11-15-2013

Real-time qPCR Assay Development for Detection of Bacillus thuringiensis and Serratia marcescens DNA and the Influence of Complex Microbial Community DNA on Assay Sensitivity

Jonathan Segal
jonathansegal89@gmail.com

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

REAL-TIME qPCR ASSAY DEVELOPMENT FOR DETECTION OF *BACILLUS THURINGIENSIS* AND *SERRATIA MARCESCENS* DNA AND THE INFLUENCE OF COMPLEX MICROBIAL COMMUNITY DNA ON ASSAY SENSITIVITY

A thesis submitted in partial fulfillment of

the requirements for the degree of

MASTER OF SCIENCE

in

FORENSIC SCIENCE

by

Jonathan Segal

2013

To: Dean Kenneth G. Furton
College of Arts and Sciences

This thesis, written by Jonathan Segal, and entitled Real-time qPCR Assay Development for Detection of *Bacillus thuringiensis* and *Serratia marcescens* DNA and the Influence of Complex Microbial Community DNA on Assay Sensitivity, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

_____
DeEtta Mills

_____
Giri Narasimhan

_____
Kalai Mathee, Major Professor

Date of Defense: November 15, 2013

The thesis of Jonathan Segal is approved.

_____
Dean Kenneth G. Furton
College of Arts and Sciences

_____
Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2013

DEDICATION

I dedicate this thesis to my extraordinary parents, whose selfless support and unconditional love is beyond words.

ACKNOWLEDGMENTS

I want to express the deepest gratitude to my major professor, Dr. Kalai Mathee, who graciously took me under her wing. Her unwavering confidence in my abilities provided me with motivation to be a better scientist and a better human being. If I have my way, she will be a life-long mentor and will be involved in my future endeavors, scientific or otherwise.

I wish to acknowledge my committee members, Dr. DeEtta Mills and Dr. Giri Narasimhan, for their commendable patience and help throughout this project. Specifically, I would like to thank Dr. DeEtta Mills for the topic-inspiring brainstorming sessions early in my graduate career and Dr. Giri Narasimhan for his calm, collected voice of reason and helpful comments.

Finally, I want to thank the past and present members of the Mathee lab crew, who welcomed me with open arms and taught me so much about science and research.

ABSTRACT OF THE THESIS

REAL-TIME qPCR ASSAY DEVELOPMENT FOR DETECTION OF *BACILLUS THURINGIENSIS* AND *SERRATIA MARCESCENS* DNA AND THE INFLUENCE OF COMPLEX MICROBIAL COMMUNITY DNA ON ASSAY SENSITIVITY

by

Jonathan Segal

Florida International University, 2013

Miami, Florida

Professor Kalai Mathee, Major Professor

Real-time quantitative polymerase chain reaction (real-time qPCR) assays are an effective technique to detect biological warfare agents and surrogate organisms. In my study, primers were designed to detect chromosomal DNA of biological warfare agent surrogates *B. thuringiensis* and *S. marcescens* (representing *B. anthracis* and *Y. pestis*, respectively) via real-time qPCR. Species-level specificity of the primers was demonstrated through comparisons with a bacterial strain panel and corroborated by qPCR data. Additionally, the primer efficacy was tested when template DNA was spiked into metagenomic DNA extracted from clinical lung microbiome samples. The results showed that while detection of *B. thuringiensis* or *S. marcescens* was still largely successful, the addition of metagenomic DNA did significantly inhibit amplification in most cases. The present study is significant not only for the design of multiple novel primer pairs able to detect bacterial agents in metagenomic DNA, but also the quantitative insight to the influence of background DNA on single species detection at low DNA concentrations.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

INTRODUCTION

The ubiquity of deoxyribonucleic acid (DNA) and its vital role within living systems makes it the single most influential molecule in biology. The information carried by DNA is the fundamental guide for every living organism, from single-cell bacteria to complex multi-cellular mammals. The uniqueness and universality of DNA has predicated its usage in many scientific fields. DNA is used in forensic science and microbial forensics as a means to experimentally individualize the perpetrator of a crime or to aid in identification of an unknown biological agent. The introduction will discuss the role of DNA analysis in forensic science and microbial forensics, biological agents and surrogate organisms, as well as detection techniques.

I. DNA Analysis and Forensic Science

The advent of DNA analysis in the late 20th century has provided the framework for significant advancements in biology and science. One field that benefitted tremendously from the arrival of DNA analysis is forensic science, or the practical application of science to legal investigations and matters of the law. While the pre-DNA era of forensic science was limited to chemical reactivity, physical evidence examination, and serology (Li 2008; Butler 2010), modern human DNA typing has provided the tools to individualize beyond a reasonable doubt. DNA analysis in forensic investigation is now standard operating procedure (SOP) in crime laboratories, and has gained reputation in the courtroom as highly compelling evidence (Carey and Mitnik 2002). DNA evidence

has resulted not only in countless convictions, but also exoneration of the guiltless, epitomized by the work of the 'The Innocence Project' (3).

Evolution of human DNA typing. Forensic DNA typing did not develop in a vacuum. It is the product of efforts and scientific advances that continue into the present day. Early forensic biology analyses consisted largely of blood antigen typing and measuring protein interactions (Li 2008; Butler 2010). Despite these not using DNA sequence, they exploit the basic concept of polymorphism and provide some power of identification. While they were fairly successful and can still be used as presumptive tests today, the probability of obtaining false positives is higher than desired. DNA-based typing was pioneered by Sir Alec Jeffreys, who developed the variable number tandem repeat (VNTR) technique in 1984 and used it as forensic evidence in 1985 (Jeffreys *et al.* 1985). Jeffreys found the human genetic code to be far more polymorphic than any protein, thus imparting a greater ability to differentiate biological samples (Li 2008). Specifically, he targeted repeating regions of the DNA that were polymorphic in length rather than in sequence. The occurrence of length differences in various repeat sequences is the basis of most DNA typing applications, including Restriction Fragment Length Polymorphism (RFLP), which relies on restriction enzymes and probes, and techniques using Polymerase Chain Reaction (PCR) (Reynolds *et al.* 1991).

Few years after the development of the polymerase chain reaction (PCR) by Kary Mullis and predecessors in 1993 (Saiki *et al.* 1988; Mullis 1990; Mullis *et al.* 1992), PCR-based typing techniques such as short tandem repeat (STR) assays became the norm for effective analysis of polymorphic DNA markers (Li 2008). The development of

techniques such as mitochondrial DNA (mtDNA) typing (Wrischnik *et al.* 1987; Piercy *et al.* 1993), single nucleotide polymorphism (SNP) (Landegren *et al.* 1998; Wang *et al.* 1998) assays, and next generation sequencing (NGS) technologies (Mardis 2008; Schuster 2008; Shendure and Ji 2008; Metzker 2010) are providing even ability to match crime scene evidence to a perpetrator and represent the progression of the discipline.

The Scientific Working Group on DNA Analysis Methods (SWGDAM) and Combined DNA Index System (CODIS). Standard Operating Procedures (SOPs) for modern human DNA typing are developed by the Federal Bureau of Investigation (FBI) laboratory and maintained for quality assurance by The Scientific Working Group on DNA Analysis Methods (SWGDAM). SWGDAM works in conjunction with the FBI for establishing standards and guidelines for DNA analysis (Budowle *et al.* 1998). Most notable was the development of the Combined DNA Index System (CODIS) database, which led the launch of the National DNA Index System (NDIS) under the DNA Identification Act of 1994 (U.S.C.-14132 1994). CODIS serves as an umbrella term for all DNA reference repositories at different levels of government, including Local DNA Index System (LDIS), State DNA Index System (SDIS), and the National DNA Index System (NDIS). DNA profiles from many sources can be deposited, including convicted offenders, arrestees, forensic samples, and remains (U.S.C.-14132 1994).

With the exception of mitochondrial DNA (mtDNA), CODIS currently stores DNA profile information from the 13 core CODIS STR loci. The combined use of these 13 nuclear STR loci (CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, THO1, TPOX, vWA) (Table 1A) and amelogenin for

sex determination allow for standardized and powerful human identification via allelic discrimination (Budowle *et al.* 1998; Butler and Reeder 2013). Biotechnology firms have created comprehensive kits to amplify all the loci (e.g., AmpFLSTR® Profiler Plus® and COfiler® kits, Applied Biosystems, Madison, WI) and computer software to streamline allelic calling (e.g., GeneMapper®, Applied Biosystems, Madison, WI).

Recently, additional loci have also been identified and used to increase discrimination power and provide additional coverage of the genome in the case of degraded or limited DNA. The supplementary coverage generally involves the addition of more loci, such as two pentanucleotide STR loci Penta E and Penta D (Table 1B), which are now incorporated into commercial kits such as the PowerPlex 16 System (Promega, Madison, WI) for use in forensic laboratories after proper validation.

Forensic analysis of non-human DNA. DNA from a variety of non-human sources, including plants, animals, and microorganisms, have proven to be useful sources of evidence. Documented usage of animal DNA in cases such as the Beamish murder trial (Menotti-Raymond *et al.* 1997) and *State of Vermont v. Michael J. Demers* (Supreme Court of Vermont 1997) indicate the value of animal DNA in cases, ranging from murders to poaching. Plant DNA has also been used as scientific evidence in court, such as in *State of Arizona v. Mark Alan Bogan* (Court of Appeals of Arizona 1995), where plant matter helped place connect a murder suspect to the crime scene. Microorganisms can also play a large role in forensics, notably in soil analysis. DNA profiling of microbial communities in soil can distinguish soils better than chemical analyses in some instances (Moreno 2005; Moreno *et al.* 2006).

Table 1. STR Loci. A) The 13 core loci of the Combined Index Database System (CODIS). B) Two additional loci for increased statistical power. All data sourced from STRBase (Butler and Reeder 2013).

A)

Locus Name	Repeat Unit	Chromosomal Location	Number of Alleles Known
CSF1PO	TAGA	5q33.1	20
FGA	CTTT	4q31.3	80
TH01	TCAT	11p15.5	20
TPOX	GAAT	2p25.3	15
VWA	[TCTG][TCTA]	12p13.31	28
D3S1358	[TCTG][TCTA]	3p21.31	24
D5S818	AGAT	5q23.2	15
D7S820	GATA	7q21.11	30
D8S1179	[TCTA][TCTG]	8q24.13	17
D13S317	TATC	13q31.1	17
D16S539	GATA	16q24.1	19
D18551	AGAA	18q21.33	51
D21511	Complex [TCTA] [TCTG]	21q21.1	82

B)

Locus Name	Repeat Unit	Chromosomal Location	Number of Alleles Known
Penta D	AAAGA	21q22.3	14
Penta E	AAAGA	15q26.2	21

II. Microbial Forensics and Biological Warfare Agents

The exploitation of biological agents to inflict harm has persisted since antiquity. Modern biological warfare (20th century – present) is initially thought of in the context of military application of biological and chemical agents, which began on a large scale in World War I and II (Tucker and Koblentz 2009). Concerned groups, primarily governments and regimes, concentrated efforts on research and development of biological agents. The Biological Weapons Convention of 1972 then banned the accumulation and storage of bioagents internationally (Tucker and Koblentz 2009). Since the treaty, research has gravitated towards biodefense in light of the potential usage of biological weapons by terrorists or undesirables.

The field of microbial forensics developed as a niche subfield that examines biological agents related to bioterrorism, biological warfare, biosecurity, or accidental discharge (Budowle *et al.* 2005; Tucker and Koblentz 2009). Typically, the investigations intend to identify or confirm the identity of an unknown agent as well as to establish etiology (Budowle *et al.* 2005). In its infancy, microbial forensics was dependent on culturing bacteria and protein analyses for identification of a microorganism (Budowle 2011). However, much like human DNA typing, the advents of PCR and sequence-based techniques transformed the field. While microbiology techniques are still valuable and useful, molecular analyses and genetics supply the tools to look past culturable phenotypes and characterize microorganisms more accurately (Handelsman 2004; Schmeisser *et al.* 2007).

The Center for Disease Control categorizes known biological warfare agents on the basis of several factors such as ease of dissemination, mortality, and potential for person-to-person transmission (Table 2). Category A agents are considered more threatening that Category B or Category C agents. The current list of Category A agents consists of *Bacillus anthracis, Yersinia pestis, Clostridium botulinum, Francisella tularensis*, *Variola major*, hemorrhagic fever filoviruses (Ebola and Marburg fever), and hemorrhagic arenaviruses (Lassa fever, Junin, Machupo) (Budowle *et al.* 2005). These agents and their associated dangers have prompted scientists to concentrate on those that pose the highest risk. Research is focused on developing methods to identify these agents, as rapid and sensitive detection is of utmost importance to minimize potential damage caused by these agents in the event of an outbreak or attack.

III. *Bacillus anthracis*, the Causative Agent of Anthrax

Microbial forensics and *B. anthracis* gained public fame following the 'Anthrax-Letter Event' bioterrorism attack in 2001 and its subsequent microbiological investigation (Justice 2010). Between September and October 2001, four confirmed (five suspected) envelopes with *B. anthracis* spores were mailed to New York Post, NBC new anchor Tom Brokaw in New York City, the offices of Senator Tom Daschle and Senator Patrick Leahy in Washington D.C., and American Media International offices in Boca Raton, Florida (Rasko *et al.* 2011). Although none of the intended targets were reached, 11 individuals contracted inhalation anthrax and 11 others contracted cutaneous anthrax infections, resulting in five deaths (Rasko *et al.* 2011). The victims were mail handlers or

Table 2. Description of Center of Disease Control (CDC) Category A and Category B Bioterrorism Agents. (Budowle *et al.* 2005)

Name	Threat	Type
CATEGORY A		
Bacillus anthracis	**Anthrax**	**Bacterial**
***Clostridium botulinum* toxin**	**Botulism**	**Bacterial toxin**
Yersinia pestis	**Plague**	**Bacterial**
Variola major virus	**Smallpox**	**Viral**
Francisella tularensis	**Tularemia**	**Bacterial**
Arenaviruses, Ebola virus, Lassa virus, Marburg virus	**Viral hemorrhagic fever**	**Viral**
CATEGORY B		
***Brucella* spp.**	**Brucellosis**	**Bacterial**
***Clostridium perfringens* toxin**	**Epsilon toxin**	**Bacterial toxin**
***Salmonella* spp., *Shigella* spp., *Escherichia coli* 0157:H7**	**Food safety threats**	**Bacterial**
Burkholderia mallei	**Glanders**	**Bacterial**
Burkholderia pseudomallei	**Meliodosis**	**Bacterial**
Chlamydia psittaci	**Psittacosis**	**Bacterial**
Coxiella burnetti	**Q Fever**	**Bacterial**
***Ricinus communis* toxin**	**Ricin**	**Plant toxin**
Staphylococcal enterotoxin B	**Enterotoxin**	**Bacterial toxin**
Rickettsia prowazekki	**Typhus fever**	**Bacterial**
Venezuelan equine virus, Eastern equine virus, Western equine virus	**Viral encephalitis**	**Viral**
Vibrio cholerae*, *Cryptosporidium parvum	**Water safety threats**	**Bacterial, Protozal**

processors, exposed to the anthrax spores at their respective worksites (Jernigan *et al.* 2001; Jernigan *et al.* 2002).

The FBI formed a task force and began an in-depth investigation to discover the responsible party. The investigation, code-named 'Amerithrax', was a multi-faceted inquiry that included genetic analysis of the sporulated *B. anthracis* strains found in the envelopes (U.S.D.O.J 2010). The microbiological and genetic investigation compared the strains found in all of the envelopes to reference strains and too each other. By comparison of specific mutations, the analysis revealed a single type 'Ames' strain in all of the envelopes, suggesting that all the *B. anthracis* came from a single source (U.S.D.O.J 2010).

Prior to the Amerithrax investigation, governments and intelligence agencies had been dealing with biological warfare threats for decades. One such event that gained little media coverage in the U.S. was the Kameido Anthrax Incident of 1993 (Takahashi *et al.* 2004). The Aum Shinrikyo, a radical Japanese sect, aerosolized a liquid suspension of *B. anthracis* spores in Tokyo in a rather understated act of terrorism (Budowle *et al.* 2005). No people were reported sick or dead from anthrax exposure (Takahashi *et al.* 2004). As a result of the event's inconspicuous nature, only retroactive microbiological examination in 2000 identified the aerosol liquid as sporulated *B. anthracis* (Keim *et al.* 2001). Using multiple locus variable number tandem repeat analysis (MLVA), a PCR based multi locus variant of VNTR, analysts identified the isolate as the Sterne strain of *B. anthracis*, an attenuated strain commonly used for veterinary vaccinations (Budowle 2011).

The weaponization of *B. anthracis* was not limited to terrorist groups. Weaponization of biological agents for warfare is evidenced by events such as the

Gruinard Island experiments conducted by the British government in 1942 and 1943 (Manchee and Stewart 1988; Inglesby *et al.* 1999; Spencer 2003). These experiments involved live testing of 'N-bombs' containing sporulated *B. anthracis*, as well as development of anthrax-infested cattle cakes intended to wipe out food supplies of Nazi Germany (Spencer 2003). Additionally, a 1979 outbreak of anthrax poisoning occurred around a military facility in Sverdlovsk, USSR, killing upwards of 66 people (Meselson *et al.* 1994; Jackson *et al.* 1998; Spencer 2003).

***B. anthracis* pathogenicity.** *Bacillus anthracis* is a Gram-positive spore-forming species that operates as an obligate aerobe (Sacchi *et al.* 2002; Spencer 2003). As a member of the *B. cereus* group of bacteria, it is notoriously difficult to distinguish *B. anthracis* from other group members by phenotype or 16S rRNA sequencing (Ash *et al.* 1991; Sacchi *et al.* 2002; Hadjinicolaou *et al.* 2009). Conclusive differentiation can be obtained with amplified fragment length polymorphism (AFLP) (Spencer 2003) or genome sequencing (Rasko *et al.* 2011).

Pathogenicity of *B. anthracis* stems largely from two virulence factor plasmids, pXO1 and pXO2 (Spencer 2003) (Figure 1). Three components are encoded on the 184.5 kb pXO1 plasmid: an 89-kDa edema factor (EF), an 83-kDa calmodulin-dependent adenylate cyclase lethal factor (LF), and an 85-kDa protective antigen (PA) that binds to host cell receptors and facilitates entry into the cell (Little and Ivins 1999; Spencer 2003). Lethal factor (LF) and edema factor (EF) competitively bind to the protective antigen (PA), forming lethal toxin or edema toxin, respectively. (Little and Ivins 1999). Research with mice demonstrates that the lethal toxin is more crucial in virulence, but a synergistic

effect is observed when both toxins are present (Pezard *et al.* 1991). Additionally, regulation of virulence is primarily mediated by *atxA*, a gene found on pXO1 (Figure 1) that is required for the transcription of the toxin genes (Bourgogne *et al.* 2003).

The second 95.3-kb plasmid pXO2 acts in a complementary fashion to pXO1. The pXO2 plasmid harbors three genes (*capA*, *capB*, *capC*; Figure 1) that are responsible for production of a protective polyglutamyl capsule that surrounds that bacterium (Makino *et al.* 1989; Spencer 2003). The edema, hemorrhage, and necrosis caused by the pXO1 toxin compounds is exacerbated by the encapsulation and the subsequent inability of the leukocytes to phagocytose the bacteria. Complete virulence requires both pXO1 and pXO2, but attenuated strains can be achieved through loss of one or both of the plasmids (Spencer 2003).

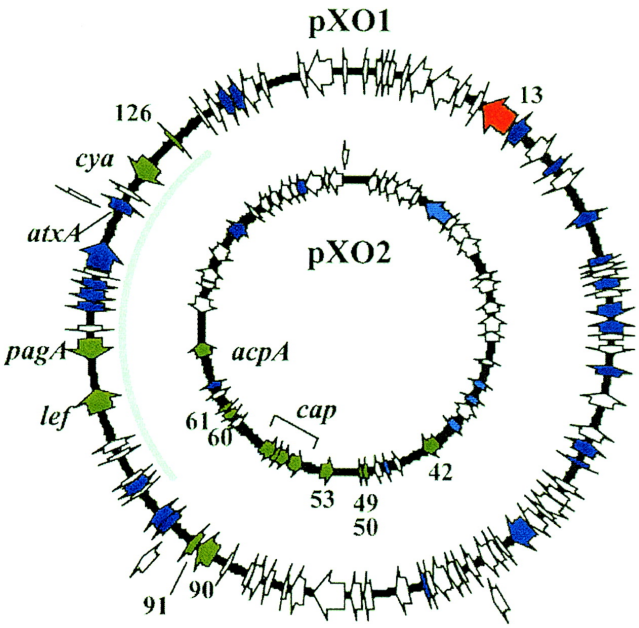

Figure 1. *B. anthracis* pXO1 and pXO2 plasmids. Green arrows indicate positive regulation by *atxA*, red arrows show negative regulation, and blue arrows indicate genes that are not regulated by *atxA*. White arrows indicate genes not expressed under the growth conditions of the study (Bourgogne *et al.* 2003).

Anthrax infections. *B. anthracis* can be introduced into the body through various routes, including the skin, gastrointestinal tract, or the lungs, resulting in cutaneous anthrax, gastrointestinal anthrax, or pulmonary anthrax poisoning respectively (Spencer 2003). Cutaneous anthrax is the most common type of infection and also the mildest (Little and Ivins 1999; Spencer 2003). The infection results in establishment of a papule, formation of edema, and eventual escharing of the tissue (Figure 2).

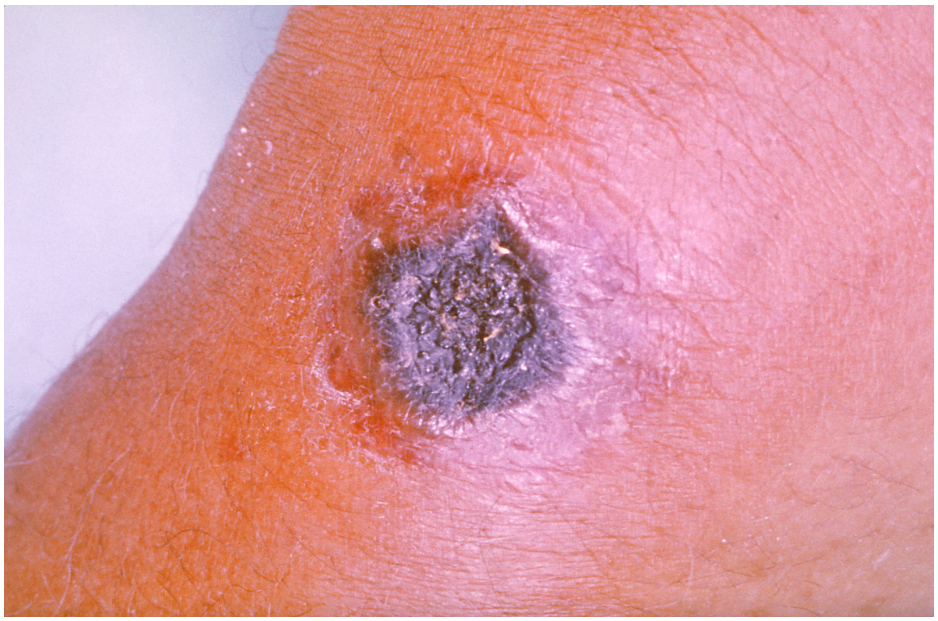

Figure 2. Cutaneous anthrax lesion. The dark eschared tissue is typical of a late-stage anthrax lesion (Steele 1962).

Cutaneous infection can be effectively treated using antibiotics (Little and Ivins 1999). Infections resulting from inhalation or ingestion of anthrax are considered extremely severe with high risk of septicemia and toxemia (Little and Ivins 1999). Although uncommon, gastrointestinal anthrax (associated with ingestion of undercooked or contaminated meat) results in the formation of a papule on the wall of the esophagus,

stomach, duodenum or ileum (Spencer 2003). Manifesting in severe edema, necrosis, and the enlargement of mesenteric lymph nodes, the gastrointestinal infection results in septicemia and death if not treated promplty with antibiotics (Little and Ivins 1999; Spencer 2003).

Perhaps the most serious and infamous type of anthrax infection is pulmonary anthrax, initiated with the deposition of sporulated *B. anthracis* within the alveolar spaces of the lungs (Spencer 2003). The spores are unintentionally transported to lymph nodes by phagocytic leukocytes (the protective capsule prevents destruction within the phagocyte) and germinate within the lymph nodes (Little and Ivins 1999; Spencer 2003). Once in the vegetative state, the bacterial load increases and bacteria can eventually enter the bloodstream, resulting in severe septicemia and death unless treated early powerful antibiotics (Little and Ivins 1999; Spencer 2003).

IV. *Yersinia pestis*, the Causative Agent of Plague

Yersinia pestis, the causative agent of plague, can be transferred from infected rodents to humans via flea vectors. Outbreaks of bubonic plague and pneumonic plague have proven fatal to countless people throughout history. Although the numbers are uncertain, some estimate that 200 million people have died during the course of recorded history as a result of infection by *Y. pestis* (Perry and Fetherston 1997). Responsible for most of these deaths were several massive pandemics, including the Justinian plague (6^{th}-8^{th} century), the 'Black Death' plague (14^{th}-19^{th} century), and the unnamed third

pandemic, beginning in the Yunnan province of China in 1855 (19^{th}-20^{th} century) (Perry and Fetherston 1997; Parkhill *et al.* 2001; Haensch *et al.* 2010).

The wrath of plague is not purely historical; recent events have again shown the remarkable ability of *Y. pestis* to infect humans. In 1994, a major outbreak of pneumonic plague occurred in Surat, India claimed the lives of 56 people (Dennis 1994). The plague was believed to have originated from less severe bubonic cases reported to the World Health Organization (WHO) in neighboring Maharashtra (Dennis 1994). The outbreak was eventually controlled when treatment with tetracycline hydrochloride proved effective and was administered to patients *en masse* (Nandan 1994). Other outbreaks arose as recently as 2003 in Algeria (Bertherat *et al.* 2007) and 2009 in Libya (Cabanel *et al.* 2013). Although outbreaks like these have occurred recently and frequently, the availability of effective antibiotics in the past several decades has significantly reduced the impact of the disease on human populations (Cabanel *et al.* 2013). As a consequence of the deadly effects of plague infection in humans, *Y. pestis* is also considered to have dangerous potential as a biological warfare agent (Prentice and Rahalison 2007). One example of such activity include the conduction of *Y. pestis* –infected flea bomb trials by the Japanese military on prisoners of war from 1930s until World War II (Prentice and Rahalison 2007). Moreover, reports of former Soviet scientists alluded to an attempt by the USSR to produce multi-drug resistant *Y. pestis* for use as a biological warfare agent during the Cold War era (Kadlec *et al.* 1997).

***Y. pestis* pathogenicity.** *Yersinia pestis* is a Gram-negative non-spore forming *coccobacillus* that operates as a facultative anaerobe (Perry and Fetherston 1997). As a

member of both the *Enterobacteriaceae* family and *Yersinia* genus, it is closely related to other pathogenic *Yersinia* species such as *Y. pseudotuberculosis*, and *Y. enterocolitica*. *Y. pestis* also has a unique set of metabolic growth requirements that require a semi-parasitic relationship with a host (Perry and Fetherston 1997).

The pathogenicity of *Y. pestis* is well characterized. Its facultative parasitic lifestyle lends to its ability to successfully evade host defenses in many complex ways (Parkhill *et al.* 2001; Prentice and Rahalison 2007). Because of its complexity, only select virulence factors will be discussed here. *Y. pestis* derives is pathogenicity principally from three plasmids (Figure 3), the 100–110 kb pMT1 (or pFra), the 70-kb pCD1 and the 9.5 kb pesticin plasmid (pPCP1) (Kingston *et al.* 2007). The pMT1 and pCD1 plasmids harbor genes for two antigens, Fraction 1 (F1) and LcrV (V) respectively, which are produced once the bacterium encounters human body temperature of 37˚C (Kingston *et al.* 2007). The F1 antigen has anti-phagocytic functions while V antigen is thought to function as a key regulator of a Type 3 Secretion System (T3SS) (Kingston *et al.* 2007). The T3SS is responsible for injecting *Yersinia* outer proteins (YopB, YopD, YopO, YopH, YopM, YopT, YopJ, and YopE) that function to form pores within the wall of a host immune cell and interfere with immune functions. The Yop proteins can hinder phagocytosis, stop the release of cytokines, or even induce apoptosis of host immune cells (Viboud and Bliska 2005). The impairment of the immune system allows for the proliferation of *Y. pestis* within lymph nodes, causing severe lymphadenopathy and eventual septicemia once bacteria reach critical levels in the bloodstream (Prentice and Rahalison 2007).

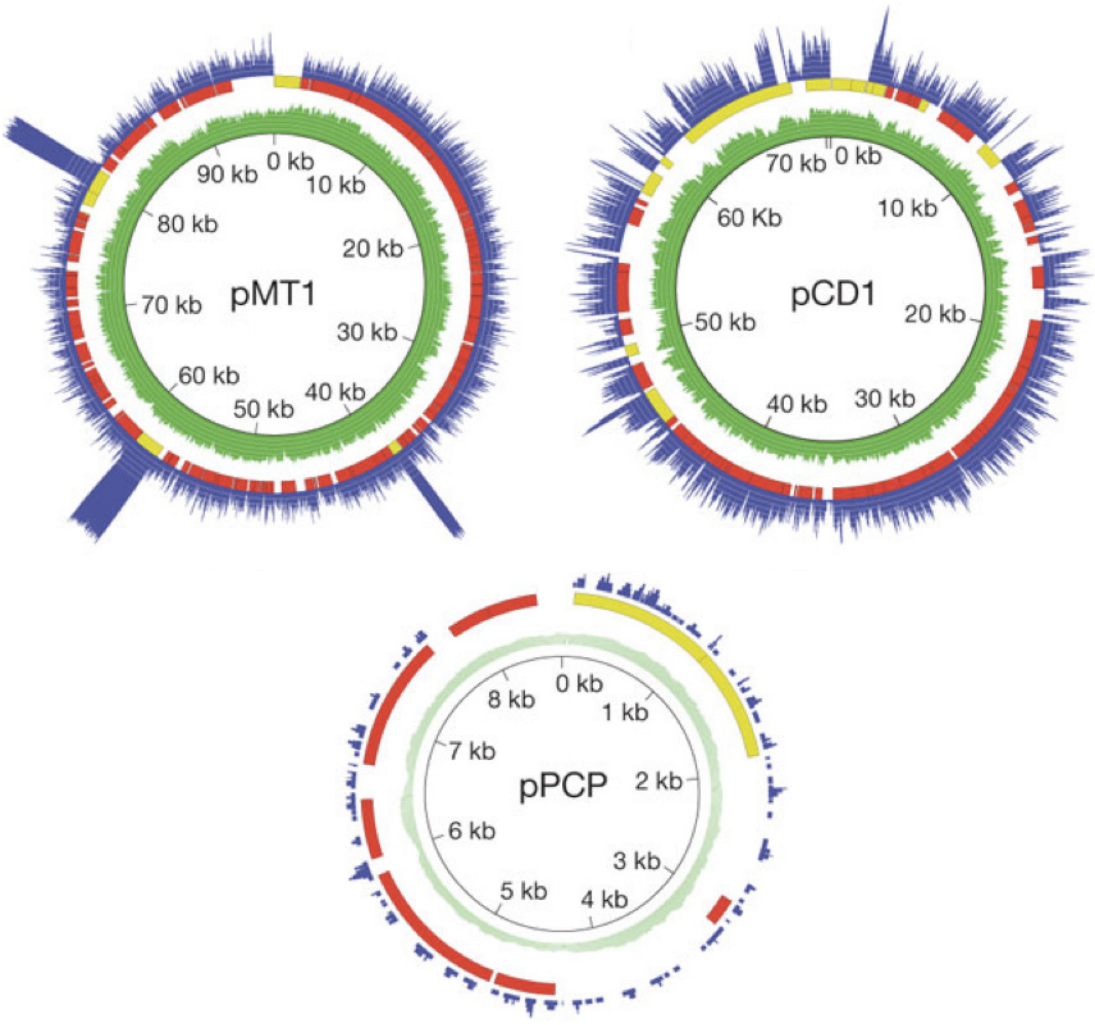

Figure 3. *Y. pestis* virulence plasmids. Sequencing read coverage indicated in blue, GC content in green, shown on a relative scale. Red regions designate coding regions and yellow regions designate mobile elements (Bos *et al.* 2011).

Other factors promote virulence, including the expression of the *yadBC* gene, which encodes an adhesion protein allowing for invasion of epithelial cells in bubonic plague (Forman *et al.* 2008). In addition, a plasminogen activator Pla can be expressed, promoting the fibrinolysis of blood clots and allowing for the bacteria to reside and multiply uninhibited within the bloodstream (Lathem *et al.* 2007).

Disease characteristics of plague. Several types of plague exist, the most common varieties being bubonic, pneumonic, and septicemic plague (Prentice and Rahalison 2007). Bubonic plague is distinguished by the presentation of swollen lymph nodes, or buboes (Figure 4), which accompany high fever, headache, and malaise (Teh 1922). Pneumonic plague is characterized by person-to-person transfer via respiratory droplets and presents with shortness of breath, chest pain, bloody cough, and high fever (Ratsitorahina *et al.* 2000). In addition, septicemic plague can present with symptoms similar to bubonic plague, but without the appearance of buboes, making it much harder to diagnose, resulting in high mortality rates (Prentice and Rahalison 2007). Without treatment before several days of incubation, potentially fatal infections such as meningitis, pneumonia, and septicemia will occur (Prentice and Rahalison 2007).

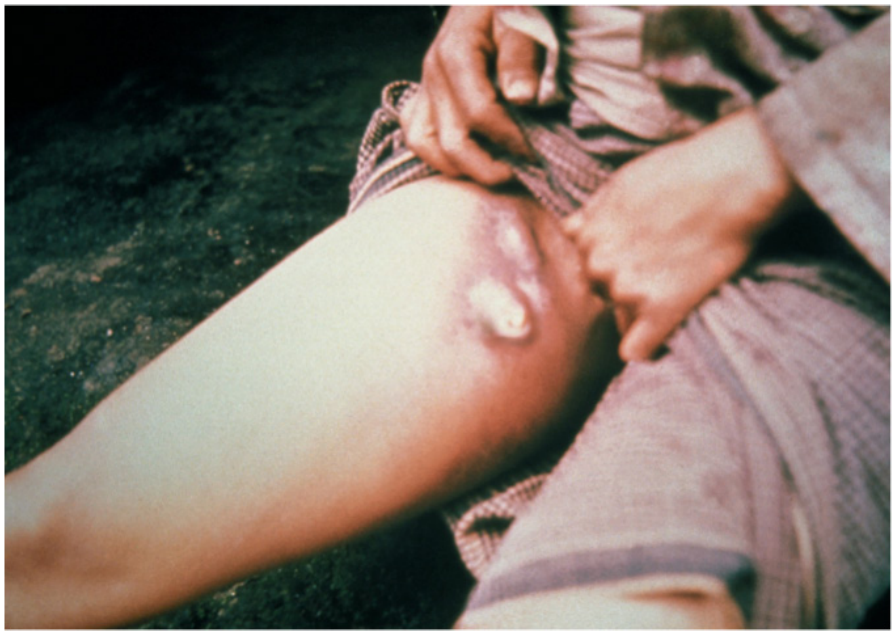

Figure 4. Inguinal buboe caused by the rupture of a lymph node (Unknown 1993).

V. Molecular Markers and Techniques

Traditional identification of bacteria is performed using microbiological methods, involving culturing on selective media and observation of various phenotypes and morphologies. Although well developed, these methods are laborious and can result in false positives or negatives. Molecular markers rely on the genetic material within the microorganism for differentiation. A good universal molecular marker is conserved throughout all organisms, yet provides enough variation to discriminate between organisms. Several examples of good molecular markers are essential to the organism, such as RecA and GyrB, or 16S rRNA (Moreno 2005). 16S rRNA genes are likely the most commonly used molecular marker for prokaryotes to date.

16S rRNA as a molecular marker. The 16S rRNA exist in every prokaryote. and makes up part of the 30S small subunit of the ribosome, responsible for scaffolding and structural integrity of the ribosome (Evguenieva-Hackenberg 2005). 16S rRNA genes are a unique marker because it is completely conserved in all prokaryotes, but has hypervariable (V) regions that vary between organisms (Figure 5A, B). This allows for 16S rRNA genes to be used as phylogenetic tool to allow for classification of bacteria and archaea based on the sequence of the 16S rRNA genes (Evguenieva-Hackenberg 2005). The most effective way to target different V regions within the 16S rRNA genes for sequencing is through PCR amplification with designed primers specific for the conserved regions.

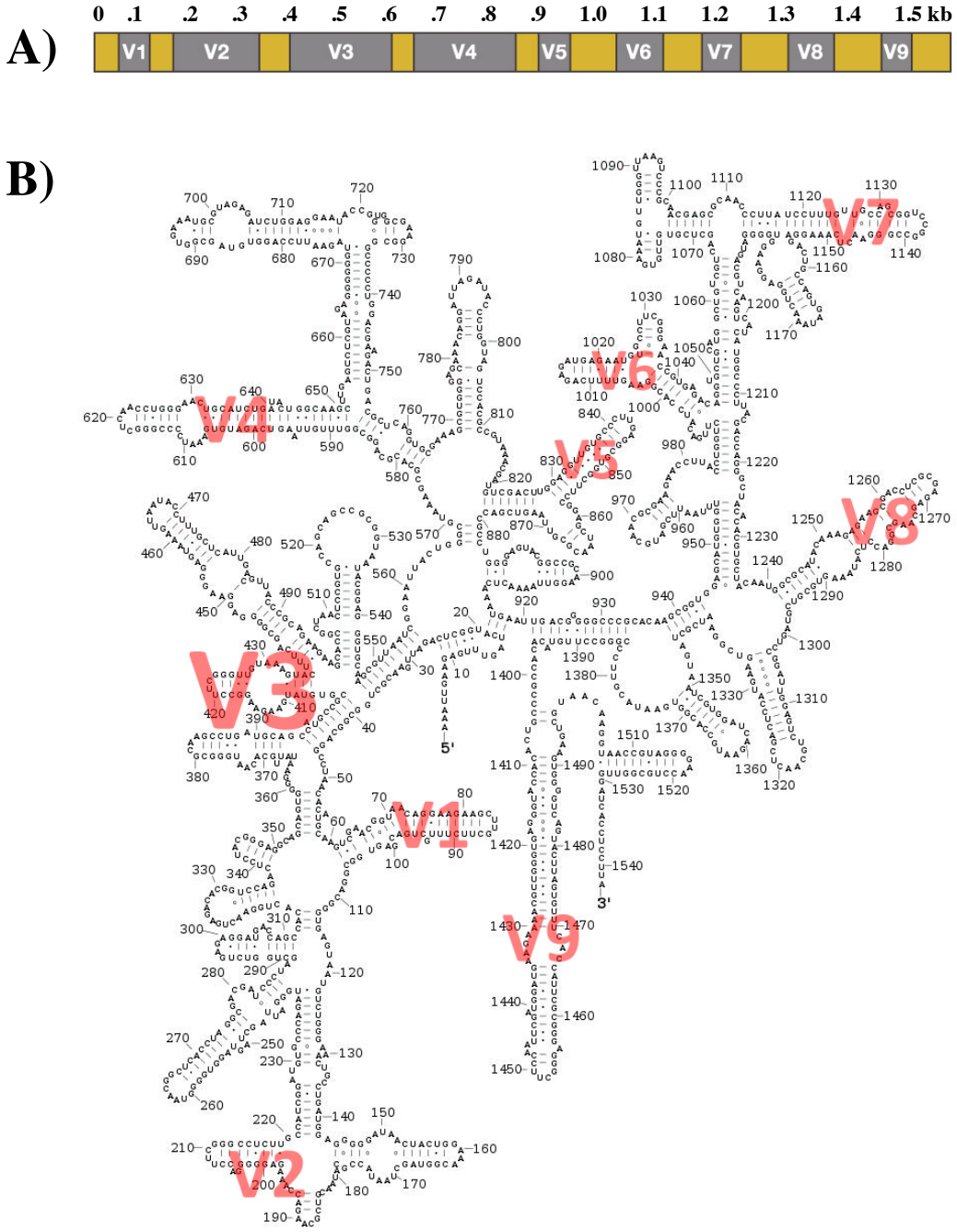

Figure 5. 16S rRNA. A) Delineation of hypervariable (V) regions in 16S rDNA. V1-V9 represents the hypervariable regions, linked by conserved orange regions. Base pair sizes are marked. B) 16S rRNA secondary structure. V1-V9 locations are indicated with size-scaled labels. Modified from Lambert *et al.* (Lambert *et al.* 2006)

Polymerase Chain Reaction (PCR). PCR is capable of exponentially amplifying billions of copies of a particular gene or DNA sequence of interest by exploiting the stability changes of double stranded DNA at different temperatures. The PCR requires a number of components: a template strand of DNA, a polymerase enzyme, ample amount of nucleotides for polymerization, and short oligonucleotide primers to guide the amplification process (Mullis 1990; Arnheim and Erlich 1992; Mullis *et al.* 1992). In brief, the process involves cycling between temperatures: a high temperature will denature the dsDNA into ssDNA, a lower temperature will allow primer annealing, and a slightly higher temperature for strand extension by the polymerase (Figure 6).

Primer design. The primers are largely responsible for the success of a PCR by binding to particular DNA sequences and directing the reaction. Variation in the design of the primers dictates the level of specificity of the amplification reaction. The design is guided by the goal of the experiment, whether it is to amplify DNA from every microorganism in a community or to amplify a DNA sequence from one species.

The most effective way to designing primers is through the use of bioinformatics tools. Primer design programs such as Primer3 (Rosen and Skaletsky 1998), sequence comparison programs such as BLAST (Altschul *et al.* 1990), or combination programs such as primerBLAST (Ye *et al.* 2012) allow for a researcher to design and predict the specificity of primers. Universal primers are designed to bind to conserved regions and amplify regions that vary in organisms by using degenerate sequences (Jaric *et al.* 2013). This is generally the case for most primers designed to target 16S rDNA for phylogenetic or metagenomic studies (Wang and Qian 2009). Alternatively, primers

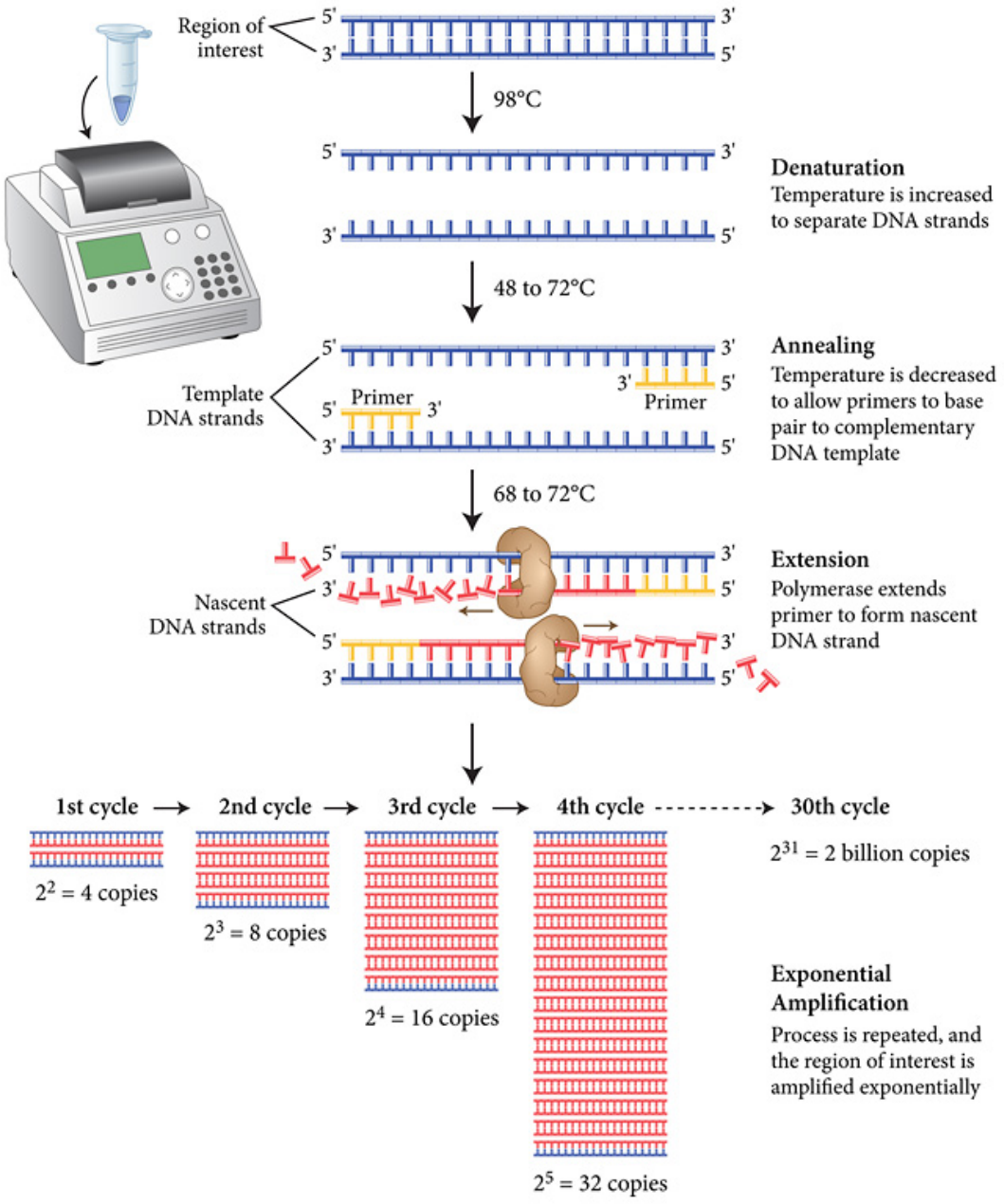

Figure 6. Schematic of PCR. The cycling of the denaturation, annealing, and extension steps results in amplification of the desired product. Exact temperatures for each cycle depend on the template sequence, the primer sequences, and the optimal working temperature for the polymerase. (http://www.neb.com)

designed for amplifying sequences from a specific organism must bind only to DNA sequences unique to that organism**.** Species or genus-specific primers are used in many applications, including the detection of pathogenic agents in the environment. In some cases these primers can amplify 16S rDNA sequences or *gyrB* sequences for downstream sequencing, but chromosomal signature sequences must be found if no sequencing step is used.

VI. Detection of Biological Warfare Agents

Real time qPCR. The weaponization of biological warfare agents has heightened the importance of finding quick and sensitive detection assays. Real-time quantitative PCR (qPCR) assays are a common option for detecting biological warfare agents by nucleic acid sequence (Thavaselvam and Vijayaraghavan 2010). The two most common types of qPCR assays are TaqMan and SYBR-Green assays (Life Technologies, Grand Island, NY). Although based on the same principle of integrating PCR with fluorescence, TaqMan assay relies on primers and a fluorescent probe that bind to the template, while SYBR-Green assay utilizes primers and a fluorescent dye that intercalates with dsDNA (Figure 7).

Regardless of the method used, targeting specific DNA sequences that are unique can provide a sensitive and specific way to test for the presence that organism (Arikawa *et al.* 2008). A number of studies have developed real-time qPCR assays that can identify important biological warfare agents such as *B. anthracis, F. tularensis* and *Y. pestis*

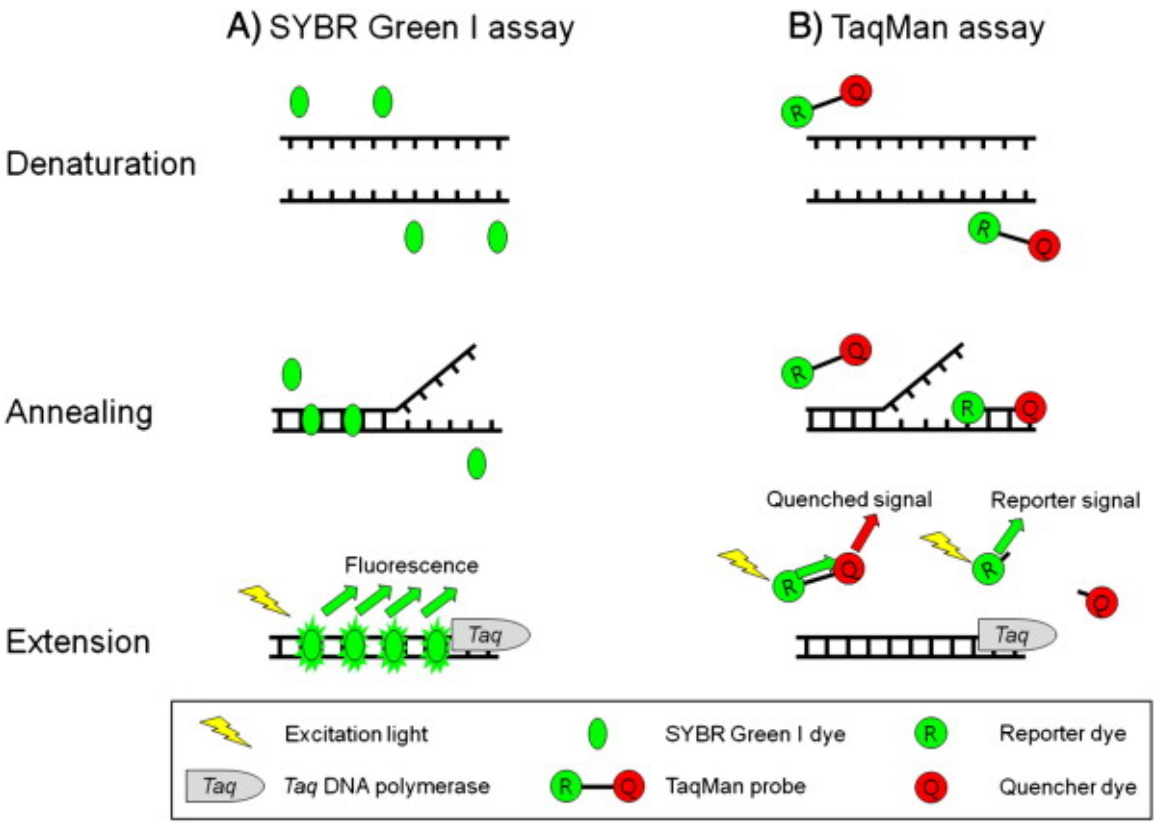

Figure 7. Comparison of SYBR-Green and TaqMan qPCR assays. (Kim *et al.* 2013)

(Skottman *et al.* 2007; Janse *et al.* 2010). *B. anthracis* qPCR assays commonly target pXO1 and pXO2 virulence plasmids (Janse *et al.* 2010) or species- specific sequences, such as those in the *pag* and *cap* genes (Skottman *et al.* 2007). Similarly, qPCR assays designed to detect *Y. pestis* frequently target sequences in the *pla* gene (Skottman *et al.* 2007).

Microbiological techniques. Other detection methods aside from real-time qPCR have been developed and will be summarized here. The first and oldest detection method is performed by observing culture phenotype and morphology and the results of various biochemical challenges (Thavaselvam and Vijayaraghavan 2010). For example, *B. anthracis* can be identified through the following observations: visually matte, flat, white cultures on nutrient agar, non-hemolytic, non-motile, penicillin sensitive and gamma phage sensitive with encapsulation observed on nutrient agar with 0.7% bicarbonate incubated in 5-20% CO_2 environment (Turnbull 1999). The disadvantage of culturing is that live cells are required, and any genetic changes or mutations within the bacterium could skew the observations and result in misidentification. This is especially poignant in the case of *B. anthracis*, which is very hard to distinguish from other *B. cereus* group organisms (Ash *et al.* 1991; Sacchi *et al.* 2002). Moreover, this method requires time, specialized equipment, and expertise to be performed properly.

Immunological assays. Another option for detection is using antibody-antigen assays such as enzyme-linked immunosorbent assay or ELISA (Andreotti *et al.* 2003; Thavaselvam and Vijayaraghavan 2010). Immunoassay methods are quite highly developed for detection of *B. anthracis, Y. pestis, Clostridium botulinum, Brucella* spp., *Burkholderia mallei* and *Burkholderia pseudomallei* (Thavaselvam and Vijayaraghavan 2010). The strength and specificity of ELISA assays are directly related to the affinity of the antibody-antigen interaction. This can be problematic if the affinity is low, because more substance must to be present as compared to other methods (Thavaselvam and Vijayaraghavan 2010).

Emerging technologies. The current trajectory of efforts seems to be towards speed enhancement and miniaturization for field use without sacrificing specificity and sensitivity. Miniaturization of nucleic acid detection systems (i.e., qPCR) in cartridges or glass capillaries are being developed, some already commercially available (Thavaselvam and Vijayaraghavan 2010). Also on the rise is the use of 'biosensors', or a physical component that is integrated into a biological system to aid in transmission of a measurable signal. A prime example of biosensor integration is the Autonomous Pathogen Detection System (APDS) (Lawrence Livermore National Laboratory), which utilizes Luminex beads, conjugated to specific antibodies for particular biological warfare agents, to be detected via flow cytometry followed by a secondary PCR confirmatory assay (Hindson *et al.* 2005). According to the developers, the system has been proven to be autonomous and accurate in field-tests performed around the country. Implementation of such technology could greatly improve public safety without requiring regular monitoring and testing by laboratories unless necessary.

VII. Surrogate Organisms of Biological Warfare Agents

Surrogates play an important role in research of biological warfare agents and highly pathogenic organisms. They allow for research to be conducted in laboratories not properly equipped or certified to experiment with highly virulent organisms (Saikaly *et al.* 2007). Ideally, organisms used as surrogates are nearly identical to the target organism in a variety of parameters such as morphology, genetic similarity, physiology, and environmental stimuli response (Greenberg *et al.* 2010). *Bacillus thuringiensis* and

Serratia marcescens have previously been used as surrogates for biological warfare agents *B. anthracis* and *Y. pestis*, respectively (Saikaly *et al.* 2007; Greenberg *et al.* 2010), and are used in my research.

B. thuringiensis. *Bacillus thuringiensis* is a Gram-positive spore-forming bacterium that is found in a variety of environments (Schnepf *et al.* 1998). *B. thuringiensis* is also a member of the *B. cereus* group of organisms, a phylogenetically and functionally similar group notoriously difficult to differentiate. Although other *Bacilli* spp. such as *Bacillus atrophaeus* and *Bacillus cereus* have been used as surrogates for *B. anthracis*, research has demonstrated *B. thuringiensis* closely resembles *B. anthracis* with regard to physiology, spore formation, and overall similarity (Greenberg *et al.* 2010; Justice 2010).

S. marcescens. *S. marcescens* is a ubiquitous Gram-negative bacterium characterized by red-colored colonies (Hejazi and Falkiner 1997). *S. marcescens* is an *Enterobacteriaceae* member that is occasionally implicated in nosocomial infections of the urinary tract, open wounds, and orbital cavities (Hejazi and Falkiner 1997). Its phylogenetic similarity to the *Yersinia* genus has resulted in its usage as a surrogate organism for *Y. pestis* for some time (Saikaly *et al.* 2007). Recently, *S. marcescens* has been utilized as a *Y. pestis* surrogate in studies investigating persistence in environments such as indoor air circulation systems, building debris leachate, and landfill leachate (Saikaly *et al.* 2007; Saikaly *et al.* 2010).

VIII. qPCR Identification of Biological Warfare Agent Surrogates

Real-time quantitative PCR (real-time qPCR) is a common method for detecting biological warfare agents and surrogate organisms. Assays can be designed to target plasmid-bound sequences, chromosomal sequences, or both (Qi *et al.* 2001). Plasmid-bound sequences are less suited for precise quantification due to variation in plasmid copy numbers between bacteria of the same species. However, targeting chromosomal markers allow for more exact quantification due to stability of copy number between organisms of the same species (Qi *et al.* 2001). Furthermore, attenuated or altered strains of pathogenic bacteria can be lacking in one or more of the virulence plasmids, yet still retain significant pathogenic and infectious capabilities (Ivins *et al.* 1986; Hadjinicolaou *et al.* 2009).

Experimental designs for specific identification of *B. thuringiensis* have primarily centered on detection of plasmid-bound *cry* and *cyt* genes, which are responsible for production of parasporal crystals and unique to *B. thuringiensis* (Porcar and Juarez-Perez 2003). Though chromosomal-bound markers are used for qPCR assays targeting *B. thuringiensis*, none provided sufficient distinction of *B. thuringiensis* from other closely related *B. cereus* group organisms (Bavykin *et al.* 2004). Previous studies have identified putative species-specific primers targeting the chromosome of *S. marcescens* in *gyrB*, *wzm,* and *recA* genes (Saikaly *et al.* 2007), but the specificity of the primers and probes is largely untested.

IX. Complex Environmental and Clinical Samples

Bacteria do not live in solitude. Environmental and clinical samples are often complex mixtures, containing a multitude of bacteria, archaea, viruses, and fungi in one sample. As such, assays used to identify biological warfare agents must be effective when the target organism is in a complex mixture or matrix. Basic probability and statistical models have shown that the 'quality' of a designed signature sequence varies from sample to sample, introducing an inherent but unavoidable bias (Reed *et al.* 2007). Additionally, elements such as mutation rate, size, and diversity of microbial background can potentially impose interferences effecting detection of a single species (Reed *et al.* 2007). Minimal research has been performed on the subject, however one study reports that increasing concentration of non-target DNA (from *Pseudomonas fluorescens)* has an inhibitory effect on amplification of target DNA from *E. coli* (Ludwig and Schleifer 2000). Moreover, substances that inhibit PCR can be present in environmental or clinical samples and have an effect on downstream assays (Opel *et al.* 2010). Although various extraction protocols can deal with many of these inhibitors, some can persist to disrupt experiments and skew the outcome.

The biases described above could be negligible in many cases. However, in instances of low amounts of template DNA concentration or significant interference, this bias could result in false negative detection. This bias can be evaluated by testing the specificity and selectivity of an assay designed to detect a single species in isolation compared to environments with high levels of potential interferences, such as metagenomic DNA extracted from a microbial community.

X. Chronic Obstructive Pulmonary Disease Lung Microbiome as a Model System

Chronic obstructive pulmonary disease, or COPD, is a lung disease typified by emphysema-like destruction of alveoli, airway inflammation, airway fibrosis, and increased production of mucus (Sethi 2000; Sethi and Murphy 2001; Sethi *et al.* 2009; Erb-Downward *et al.* 2011). COPD can be fatal, responsible for the deaths of approximately 2.5 million people worldwide in 2000 (Murray *et al.* 2001). The disease can be exacerbated by smoking or air pollution, leading to impaired mucocilliary clearance and weakened immune response (Lopez *et al.* 2006).

While the microbiome of the COPD lung is still being characterized (Erb-Downward *et al.* 2011), it is clear that a significant exacerbation is correlated with bacterial infection. The damaged and unnatural environment associated with COPD lungs has the propensity to promote bacterial colonization. The enhanced growth of bacteria within the lungs of COPD patients is unique, as a healthy lower respiratory tract is generally sterile, or colonized by non-pathogenic microorganisms (Sethi 2000; Sethi and Murphy 2001).

From a forensics standpoint, the COPD lung microbiome is a complex, heterogeneous community that can be used as testing grounds for detection assays of biological warfare agents. The lung is the main route-of-entry for important biological warfare agents such as *B. anthracis* and *Y. pestis*, giving real-world significance to the usage of the COPD lung microbiome DNA as a contaminant. Assay specificity, sensitivity, and influence of the background can all be effectively tested using a COPD lung microbiome model.

XI. Aims

Real-time quantitative polymerase chain reaction (real-time qPCR) has been used to detect biological warfare (BW) agents such as *B. anthracis* and *Y. pestis* in a variety of environments (Skottman *et al.* 2007; Janse *et al.* 2010). However, little is known about the effect of the immediate environment on the efficacy of the qPCR, particularly if these bacteria are amongst a complex microbial community, such as bronchoalveolar lavages (BALs) from the lungs of COPD patients.

In this research, we seek to explore the potential role of the heterogeneous microbial community DNA on the sensitivity of detecting biological warfare agent using surrogate organisms of *B. anthracis* and *Y. pestis, B. thuringiensis* and *S. marcescens,* respectively. The study will be organized into three specific aims:

1. Design species-specific primers targeting *B. thuringiensis* and *S. marcescens* chromosomal signature sequences
2. Test the specificity of the designed primers
3. Examine sensitivity and efficacy of primers when amplifying template DNA in the presence of DNA from COPD BALs via qPCR

Evaluation of the newly designed primers and amplification between conditions in a quantitative manner could provide some insight to the practical importance of background composition on single species detection, particularly in instances of low template DNA concentration or high amounts of non-target DNA.

XII. Hypothesis

Real-time qPCR detection of *B. thuringiensis* and *S. marcescens* will be significantly inhibited by the presence of sufficient pulmonary microbiome DNA.

METHODS

I. *In silico* Primer Design for *B. thuringiensis* and *S. marcescens*

Candidate marker sequences. For *B. thuringiensis* serovar *kurstaki* HD-73, the search for markers began by identifying DNA signatures using the Insignia program (Figure 8), created by the Center for Bioinformatics and Computational Biology at the University of Maryland (Phillippy *et al.* 2009). The open-source program allows for searching of signature sequences by comparing the genome of the desired organism with its large database of sequences from many pathogenic bacteria and microorganisms. Search parameters were set to exclude short sequences and include larger signature chains (<100 bp), as longer sequences often result in more dissimilarity of a reference genome to the background.

Figure 8. Primer Design Workflow

For *S. marcescens* strain ATCC 13880T, the exploration for markers began by obtaining the complete genome sequence of *S. marcescens* Db11 (Accession # PRJEB2662, The European Bioinformatics Institute). The sequence is publicly available in an unannotated shotgun-library format. General areas of interest (such as *gyrB*) were

manually located, divided into arbitrary sequence segments, and used as input for the next step. Previous research has shown *gyrB* to be more effective in discriminating between *Serratia* species than 16S rDNA or other markers (Dauga 2002).

Primer design and analysis. From both search methods, candidate sequences of large size (300 bp or over) were screened using the NCBI primerBLAST program (Ye *et al.* 2012) (Figure 8). The screening process fulfilled several essential functions. It identifies primer pairs using the Primer3 software while simultaneously testing for selectivity of the primer on the target genome (≤ 3 base pair mismatches for a 20-mer primer pair). The primer pairs are automatically input into NCBI BLAST, which compares the sequences against the GenBank database to produce a list of potential non-specific binding issues indicative of potential specificity loss. The primer pairs were then analyzed using the Beacon Designer Free Edition (Premier Biosoft, Palo Alto, CA). The program calculates and visually presents attributes such as T_m, GC %, probability of self-dimerization, cross dimerization, and hairpin formation, allowing quantitative and qualitative comparison of candidate primer pairs.

Universal primers. In addition to specific oligonucleotide primers, high degeneracy universal primers were used as controls (Jaric *et al.* 2013). The universal primer pair (designed by Jaric and Narasimhan, BIORG, Florida International University) used targets 16S ribosomal RNA [rRNA] hypervariable regions of the genome, specifically the region spanning V6 –V8 regions, which were experimentally determined

to provide the most coverage and universality (Jaric *et al.* 2013). The finalized primer pairs and appropriate information can be found in Table 4.

II. Strains and Culturing

The bacterial strains used in the study are listed in Table 3. The strains were chosen based on availability and similarity to the two primary strains, *B. thuringiensis* serovar *kurstaki* HD-73 and *S. marcescens* ATCC 13880T. *Bacillus cereus* UW85, *B. megaterium* ATCC 14581, *B. subtilis* ATCC 23857 were selected due to their close relationship with *B. thuringiensis,* as all are members of the aforementioned *B. cereus* group organisms (Bavykin *et al.* 2004). Additionally, *P. aeruginosa* PA01, and *H. influenzae* ATCC 51907 were selected for their similarity to *S. marcescens*. Both Gram-negative bacteria commonly found in lungs and are capable of causing infection (Musser and Beamer 1961; Musher 1983). All strains were grown in standard Luria-Bertani (LB) Media (Fisher, Pittsburgh, PA) and cultured using aseptic techniques. Cultures for *S. marcescens*, *B. megaterium*, *B. cereus*, and *B. thuringiensis* were grown at 30°C while *P. aeruginosa* and *B. subtilis* were grown at 37°C. Sources for strains are noted in Table 4. Freezer stocks (-80°C) for all strains were maintained in 1:1 LB/glycerol with the exception of *P. aeruginosa*, which was maintained in 10% skim milk.

Table 3. Organisms and strains utilized in the study.

Organism	Strain I.D.	Source
Serratia marcescens	**ATCC 13880**	**Lab available**
Bacillus thuringiensis	**HD-73, ATCC 35866**	**Purchased from ATCC**
Bacillus megaterium	**ATCC 14851**	**Dr. Barry Rosen's Laboratory**
Bacillus subtilis	**ATCC 23857**	**Dr. Barry Rosen's Laboratory**
Bacillus cereus	**UW85**	**Dr. Barry Rosen's Laboratory**
Pseudomonas aeruginosa	**PAO1**	**Lab available**
Haemophilus influenzae (genomic DNA)	**ATCC 51097**	**Lab available**

III. Primer Specificity

Single colonies from both *B. thuringiensis* and *S. marcescens* (and all other strains used) were subjected to phenol/chloroform genomic extraction protocols. Amplification using all the primers was tested on other closely related bacteria and clinically relevant bacteria in the strain panel (Table 3) All extracted DNA from all the strains was subjected to PCR with chosen primers (Table 4), and resulting amplification was visualized on an agarose gel.

IV. Chromosomal DNA Extraction

The strain of interest was grown overnight in 5 mL of LB broth. To ensure highest possible yield, a fresh culture was grown to an optical density (OD) of 0.8 for use in the extraction process. One mL of fresh liquid culture was moved into a 2-mL

Table 4. Primer sequences used in this study.

Target gene	Primer	Primer sequence (5'-3')	Tm (° C)	%GC	Amplicon length (bp)	Accession number (position)	Source
B. thuringiensis hypothetical protein	BTK_1F	GCAAGCAAGGCTGATGGGGCA	63.4	61.9	138	NZ_CM000751 (2154-2291)	This study
	BTK_1R	CCACCAGGTGTCGTCTCGGGAT	63	63.6			
B. *thuringiensis* FAD - dependent oxidase	BTK_2F	TGGGCAGAAGCAGCAGCGAA	63.1	60.0	194	NZ_CM000751 (250-444)	This study
	BTK_2R	ACCTCAGCAACAGCACCGCC	63.7	65.0			
S. marcescens partial DNA gyrase B subunit	SM_1F	CCGCGAAGGCAAAGTGCACGAACA	64.5	58.3	113	AJ300536.1 (101-224)	This study
	SM_1R	G(G/C) CTTGGCCAGAAGCGCACCATAG	64.5	62.5			
S. marcescens partial DNA gyrase B subunit	SM _2F	AGTGCACGAACAAACTTACAG	53.4	42.8	138	AJ300536 (113–251)	Saikaly et al., 2007
	SM_2R	GTCGTACTCGAAATCGGTCACA	56.7	50.0			
Universal 16s rRNA V6-V8	MJ_68F	TGCATGGWWGTCGTCAGC	58.6	61.1	320	N/A	Jaric et al., 2013
	MJ_68R	TGTGTACAAGWCCCGWGAACG	59.4	57.1			

microcentrifuge tube and centrifuged for 5 minutes at 13,300 rpm. The supernata removed and the pellet was resuspended in 0.5 mL Bacterial Lysis Buffer B1 (5 Tris•Cl pH 8.0; 50 mM EDTA pH 8.0; 0.5% Tween 20; 0.5% Triton-X100, Qiagen was followed by the addition of 2 µL RNase A, 20 µL of lysozyme (100 mg/mL) a µL Proteinase K (20 mg/mL), then incubated at 37°C for 30 minutes. Then, 175 Bacterial Lysis Buffer B2 (3 M guanidine hydrochloride, 20% Tween 20, Qiager added to the sample, mixed by inversion, and incubated at 50°C for another 30 minu

To extract the bacterial DNA, an equal volume of phenol-chloroform so (742 µL) was added to the lysed bacteria and vortexed for 10 seconds. The tub centrifuged for 15 minutes at 13,300 rpm to fully separate the biphasic mixtur isolate the nucleic acids from protein, extraneous RNA, and lipid mole Subsequently, the aqueous phase (top layer) containing the nucleic acids was trans to a new 2-mL microcentrifuge tube. To precipitate the DNA, $1/10^{th}$ of the volum M sodium acetate (NAOAc, CH_3COONa) was added, followed by an addition volumes of ice-cold ethanol. The tube was vortexed thoroughly and placed in overnight.

The tube was centrifuged for 5 minutes at 13,300 rpm. The supernatan discarded, the pellet was washed with 70% ethanol, and then fully dried using a va centrifuge. The pellet containing DNA was subsequently resuspended in 50 µL of 1 Tris elution buffer (EB) and stored at 4°C or -20°C for downstream use. The DN quantified via UV spectrophotometry on the BioTek Synergy HT Multi-Det Microplate Reader (Winooski, VT) and a Take 3 microplate (BioTek, Winooski, VT

V. Species-Specific PCR Amplification of *B. thuringiensis* and *S. marcescens*

A single PCR was designed and optimized to amplify genomic DNA from *B. thuringiensis* or *S. marcescens* using the putative species-specific primers in Table 4. As such, the PCR reaction also served to test for non-specific amplification against gDNA from all other available strains (Table 4). Final concentrations and volumes of the reaction mixture are as follows: 2 µL of 10X PCR Buffer (15 mM $MgCl_2$) (Qiagen), 0.4 µL of 10 µM dNTPs, 0.4 µL of forward and reverse primers (10 µM), 0.06 µL HotStarTaq *Plus* (Qiagen), 1 µL of 10 ng/µL DNA template, and 16 µL of PCR water.

All amplifications were performed on either a MJ Research PTC-200 Peltier Thermal Cycler (MJ Research, Waltham, MA) or a Techne Touchgene Gradient PCR Thermal Cycler (Techne, Burlington, NJ). The amplification parameters used were as follows: 95°C for 5 minutes initial denaturation and HotStarTaq *Plus* activation, followed by 30 cycles of denaturation at 95°C for 20 seconds, primer annealing at 50°C for 30 seconds, and extension at 72°C for 1 minute, followed by a final elongation step at 72°C for 10 minutes. Negative controls were performed without addition of template DNA to observe possible contamination. Positive control amplification of *P. aeruginosa* with 16S universal primers (Table 4) was used to confirm amplification success.

VI. Agarose Gel Electrophoresis

A 1% or 2% (weight/volume) agarose electrophoresis gel (M.G. Scientific J.T. Baker® Agarose, Standard, Low Electroendosmosis, Pleasant Prairie, WI) was run to

determine the amplification efficacy of the PCR reaction. Electrophoresis was carried out in 0.5X Tris-Borate-EDTA (TBE) buffer (0.89M Tris-Borate, 0.02 M EDTA, pH 8.0). A 100-bp DNA ladder was loaded as size reference on all gels (Promega, Madison, WI). Gels were stained with Invitrogen SYBR® Safe DNA Gel Stain (Life Technologies, Grand Island, NY) and viewed and photographed under UV trans-illumination on a KODAK Gel Logic 2200 Imaging System (Carestream Health, Rochester, NY).

VII. Bronchoalveolar Lavage (BAL) DNA Samples

Total DNA was isolated from bronchoalveolar lavage (BAL) samples from five subjects collected under the IRB approved protocol #062810-01. Protocol #062810-01 is a collaborative project with Dr. Adam Wanner and colleagues at the University of Miami (UM). All subject interaction (recruitment, consent, BAL sampling) was performed at the UM site. Samples for this project were spare samples that would otherwise have been discarded, providing no barrier to the completion of the aforementioned project. BAL samples from five subjects were selected for this project.

BAL DNA extraction. Total metagenomic DNA from BAL samples was extracted using a FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, OH) and a modified version of the complementary protocol. First, 978 µl of BAL sample was transferred to a Lysing Matrix E tube and 122 µl MT Buffer (90-100% H_2O, 1-5% $C_{10}H_{14}N_2Na_2O_8 \cdot 2H_2O$, 1-5% $NaC_{12}H_{25}SO_4$, 1-5% $(C_6H_9NO)_n$, <1% Proprietary inorganic salts, MP Biomedicals) was added. If the sample had less than 978 µl, the

remaining volume was supplemented with sodium phosphate buffer (90-100% H_2O, 1-5% NaH_2PO_4) provided in the kit. The sample was homogenized in the FastPrep instrument for 40 seconds at a speed setting of 6.0, then centrifuged at 13,300 rpm for 10 minutes to pellet the cellular debris. The supernatant was transferred to a clean 2-mL microcentrifuge tube, then 250 µl protein precipitation solution (60-70% H_2O, 30-40% NaOAc, 1-5% $C_2H_4O_2$, MP Biomedicals) was added and mixed by shaking the tube 10 times. The sample was centrifuged again at 13,300 rpm for five minutes to pellet the protein precipitate. The supernatant was transferred to a clean 15-mL polypropylene tube.

To isolate the DNA, 1 mL of resuspended Binding Matrix suspension solution (60-70% $C_2H_6N_4S$, 10-20% H_2O, 10-20% MP Silica, MP Biomedicals) was added to the supernatant. To allow sufficient binding of DNA to the silica matrix, the tube was placed on a rotator at slow-medium speed for two minutes and the tube was placed in a rack for three additional minutes. The mixture was then re-suspended and 600 µl was added to a SPIN Filter and Catch Tube apparatus and centrifuged at 13,300 rpm for one minute. The contents of the catch tube were discarded: this cycle of adding mixture in 600 µl increments, centrifuging, and emptying the catch tube continued until all of the mixture has been passed through the filter.

A 500 µl aliquot of prepared Salt/Ethanol Wash Solution (SEWS-M) (80-90% H_2O, 10-20% Tris-HCl, MP Biomedicals) was added and the pellet above the SPIN filter was gently re-suspended using the force of the liquid from the pipet tip. The tube was centrifuged from one minute at 13,300 rpm and the catch tube was emptied. To ensure complete ‘drying’ of the matrix of residual wash solution, the tube was centrifuged again for two minutes at 13,300 rpm. However, now the catch tube is discarded and replaced

with a new, clean catch tube. The SPIN filter was air-dried for 5 minutes at room temperature; the DNA/Binding matrix mixture above the SPIN filter was re-suspended in 100 µl of DES (DNase/Pyrogen-Free Water). To increase yield and elution efficiency of DNA from the Binding matrix, the tube was incubated for five minutes at 55˚C in heat block or water bath.

The tube was centrifuged at 13,300 rpm for one minute to bring eluted DNA into the clean catch tube. The resulting DNA was ready for PCR and other downstream applications and stored at -20˚C for extended periods or 4˚C for shorter periods. The DNA was quantified via UV spectrophotometry on the BioTek Synergy HT Multi-Detection Microplate Reader (Winooski, VT).

VIII. Real-time qPCR Sensitivity Experiments

The sensitivity and specificity of the primers were tested using pure template DNA and DNA spiked with metagenomic DNA extracted from five BAL samples. All qPCR experiments were performed using the Applied Biosystems 7500 Real-Time PCR System (Madison, WI). Reaction mixtures for each 20 µl volume contained 7 µl MilliQ H_2O, 10 µl 2x SYBR® Green PCR Master Mix, 1 µl template DNA, 1 µl of 10 nM forward primer, and 1 µl of 10 nM reverse primer. The cycling parameters were: 50˚C for 2 minutes, 95˚C for 30 seconds (initial denaturation), 40 cycles of (denaturation at 95˚C for 15 seconds and primer annealing/extension at 55˚C for 1 minute), 95˚C for 1 minute. Fluorescence data was automatically collected after each amplification cycle. Additionally, a melt curve analysis was performed immediately following the

amplification, beginning at 60˚C, increasing by 0.5˚C increments every 30 seconds until 95˚C. Fluorescence data was collected at every temperature increment increase.

***B. thuringiensis* gDNA spiking experiments.** The C_t values were determined for amplification of three concentrations of *B. thuringiensis* (0.2, 0.04, 0.008 ng/µL) of genomic DNA (gDNA) with the BTK1_FR primer pair when BAL DNA extracts (BAL 1-5) were added. BAL DNA concentration was normalized so that approximately 200 ng was present in each sample. All experiments were run in triplicates. A standard curve was produced using 1:5 dilutions of *B. thuringiensis* gDNA (1, 0.2, 0.04, 0.008, and 0.002 ng/µL concentrations). All runs included negative controls.

***S. marcescens* gDNA spiking experiments.** Similarly, C_t values were determined for amplification of three concentrations of *S. marcescens* (0.2, 0.04, 0.008, ng/µL) gDNA with the SM1_FR primer pair when BAL DNA extracts (BAL 1-5) were added. BAL DNA concentration was normalized so that approximately 200 ng was present in each sample. All experiments were run in triplicates. A standard curve was produced using 1:5 dilutions of *S. marcescens* gDNA (1, 0.2, 0.04, 0.008, and 0.002 ng/µL). All runs included negative controls.

IX. Data Analysis

All data were analyzed using both AB 7500 Software (v. 2.0.4) and R- language programming scripts. The R scripts used AB 7500 Software raw data output as input data

to visualize results using boxplots and test for significance using Analysis of Variance (ANOVA) and post hoc Tukey Honestly-Significant Difference (HSD) tests. A template R script was constructed to facilitate easy manipulation for each data set (Figure 9). Results from all assays were compared to determine any significant differences in performance with different backgrounds (mispriming, non-specific amplification, or concentration dependent effects).

```
# Read in .csv data in specified manner
# BTK1_OneNG = Representative title of dataset
# One.ng.BT.DNA.Spike = Representative .csv column title
# Ct.Value = Representative .csv column title
# MIC 24 = BAL 1, MIC 32 = BAL 2, MIC 27 = BAL 3, MIC 50 = BAL 4, MIC 55 = BAL 5
BTK1_OneNG<-read.csv(file.choose(), header=T)
# Reorder the x axis data and labels from default
BTK1_OneNG$One.ng.BT.DNA.Spike<-factor(BTK1_OneNG$One.ng.BT.DNA.Spike, levels=c("STD",
 "BAL 1", "BAL 2", "BAL 3", "BAL 4", "BAL 5"), labels =c("STD", "BAL 1", "BAL 2", "BAL
 3", "BAL 4", "BAL 5"))
# Boxplot
boxplot(BTK1_OneNG $Ct.Value~BTK1_OneNG $One.ng.BT.DNA.Spike)
# Addition of titles and axis labels
title(main="BTK1_FR 1 ng B.thuringiensis Spike", xlab="Sample", ylab="Mean Cycle
 Threshold Value (Ct)")
# ANOVA
BTK1_OneNG.anova<-aov(BTK1_OneNG $Ct.Value~BTK1_OneNG $One.ng.BT.DNA.Spike)
# Visualize ANOVA results
summary(BTK1_OneNG.anova)
# Post Hoc Confirmation
TukeyHSD(BTK1_OneNG.anova)
```

Figure 9. Sample R-language Script. (#) indicates commentary, not executable script.

RESULTS

I. Primer Design

For species-specific PCR amplification of *B. thuringiensis* and *S. marcescens*, primers were either designed using the workflows described in the methods or identified in existing literature (Table 4).

A PCR was performed to evaluate the ability of the primer pairs to amplify their intended targets. A template mixture of 5 ng of gDNA from *S. marcescens*, *B. megaterium*, *B. cereus*, and *B. thuringiensis, B. subtilis* was created as the template for all reactions. Electrophoresis using a 2% agarose gel was run to visualize the amplification products (Figure 10).

Serratia marcescens DNA was successfully amplified using both the SM1_FR and SM2_FR primer pairs, as evidenced by the single product bands at the expected amplicon sizes, 113 bp and 138 bp, respectively (Figure 10, Lane 2 & 3). *Bacillus thuringiensis* DNA was also effectively amplified using BTK1_FR and BTK2_FR primer pairs, indicated by the major product bands at 138 bp and 194 bp, respectively (Figure 10, Lane 4 & 5). As expected, the 16S MJ68_FR primers amplified *B. thuringiensis* very successfully as indicated by the major product band at 320 bp (Figure 10, Lane 6). No-template control containing MJ68_FR produced no amplification products (Figure 10, Lane 7). No primer dimers were formed, as evidenced by the lack of any bands in the 30-50 bp range.

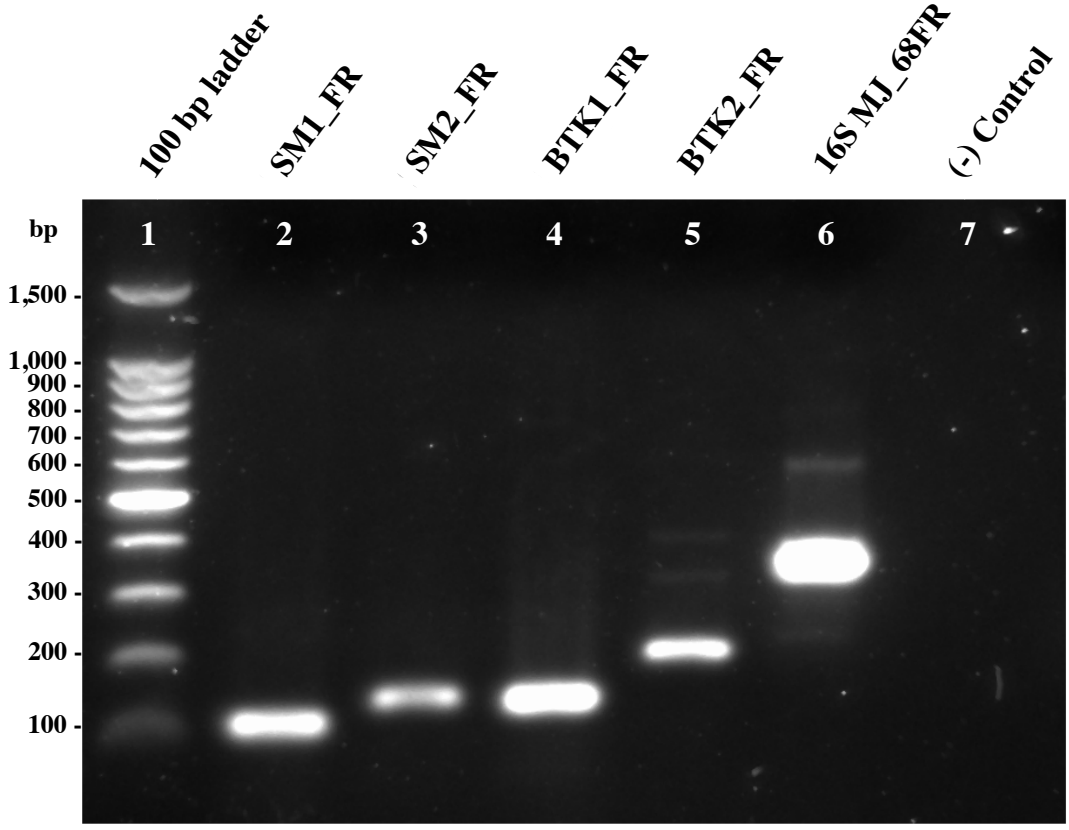

Figure 10. Gel electrophoresis of PCR products for each primer pair (Table 4). Template was a mixture of 5 ng/µL of gDNA extracted from each panel member. The DNA mixture was amplified using the respective species-specific primers (Table 4) and run on a 2% TBE agarose electrophoresis gel. Lane 1: 100 bp ladder, Lane 2: *S. marcescens* amplification with SM1_FR primers, Lane 3: *S. marcescens* amplification with SM2_FR primers, Lane 4: *B. thuringiensis* amplification with BTK1_FR primers, Lane 5: *B. thuringiensis* amplification with BTK2_FR primers, Lane 6: Mixed gDNA amplification with 16S MJ68_FR primers, Lane 7: MJ68_FR primers with no template DNA (- control).

II. Primer Specificity

The BTK1_FR, BTK2_FR, SM1_FR, and SM2_FR primer pairs were tested for specificity against a small panel of strains (Table 3). All strains were subjected to PCR amplification with each primer pair, including 16S primer pair MJ68_FR (Figure 11). The products were run on a 2% TBE agarose gel with a 100-bp ladder for size reference, and subsequently viewed under UV trans-illumination.

The BTK1_FR and BTK2_FR primer pairs both preferentially amplified *B. thuringiensis*, as noted by single product bands at the predicted amplicon sizes of 138 bp and 194 bp respectively (Figure 11, Lanes 22, 31). Moreover, BTK1_FR nor BTK2_FR primer pairs produced any amplification of DNA from *B. cereus, B. megaterium, B. subtilis, S. marcescens, P. aeruginosa, or H. influenzae* (Figure 11, Lanes 17-21, 23-30).

Similarly, the SM1_FR and SM2_FR primer pairs also amplified *S. marcescens* preferentially. Single product bands are clearly distinguishable at predicted amplicon sizes of 113 bp and 138 bp for SM1_FR and SM2_FR primer pairs, respectively (Figure 11, Lanes 2, 9). No amplification of DNA from *B. cereus, B. megaterium, B. subtilis, S. marcescens, P. aeruginosa, or H. influenzae* was observed with SM1_FR or SM2_FR primer pairs (Figure 11, Lanes 3-8, 10-16).

Additionally, the positive control primer pair (16S MJ68_FR) successfully amplified DNA from all strains (Figure 11, Lanes 33-39), indicating that the extracted DNA from each panel member was amplifiable. The negative no-template controls (NTC) for each primer pair produced no measurable amplification (Figure 11, Lanes 41-45)

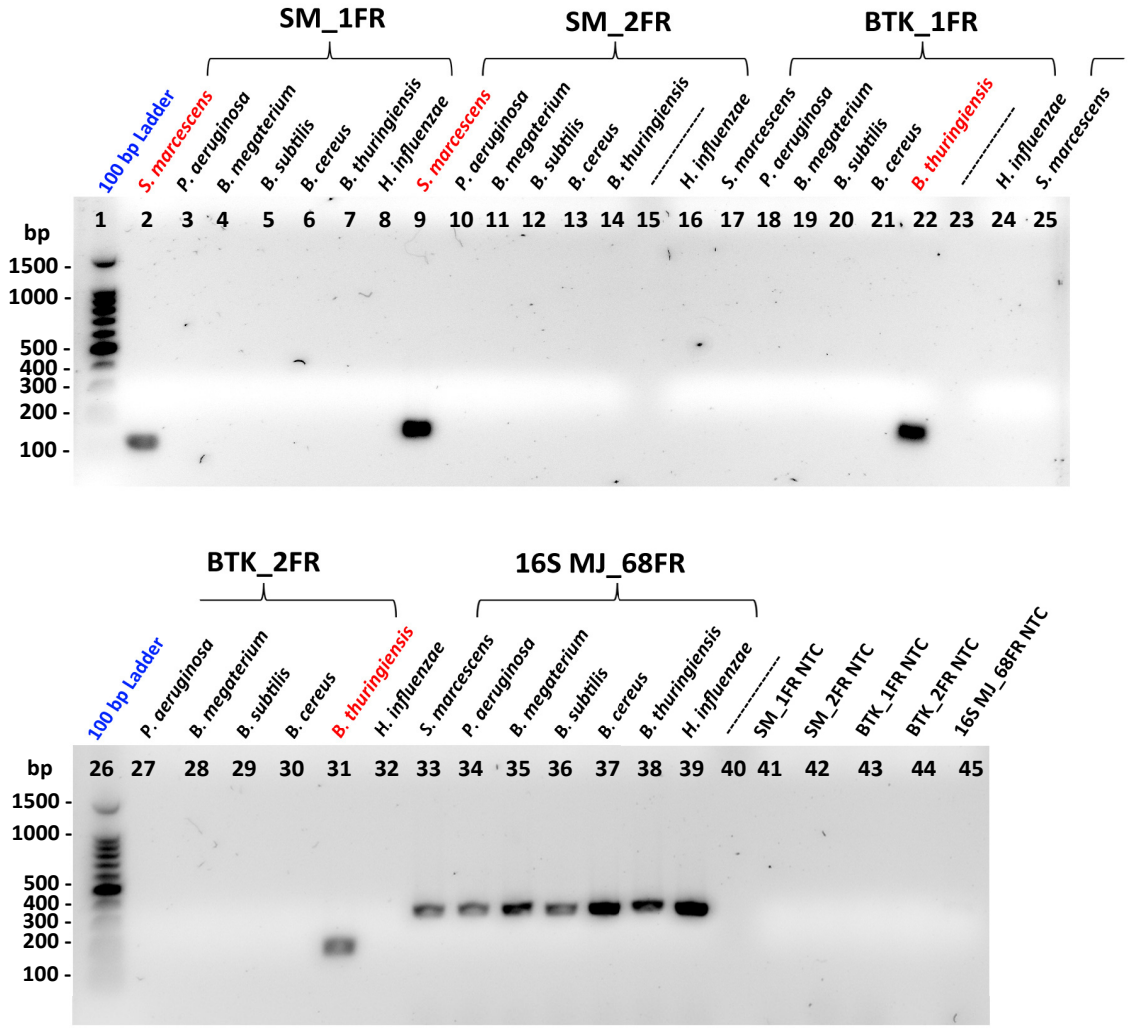

Figure 11. Determination of primer specificity. *B. thuringiensis (*BTK1_FR, BTK2_FR*)* and *S. marcescens* (SM1_FR, SM2_FR) primer pair specificity is tested against related and common lung microbes (Table 3). PCR products were visualized using gel electrophoresis (2% TBE agarose gel). Amplification with the universal primer pair 16S MJ_68FR was used as positive control.

III. Comparison of *B. thuringiensis*-specific Primers

Previous results demonstrated that *B. thuringiensis* DNA can be successfully amplified using the BTK1_FR or BTK2_FR primer pairs (Figure 10). However, the amplification with BTK1_FR (Figure 10, Lane 4) produces a thicker band as compared to BTK2_FR (Figure 10, Lane 5). In addition to the 194-bp amplicon, the amplification with BTK2_FR produced faint bands at 320 bp and 400 bp (Figure 10, Lane 5).

For a precise comparison of the BTK1_FR and BTK2_FR primer pairs, triplicate SYBR-green qPCR assays were performed using two different concentrations of template DNA (10 and 1 ng/µL) (Figure 12). Mean C_t values were determined at the point of intersection with the threshold value line (Figure 12) and reported in Table 5, along with threshold value. At 10 ng of *B. thuringiensis* DNA as the template, the amplification with BTK1_FR primers (C_t = 27.68) (Table 5) crossed threshold approximately five cycles before the amplification with BTK2_FR primers (C_t = 32.49) (Table 5).

Table 5. Comparison of BTK1_FR and BTK2_FR primers.

Primers	T_m(°C)	Template (ng/µL)	Mean C_t Value	Threshold (ΔRn)	Melt Curve Peak Height (-Rn)
BTK1_FR	77.69	10	27.683	.2	2.9-3
		1	31.033		2.5-2.6
BTK2_FR	77.5	10	32.494	.2	1.6-1.8
		1	36.64		0.7-0.9

* C_t values were calculated from amplification data, threshold values were set at 0.2, and T_m and peak heights were determined via software calculation and melt curve (Figure 13).

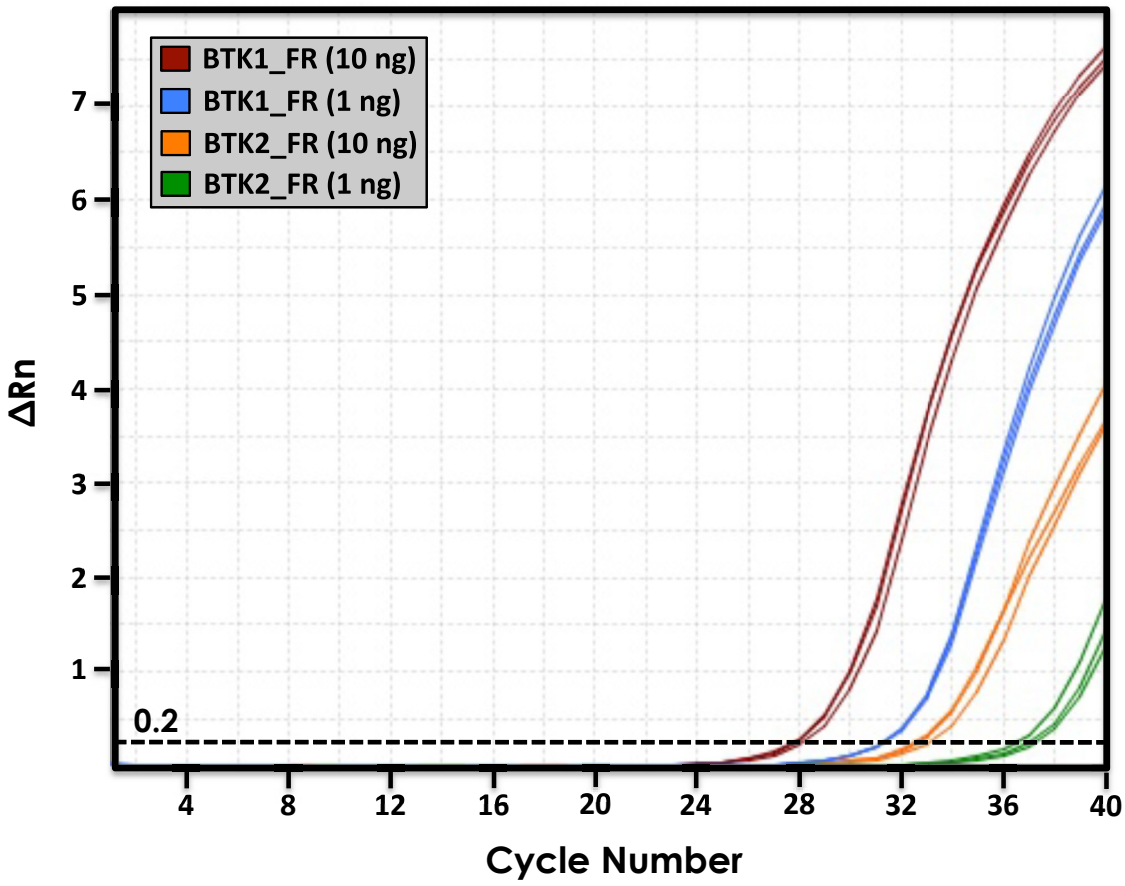

Figure 12. Sensitivity of the BTK1_FR and BTK2_FR primer pairs. Two different amounts (1 or 10 ng) of purified *B. thuringiensis* gDNA were used as the template for amplification. C_t values were determined as the intercept of the magnitude of the fluorescence signal (ΔRn) with the calculated threshold line (Table 5). The reported mean C_t values for a given sample are from triplicate runs.

Using a smaller amount of *B. thuringiensis* DNA as template (1 ng/µL) yielded comparable results, with the amplification with BTK1_FR primers (C_t = 31.03) (Table 5) reaching threshold over five cycles before the amplification with BTK2_FR primers (C_t = 36.64) (Table 5). At both concentrations, the differences in C_t values were found to be statistically significant via one-tail t-test ($p < .05$).

Post-amplification melt curve analyses were performed for both BTK1_FR and BTK2_FR primer pairs (Figure 13) with 10 and 1 ng/µL *B. thuringiensis* gDNA. The analysis determined T_m values and peak heights (Table 5) for each condition, representative of the temperature of amplicon dissociation and amount of fluorescence lost as a result. The T_m for BTK1_FR and BTK2_FR amplification products were similar, 77.69 ˚C and 77.5˚C, respectively (Table 5). Only single peaks associated with the T_m were observed, indicating the production of a single PCR product At a higher concentration of *B. thuringiensis* DNA (10 ng/µL), melt curve peak heights for BTK1_FR amplification products (Figure 13) are approximately two times higher than that of the BTK2_FR (Table 5). The difference in peak heights is increased for the lower *B. thuringiensis* DNA concentration (1 ng/µL) (Figure 13), with observed heights nearly three times higher for the BTK1_FR than the BTK2_FR amplification (Table 5).

IV. Comparison of *S. marcescens*-specific Primers

The effectiveness of SM1_FR and SM2_FR primer pairs (Table 4) amplifying *S. marcescens* gDNA was compared with PCR. A 2% agarose gel was run to visualize the

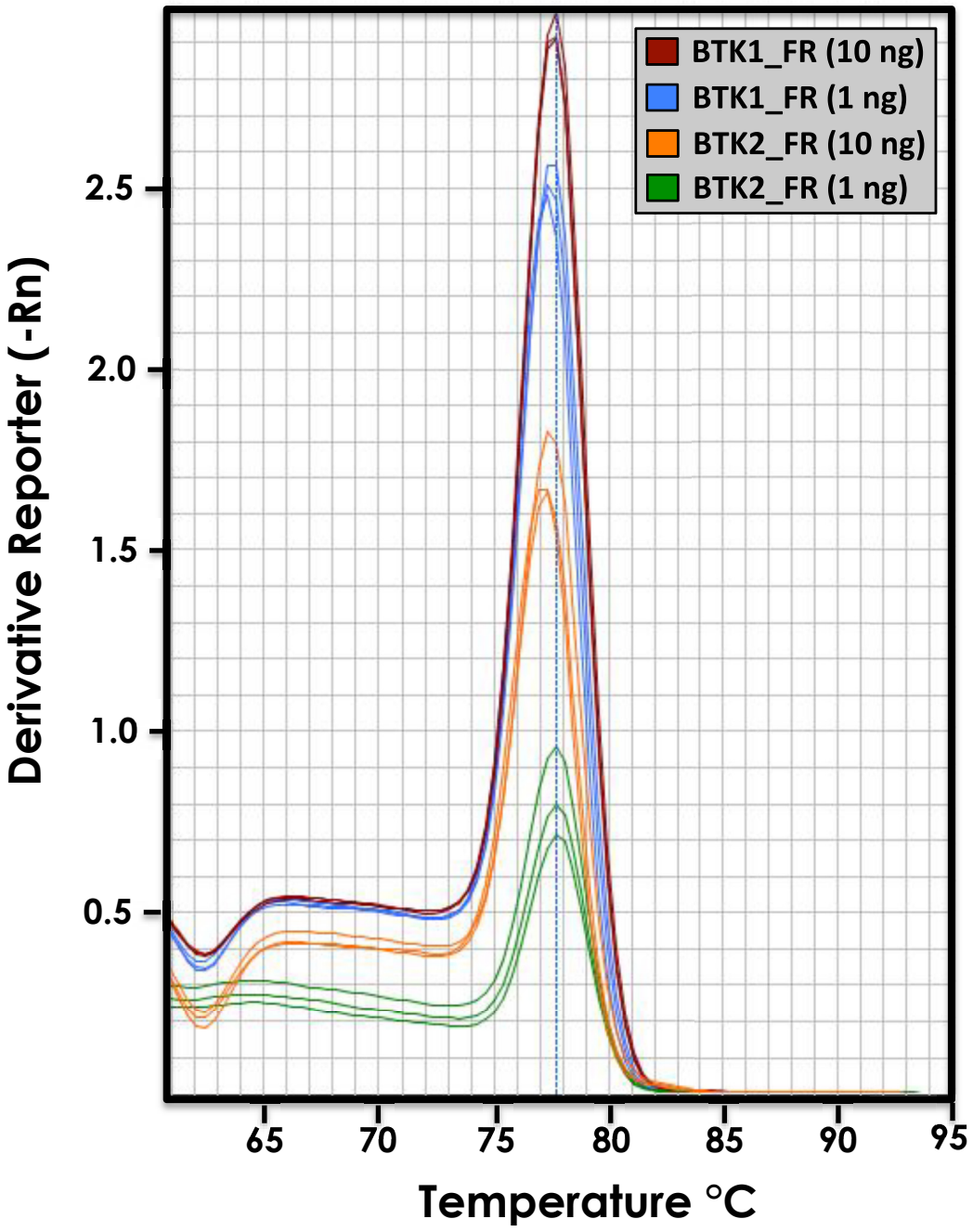

Figure 13. Melt-curve analysis using BTK1_FR and BTK2_FR primer pairs. *B. thuringiensis* species-specific primers BTK1_FR and BTK2_FR amplified two different amounts of *B. thuringiensis* gDNA (1 or 10 ng). The negative value of the change in rate of fluorescence (-Rn') is plotted against temperature. The curves from three replicates from each primer pair and template concentration are shown.

amplification products (Figure 9\10). Previous results in this study have shown that *S. marcescens* gDNA is successfully amplified using the SM1_FR and SM2_FR primer pairs (Figure 10, Lane 2-3), although amplification with SM1_FR produces a slightly denser band. For both SM1_FR and SM2_FR primer pairs, only the expected amplicons are present (113 and 138, respectively) (Figure 10, Lane 2-3; Figure 11, Lane 2,9).

For a more accurate comparison of the SM1_FR and SM2_FR primer pairs, triplicate SYBR-green qPCR assays were performed using two different concentrations of template DNA (10 and 1 ng/µL) (Figure 14). Mean C_t values were determined at the point of intersection with the calculated threshold value line (Figure 14; Table 6). Using 10 ng/µL of *S. marcescens* DNA as the template, amplification using the SM1_FR (C_t = 16.66) and SM2_FR primers (C_t = 16.66) (Table 6) reached threshold at the same cycle (Figure 14).

Table 6. Comparison of SM1_FR and SM2_FR primers.

Primers	T_m (°C)	Template (ng/µL)	Mean C_t Value	Threshold (ΔRn)	Melt Curve Peak Height (-Rn)
SM1_FR	86.33	10	16.66	.2	3.2
		1	20.294		3-3.1
SM2_FR	86.52	10	16.66	.2	2.8-3
		1	20.096		3-3.1

* C_t values were calculated from amplification data, the AB software automatically calculated threshold values, and T_m and peak heights were determined via the melt curve (Figure 15).

A smaller amount of *S. marcescens* gDNA as template (1 ng/µL) yielded comparable results, with the amplification using SM1_FR (C_t = 22.29) and SM2_FR

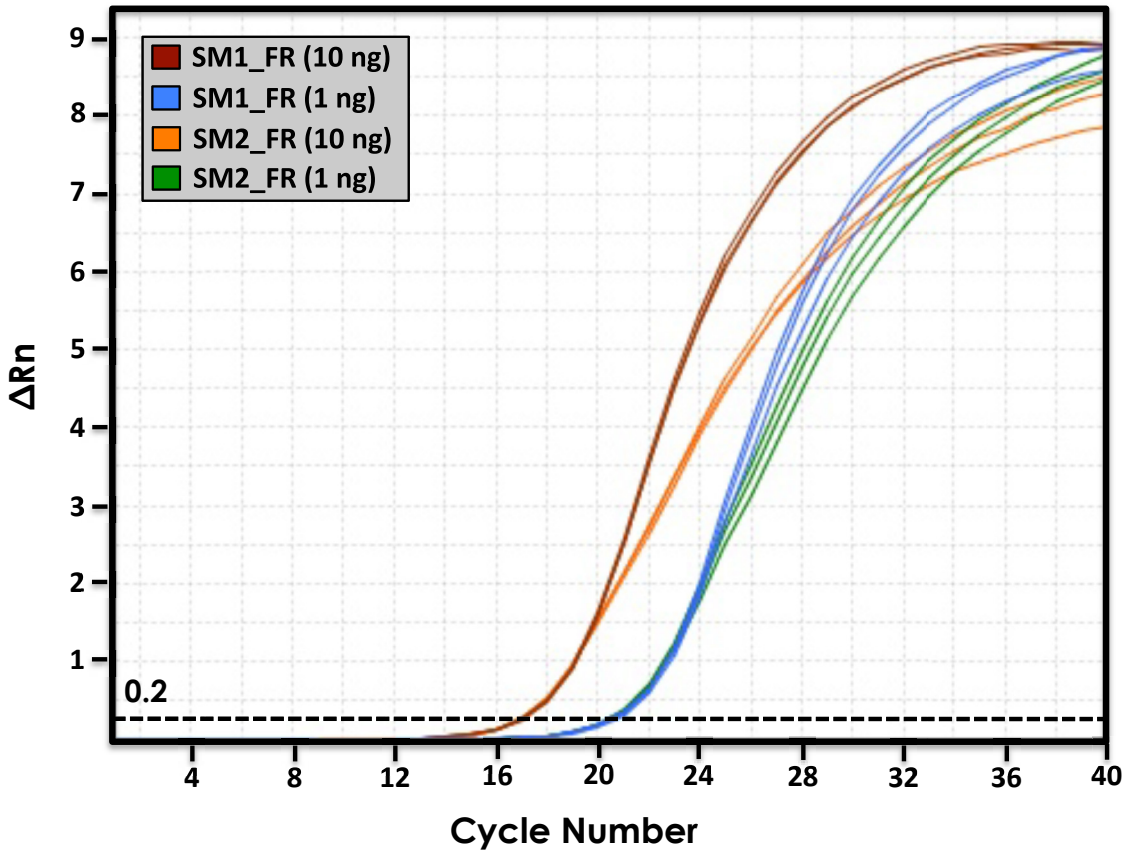

Figure 14. Sensitivity of the SM1_FR and SM2_FR primer pairs. Two different amounts (1 or 10 ng) of purified *S. marcescens* gDNA were used as the template for amplification. C_t values were determined as the intercept of the magnitude of the fluorescence signal (ΔRn) with the calculated threshold line (Table 6). The reported mean C_t values for a given sample are from triplicate runs.

primers (C_t = 22.10) (Table 6) reaching threshold at approximately the same cycle. At both concentrations, the differences in C_t values are not significant ($p < .05$).

Post-amplification melt curve analyses were performed for both SM1_FR and SM2_FR primer pairs (Figure 15) with 10 and 1 ng/µL *S. marcescens* DNA. T_m values and peak heights are reported in Table 6. The T_m value for SM1_FR and SM2_FR amplification products were 86.33 ˚C and 86.52 ˚C, respectively (Table 6). Similar results were observed at a higher concentration of *S. marcescens* DNA (10 ng/µL). The similarity in peak heights is conserved for the lower *S. marcescens* DNA concentration (1 ng/µL) (Figure 15), with observed heights nearly identical for both the primer pairs (Table 6). However, undulations in the curve at lower temperature than the T_m are observed in SM2_FR amplification of both template concentrations (Figure 14). The presence of a small shoulder peak at a T_m of approximately 81˚C suggests the presence of a small amount of an additional second product (Figure 15). The second product is not observed in the SM1_FR melt curve (Figure 15).

V. *B. thuringiensis* qPCR Assay Using BTK1_FR Primers

To investigate the effect of metagenomic DNA on amplification of *B. thuringiensis* DNA, a SYBR-green qPCR assay was performed using the spiking scheme outlined in the methods. In brief, the BTK1_FR primer pair was used to detect *B. thuringiensis* DNA of varying concentrations when spiked with 200 ng of non-target DNA from five BAL DNA extracts. Purified *B. thuringiensis* gDNA were used to construct calibration curves (Figure 16) to assess the quality of the DNA.

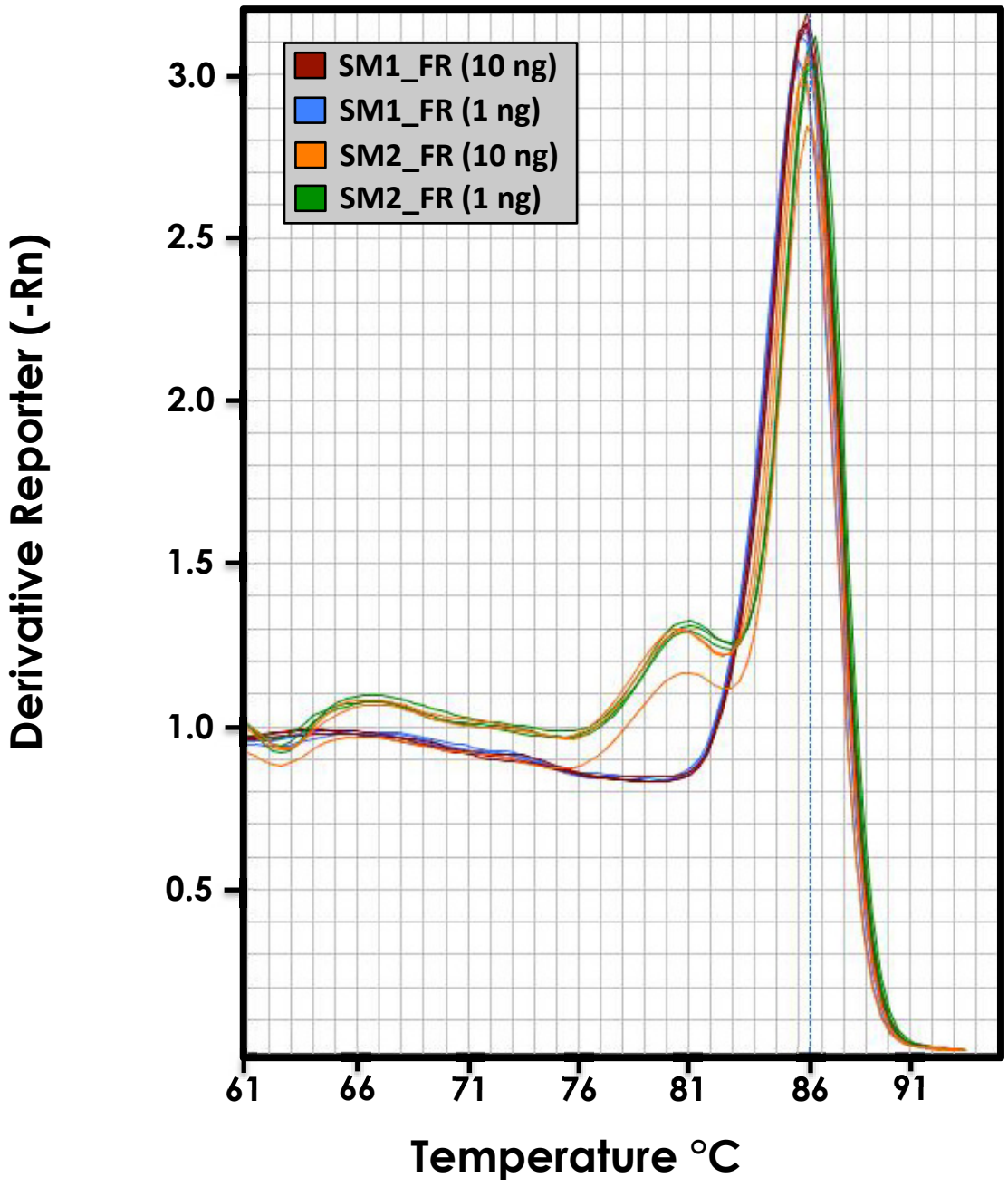

Figure 15. Melt-Curve analysis for SM1_FR and SM2_FR primer pairs. *S. marcescens* species-specific SM1_FR and SM2_FR primer pairs amplified two different amounts (1 or 10 ng) of *S. marcescens* gDNA. The negative value of the change in rate of fluorescence (-Rn') is plotted against temperature. The curves from three replicates from each primer pair and template concentration are shown.

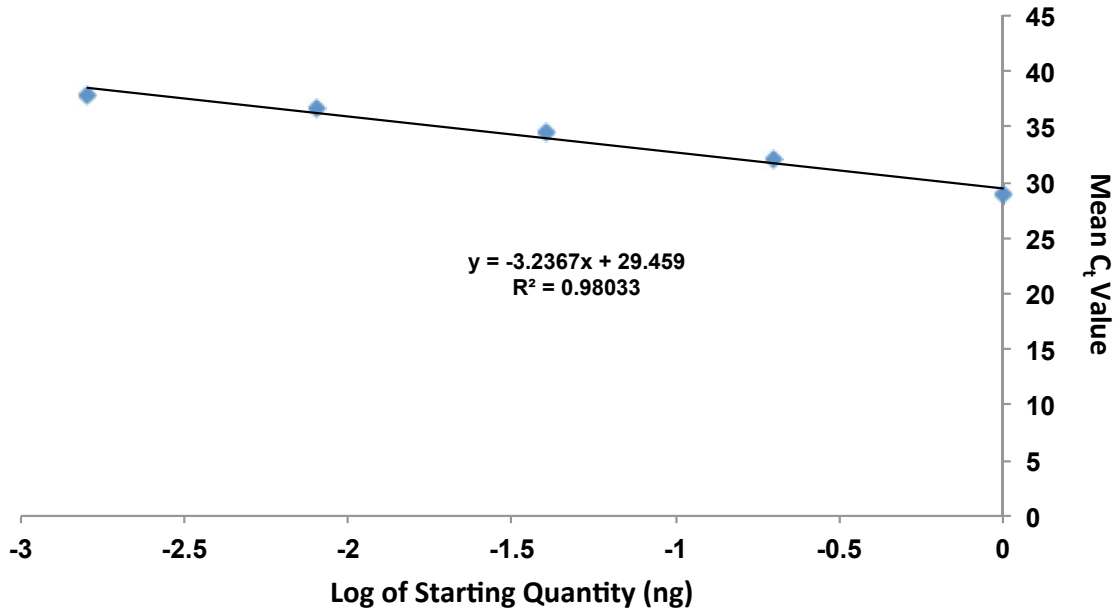

Figure 16. Calibration curve for *B. thuringiensis* DNA amplification with BTK1_FR primers. Log of the starting quantities is plotted against the mean C_t values of BTK1_FR amplification of purified *B. thuringiensis* gDNA. The blue symbols represent the triplicate PCR amplification of the standard DNA samples (1, 0.2, 0.04, 0.008, and 0.002 ng/µL concentrations). Linear regression analysis yielded the reported slope and R^2 value.

The C_t values varied with the amount of template DNA in a dose-dependent manner. Due to the complex nature of the qPCR assay, it was determined that a box plot analysis (accompanied by appropriate statistical testing) for each concentration would represent the data most effectively (Figure 17). Upon visual inspection, the lower template concentration (Figure 17, Panel A) has a higher mean C_t values compared to the higher template concentration (Figure 17, Panel B), which is an expected dose-dependent effect. More interestingly, in four out of five cases, DNA spiked with BAL DNA (with the exception of BAL 2) had different mean C_t values as compared to the unspiked gDNA. To determine if these differences are significant, statistical tests were performed.

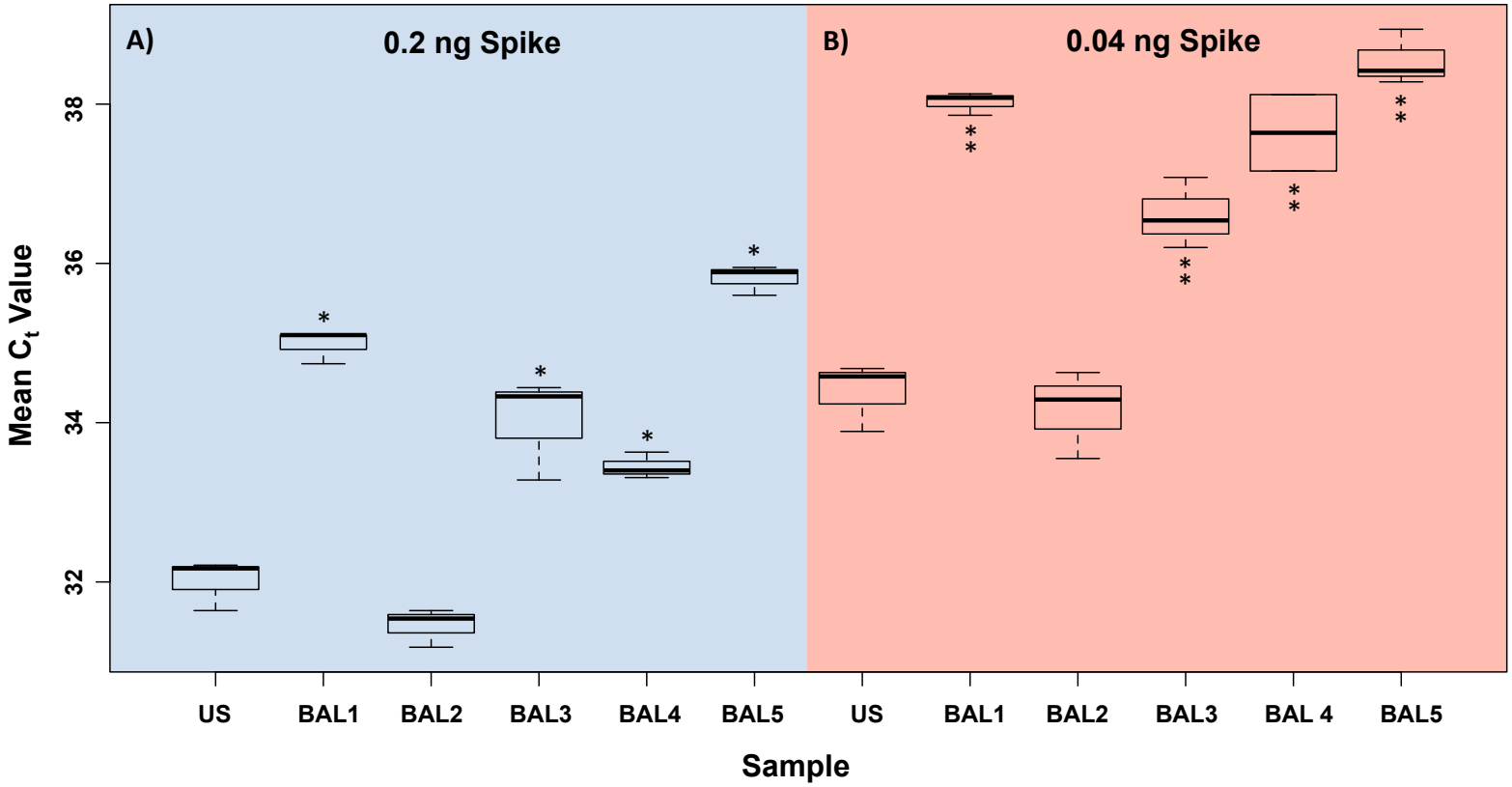

Figure 17. Box plot of C_t values for spiked and unspiked *B. thuringiensis* DNA. C_t values were determined for amplification of two amounts of *B. thuringiensis* (A) 0.2 ng and (B) 0.04 ng of genomic DNA unspiked (US) spiked with BAL DNA (BAL 1-5) using the BTK_1FR primer pair. Approximately 200 ng of DNA is present in each BAL sample. amplifications with 0.008 ng of template failed to reach threshold and were excluded. * Significantly different from 0.2 ng control ($p < 0.05$), ** Significantly different from the 0.04 ng control ($p < 0.05$).

A one-way ANOVA was conducted to compare C_t values of the unspiked control (0.2 ng *B. thuringiensis* DNA) and the spiked samples (0.2 ng *B. thuringiensis* gDNA + BAL DNA) using the BTK1_FR primers. A significant effect on the mean C_t values was seen between almost all samples ($p < .05$). A Post-hoc Tukey HSD test indicated that the mean C_t value of the unspiked 0.2 ng control (32.01 ± 0.32) was significantly lower than the mean C_t value of samples spiked with DNA from BAL 1 (34.98 ± 0.21, *p*-value = 1.7×10^{-6}), BAL 3 (34.02 ± 0.64, *p*-value = .0001), BAL 4 (33.45 ± 0.17, *p*-value = 0.0021), and BAL 5 (35.81 ± .19, *p*-value = 1×10^{-7}). No significant difference was observed between the mean C_t values of the 0.2 ng control (32.01 ± 0.32) and the BAL 2 spiked sample (31.45 ± 0.24, *p*-value = 0.386).

Similarly, a one-way ANOVA was conducted to compare mean C_t values of the unspiked control (0.04 ng *B. thuringiensis* gDNA) and the spiked samples (0.04 ng *B. thuringiensis* gDNA + BAL DNA). A significant effect on the mean C_t values was seen between most samples ($p < 0.05$). A Post-hoc Tukey HSD test indicated that the mean C_t value of the 0.04 ng control (34.48 ± 0.43) was significantly lower when unspiked as compared to samples spiked with BAL 1 (38.02 ± 0.14, *p*-value = 7.2×10^{-6}), BAL 3 (36.61 ± 0.44, *p*-value = 0.00071), BAL 4 (37.64 ± 0.68, *p*-value = 6.21×10^{-5}), and BAL 5 (38.55 ± .35, $p=1.8 \times 10^{-6}$) samples. However, no significant difference was observed between the mean C_t values of the 0.04 ng control (34.48 ± 0.43) and the BAL 2 spiked sample (34.16 ± .56, *p*-value = 0.9859).

A melt curve analysis was performed post amplification (Figure 18). The curve indicates the major amplicon disassociation of all samples to be at T_m = 78.28˚C. Some lingering fluorescence occurs after the main amplicon dissociation T_m (Figure 18).

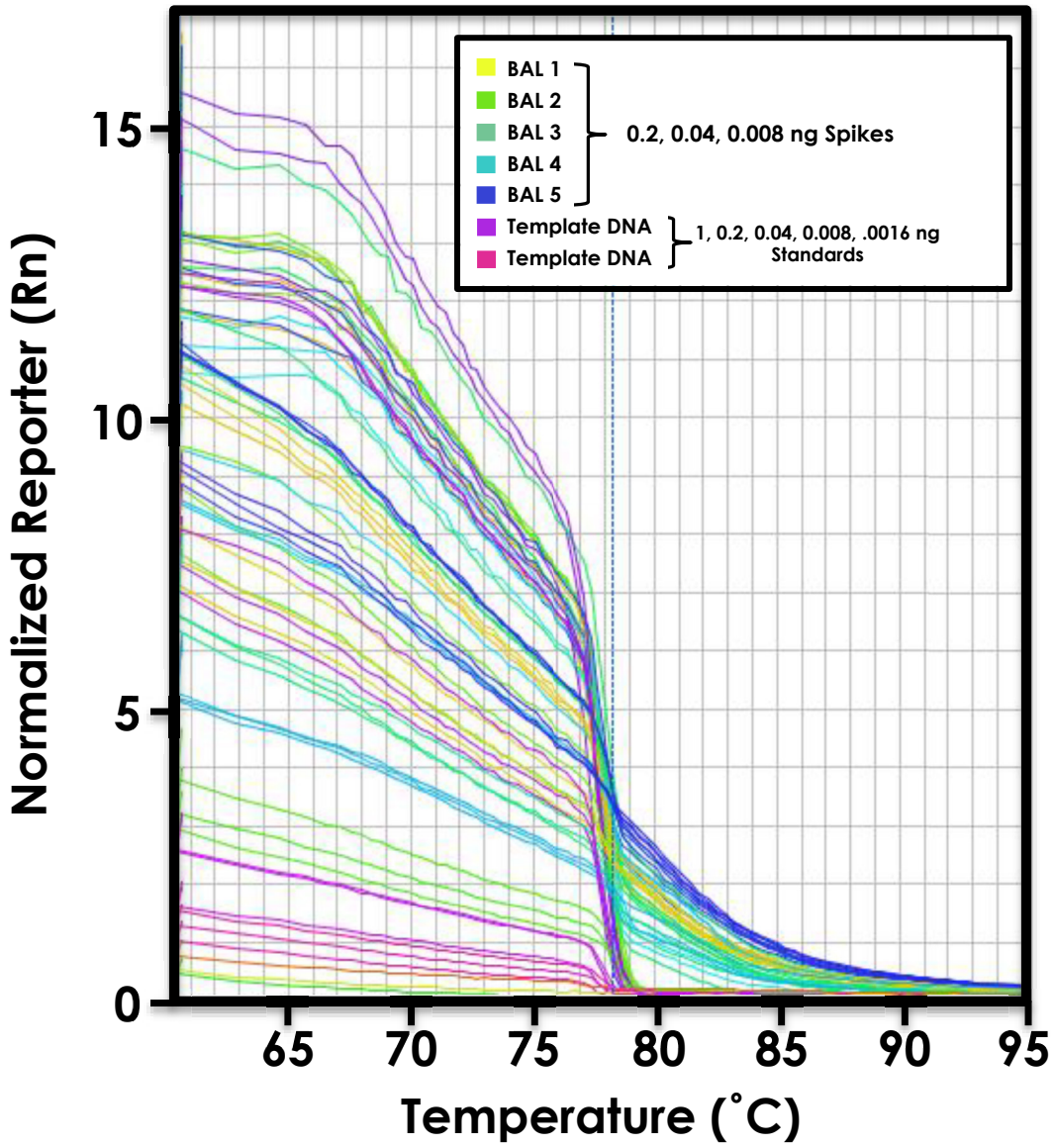

Figure 18. Melt-Curve analysis using BTK1_FR primers and spiked *B. thuringiensis* DNA. The normalized reporter fluorescence (Rn) values are plotted against temperature.

VI. *S. marcescens* qPCR Assay Using SM1_FR Primers

To investigate the effect of metagenomic DNA on amplification of *S. marcescens* DNA, a qPCR assay was performed using metagenomic DNA from BAL samples spiked with template gDNA. In brief, the SM1_FR primer pair was used to detect *S. marcescens* gDNA of varying concentrations when spiked with 200 ng of non-target DNA from five BAL DNA extracts. Purified *S. marcescens* gDNA was used in the construction of calibration curves to assess the quality of the DNA (Figure 19).

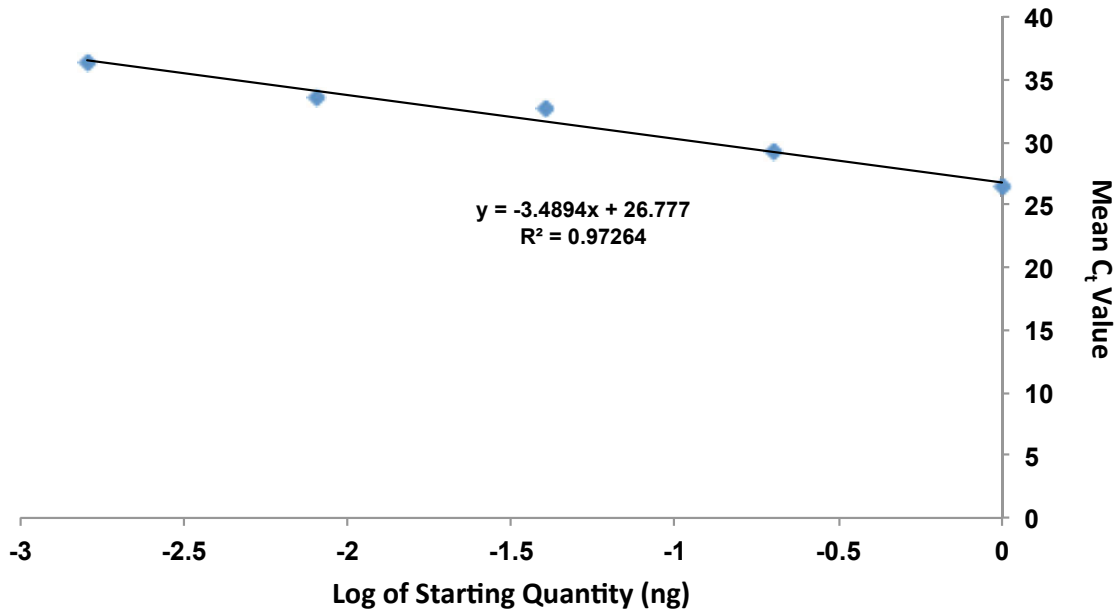

Figure 19. Calibration curve for *S. marcescens* DNA amplification using SM1_FR primers. Log of starting quantities is plotted against mean C_t values of amplification of purified *S. marcescens* gDNA. The blue symbols represent the triplicate PCR amplification of the standard DNA samples (1, 0.2, 0.04, 0.008, and 0.002 ng/µL concentrations). Linear regression analysis yielded the reported slope and R^2 values.

The C_t values varied with the amount of template DNA added. Figure 20 illustrates the results of qPCR amplification of each sample spiked with BAL at a given concentration of *S. marcescens* gDNA using the SM1_FR primer pair. Upon visual inspection, the lower template concentration (Figure 20, Panel C) seems to have generally higher mean C_t values compared to the higher concentration (Figure 20, Panel A/B), as expected. Additionally, all the samples spiked with BAL DNA (with the exception of BAL 2 and BAL 3 in panel B) seem to have different mean C_t values as compared to the control. To determine the significance of these differences, statistical tests were performed.

A one-way ANOVA was conducted to compare mean C_t values of the unspiked control and the samples spiked with BAL DNA. A significant effect on the mean C_t values was seen between almost all samples ($p < 0.05$). Post-hoc comparisons using the Tukey HSD test indicated that the mean C_t value of the 0.2 ng control (29.14 ± 0.32) was significantly lower than the mean C_t value of sample spiked with DNA from BAL 1 (30.99 ± 0.28, *p*-value = 4.12×10^{-5}), BAL 3 (31.99 ± 0.06, *p*-value = 4×10^{-7}), BAL 4 (30.46 ± 0.37, *p*-value = 0.001), and BAL 5 (32.22 ± .18, p= 2×10^{-7}). However, no statistically significant difference was observed between the mean C_t values of the 0.2 ng control (29.14 ± 0.32) and the BAL 2 spiked sample (28.66 ± 0.34, *p*-value = 0.34).

For 0.04 ng of *S. marcescens* DNA, a significant effect on the mean C_t values was seen between almost all samples ($p < .05$). Post-hoc comparisons using the Tukey HSD test indicated that the mean C_t value of the 0.04 ng control (31.88 ± 0.18) was significantly lower than the mean C_t value of samples spiked with DNA from BAL 3 (35.16 ± 0.22, *p*-value = 0.00014), BAL 4 (33.21 ± 0.28, *p*-value = 0.0101), and BAL 5

(35.36 ± 0.16, p= 8.09 x 10^{-5}). However, no statistically significant difference was observed between the mean C_t values of the 0.04 ng control (31.88 ± 0.18) and samples spiked with BAL 1 (33.15 ± 1.001, *p*-value = 0.113) or BAL 2 (31.68 ± 0.34, *p*-value = 0.9969).

The smallest amount *S. marcescens* DNA template (0.008 ng), the same significant effect on the mean C_t values was seen between almost all samples ($p < 0.05$). Post-hoc comparisons using the Tukey HSD test indicated that the mean C_t value of the 0.008 ng unspiked control (33.59 ± .024) was significantly lower than the mean C_t value of samples spiked with DNA from BAL 1 (36.49 ± 0.05, *p*-value = 0.00), BAL 3 (37.06 ± 0.08, *p*-value = 0.00), BAL 4 (35.94 ± 0.32, *p*-value = 0.00), and BAL 5 (37.97 ± 0.16, *p*-value = 0.00). However, similarly to the higher concentrations, no significant difference was observed between the mean C_t values of the 0.008 ng control (29.14 ± .032) and the BAL 2 spiked sample (33.80 ± 0.21, *p*-value = 0.6497).

A melt curve analysis was also performed post-amplification (Figure 21). The curve indicates the major amplicon disassociation of all samples to be at Tm = 86.66˚C. No lingering fluorescence is detected after the major amplicon dissociation T_m (Figure 21).

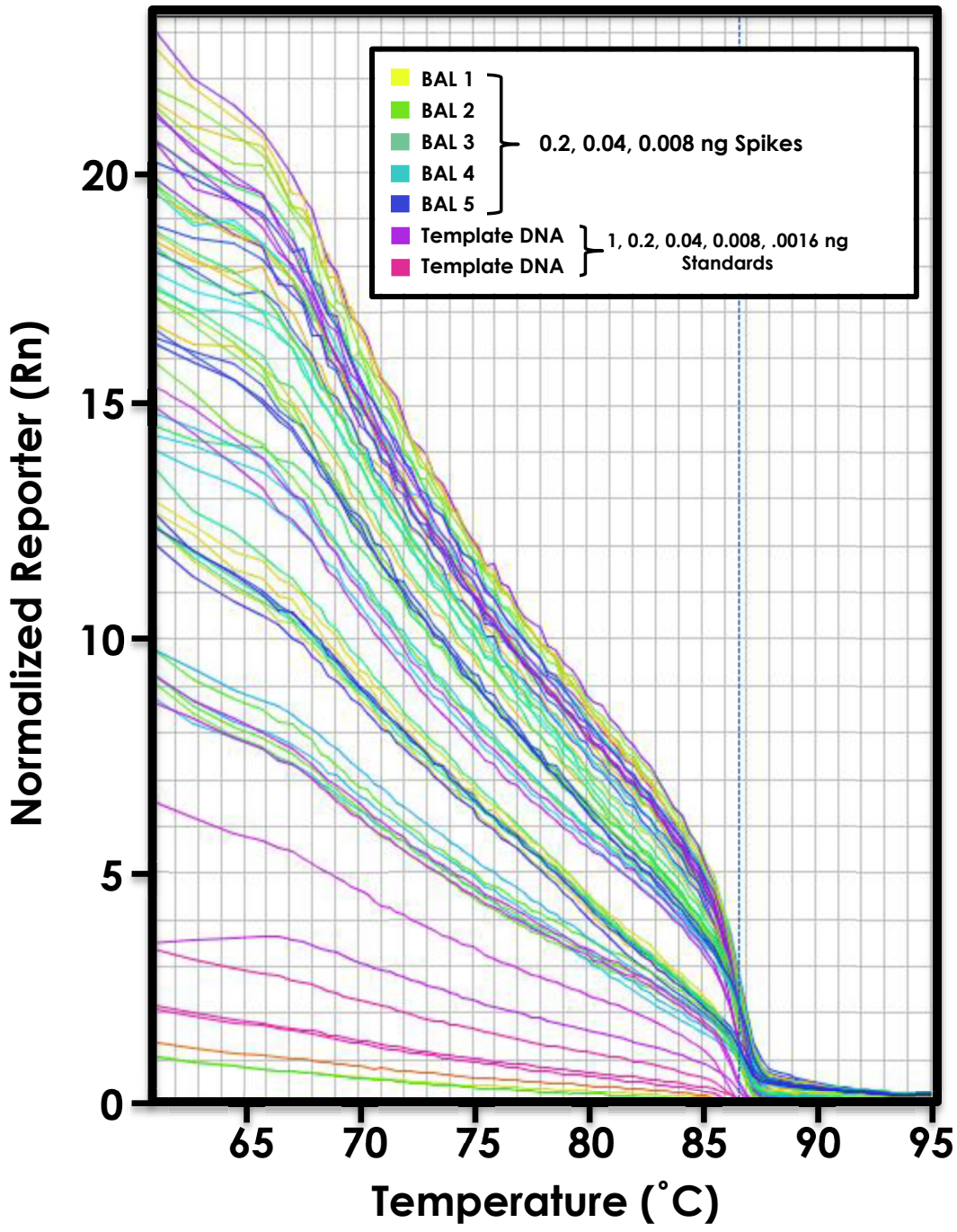

Figure 21. Melt-curve analysis using SM1_FR primers and spiked *S. marcescens* gDNA. The normalized reporter fluorescence (Rn) values are plotted against temperature.

DISCUSSION

Real-time qPCR is an indispensable tool that can be used to identify the presence of specific bacteria in a variety of environments and conditions. As such, it has been a useful tool in the field of forensics to detect harmful bacterial agents. The qPCR method has been used primarily for detection of indicatory virulence plasmids of biological warfare agents such as *B. anthracis* and *Y. pestis*, as well as surrogate organisms such as *B. thuringiensis* and *S. marcescens* (Porcar and Juarez-Perez 2003; Saikaly *et al.* 2007; Skottman *et al.* 2007; Janse *et al.* 2010). The work presented here addresses the need for additional qPCR assays targeting the bacterial chromosome. In this study, I have designed species-specific primer pairs for the detection of *B. thuringiensis* and *S. marcescens* and confirmed their specificity and sensitivity.

Primer Design Workflow Yields Effective Primer Pairs

There is limited research on chromosomal targets for *B. thuringiensis* and *S. marcescens* detection. To date most species-specific detection of these strains relies on targeting markers on plasmids (Porcar and Juarez-Perez 2003; Saikaly *et al.* 2007; Skottman *et al.* 2007; Janse *et al.* 2010). This is less desirable as the copy number of plasmid can vary from strain to strain (Janse *et al.* 2010). Furthermore, in some attenuated or genetically altered strains, one or more virulence plasmids can be absent, and subsequently go undetected.

To circumvent this problem, primers targeting the chromosome were designed *de novo* for this study using the workflow as described in the materials and methods (Figure 8). Three primer pairs were designed specifically in this study: BTK1_FR, BTK2_FR, and SM1_FR. The BTK1_FR and BTK2_FR primers target genes encoding a hypothetical protein and a FAD-dependent oxidase, respectively. These two genes are several thousand base pairs apart (Table 4). The SM1_FR primer recognizes the *gyrB* gene, similar to the published SM2_FR primer (Saikaly *et al.* 2007). In fact, the amplicons overlap (Table 4). The similarity between the published primer (SM2_FR) and the *de novo* primer (SM1_FR) validated the primer design processes used in this study.

The workflow developed in this study is particularly valuable as it bypasses the need to manually acquire sequences and perform alignments by utilizing the Insignia program (Phillippy *et al.* 2009). With this program, there is no need to specify target area, rather it allows for searching of potentially new specifies-specific markers. Moreover, this design can be used to develop species-specific PCR primers for any organism that has genomic sequence information in the Insignia database.

BTK1_FR Amplifies *B.* thuringiensis gDNA Most Effectively

The BTK1_FR and BTK2_FR primer pairs were extremely specific, as they did not amplify *B. cereus*, *B. megaterium*, *B. subtilis*, *S. marcescens, P. aeruginosa, or H. influenzae* DNA (Figure 10). Although no amplification was expected in the case of *S. marcescens, P. aeruginosa, or H. influenzae*, the absence of amplification of the *B. cereus* group organisms is an important finding. Discrimination of the phylogenetically

similar *B. cereus* group organisms is generally dealt with by targeting particular plasmids, which differ between organisms, or through sequencing of conserved polymorphic markers like 16S rRNA genes (Sacchi *et al.* 2002; Bavykin *et al.* 2004). The results from this study indicate that the development of species-specific chromosomal primers for *B. thuringiensis* is feasible with careful primer design and target selection.

Both the BTK1_FR and BTK2_FR primers specifically amplify the target DNA (Figure 10). However, the latter appears to be less specific when the template concentration is increased (Figure 9) as PCR results in two different amplicons (two faint bands; Figure 9). Non-specific amplification could be due to the excessive amount of template DNA, the binding to sequences in other *Bacillus* in the mixture, or contamination of a specific aliquot of the primer mix used in that particular reaction.

The absence of non-specific amplification with BTK1_FR (Figure 10) even with very low template concentration demonstrates that this primer is more specific and sensitive than BTK2_FR. The superiority of the BTK1_FR primers is supported by the qPCR data, which shows a significant decrease in mean C_t value in comparison to BTK2_FR (Figure 11). Apart from the similar T_m values for both primer pairs, the difference in the melt curve peak heights indicates significantly decreased amount of end product in the BTK2_FR amplification as compared to the BTK1_FR. The exact cause of this sensitivity discrepancy between the two primer pairs is unclear. Most likely, at reasonable template concentrations, the BTK2_FR primers are simply be less efficient at binding their target than the BTK1_FR primers, resulting in lower C_t values and reduced total amount of amplification products.

SM1_FR Amplifies *S. marcescens* gDNA Most Effectively

SM1_FR and SM2_FR primer pairs show effective and specific amplification of *S. marcescens* gDNA (Figure 9, 10). The similarity in primer performance was also observed in the qPCR assay. At both concentrations of input template DNA, both primers had nearly identical mean C_t values (Figure 13) and T_m values (Figure 14). The similarity and the sensitivity of amplifications of *S. marcescens* gDNA with the SM1_FR and SM2_FR primer pairs target the same gene, *gyrB*, encoding DNA gyrase (Dauga 2002; Saikaly *et al.* 2007). A closer look at the melting curve reveals some differences in the curve shape, particularly a shoulder in the curve for the SM2_FR amplification products (Figure 14). This small shoulder (consistent between both template concentrations) could indicate the presence of a second amplicon, due to its small size and its T_m occurring at a lower temperature than the main amplicon. This peak is absent in the SM1_FR melt curve, suggesting the lack of any secondary product formation. Thus, SM1_FR primer pair designed in this study is superior to the published SM2_FR primer pair.

The Presence of Metagenomic DNA Reduces Sensitivity of qPCR Assays

The sensitivity of the primer pairs was tested by spiking metagenome from BAL samples with varying amount of target genomic DNA. Although the composition of metagenome in each BAL sample is unknown, the amount of DNA was normalized to 200 ng. The sensitivity was reduced in the presence four out of the five BAL samples in

every spike concentration *B. thuringiensis* and *S. marcescens*. As previously shown where increasing concentrations of background DNA affected amplification performance by hindering access to the target sequences (Ludwig and Schleifer 2000). This effect of 'crowding' the polymerase with non-target DNA may result in decreased efficiency due to stochastic effects, effectively reducing the probability of target sequences being found. The ratio of the template to the background DNA was crucial to observe the changes in assay sensitivity reported in this study. If the ratio of spike to background DNA was small (i.e. low [template DNA] and low [background DNA] or visa versa), there was no significant loss in sensitivity, suggesting that the background was not sufficient to influence template amplification. Alternatively, with a large ratio of spike to background DNA (i.e. low [template DNA] and high [background DNA]), the difference in sensitivity becomes more significant.

The significant decrease in sensitivity observed could be due to the presence of PCR inhibitors in the metagenomic DNA, which was not directly tested in this study. However, the likelihood of potential PCR inhibitors carrying over through the BAL DNA extraction process is very small, due to the silica-bead based extraction method used. This technique, which allows binding of DNA to beads and subsequent washing of contaminants, retains fewer PCR inhibitors than other methods (e.g. organic phenol/chloroform extraction, chelex extraction). Additionally, the presence of PCR inhibitors would have likely caused shift in the melt curve analyses, which was not observed. The BAL DNA from all five samples was successfully amplified using 16S MJ68_FR primers in another study (data not shown), indicating that the DNA was amplifiable and did not contain inhibitory compounds.

Of the five BAL samples, BAL2 consistently showed no difference in amplification sensitivity compared to the standards. This was conserved through both the *B. thuringiensis* and *S. marcescens* assays. A number of issues could account for this lack of difference. Although unlikely, the quantification of the DNA could have been incorrect, presumably reporting higher values than the actual concentration. Although the method of quantification used is not the most accurate or sensitive (UV spectrophotometry), it is rarely an issue at such high concentrations of DNA using proper blanks. Quantification errors would not only affect BAL2, as all other DNA quantification was performed with the same method.

BAL2 could also potentially contain DNA from *B. thuringiensis* and *S. marcescens*, increasing the overall concentration of template DNA in the mixed sample. This is highly doubtful due to the low probability of finding *B. thuringiensis* and *S. marcescens* in the lungs, much less in the same lung sample. Alternatively, non-specific binding could be occurring due to DNA sequences present within the extracted BAL DNA. Although no obvious additional amplicons were seen, (Figure 18, Figure 22) repeated analysis of the melt curve did not reveal non-specific amplification.

Degradation of BAL2 DNA could offset the inhibitory effect observed with other spiked BAL samples. Repeated freeze-thaw cycles during experimentation or the incomplete removal of exonuclease enzymes during extraction could cause such degradation (Grecz *et al.* 1980; Cohen 1993). Research shows that while the degradation of DNA does skew some methods of DNA quantification, spectrophotometric measurements of DNA content are not effected (Sedlackova *et al.* 2013). The physical hindrance of PCR primers and polymerase from binding their target sequences could be

decreased with smaller metagenomic BAL2 DNA fragments, thereby not affecting sensitivity, as was observed.

B. thuringiensis and *S. marcescens* Are Detectable in the Complex Mixture

The BAL DNA caused a significant reduction in sensitivity in four out of the five spiked samples, indicated by a significant increase in C_t values (except BAL 2) compared to the unspiked control (Figure 21). Although the assay sensitivity decreased considerably, both *B. thuringiensis* and *S. marcescens* were successfully amplified, suggesting the primer-pairs designed are very specific. *B. thuringiensis* DNA was detected even at 0.04 ng of template (equivalent to approximately 2000 cells) (Figure 16, 17). The absence of reliable amplification with 0.008 ng of *B. thuringiensis* DNA is likely due to the lower sensitivity as compared to the SM1_FR *S. marcescens* assay. The small amount of non-specific amplification, indicated by the continued melting of several samples after the T_m in the dissociation curve (Figure 18), could also decrease sensitivity. The sensitivity of the SM1_FR primer pair is corroborated by the successful amplification of *S. marcescens* DNA at 0.008 ng of template (approximately 1000 cells), even in the presence of BAL DNA.

Concluding Remarks

The research reported in this thesis includes several valuable and interesting findings. Notably, the developed global approach to primer design can be used in a

number of different applications to develop species-specific primers for many organisms. Additionally, the establishment of novel species-specific primers for *B. thuringiensis* and *S. marcescens* provides additional options for researchers and scientists using PCR for bacterial detection. This work is particularly pertinent due to the use of clinical lung sample DNA as the background matrix, which is a main route-of-entry of dangerous biological warfare agents like *B. anthracis* and *Y. pestis*. The application of the assays extends past microbial forensics, as any field where detection of bacterial either species in a complex matrix is required can benefit directly from these assays.

Furthermore, the biases introduced by complex background DNA were quantitatively tested in this research. While this work dealt with detection of specific bacteria in a specific type of background, the knowledge of the potential biases introduced by environmental or clinical samples is important, and should taken into account when designing assays. It is also crucial to note the situations in which these reported sensitivity biases are significant. The discrepancy in sensitivity could be irrelevant when large amounts of template sample are present. Contrastingly, the bias could be relevant in clinical applications such as early testing for the presence of *B. anthracis* within the lung, before they present in the bloodstream. In such an instance, the resident microbial community would surround samples containing the target organism. Environmental applications such as bio-monitoring of water or air supplies could be affected as well. If the background matrix could significantly inhibit detection of the target, the pathogen could go undetected as a false negative.

Future Directions

Sequencing of PCR products. Beyond the melt curve analysis, definitive verification of products produced from both the *B. thuringiensis* and *S. marcescens* could be accomplished using direct sequencing or next-generation sequencing methods. In this study, sequencing could be used in two ways. First, sequencing the products from amplifications with pure genomic DNA as the template can unequivocally confirm the amplification of the desired product. Sequencing of the PCR products from the spiked DNA templates would help identify potential unintended amplicons, particularly in the case of amplifications with BTK1_FR primers.

Expanded strain panel. Future experiments or additions could be performed to improve and corroborate this work. Firstly, a larger panel of strains for primer specificity testing could be assembled. A more extensive panel would likely include several *Serratia* spp., alternative strains of *B. thuringiensis* and *S. marcescens*, as well as more common lung bacteria, such as *Staphylococcus*, *Streptococcus*, *Haemophilus*, and *Prevotella* spp. (Erb-Downward *et al.* 2011). The testing of the designed primers against this larger panel would allow for more confidence in the specificity of the primers reported in this study.

Additional background samples. Increasing the number of BAL samples used could also afford more latitude to the research. The five BAL samples used were randomly selected from a cohort of 60 subjects. Increasing the number of BAL samples used to 15 or 20, could improve the study. Knowing the microbial composition of the

BAL samples can reveal potential effects of specific groups of bacteria on the sensitivity of the designed assays. Additionally, the inclusion of environmental samples such as soil samples could provide an even more rigorous test for the assays, as soil could contain other types of organisms more likely to cross-react with the designed primers.

Whole-cell spiking experiment. Lastly, for this research, DNA was extracted from the individual bacteria separately from the BAL samples, and then mixed together. This method was chosen due to the ability to have greater control over the exact DNA concentrations in the final mixture. However, experiments that involved mixing whole bacterial cells with BAL samples, then extracting total DNA would be more analogous to a real life sample testing. A cell level spiking experiment could also account for differences in extraction efficiencies by using internal controls.

REFERENCES

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol 215: 403-410.

Andreotti, P. E., G. V. Ludwig, A. H. Peruski, J. J. Tuite, S. S. Morse and L. F. Peruski, Jr. (2003). "Immunoassay of infectious agents." Biotechniques 35: 850-859.

Arikawa, E., Y. Sun, J. Wang, Q. Zhou, B. Ning, S. L. Dial, L. Guo and J. Yang (2008). "Cross-platform comparison of SYBR Green real-time PCR with TaqMan PCR, microarrays and other gene expression measurement technologies evaluated in the MicroArray Quality Control (MAQC) study." BMC Genomics 9: 328.

Arnheim, N. and H. Erlich (1992). "Polymerase chain reaction strategy." Annu Rev Biochem 61: 131-156.

Ash, C., J. A. E. Farrow, M. Dorsch, E. Stackebrandt and M. D. Collins (1991). "Comparative analysis of *Bacillus anthracis*, *Bacillus cereus*, and related species on the basis of reverse-transcriptase sequencing of 16s rRNA." Int J Syst Bacteriol 41: 343-346.

Bavykin, S. G., Y. P. Lysov, V. Zakhariev, J. J. Kelly, J. Jackman, D. A. Stahl and A. Cherni (2004). "Use of 16S rRNA, 23S rRNA, and gyrB gene sequence analysis to determine phylogenetic relationships of *Bacillus cereus* group microorganisms." J Clin Microbiol 42: 3711-3730.

Bertherat, E., S. Bekhoucha, S. Chougrani, F. Razik, J. B. Duchemin, L. Houti, L. Deharib, C. Fayolle, B. Makrerougrass, R. Dali-Yahia, R. Bellal, L. Belhabri, A. Chaieb, E. Tikhomirov and E. Carniel (2007). "Plague reappearance in Algeria after 50 years, 2003." Emerg Infect Dis 13: 1459-1462.

Bos, K. I., V. J. Schuenemann, G. B. Golding, H. A. Burbano, N. Waglechner, B. K. Coombes, J. B. McPhee, S. N. DeWitte, M. Meyer, S. Schmedes, J. Wood, D. J. Earn, D. A. Herring, P. Bauer, H. N. Poinar and J. Krause (2011). "A draft genome of *Yersinia pestis* from victims of the Black Death." Nature 478: 506-510.

Bourgogne, A., M. Drysdale, S. G. Hilsenbeck, S. N. Peterson and T. M. Koehler (2003). "Global effects of virulence gene regulators in a *Bacillus anthracis* strain with both virulence plasmids." Infect Immun 71: 2736-2743.

Budowle, B. (2011). Microbial Forensics. Burlington, MA, Elsevier Academic.

Budowle, B., T. R. Moretti, S. J. Niezgoda and B. L. Brown (1998). "CODIS and PCRbased short tandem repeat loci: Law enforcement tools". Second European Symposium on Human Identification, Madison, WI, Promega Corporation.

Budowle, B., R. Murch and R. Chakraborty (2005). "Microbial forensics: the next forensic challenge." Int J Legal Med 119: 317-330.

Butler, J. M. (2010). Fundamentals of forensic DNA typing. Amsterdam ; Boston, Academic Press/Elsevier.

Butler, J. M. and D. J. Reeder (2013). "STRBase". NIST Standard Reference Database: SRD 130, National Institute of Standards and Technology.

Cabanel, N., A. Leclercq, V. Chenal-Francisque, B. Annajar, M. Rajerison, S. Bekkhoucha, E. Bertherat and E. Carniel (2013). "Plague outbreak in Libya, 2009, unrelated to plague in Algeria." Emerg Infect Dis 19: 230-236.

Carey, L. and L. Mitnik (2002). "Trends in DNA forensic analysis." Electrophoresis 23: 1386-1397.

Cohen, J. J. (1993). "Apoptosis." Immunol Today 14: 126-130.

Court of Appeals of Arizona (1995). State of Arizona v. Mark Alan Bogan. 183: 515.

Dauga, C. (2002). "Evolution of the *gyrB* gene and the molecular phylogeny of *Enterobacteriaceae*: a model molecule for molecular systematic studies." Int J Syst Evol Microbiol 52: 531-547.

Dennis, D. T. (1994). "Plague in India." Brit Med J 309: 893-894.

Erb-Downward, J. R., D. L. Thompson, M. K. Han, C. M. Freeman, L. McCloskey, L. A. Schmidt, V. B. Young, G. B. Toews, J. L. Curtis, B. Sundaram, F. J. Martinez and G. B. Huffnagle (2011). "Analysis of the lung microbiome in the "healthy" smoker and in COPD." PLoS One 6: e16384.

Evguenieva-Hackenberg, E. (2005). "Bacterial ribosomal RNA in pieces." Mol Microbiol 57: 318-325.

Forman, S., C. R. Wulff, T. Myers-Morales, C. Cowan, R. D. Perry and S. C. Straley (2008). "*yadBC* of *Yersinia pestis*, a new virulence determinant for bubonic plague." Infect Immun 76: 578-587.

Grecz, N., T. L. Hammer, C. J. Robnett and M. D. Long (1980). "Freeze-thaw injury: evidence for double strand breaks in *Escherichia coli* DNA." Biochem Biophys Res Commun 93: 1110-1113.

Greenberg, D. L., J. D. Busch, P. Keim and D. M. Wagner (2010). "Identifying experimental surrogates for *Bacillus anthracis* spores: a review." Investig Genet 1: 4-16.

Hadjinicolaou, A. V., V. L. Demetriou, J. Hezka, W. Beyer, T. L. Hadfield and L. G. Kostrikis (2009). "Use of molecular beacons and multi-allelic real-time PCR for detection of and discrimination between virulent *Bacillus anthracis* and other *Bacillus* isolates." J Microbiol Methods 78: 45-53.

Haensch, S., R. Bianucci, M. Signoli, M. Rajerison, M. Schultz, S. Kacki, M. Vermunt, D. A. Weston, D. Hurst, M. Achtman, E. Carniel and B. Bramanti (2010). "Distinct clones of *Yersinia pestis* caused the black death." PLoS Pathog 6: e1001134.

Handelsman, J. (2004). "Metagenomics: application of genomics to uncultured microorganisms." Microbiol Mol Biol Rev 68: 669-685.

Hejazi, A. and F. R. Falkiner (1997). "*Serratia marcescens*." J Med Microbiol 46: 903-912.

Hindson, B. J., A. J. Makarewicz, U. S. Setlur, B. D. Henderer, M. T. McBride and J. M. Dzenitis (2005). "APDS: the Autonomous Pathogen Detection System." Biosens Bioelectron 20: 1925-1931.

New England Biosystems Inc. DNA Amplification and PCR. http://www.neb.com.

Inglesby, T. V., D. A. Henderson, J. G. Bartlett, M. S. Ascher, E. Eitzen, A. M. Friedlander, J. Hauer, J. McDade, M. T. Osterholm, T. O'Toole, G. Parker, T. M. Perl, P. K. Russell, K. Tonat and W. G. C. Biodefense (1999). "Anthrax as a biological weapon - Medical and public health management." Jama-J Am Med Assoc 281: 1735-1745.

Ivins, B. E., J. W. Ezzell, Jr., J. Jemski, K. W. Hedlund, J. D. Ristroph and S. H. Leppla (1986). "Immunization studies with attenuated strains of *Bacillus anthracis*." Infect Immun 52: 454-458.

Jackson, P. J., M. E. Hugh-Jones, D. M. Adair, G. Green, K. K. Hill, C. R. Kuske, L. M. Grinberg, F. A. Abramova and P. Keim (1998). "PCR analysis of tissue samples from the 1979 Sverdlovsk anthrax victims: the presence of multiple *Bacillus anthracis* strains in different victims." Proc Natl Acad Sci U S A 95: 1224-1229.

Janse, I., R. A. Hamidjaja, J. M. Bok and B. J. van Rotterdam (2010). "Reliable detection of *Bacillus anthracis*, Francisella tularensis and *Yersinia pestis* by using multiplex qPCR including internal controls for nucleic acid extraction and amplification." BMC Microbiol 10: 314.

Jaric, M., J. Segal, E. Silva-Herzog, L. Schneper, K. Mathee and G. Narasimhan (2013). "Better Primer Design for Metagenomics Applications by Increasing Taxonomic Distinguishability." BMC Proceedings. In press.

Jeffreys, A. J., V. Wilson and S. L. Thein (1985). "Hypervariable 'minisatellite' regions in human DNA." Nature 314: 67-73.

Jernigan, D. B., P. L. Raghunathan, B. P. Bell, R. Brechner, E. A. Bresnitz, J. C. Butler, M. Cetron, M. Cohen, T. Doyle, M. Fischer, C. Greene, K. S. Griffith, J. Guarner, J. L. Hadler, J. A. Hayslett, R. Meyer, L. R. Petersen, M. Phillips, R. Pinner, T. Popovic, C. P. Quinn, J. Reefhuis, D. Reissman, N. Rosenstein, A. Schuchat, W. J. Shieh, L. Siegal, D. L. Swerdlow, F. C. Tenover, M. Traeger, J. W. Ward, I. Weisfuse, S. Wiersma, K. Yeskey, S. Zaki, D. A. Ashford, B. A. Perkins, S. Ostroff, J. Hughes, D. Fleming, J. P. Koplan, and J. L. Gerberding (2002). "Investigation of bioterrorism-related anthrax, United States, 2001: epidemiologic findings." Emerg Infect Dis 8: 1019-1028.

Jernigan, J. A., D. S. Stephens, D. A. Ashford, C. Omenaca, M. S. Topiel, M. Galbraith, M. Tapper, T. L. Fisk, S. Zaki, T. Popovic, R. F. Meyer, C. P. Quinn, S. A. Harper, S. K. Fridkin, J. J. Sejvar, C. W. Shepard, M. McConnell, J. Guarner, W. J. Shieh, J. M. Malecki, J. L. Gerberding, J. M. Hughes, and B. A. Perkins (2001). "Bioterrorism-related inhalational anthrax: the first 10 cases reported in the United States." Emerg Infect Dis 7: 933-944.

Kadlec, R. P., A. P. Zelicoff and A. M. Vrtis (1997). "Biological weapons control. Prospects and implications for the future." JAMA 278: 351-356.

Keim, P., K. L. Smith, C. Keys, H. Takahashi, T. Kurata and A. Kaufmann (2001). "Molecular investigation of the Aum Shinrikyo anthrax release in Kameido, Japan." J Clin Microbiol 39: 4566-4567.

Kim, J., J. Lim and C. Lee (2013). "Quantitative real-time PCR approaches for microbial community studies in wastewater treatment systems: Applications and considerations." Biotechnol Adv 31: 1358-1373.

Kingston, R., F. Burke, J. H. Robinson, P. A. Bedford, S. M. Jones, S. C. Knight and E. D. Williamson (2007). "The fraction 1 and V protein antigens of *Yersinia pestis* activate dendritic cells to induce primary T cell responses." Clin Exp Immunol 149: 561-569.

Lambert, M. N., E. Vocker, S. Blumberg, S. Redemann, A. Gajraj, J. C. Meiners and N. G. Walter (2006). "Mg^{2+}-induced compaction of single RNA molecules monitored by tethered particle microscopy." Biophys J 90: 3672-3685.

Landegren, U., M. Nilsson and P. Y. Kwok (1998). "Reading bits of genetic information: methods for single-nucleotide polymorphism analysis." Genome Res 8: 769-776.

Lathem, W. W., P. A. Price, V. L. Miller and W. E. Goldman (2007). "A plasminogen-activating protease specifically controls the development of primary pneumonic plague." Science 315: 509-513.

Li, R. (2008). Forensic biology. Boca Raton, CRC Press/Taylor & Francis.

Little, S. F. and B. E. Ivins (1999). "Molecular pathogenesis of *Bacillus anthracis* infection." Microbes Infect 1: 131-139.

Lopez, A. D., K. Shibuya, C. Rao, C. D. Mathers, A. L. Hansell, L. S. Held, V. Schmid and S. Buist (2006). "Chronic obstructive pulmonary disease: current burden and future projections." Eur Respir J 27: 397-412.

Ludwig, W. and K. H. Schleifer (2000). "How quantitative is quantitative PCR with respect to cell counts?" Syst Appl Microbiol 23: 556-562.

Makino, S., I. Uchida, N. Terakado, C. Sasakawa and M. Yoshikawa (1989). "Molecular characterization and protein analysis of the cap region, which is essential for encapsulation in *Bacillus anthracis*." Journal of Bacteriology 171: 722-730.

Manchee, R. J. and W. D. P. Stewart (1988). "The decontamination of Gruinard Island." Chem Brit 24: 690-692.

Mardis, E. R. (2008). "The impact of next-generation sequencing technology on genetics." Trends Genet 24: 133-141.

Menotti-Raymond, M. A., V. A. David and S. J. O'Brien (1997). "Pet cat hair implicates murder suspect." Nature 386: 774.

Meselson, M., J. Guillemin, M. Hugh-Jones, A. Langmuir, I. Popova, A. Shelokov and O. Yampolskaya (1994). "The Sverdlovsk anthrax outbreak of 1979." Science 266: 1202-1208.

Metzker, M. L. (2010). "Sequencing technologies - the next generation." Nat Rev Genet 11: 31-46.

Moreno, L. (2005). Molecular and chemical characterization of three Miami-Dade soil types for forensic comparison. Miami, Florida International University. Masters of Science in Forensic Science: Dr. Kalai Mathee.

Moreno, L. I., D. K. Mills, J. Entry, R. T. Sautter and K. Mathee (2006). "Microbial metagenome profiling using amplicon length heterogeneity-polymerase chain reaction proves more effective than elemental analysis in discriminating soil specimens." J Forensic Sci 51: 1315-1322.

Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich (1992). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986." Biotechnology 24: 17-27.

Mullis, K. B. (1990). "Target amplification for DNA analysis by the polymerase chain reaction." Ann Biol Clin (Paris) 48: 579-582.

Murray, C. J. L., A. D. Lopez, C. D. Mathers and C. Stein (2001). The global burden of disease 2000 project: Aims, methods and data sources. Global Programme on Evidence for Health Policy Discussion. Geneva, World Health Organization. 36: 1-57.

Musher, D. M. (1983). "*Haemophilus influenzae* infections." Hosp Pract (Off Ed) 18: 158-161, 166, 169-170.

Musser, A. W. and P. R. Beamer (1961). "Infections caused by *Pseudomonas aeruginosa*." J Indiana State Med Assoc 54: 1627-1634.

Nandan, G. (1994). "Plague spreads in India but is under control." Brit Med J 309: 897-897.

Opel, K. L., D. Chung and B. R. McCord (2010). "A study of PCR inhibition mechanisms using real time PCR." J Forensic Sci 55: 25-33.

Parkhill, J., B. W. Wren, N. R. Thomson, R. W. Titball, M. T. Holden, M. B. Prentice, M. Sebaihia, K. D. James, C. Churcher, K. L. Mungall, S. Baker, D. Basham, S. D. Bentley, K. Brooks, A. M. Cerdeno-Tarraga, T. Chillingworth, A. Cronin, R. M. Davies, P. Davis, G. Dougan, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Leather, S. Moule, P. C. Oyston, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead and B. G. Barrell (2001). "Genome sequence of *Yersinia pestis*, the causative agent of plague." Nature 413: 523-527.

Perry, R. D. and J. D. Fetherston (1997). "*Yersinia pestis* - Etiologic agent of plague." Clin Microbiol Rev 10: 35-&.

Pezard, C., P. Berche and M. Mock (1991). "Contribution of individual toxin components to virulence of *Bacillus anthracis*." Infect Immun 59: 3472-3477.

Phillippy, A. M., K. Ayanbule, N. J. Edwards and S. L. Salzberg (2009). "Insignia: a DNA signature search web server for diagnostic assay development." Nucleic Acids Res 37: W229-234.

Piercy, R., K. M. Sullivan, N. Benson and P. Gill (1993). "The application of mitochondrial DNA typing to the study of white Caucasian genetic identification." Int J Legal Med 106: 85-90.

Porcar, M. and V. Juarez-Perez (2003). "PCR-based identification of *Bacillus thuringiensis* pesticidal crystal genes." FEMS Microbiol Rev 26: 419-432.

Prentice, M. B. and L. Rahalison (2007). "Plague." Lancet 369: 1196-1207.

Qi, Y., G. Patra, X. Liang, L. E. Williams, S. Rose, R. J. Redkar and V. G. DelVecchio (2001). "Utilization of the *rpoB* gene as a specific chromosomal marker for real-

time PCR detection of *Bacillus anthracis*." Appl Environ Microbiol 67: 3720-3727.

Rasko, D. A., P. L. Worsham, T. G. Abshire, S. T. Stanley, J. D. Bannan, M. R. Wilson, R. J. Langham, R. S. Decker, L. Jiang, T. D. Read, A. M. Phillippy, S. L. Salzberg, M. Pop, M. N. Van Ert, L. J. Kenefic, P. S. Keim, C. M. Fraser-Liggett and J. Ravel (2011). "*Bacillus anthracis* comparative genome analysis in support of the Amerithrax investigation." Proc Natl Acad Sci U S A 108: 5027-5032.

Ratsitorahina, M., S. Chanteau, L. Rahalison, L. Ratsifasoamanana and P. Boisier (2000). "Epidemiological and diagnostic aspects of the outbreak of pneumonic plague in Madagascar." Lancet 355: 111-113.

Reed, C., V. Fofanov, C. Putonti, S. Chumakov, T. Slezak and Y. Fofanov (2007). "Effect of the mutation rate and background size on the quality of pathogen identification." Bioinformatics 23: 2665-2671.

Reynolds, R., G. Sensabaugh and E. Blake (1991). "Analysis of genetic markers in forensic DNA samples using the polymerase chain reaction." Anal Chem 63: 2-15.

Rosen, S. and H. J. Skaletsky (1998). "Primer3." Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html.

Sacchi, C. T., A. M. Whitney, L. W. Mayer, R. Morey, A. Steigerwalt, A. Boras, R. S. Weyant and T. Popovic (2002). "Sequencing of 16S rRNA gene: a rapid tool for identification of *Bacillus anthracis*." Emerg Infect Dis 8: 1117-1123.

Saikaly, P. E., M. A. Barlaz and F. L. de Los Reyes, 3rd (2007). "Development of quantitative real-time PCR assays for detection and quantification of surrogate biological warfare agents in building debris and leachate." Appl Environ Microbiol 73: 6557-6565.

Saikaly, P. E., K. Hicks, M. A. Barlaz and F. L. de Los Reyes, 3rd (2010). "Transport behavior of surrogate biological warfare agents in a simulated landfill: effect of leachate recirculation and water infiltration." Environ Sci Technol 44: 8622-8628.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich (1988). "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase." Science 239: 487-491.

Schmeisser, C., H. Steele and W. R. Streit (2007). "Metagenomics, biotechnology with non-culturable microbes." Appl Microbiol Biotechnol 75: 955-962.

Schnepf, E., N. Crickmore, J. Van Rie, D. Lereclus, J. Baum, J. Feitelson, D. R. Zeigler and D. H. Dean (1998). "*Bacillus thuringiensis* and its pesticidal crystal proteins." Microbiol Mol Biol Rev 62: 775-806.

Schuster, S. C. (2008). "Next-generation sequencing transforms today's biology." Nat Methods 5: 16-18.

Sedlackova, T., G. Repiska, P. Celec, T. Szemes and G. Minarik (2013). "Fragmentation of DNA affects the accuracy of the DNA quantitation by the commonly used methods." Biol Proced Online 15: 1-8.

Sethi, S. (2000). "Bacterial infection and the pathogenesis of COPD." Chest 117: 286-291.

Sethi, S., P. Mallia and S. L. Johnston (2009). "New paradigms in the pathogenesis of chronic obstructive pulmonary disease II." Proc Am Thorac Soc 6: 532-534.

Sethi, S. and T. F. Murphy (2001). "Bacterial infection in chronic obstructive pulmonary disease in 2000: a state-of-the-art review." Clin Microbiol Rev 14: 336-363.

Shendure, J. and H. Ji (2008). "Next-generation DNA sequencing." Nat Biotechnol 26: 1135-1145.

Skottman, T., H. Piiparinen, H. Hyytiainen, V. Myllys, M. Skurnik and S. Nikkari (2007). "Simultaneous real-time PCR detection of *Bacillus anthracis*, *Francisella tularensis* and *Yersinia pestis*." Eur J Clin Microbiol Infect Dis 26: 207-211.

Spencer, R. C. (2003). "*Bacillus anthracis*." J Clin Pathol 56: 182-187.

Steele, J. H. (1962). Anthrax lesion on the skin of the forearm caused by the bacterium *Bacillus anthracis*. Center for Disease Control and Prevention: Public Health Image Library.

Supreme Court of Vermont (1997). State of Vermont v. Michael J. Demers. 96: 452.

Takahashi, H., P. Keim, A. F. Kaufmann, C. Keys, K. L. Smith, K. Taniguchi, S. Inouye and T. Kurata (2004). "*Bacillus anthracis* incident, Kameido, Tokyo, 1993." Emerg Infect Dis 10: 117-120.

Teh, W. L. (1922). "Plague in the Orient with special reference to the Manchurian outbreaks." J Hyg (Lond) 21: 62-76.

Thavaselvam, D. and R. Vijayaraghavan (2010). "Biological warfare agents." J Pharm Bioallied Sci 2: 179-188.

The United States Department of Justice. (2010). Amerithrax Investigative Summary.

Tucker, J. B. and G. D. Koblentz (2009). "The four faces of microbial forensics." Biosecurity and Bioterrorism-Biodefense Strategy Practice and Science 7: 389-397.

Turnbull, P. C. (1999). "Definitive identification of *Bacillus anthracis*--a review." J Appl Microbiol 87: 237-240.

U.S.C.-14132 (1994). DNA Identification Act, U.S.C. 14132: 7685-7686.

Unknown (1993). "This plague patient is displaying a swollen, ruptured inguinal lymph node, or buboe". Center for Disease Control and Prevention: Public Health Image Library.

Viboud, G. I. and J. B. Bliska (2005). "*Yersinia* outer proteins: role in modulation of host cell signaling responses and pathogenesis." Annu Rev Microbiol 59: 69-89.

Wang, D. G., J. B. Fan, C. J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Hudson, R. Lipshutz, M. Chee and E. S. Lander (1998). "Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome." Science 280: 1077-1082.

Wang, Y. and P. Y. Qian (2009). "Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies." PLoS One 4: e7401.

Wrischnik, L. A., R. G. Higuchi, M. Stoneking, H. A. Erlich, N. Arnheim and A. C. Wilson (1987). "Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA." Nucleic Acids Res 15: 529-542.

Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen and T. L. Madden (2012). "Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction." BMC Bioinformatics 13: 1-11.