical protein ICU_03857	Bacillus cereus BAG2X1-1	2.9E-23
70	37290	37670	14.79	hypothetical protein CmalA3_01914	Carnobacterium maltaromaticum ATCC 35586	6.8 E-24

71	37667	37906	9.00	hypothetical protein S23_41210	Bradyrhizobium sp. S23321	6.1E-9
72	37919	38449	20.37	recombination protein U	Bacillus cereus BAG3X2-2	0.0E0
73	38450	38725	10.60	hypothetical protein bthur0005_56060	Bacillus thuringiensis serovar pakistani str. T13001	8.8E-12
74	38758	38862				
75	38852	39502	25.56	dUTPase	Geobacillus thermoglucosidasius C56-YS93	4.6E-41
76	39791	39579		hypothetical protein BRLA_c33790	Brevibacillus laterosporus LMG 15441	4.6E-15
77	39857	40327	18.02	phage N-6-adenine methyltransferase	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
78	40386	40661	10.23	catechol dioxygenase, putative	Metarhizium anisopliae ARSEF 23	0.63
79	40658	40912	9.66	RNA polymerase	Cyanophage 9515-10a	0.85
80	40912	41133	7.07	hypothetical protein TBLA_0C06770	Tetrapisispora blattae CBS 6284	1.44
81	41169	41363	6.03	hypothetical protein PelgB_33761	Paenibacillus elgii B69	5.0 E-21
82	41467	41862	15.71	hypothetical protein KSO_07894	Bacillus amyloliquefaciens IT-45	5.7 E-23
83	41877	42359	11.48	hypothetical protein PMI05_01596	Brevibacillus sp. BC25	1.0E-10
84	42177	42359	7.09	hypothetical protein CC1G_09777	Coprinopsis cinerea okayama7#130	0.78
85	42337	42675	13.16	hypothetical protein BATR1942_07635	Bacillus atrophaeus 1942	1.9 E-41
86	42709	43374	27.27	hypothetical protein PAV_1c09130	Paenibacillus alvei DSM 29	2.3E-25
87	43358	43600	9.34	hypothetical protein BBR47_35560	Brevibacillus brevis NBRC 100599	1.1E-9
88	43579	43848	10.43	putative MarR family regulatory protein	Pseudomonas fluorescens SBW25	4.34
89	43937	44446	19.54	RNA polymerase, sigma-24 subunit,	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0

				ECF subfamily protein		
90	44518	44856	12.69	hypothetical protein Spirs_2785	Spirochaeta smaragdinae DSM 11293	5.9 E-37
91	44929	45249	12.47	hypothetical protein BRLA_c22590	Brevibacillus laterosporus LMG 15441	5.1 E-26
92	45283	45534	8.97	hypothetical protein bcere0002_54360	Bacillus cereus ATCC 10876	0.17

TABLE 07 – Detailed List of Davies Genes

Gene	Start Site	Stop Site	Molecular mass of protein (kDa)	Function	Homologue	E -Value
1	39	476	16.17	terminase small subunit G1p	Bacillus clausii KSM-K16	0.0E0
2	463	1719	8.01	pbsx family phage terminase large subunit	Paenibacillus mucilaginosus 3016	0.0E0
3	1732	3189	56.35	phage portal protein, SPP1 family	Desulfosporosinus youngiae DSM 17734	0.0E0
4	3186	4226	40.28	phage putative head morphogenesis protein, SPP1 gp7 family	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
5	4305	4931	22.64	phage minor structural GP20	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
6	4922	5260	11.84	prophage protein	Lactobacillus pentosus MP-10	5.1 E-27
7	5276	6301	38.25	phage protein	Enterococcus sp. C1	0.0E0
8	6316	6618	9.29	hypothetical protein nfa15290	Nocardia farcinica IFM 10152	1.9E-5
9	6611	6976	13.57	Phage QLRG family, putative DNA packaging protein	Desulfosporosinus youngiae DSM 17734	1.7E-34
10	6970	7332	13.52	hypothetical protein Desde_1086	Desulfitobacterium dehalogenans ATCC 51507	2.4E-42
11	7332	7763	16.90	hypothetical protein PaelaDRAFT_2404	Paenibacillus lactis 154	0.0E0
12	7750	8184	17.89	hypothetical protein DesyoDRAFT_1114	Desulfosporosinus youngiae DSM 17734	5.6E-45
13	8177	8353	6.77	hypothetical protein PAV_11c00640	Paenibacillus alvei DSM 29	3.4E-8
14	8355	9668	47.61	phage tail sheath protein	Paenibacillus alvei DSM 29	0.0E0
15	9669	10130	17.13	core tail protein	Clostridium botulinum A2 str. Kyoto	0.0E0
16	10316	10417				

17	10663	11073	15.37	Phage XkdN-like protein	Desulfitobacterium dehalogenans ATCC 51507	0.0E0
18	11133	11243				
19	11285	12004	25.63	extracellular solute-binding protein	Nitrosococcus watsonii C-113	0.11
20	11994	14099	75.49	hypothetical protein Plarl_07000	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
21	14112	14393	10.58	GK18909	Drosophila willistoni	0.30
22	14386	15063	25.45	uncharacterized protein PPOP_1629	Paenibacillus popilliae ATCC 14706	0.0E0
23	15078	16067	37.67	hypothetical protein Plarl_13404	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
24	16051	16410	13.23	Protein of unknown function (DUF2577)	Desulfosporosinus youngiae DSM 17734	2.4E-30
25	16407	16589	7.20	Polyprotein	Hepatitis C virus	0.86
26	16586	16984	15.10	Protein of unknown function (DUF2634)	Desulfosporosinus youngiae DSM 17734	2.1E-35
27	16981	18051	39.46	putative phage protein	Paenibacillus alvei DSM 29	0.0E0
28	18044	18051	26.20	phage-like element PBSX protein	Paenibacillus sp. JC66	0.0E0
29	18705	18983	10.16	hypothetical protein Desde_1343	Desulfitobacterium dehalogenans ATCC 51507	8.8E-5
30	18987	19802	29.64	Kelch repeat type 1-containing protein	Paenibacillus larvae subsp. larvae B-3650	1.0E-12
31	19817	20407	20.93	hypothetical protein GY4MC1_0642	Geobacillus sp. Y4.1MC1	2.6E-8
32	20425	20814	14.82	hypothetical protein	Geobacillus thermoleovorans CCB_US3_UF5	2.5
33	20818	20952	5.32	hypothetical protein HMPREF9469_05014	Clostridium citroniae WAL-17108	0.16
34	21045	21311	10.28	hypothetical protein BRLA_c28520	Brevibacillus laterosporus LMG 15441	5.1 E-40

35	21316	21558	8.21	hypothetical protein BRLA_c34210	Brevibacillus laterosporus LMG 15441	8.0E-31
36	21555	22178	22.64	Exo-glucosaminidase LytG	Brevibacillus laterosporus LMG 15441	0.0E0
37	22603	30891				
38	23373	22687	24.10	hypothetical protein BLGI_3418	Brevibacillus laterosporus GI-9	3.4 E-42
39	23645	23983	13.56	yolD-like family protein	Brevibacillus laterosporus GI-9	0.0E0
40	24229	24065	6.33	prophage antirepressor	Bacillus sp. M 2-6	4.6E-15
41	24417	24241	4.55	CRISPR-associated helicase Cas3	Leptospira noguchii str. 2006001870	12.30
42	25948	24758	45.99	Integrase	Geobacillus sp. Y4.1MC1	0.0E0
43	26256	26915	25.25	putative membrane protein	Brevibacillus laterosporus GI-9	1.0 E-23
44	26992	27660	25.62	hypothetical protein BRLA_c33170	Brevibacillus laterosporus LMG 15441	5.7 E-35
45	27653	28171	19.58	accessory gene regulator B family protein	Brevibacillus laterosporus GI-9	1.5E-28
46	28168	28296	4.21	hypothetical protein BRLA_c33190	Brevibacillus laterosporus LMG 15441	0.26
47	28306	38671	13.79	hypothetical protein BRLA_c33200	Brevibacillus laterosporus LMG 15441	1.3E-11
48	28744	28890				
49	28906	29118	5.54	Phage terminase, small subunit, putative, P27	Rhodospirillum photometricum DSM 122	0.71
50	29132	29503	13.53	hypothetical protein HMPREF0987_01484	Lachnospiraceae bacterium 9_1_43BFAA	2.8E-15
51	30177	29545	23.98	putative phage repressor	Clostridium difficile ATCC 43255	2.7E-21
52	30332	30574	9.02	hypothetical protein BCAH187_A0631	Bacillus cereus AH187	5.4E-14
53	30772	30596	6.66	hypothetical protein Plarl_22353	Paenibacillus larvae subsp. larvae BRL-230010	8.8E-10

54	30872	30753	4.36	putative transcriptional regulator	Rhizobium sp. CF142	0.84
55	31136	30891		hypothetical protein CLOLEP_01249	Clostridium leptum DSM 753	2.7E-10
56	31271	31480	7.96	helix-turn-helix domain protein	Pelosinus fermentans JBW45	5.9E-13
57	31483	31680	7.38	hypothetical protein HBHAL_4715	Halobacillus halophilus DSM 2266	4.8E-15
58	31694	31816	4.72	homoserine kinase	Lactobacillus kisonensis F0435	3.08
59	31813	32061	9.05	XRE family transcriptional regulator	Paenibacillus polymyxa SC2	1.8E-27
60	32075	32272	7.46			
61	32277	32450	6.74	hypothetical protein CCM_04403	Cordyceps militaris CM01	14.57
62	32499	32747	9.86	hypothetical protein CHGG_09697	Chaetomium globosum CBS 148.51	0.82
63	32731	32988	9.96	hypothetical protein BRLA_c33620	Brevibacillus laterosporus LMG 15441	0.72
64	32985	33470	18.15	gp157-like protein	Deep-sea thermophilic phage D6E	1.6E-34
65	33481	34068	21.80	hypothetical protein PAV_5c00050	Paenibacillus alvei DSM 29	0.0E0
66	34061	34471	15.02	single-stranded DNA-binding protein	Brevibacillus laterosporus LMG 15441	0.0E0
67	34485	34832	13.10	hypothetical protein BBR47_35660	Brevibacillus brevis NBRC 100599	2.8 E-38
68	34852	35913	39.67	putative prophage LambdaCh01, replication protein O	Bacillus methanolicus PB1	8.3E-17
69	35903	36601	27.36	DNA replication protein	Paenibacillus popilliae ATCC 14706	0.0E0
70	36594	36863	10.74	DNA repair protein RecN	Nodularia spumigena CCY9414	0.47

71	36876	37610	28.44	hypothetical protein CBCST_07962	Clostridium botulinum C str. Stockholm	0.0E0
72	37614	37823	8.17	unnamed protein product	Tetraodon nigroviridis	0.14
73	37792	38286	19.50	hypothetical protein IGO_05662	Bacillus cereus HuB5-5	1.6E-18
74	38274	38684	16.08	hypothetical protein CmalA3_01914	Carnobacterium maltaromaticum ATCC 35586	1.3E-19
75	38681	39232	20.23	recombination protein U	Bacillus cereus BAG3X2-2	0.0E0
76	39233	39508	10.35	hypothetical protein PDENDC454_04229	Paenibacillus dendritiformis C454	1.8E-14
77	39533	39640		tonb-dependent receptor	Leadbetterella byssophila DSM 17132	2.58
78	39630	40268	24.90	dUTPase	Geobacillus thermoglucosidasius C56-YS93	0.0E0
79	40265	40351	4.73			
80	40348	41175	31.82	Site-specific DNA methylase	Bacillus subtilis BSn5	0.0E0
81	41361	41125		hypothetical protein BRLA_c33790	Brevibacillus laterosporus LMG 15441	6.6E-14
82	41341	41685	13.07	hypothetical protein TCSYLVIO_006803	Trypanosoma cruzi	0.03
83	41726	42121	15.63	hypothetical protein KSO_07894	Bacillus amyloliquefaciens IT-45	1.1 E-23
84	42136	42429	11.48	hypothetical protein PMI05_01596	Brevibacillus sp. BC25	1.0 E-10
85	42436	42934	7.09	hypothetical protein CC1G_09777	Coprinopsis cinerea okayama7#130	0.78
86	42596	42934	13.16	hypothetical protein BATR1942_07635	Bacillus atrophaeus 1942	1.9E-41
87	42968	43633	27.27	hypothetical protein PAV_1c09130	Paenibacillus alvei DSM 29	2.3E-25
88	43617	43859	9.34	hypothetical protein BBR47_35560	Brevibacillus brevis NBRC 100599	1.1E-9
89	43838	44107	10.43	putative MarR family regulatory protein	Pseudomonas fluorescens SBW25	4.34

90	44196	44705	19.54	RNA polymerase, sigma-24 subunit, ECF subfamily protein	Paenibacillus larvae subsp. larvae BRL-230010	0.0E0
91	44777	45115	12.69	hypothetical protein Spirs_2785	Spirochaeta smaragdinae DSM 11293	5.9E-37
92	45188	45508	12.47	hypothetical protein BRLA_c22590	Brevibacillus laterosporus LMG 15441	5.1E-26
93	45542	45793	8.97	hypothetical protein bcere0002_54360	Bacillus cereus ATCC 10876	0.17

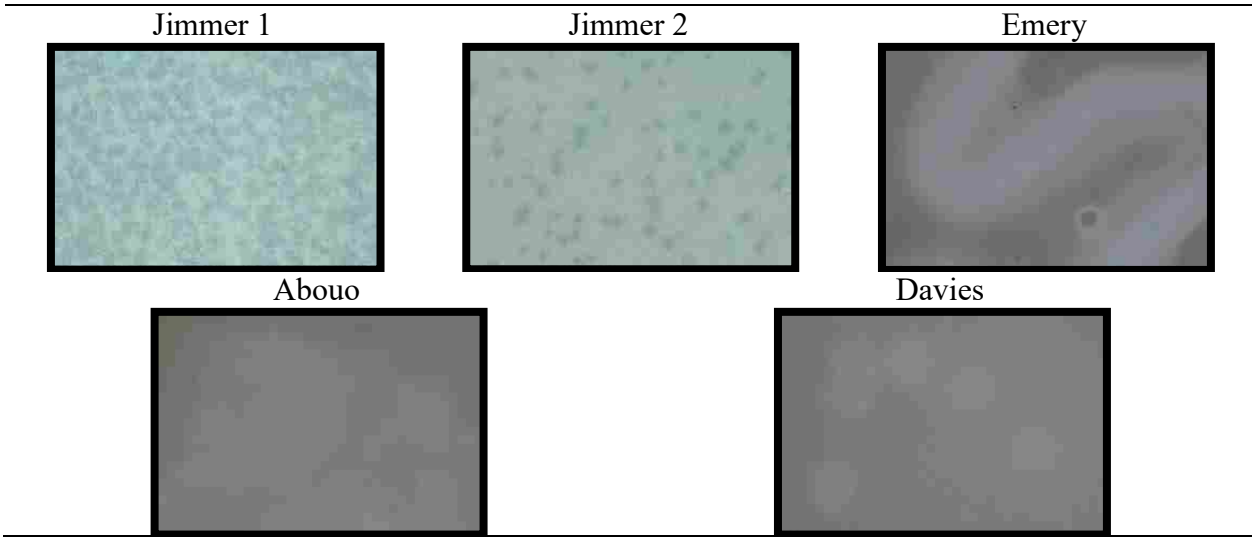

FIGURE 01 – Plaque Morphologies of Five *Brevibacillus* Bacteriophage

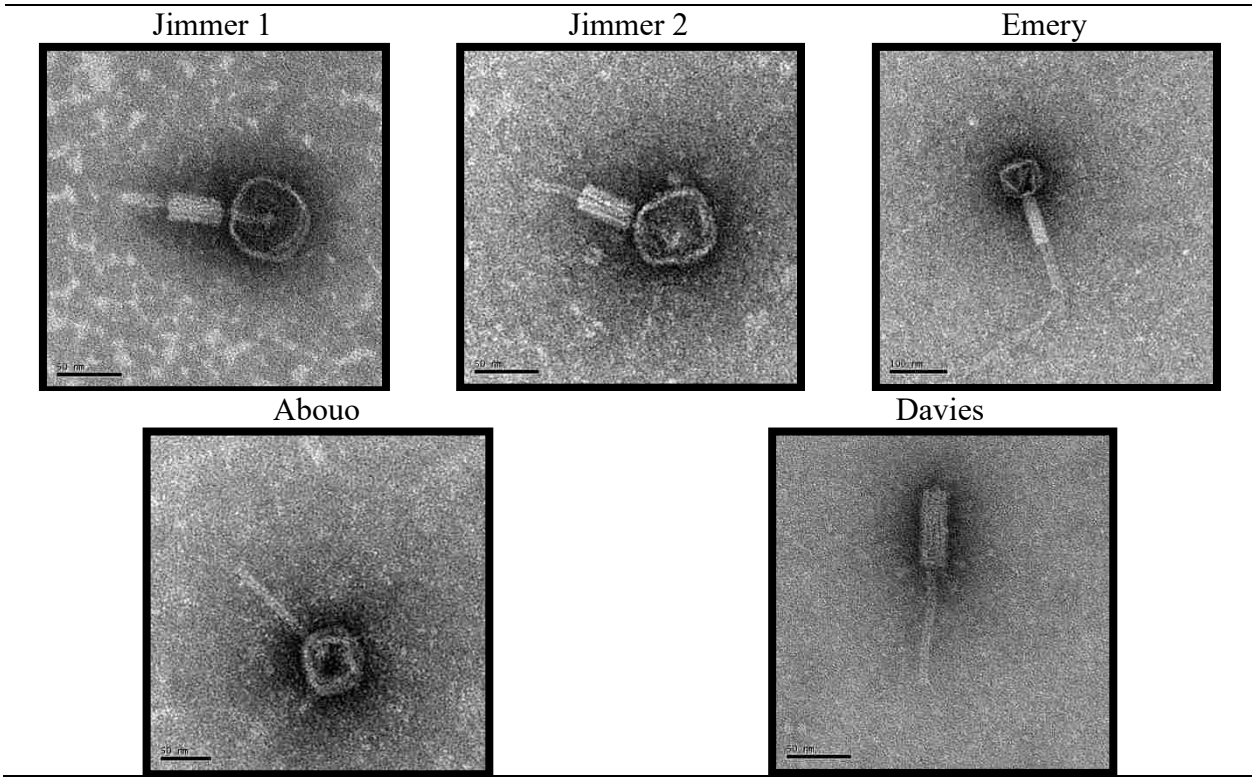

FIGURE 02 – Electron Microscope Images of Five *Brevibacillus* Bacteriophage

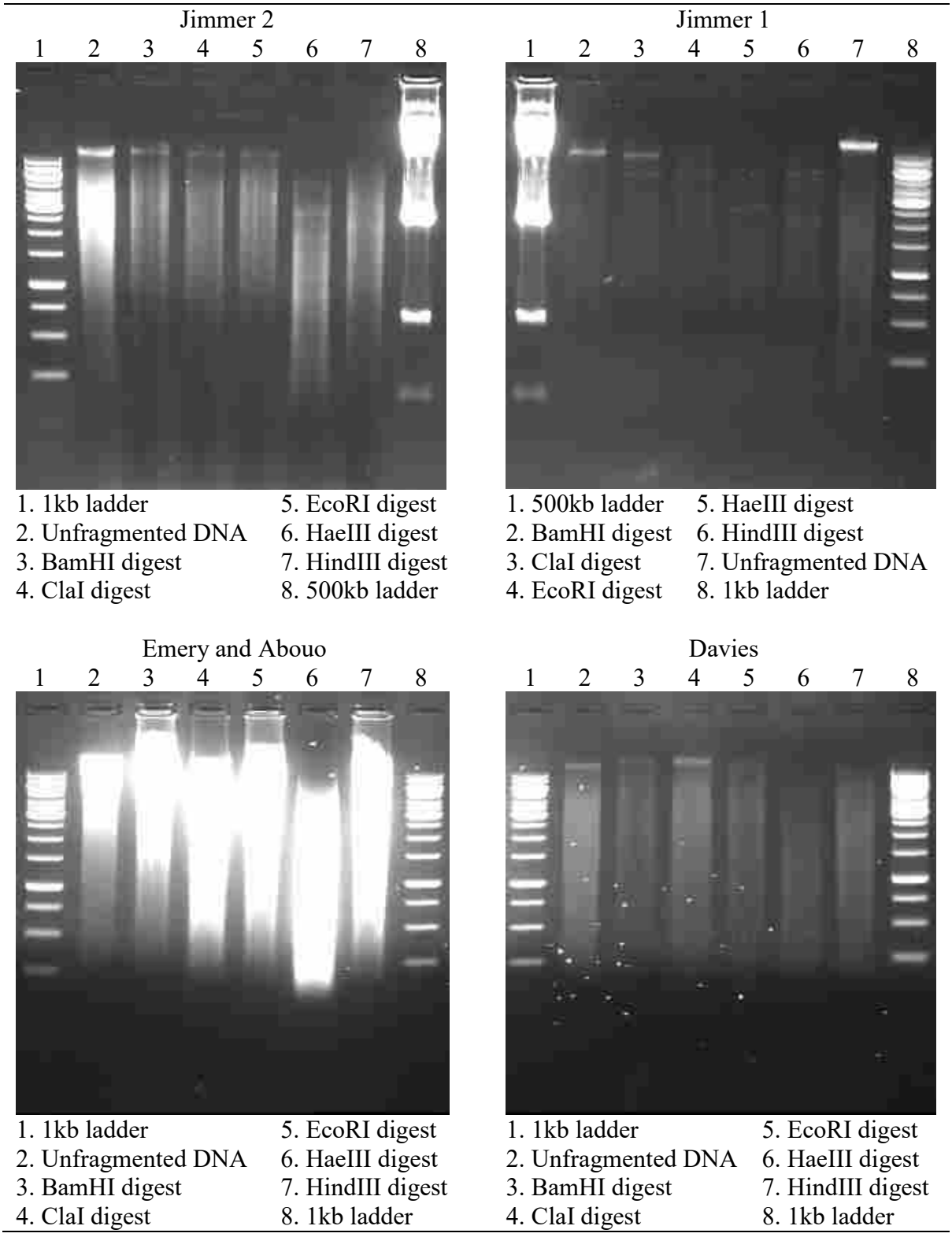

FIGURE 03 – Restriction Endonuclease Digest of Five *Brevibacillus* Bacteriophage

BIBLIOGRAPHY

Alippi, A.M. (1995). “Detection of *Bacillus larvae* spores in Argentinian honeys by using a semi-selective medium.” Microbiologia 11: 343-350.

Ackermann, H.W. (1998). “Tailed bacteriophages: The order caudovirales.” Advances in virus research 51: 135-201.

Antúnez, K., M. Anido, et al. (2012). "American Foulbrood in Uruguay: Twelve years from its first report." Journal of Invertebrate Pathology 110(1): 129-131.

Antúnez, K., C. Piccini, et al. (2007). "Phenotypic and genotypic characterization of *Paenibacillus larvae* isolates." Veterinary Microbiology 124(1-2): 178-183.

Bakhiet, N. and D.P. Stahly (1985). “Studies on transfection and transformation of protoplasts of *Bacillus larvae*, *Bacillus subtilis*, and *Bacillus popilliae*.” Applied and Enviromental Microbiology 49(3): 577-581.

Bakonyi, T., I. Derakhshifar, et al. (2003). "Development and evaluation of PCR assays for the detection of *Paenibacillus larvae* in honey samples: Comparison with isolation and biochemical characterization." Applied and Environmental Microbiology 69(3): 1504-1510.

Brussow, H. (2013). “Bacteriophage-host interaction: from splendid isolation into a messy reality.” Current Opinion in Microbiology 16 (4): 500-506.

Cairns, B. J., A. R. Timms, et al. (2009). "Quantitative models of in vitro bacteriophage-host dynamics and their application to phage therapy." PLoS Pathog 5(1): e1000253.

Capparelli, Rosanna; et al. (2007). "Experimental phage therapy against Staphylococcus aureus in mice." *Antimicrobial Agents and Chemotherapy* 51 (8): 2765–2773.

Casjens, S.R. and P.A. Thuman-Commike (2011). “Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly.” Virology 411: 393-415.

Chan, Q. W. T., R. S. Cornman, et al. (2011). "Updated genome assembly and annotation of *Paenibacillus larvae*, the agent of American foulbrood disease of honey bees." BMC Genomics 12(1): 450.

de Graaf, D. C., A. M. Alippi, et al. (2006). "Diagnosis of American foulbrood in honey bees: a synthesis and proposed analytical protocols." Letters in Applied Microbiology 43(6): 583-590.

de Graaf, D. C., P. De Vos, et al. (2006). "Identification of *Paenibacillus larvae* to the subspecies level: An obstacle for AFB diagnosis." Journal of Invertebrate Pathology 91(2): 115-123.

Deresinski, S. (2009). "Bacteriophage Therapy: Exploiting Smaller Fleas." Clinical Infectious Diseases 48(8): 1096-1101.

Di Pinto, A., L. Novello, et al. (2011). "ERIC-PCR genotyping of *Paenibacillus larvae* in southern Italian honey and brood combs." Curr Microbiol.

Dingman, D. W., Bakhiet, N., et al. (1984). "Isolation of two bacteriophages from *Bacillus larvae*, PBL1 and PBL0.5, and partial characterization of PBL1." Journal General Virology 65: 1101-1105.

Dingman, D. W. (2012). "*Paenibacillus larvae* 16S–23S rDNA intergenic transcribed spacer (ITS) regions: DNA fingerprinting and characterization." Journal of Invertebrate Pathology 110(3): 352-358.

Eischen, F. A., R. H. Graham, et al. (2005). "Regional distribution of *Paenibacillus larvae* subspecies *larvae*, the causative organism of American Foulbrood, in honey bee colonies of the Western United States." Journal of Economic Entomology 98(4): 1087-1093.

Farrar, M.D., Howson, K.M., et al. (2007). "Genome sequence and analysis of a *Propionbacterium acnes* bacteriophage." Journal of Bacteriology 180(11):4161-4167.

Genersch, E., A. Ashiralieva, et al. (2005). "Strain- and Genotype-Specific Differences in Virulence of *Paenibacillus larvae* subsp. *larvae*, a bacterial pathogen causing American Foulbrood Disease in honeybees." Applied and Environmental Microbiology 71(11): 7551-7555.

Genersch, E. (2006). "Reclassification of *Paenibacillus larvae* subsp. *pulvifaciens* and *Paenibacillus larvae* subsp. *larvae* as *Paenibacillus larvae* without subspecies differentiation." International Journal of Systematic and Evolutionary Microbiology 56(3): 501-511.

Genersch, E. (2008). "*Paenibacillus larvae* and American Foulbrood – long since known and still surprising." Journal für Verbraucherschutz und Lebensmittelsicherheit 3(4): 429-434.

Genersch, E. (2010). "American Foulbrood in honeybees and its causative agent, *Paenibacillus larvae*." Journal of Invertebrate Pathology 103: S10-S19.

Gillard, M., J. D. Charriere, et al. (2008). "Distribution of *Paenibacillus larvae* spores inside honey bee colonies and its relevance for diagnosis." Journal of Invertebrate Pathology 99(1): 92-95.

Gouchnauer, R.W. (1954). "Recent advances in the control of American Foulbrood Disease." Report of the State Apiarist for the Year ending December 31, 1956, pp. 34-36 Des Moines, Iowa.

Gouchnauer, R. W. (1955). "The isolation of a bacteriophage (bacterial virus) from *Bacillus larvae*." Bee World 36: 101-103.

Gouchnauer, R.W. (1958). “The use of bacteriophages in the analysis of the foulbrood diseases.” Proceedings. International Congress of Entomology 4: 1091-1096.

Gouchnauer, R. W. (1970). “Some properties of a bacteriophage for *Bacillus larvae*.” Journal of Invertebrate Pathology 15(2): 149-156.

Haq, I. U. L., W. N. Chaudhry, et al. (2012). "Bacteriophages and their implications on future biotechnology: A review." Virology Journal 9(1): 9.

Hatfull, G. F., D. Jacobs-Sera, et al. (2010). "Comparative genomic analysis of 60 mycobacteriophage genomes: Genome clustering, gene acquisition, and gene size." Journal of Molecular Biology 397(1): 119-143.

Hatfull, G. F. (2014). “Molecular genetics of Mycobacteriophages.” Microbiology Spectrum 2(2): 1-36.

Henry, M., O. O'Sullivan, et al. (2010). "In silico analysis of Ardmore, a novel mycobacteriophage isolated from soil." Gene 453(1-2): 9-23.

Hershey, A. D., Chase, M. (1952). “Independent functions of viral protein and nucleic acid in growth of bacteriophage.” Journal of General Physiology 36(1): 39-52.

Hornitzky, M. A. Z. and S. Clark (1991). “Culture of *Bacillus larvae* from bulk honey samples for the detection of American foulbrood.” Journal of Apiculture Res. 30: 13-16.

Jakutyte, L., C. Baptista, et al. (2011). "Bacteriophage infection in rod-shaped gram-positive bacteria: Evidence for a preferential polar route for phage SPP1 entry in *Bacillus subtilis*." Journal of Bacteriology 193(18): 4893-4903.

Johnson, R. (2010). “Honey bee colony collapse disorder.” Congressional Research Service.

Jończyk, E., M. Kłak, et al. (2011). "The influence of external factors on bacteriophages—review." Folia Microbiologica 56(3): 191-200.

Jones, J.B., L.E. Jackson, et al (2007). “Bacteriophages for plant disease control.” Annual Review of Phytopathology 45: 245-262.

Kysela, D.T. and P.E. Turner (2007). “Optimal bacteriophage mutation rates for phage therapy.” Journal of Theoretical Biology 249: 411-421.

Lindström, A., S. Korpela, et al. (2008). "The distribution of *Paenibacillus larvae* spores in adult bees and honey and larval mortality, following the addition of American foulbrood diseased brood or spore-contaminated honey in honey bee (*Apis mellifera*) colonies." Journal of Invertebrate Pathology 99(1): 82-86.

Lauro, F. M., Favaretto, M., et al. (2003). “Rapid detection of *Paenibacillus larvae* from honey and hive samples with a novel nested PCR protocol.” International Journal of Food Microbiology 81: 195-201.

Lipsitch, M. (2002). "Antibiotics in agriculture: When is it time to close the barn door?" Proceedings of the National Academy of Sciences 99(9): 5752-5754.

Liu, Y., Han, J., et al., (2005). “Development and evaluation of 16s rDNA microarray for detecting bacterial pathogens in cerebralspinal fluid.” Society for Experimental Biology and Medicine 230:587-591. (accessed at http://ebm.sagepub.com/content/230/8/587.full.pdf+html)

Luria, S. E., Delbruck, M. (1943). “Mutations of bacteria from virus sensitivity to virus resistance.” Genetics 28 (6): 491-511.

Martínez, J., V. Simon, et al. (2010). "A real-time PCR-based strategy for the detection of *Paenibacillus larvae* vegetative cells and spores to improve the diagnosis and the screening of American foulbrood." Letters in Applied Microbiology 50(6): 603-610.

Martinez, J. L., A. Fajardo, et al. (2009). "A global view of antibiotic resistance." FEMS Microbiology Reviews 33(1): 44-65.

Merrill, B. D., J. H. Grose, et al. (2014). “Characterization of *Paenibacillus larvae bacteriophages* bacteriophages and their genomic relationships to firmicute bacteriophages.” BMC Genomics 15:745.

Morse, R. A. and N. W. Calderone (2002). “The value of honey bee pollination in the United States.” Bee Culture Magazine 128: 1-15.

Murphy, K. C. (1998). “Use of bacteriophage λ recombination functions to promote gene replacement in *Escherichia coli*.” Journal of Bacteriology 180(8): 2063-2071.

Piccini, C., D’Alessandro, B., et al. (2002). “Detection of *Paenibacillus larvae* subspecies *larvae* spores in naturally infected bee larvae and artificially contaminated honey by PCR.” World Journal of Microbiology & Biotechnology 18: 761-765.

Pohorecka, K. and A. Bober (2008). “Occurrence of *Paenibacillus larvae* spores in honey samples domestic apiaries.” Journal of Apicultural Science 52: 105-111.

Pope, W.H., Jacobs-Sera, D., et al. (2011). “Expanding the diversity of mycobacteriophages: Insights into genome architechture and evolution.” PLoS ONE 6(1): e16329.doic:10.1371/ journal.pone.0016329

Qin, X., J. D. Evans, et al. (2006). "Genome sequences of the honey bee pathogens *Paenibacillus larvae* and *Ascosphaera apis*." Insect Molecular Biology 15(5): 715-718.

Rauch, S., A. Ashiralieva, et al. (2009). "Negative correlation between individual-insect-level virulence and colony-level virulence of *Paenibacillus larvae*, the etiological agent of American Foulbrood of honeybees." Applied and Environmental Microbiology 75(10): 3344-3347.

Runckel, C., M. L. Flenniken, et al. (2011). "Temporal analysis of the honey bee microbiome reveals four novel viruses and seasonal prevalence of known viruses, Nosema, and Crithidia." PLoS One 6(6): e20656.

Ryba, S., D. Titera, et al. (2009). "A PCR method of detecting American Foulbrood (*Paenibacillus larvae*) in winter beehive wax debris." Veterinary Microbiology 139(1-2): 193-196.

Sanger, F., Air, G. M., et al (1977). "Nucleotide sequence of bacteriophage ΦX174 DNA." Nature 265: 687-695. (access at http://www.nature.com/nature/journal/v265/n5596/pdf/265687a0.pdf).

Schuch, D. M. T., Madden, R. H., et al. (2001). "An improved method for the detection and presumptive indentification of *Paenibacillus larvae* subsp. *larvae* spores in honey." Journal of Apiculture Res. 40: 59-64.

Sheflo, M. A., A. V. Gardner, et al. (2015). "Correction for Sheflo et al., complete genome sequences of five *Brevibacillus laterosporus* bacteriophages." American Society of Microbiology. Author Correction: 01 October 2015.

Stahly, D. P., A. M. Alippi, et al. (1999). "PPL1c, a virulent mutant bacteriophage useful for identification of *Paenibacillus larvae* subspecies *larvae*." Journal of Invertebrate Pathology 74(3): 295-296.

Steensma, H.Y. and J. Blok (1979). "Effect of calcium ions on the infection of *Bacillus subtilis* by Bacteriophage SF 6." Journary of General Virology 42: 305-314.

Valerianov, S. (1976). "Isolation from the soil of a bacteriophage lysing *Bacillus larvae*." Acta Microbiologica, Virologica et Immunologica 4: 81-85.

van Kessel, J.C. and G.F. Hatfull (2008). "Mycobacterial recombineering." Methods of Molecular Biology 435:203-215.

Veesler, D. and C. Cambillau (2011). "A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries." Microbiology and Molecular Biology Reviews 75(3): 423-433.

Weinbauer, M. G. (2004). "Ecology of prokaryotic viruses." FEMS Microbiology Reviews 28:127-181.

Xu, J., Hendrix, R.W., et al. (2004). "Conserved translational frameshift in dsDNA bacteriophage tail assembly genes." Molecular Cell 16:11-21.

Zhang, Y., Buchholz, F., et al (1998). “A new logic for DNA engineering using recombination in *Escherichia coli*.” Nature Genetics 20(2): 123-128. (accessed at http://www.ncbi.nlm.nih.gov/pubmed/9771703).