PATHOGEN DETECTION USING THE LUMINEX MULTI-ANALYTE SYSTEM

by

ANDREW J. SCHWEIGHARDT

A dissertation submitted to the Graduate Faculty in Criminal Justice in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2012

This manuscript has been read and accepted by the
Graduate Faculty in Criminal Justice in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy

Margaret M. Wallace, Ph.D.

Date | Chair of Examining Committee

Joshua D. Freilich, J.D., Ph.D.

Date | Executive Officer

Thomas A. Kubic, J.D., Ph.D.

Nathan H. Lents, Ph.D.

Mechthild K. Prinz, Ph.D.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

PATHOGEN DETECTION USING THE LUMINEX MULTI-ANALYTE SYSTEM

by

Andrew J. Schweighardt

Advisor: Margaret M. Wallace, Ph.D.

The causative agents of anthrax, botulism, tularemia, and plague are so destructive that they could be used maliciously by bioterrorists. Microscopical and biochemical methods of identifying pathogenic bacteria are often inaccurate and time-consuming. Genetics-based alternatives surpass traditional detection methods in terms of speed and accuracy, but often lack the capacity for multiplexing.

Luminex technology embodies all the advantages of DNA-based identification while still offering impressive multiplexing capabilities. Luminex is a liquid array platform that relies on specially labeled microscopic beads to which unique probes are covalently attached. Hybridization of the probes to the unknown target sequences is detected and quantified by the instrument.

Four core pathogenic bacterial species were selected: *Bacillus anthracis*, *Clostridium botulinum*, *Francisella tularensis* subsp. *tularensis*, and *Yersinia pestis*. Probe sequences intended for confirmatory identification of the four core pathogens were located within genes related to the toxicity of each bacterium. Additional probe sequences within the 23S ribosomal RNA gene *rrl* were selected for presumptive identification of the four core pathogens, or a relative belonging to the same genus.

Primer3 and AutoDimer software programs were used during probe and target DNA design to specify oligo parameters and minimize cross-reactivity. The specificity of all oligos was verified using whole genome searches conducted on the NCBI BLAST database. Target sequences were amplified by PCR and quantified using the Agilent 2100 Bioanalyzer. Primer specificity was confirmed by the detection of target band(s) of the requisite length.

Specificity and sensitivity were examined by introducing various increments of each target sequence to all probes. A reasonable degree of specificity was observed, although some probes were prone to false positives. The LLDs for the probes ranged from 0.1 to 10 ng. Sensitivity was improved by using lambda exonuclease to enzymatically digest the non-complementary PCR strand.

Thirty-three binary, ternary, and quaternary mixtures were examined in which all components were present in a 1:1 ratio, with an 80% success rate for identifying all components. Four binary mixtures were subjected to further study in which the two components were combined in various ratios. Both components were detected in 100% of the twenty-eight mixtures analyzed, even when the minor component was overshadowed by a ten-fold excess of the major component. False positives typically had an attenuated fluorescent response that made them easily distinguishable from a genuine positive result.

Overall, the Luminex platform is a robust method of identifying pathogenic microorganisms. This technique has a superior level of specificity and sensitivity, provided that each assay has its underpinnings in a thorough survey of all pertinent bioinformatics data. Continued research with the Luminex technology is important both for validating the utility of this technique and for satisfying legal admissibility standards.

PREFACE

It was August 2005 when I completed data collection for my senior research project at Stonehill College. I spent the rest of that summer and early autumn writing my undergraduate thesis and received my Bachelor's degree the following spring. At the time I thought that I would never again conquer a task as difficult as the one I had just overcome. Nearly six years later as I completed data collection for my graduate research at John Jay College of Criminal Justice and began to write my dissertation, I realized I was wrong to regard my undergraduate work as an accomplishment never to be surpassed.

I am glad for having accepted the challenge placed before me, but it would be disingenuous of me to say that my enthusiasm for the project was not occasionally interrupted by periods in which I regretted my decision to continue. Dr. James Watson, co-discoverer of the structure of DNA, said in the preface to his book *The Double Helix* that science seldom proceeds in the straightforward, logical manner imagined by outsiders. There are some who might regard this as an exaggerated statement, but those who have truly labored in the name of science will recognize its merit. One gains a different perspective upon the completion of a project and having the opportunity to review the results collectively, but in the midst of the actual work, the significance of the data may be nebulous and the ultimate goal seems uncertain and evasive.

Despite the inherent difficulties encountered along the journey, this project was not a futile exercise. The outcome represents a new development in the continued effort towards identifying potential bioterrorist attacks. In what is widely regarded as the oldest military treatise in the world, the ancient Chinese military strategist Sun Tzu said that

"The art of war teaches us to rely not on the likelihood of the enemy's not coming, but on our own readiness to receive him; not on the chance of his not attacking, but rather on the fact that we have made our position unassailable." Any step, no matter how small, that enables us to fortify our defense against bioterrorism is a step that is well taken.

I feel extremely fortunate to have been given the opportunity to offer my share of evidence that science can triumph over evil. Aside from that, I am also grateful that anything I accomplished prior to beginning this project did not represent the culmination of my scientific endeavors, and that I was given the curiosity and willingness to accept one more challenge. Achieving one goal does not preclude aspiring to new ones.

A. J. S.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my mentor, Dr. Margaret Wallace. To thank Dr. Wallace for her support and guidance throughout this research would be appropriate, but certainly an understatement. I have worked with Dr. Wallace for several years, even before this research project began, and I can personally attest to her selfless devotion to the students under her advisement. Aside from the dedication that Dr. Wallace possesses, she has many qualities of a fine researcher that I have grown to emulate. The greatest compliment that I can offer Dr. Wallace is to say that I aspire to match her commitment to forensic science.

I would also like to extend my gratefulness to the Director of the Forensic Science Specialty in the CUNY Criminal Justice Ph.D. program at John Jay College, Dr. Thomas Kubic, who not only introduced me to forensic science as a high school student, but helped my interest in this field to flourish during my time at John Jay. Thank you to the members of the dissertation committee: Drs. Margaret Wallace, Thomas Kubic, Nathan Lents, and Mechthild Prinz, each of whom generously gave their time to oversee the completion of this research. The late Dr. Diana Friedland made valuable contributions during the proposal phase of this research, and her input in the early stages helped to shape the direction of this research for the better. John Jay alumna Ms. Amanda Battaglia of the New York City Office of the Chief Medical Examiner conducted her own research with Luminex technology, and in doing so helped to lay a strong foundation upon which this research project was built. There were many other people of the John Jay College faculty and staff who offered some much appreciated advice and/or assistance along the way. Dr. Lawrence Kobilinsky, Dr. Anthony Carpi, Dr. Nicholas Petraco, Dr. Selman

Berger, Mr. Sing Chin, Ms. Judith Beekman, Ms. Nisha Panikkaveetil, Ms. Natasha Collier, Mr. Nikolay Azar, Mr. Peter Diaczuk, Mr. Dale Purcell, and Ms. Brooke Kammrath are a few that come to mind, but undoubtedly I have forgotten someone, so I sincerely thank the Science Department as a whole for its cooperation and assistance.

Ms. Dee Dee Aikens of the Lloyd Sealy Library at John Jay College recovered many hard-to-find scientific journal articles for me, always doing her best to find exactly what I needed no matter how many requests I made. The Biology 103 and 104 Recitation students whom I have taught over the past five years have been an extraordinary inspiration to me. They helped me to discover my fondness for teaching and always kept me mindful of the need to be a strong academic role model for them. It has been a privilege to watch many of them mature as they progressed through their studies and developed their scientific prowess. I would also like to thank Ms. Jerylle Kemp, Director of Alumni Relations at John Jay College, Mr. Michael McCann, President of the Board of Alumni, and the entire John Jay College Alumni Association Board for their encouragement and for instilling within me a sense of pride for being a John Jay alumnus.

My family sacrificed a great deal of time and energy on my behalf as I was working towards this goal; without their support the task before me would have been infinitely more difficult. I thank my parents, Rosalie and Emery, for always reminding me to appreciate the journey, no matter how appealing the final destination may seem. My grandmother Rosalie Wilson offered sympathy and encouragement on so many occasions when I was uncertain if I could continue pursuing this endeavor. As I conducted this research I watched my younger sister Katie diligently study for and successfully complete her CPA licensing exams. I thank her for allowing me to see what

it means to be dedicated to one's profession; I can only hope to match her tenacity and professional commitment. My immediate family knows how earnestly I fought to complete this project, and they often put my needs before their own just to make my life a little easier over the past few years. My extended family regularly inquired about my well-being, never letting me forget that I was not the only one anxiously anticipating my graduation. I thank my entire family for never losing their faith in me and always reminding me that I had their unconditional love and support.

I feel so fortunate to have had a strong network of friends during this project. They attentively listened to stories about my work, periodically inquired about my progress, and offered encouragement with such alacrity that I oftentimes felt that I must continue with this project, if for no other reason than not wanting to disappoint them. I thank all my friends for politely listening to me complain about the hardships of research and encouraging me to prevail. Dr. Maria Curtin was kind enough to invite me to the Shields Science Center at Stonehill College to discuss my research with science faculty and students, whereupon I received a great deal of valuable insight and suggestions. I am thankful to Mr. James Kolczynski for inviting me to speak with his forensic science class at Glen Cove High School; my visit with the students there helped me to remember what drew me to forensic science as a young person. I also had many fruitful exchanges over lunch with former Stonehill College classmate Dr. Jessica Falco, now of the Weill Cornell Medical Center. Mr. Robert Miller, Ms. Joyce Sviba, Ms. Rita Wilkes, Dr. Sharon Popper, Dr. Gary Stone, and many more friends at the Huntington YMCA kept me company as I swam laps to forget my research troubles. The congregation of the Greenlawn Presbyterian Church offered their prayers and compassion on many occasions

when I was deeply overwhelmed with this project. My longtime friend Clifford Beierle often facetiously referred to me as a career student, and this motivated me to finish my degree if for nothing else than to prove him wrong. My good friend Seamus Farrelly, with whom I worked closely at Stonehill College for several years, offered sage advice when I needed it most. He once told me that my education is like getting to the Super Bowl; it is a long journey, you have to win a lot of games, you have to be lucky, and most of all, you need to have a good team around you. I am thankful to have had the best team anyone could ask for.

The munificence of several individuals and organizations was very much appreciated in terms of meeting the fiscal demands of this research project. Partial financial assistance needed to complete this research was provided by a Forensic Science Foundation Lucas Grant, the Luminex Corporation, John Jay College of Criminal Justice, a Start-up Grant awarded by Vice Chancellor Gillian Small of The City University of New York, and Ms. Amanda Battaglia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ss……..……..single-stranded

ds……..……double-stranded

RRL…….…….....sequence originating from the *rrl* gene

TOX…….……sequence originating from a gene related to an organism's toxicity

Bac/RRL……....sequence originating from the *rrl* gene of either *B. cereus* or *B. anthracis*

Bant/TOX……..…..sequence originating from the *lef* gene of *B. anthracis*

Cdif/RRL……..sequence originating from the *rrl* gene of *C. difficile*

Cbot/RRL……...sequence originating from the *rrl* gene of *C. botulinum*

Cbot/TOX……....sequence originating from the BONT/F gene of *C. botulinum*

Fran/RRL…….sequence originating from the *rrl* gene of either *F. tularensis holarctica* or *F. tularensis tularensis*

Ftul/TOX…….....sequence originating from the FTT0435 gene of *F. tularensis tularensis*

Yent/RRL…….sequence originating from the *rrl* gene of *Y. enterocolitica*

Ypes/RRL……..sequence originating from the *rrl* gene of *Y. pestis*

Ypes/TOX……...….sequence originating from the *pla* gene of *Y. pestis*

Myc/RRL……...……....sequence originating from the *rrl* gene of *M. tuberculosis*

BC……..*Bacillus cereus*

BA…….*Bacillus anthracis*

CD…….…….*Clostridium difficile*

CB…….…….*Clostridium botulinum*

FH…….……..*Francisella tularensis* subsp. *holarctica*

FT……...……*Francisella tularensis* subsp. *tularensis*

YE…….…….*Yersinia enterocolitica*

YP…….……..*Yersinia pestis*

MT……..…….*Mycobacterium tuberculosis*

T_M……..melting temperature

NC……..……..negative control

PC……..….positive control

CHAPTER I. INTRODUCTION

Biological weapons are a perennial threat to our society. The use of biological pathogens for nefarious purposes began thousands of years ago. One of the earliest instances of biological weapons in recorded history dates back to 184 BC, when Hannibal, the great Carthaginian military commander, had his navy throw clay pots filled with poisonous snakes on to the decks of enemy naval fleets (Ghosh *et al.*, 2002). The use of biological weapons has grown more prevalent through the ages. During the medieval era, the Mongols used a trebuchet to catapult cadavers of plague victims into enemy territories (Martin *et al.*, 2007). In the 1500s, when the Spanish conquistadors came to the Americas, they gave the Native Americans goods that had been deliberately tainted with smallpox.

Instances of biological weapons became even more abundant in the 20th century. During World War II, the Japanese initiated a massive biological weapons research program, killing thousands of Chinese prisoners who were involuntarily used as test subjects (Christopher *et al.*, 1997). The United States and the Soviet Union were two of the major forces behind biological weapons development during the Korean and Cold War eras. Many of these aggressive research programs were curtailed in 1975 when the United Nations proposed the Biological Weapons Convention, a treaty initially signed by 103 nations agreeing never to research, produce, or use biological weapons (Martin *et al.*, 2007). Today, much of the threat of biological weapons stems from rogue nations and terrorist groups who have not acknowledged the disarmament. They engage in what is known as “bioterrorism”, which is the use of biological agents to achieve a political or ideological objective (Martin *et al.*, 2007).

The motives behind bioterrorism and the technologies used for growth and dispersal of biological pathogens have evolved, but the outcome has remained the same – biological weapons have the potential to cause a devastating number of illnesses and/or fatalities. Based on the fact that terrorists employ tactics that have previously demonstrated effectiveness, it is unlikely that the malicious use of biological pathogens will wane in the future. Therefore, bioterrorism is an inevitable part of the future of our international community and must be approached with preventative and responsive measures commensurate with its seriousness.

Early uses of biological pathogens fall under the umbrella of biological warfare. However, it is important to distinguish biological warfare from biological terrorism. Whereas warfare typically involves two warring nations engaging their armed forces in combat, terrorism often victimizes innocent, unarmed civilians. Terrorists are not restrained by the same ethics, international etiquette, or legal ramifications that would normally protect civilians from becoming involved in such conflicts. Success is also easier for terrorist organizations to achieve because they are not concerned with immediate victories. On the contrary, terrorists are resilient and resourceful. They will patiently and systematically attack their target until their demands are met or until they are stopped.

The fate of bioterrorism lies not in the discovery of pathogens, for it is well known that such pathogens exist and have been used with varying degrees of success. Instead, the fate of bioterrorism hinges mainly on two factors: the ability of terrorists to isolate and propagate these pathogens and their ability to overcome the dispersal dilemma. It has been difficult for terrorist organizations to safely and efficiently

propagate biological pathogens because these groups often employ amateur scientists working in sub-standard laboratories. There have been several documented incidents in which efforts to grow and manipulate biological pathogens have gone awry and have caused significant fatalities (Meselson *et al.*, 1994). It is also likely that terrorist organizations are focusing their efforts on devising methods that will effectively disperse a biological pathogen over a vast area. This will enable the terrorists to cause a catastrophe of the magnitude that will evoke the desired response from the target nation.

Although bioterrorism is a likely part of the future, there are certain ways that target nations can prepare for it. The defensive strategy involves both the prevention of future attacks as well as the development of procedures for responding to these attacks. A bioterror attack cannot be addressed until the attack has been identified as such. In other words, it must be positively ascertained that the catastrophe at hand is indeed the result of a maliciously introduced biological pathogen. Therefore, much research has sought to provide an accurate and rapid method for identifying biological pathogens so that the appropriate response can commence.

CHAPTER II. LITERATURE REVIEW

A. Non-genetic Identification of Biological Pathogens

Many researchers have generally agreed upon about a half dozen criteria that a method for detecting biological pathogens should possess (Spiro *et al.*, 2000; Vignali, 2000). These criteria include quantitative abilities, multiplexing capabilities, high specificity, and high sensitivity (even in the presence of contaminants). Many technologies satisfy some of these criteria, but very few satisfy them all.

Previous methods of identification have involved time-consuming and inefficient methods such as microscopic identification based on a comparison of morphological characteristics of the bacteria. Other methods that have been relied upon in the past are a series of enzymatic tests or staining procedures which, collectively, offer a fairly unique profile of the pathogen in question (Lim *et al.*, 2005). However, these methods provide nothing more than a presumptive identification because they do not offer enough discrimination to differentiate one closely related pathogen from another.

Aside from microscopy and enzymatic testing, pathogens can also be identified by the enzyme-linked immunosorbent assay (ELISA) or other immunoassays (Candlish, 1991). In ELISA, antigens from the unknown are immobilized on a solid surface. Specific antibodies are then washed over the antigens. If binding between the antigen and antibody occurs, this is usually detected via a fluorescent emission created by an enzyme that was attached to the antibody. ELISA is generally an inexpensive and rapid technique. The cost of the test is negligible when compared to more involved instrumental analyses. However, the results are often less quantitative and the test itself is less sensitive than

alternative analyses. Perhaps the greatest disadvantage of ELISA is that it is a monoplex technology.

There are several instrument-based methods of identification available now which have attracted the attention of more researchers. Many of these instrumental methods are variations of mass spectroscopy (MS). A mass spectrometer uses an energy source to fragment the analyte into small fractions. These particles are then ionized and separated based on their mass-to-charge ratios. There are several ways to separate the ionized particles, but many researchers have relied on the time-of-flight (TOF) method. TOF is often paired with a variation of MS called matrix assisted laser desorption and ionization (MALDI) for pathogen detection purposes (Hurst *et al.*, 1996; Taranenko *et al.*, 2002; Briese *et al.*, 2005). In MALDI-TOF, the liquefied analyte (in this case, the analyte can be a nucleic acid or any biomacromolecule from the pathogen) is first mixed with a liquid matrix compound. The analyte and the matrix co-crystallize as they dry on a plate. The plate is then transferred to a mass spectrometer where a laser pulse initiates ionization. Ionization imparts a uniform charge on the entire analyte so that any differences in separation are due to mass alone (not charge). The ions travel down the flight tube, with the smaller molecules traveling faster toward the end of the flight tube, where the detector is located. Pathogenic microorganisms can be identified using MALDI-TOF because each microorganism should have unique ionization and TOF profiles, which can be compared to known samples that have been previously analyzed.

MS-based approaches to pathogen identification such as the Ibis T5000™ Biosensor System have garnered a great deal of attention from the research community in recent years (Ecker *et al.*, 2008). One of the main benefits of MS-based technologies is

that they are rapid due to the vacuum environment in which the analysis occurs (Butler, 2005). Although the use of MS does not necessarily require any specific knowledge of the analyte (*e.g.*, nucleotide sequence), a mass spectra library of all possible analytes must be assembled prior to use. Other drawbacks are that the resolution and sensitivity may be compromised if there is excess analyte and/or salt present (Butler, 2005). Lastly, as with many instrumental methods, the start-up and long-term costs may be prohibitive for smaller laboratories operating on a limited budget.

B. Genetic Identification of Biological Pathogens

1. Overview

In contrast to many of the more traditional methods of identification, most methods in development focus on the use of genetic markers to identify biological pathogens. A genetic marker is any polymorphic region of an organism's genome that can be used for positive identification. Regions of the genome that are good candidates for markers are those that offer the ability to distinguish between very closely related microorganisms. The identification of reliable markers is the first step toward positively identifying a biological pathogen and mounting the appropriate response.

It should be noted that some research is focused on individualizing polymorphic regions of bacterial DNA for the purpose of tracing the outbreak to its original source. Although this is a worthwhile research agenda, it should not take priority over the primary objective of identifying the pathogen so that the outbreak can be mitigated. The purpose of genetic markers, in the context of this dissertation research, was to distinguish

a pathogenic microorgansm from other similar species, not to individualize the microorganism for the purpose of source attribution.

Recent genome sequencing efforts have enabled researchers to elucidate the genetic composition of many microorganisms. For example, the entire genome of *Bacillus anthracis* is now known (Read *et al.*, 2003). However, with this enlightenment comes the humbling awareness of what remains to be discovered. Startling similarities among many of the sequenced genomes have also been revealed. In other words, a particular microorganism may be difficult to identify because it has very few genetic polymorphisms that separate it from close relatives. Therefore, the development of an array capable of distinguishing even closely related pathogenic microorganisms remains an area of great research interest.

There are many ways to identify a pathogenic microorganism. Techniques exist that yield identifications based on analytes as diverse as peptides, antigens, and other ligands. The primary disadvantage of many of these methods is that they lack the specificity needed to acquire a positive identification. It is likely that any viable technique for the identification of pathogenic microorganisms will have a foundation in genetics, because an organism's DNA is its most unique, quantifiable attribute. Genetic techniques are also amenable to real-time, field deployable analyses because many involve rapid and semi-automated detections on instruments that are semi-portable.

2. Techniques Based on Length Polymorphisms: RFLP & AFLP

Polymorphism is the term used to describe the differences in the DNA from different organisms. Length polymorphisms represent differences in the size of the DNA upon fragmentation. Many early genetic techniques were based on length polymorphisms

because little knowledge of the sequence was required. Analyses based on length polymorphisms also typically require less instrumentation because, once the fragments have been separated on a gel, the differences in length can be easily visualized.

One technique that is based on length polymorphisms is called Restriction Fragment Length Polymorphism (RFLP). It begins by cutting the DNA at precise locations, resulting in a group of DNA fragments. These fragments are said to be polymorphic in length because DNA from separate origins (*i.e.*, different microorganisms) produces different sized fragments. The cleavage of the DNA is accomplished through the use of restriction enzymes, which cleave DNA molecules wherever a particular short, palindromic sequence is encountered (these locations in the genome are called "restriction sites"). Closely related microorganisms share a higher number of restriction sites, and thus their DNA is cleaved similarly when subjected to the same restriction enzyme. However, microorganisms that are distantly related have fewer restriction sites in common, thus leading to a greater degree of polymorphism in the length of fragments that are produced. The fragmented DNA is separated by electrophoresis and visualized using a fluorescent dye (usually ethidium bromide).

RFLP became the dominant method of genetic analysis in forensic biology during the 1980s (Jeffreys *et al.*, 1985). Forensic-based RFLP is somewhat more complex than conventional RFLP because it demands more knowledge of the DNA sequence. In forensic applications, double-stranded RFLP fragments called Variable Number Tandem Repeats (VNTRs) are electrophoretically separated. These fragments are converted to the single-stranded state in an alkaline environment and then blotted on to a nylon or nitrocellulose membrane. A single-stranded DNA probe (designed to hybridize with the

fragments, if present) is then introduced. Knowledge of the DNA sequence is required because the probe is designed to specifically hybridize with a particular sequence of bases in the fragmented DNA. Hybridization between the probe and the sample is detected by a radioactive or fluorescent emission, depending on how the probe was labeled.

About a decade after RFLP became the premier technique for genetic analysis, a Dutch laboratory developed a modified version of RFLP, which was called Amplified Fragment Length Polymorphism (AFLP; Vos *et al.*, 1995). In the AFLP technique, the DNA is still cut with restriction enzymes. However, it is then amplified by PCR using primers that anneal to adapters ligated to the restriction fragments. AFLP was an improvement on some RFLP-based techniques because it did not require knowledge of the sequence, except for the restriction sites. AFLP also incorporated the benefits of PCR, such as the ability to work with samples of small quantity. In many ways, AFLP has led to a renewed interest in length polymorphisms as a means of genetic identification. Researchers have used AFLP to successfully differentiate bacterial (Janssen *et al.*, 1996), parasitic (Folkertsma *et al.*, 1996), fungal (Rosendahl and Taylor, 1997), and plant specimens (Bless *et al.*, 2006). However, RFLP-based techniques are no longer in vogue for human identification in criminal investigations because they have been supplanted by modern techniques such as STR analysis (Edwards *et al.*, 1991). Techniques based on length polymorphisms are also not preferred by some researchers studying microbial forensics because of the inherent ambiguity in interpreting the electropherograms, the inability to distinguish very closely related microorganisms (Radnedge *et al.*, 2003; Valjevac, 2005), and problems with reproducibility (Fry *et al.*, 2009).

3. Techniques Based on Sequence Polymorphisms

The prior section explored the use of polymorphisms that are based on the *length* of the DNA fragments. In forensic-based RFLP, these fragments derive their unique length from the possession of a certain amount of repetitive DNA. This implies that a core sequence of nucleotides may be repeated one, two, three, or more times in any particular organism. Thus, it is the length (*i.e.*, number of repeats) of the DNA fragment that constitutes the polymorphism. In contrast, polymorphisms based on the DNA *sequence* itself can also be exploited for identification purposes. This implies that the base sequence at a particular locus may be different from one organism to the next.

One of the earliest methods of individualization used by forensic biologists that was based on sequence polymorphisms was the AmpliType® HLA-DQα kit (Walsh *et al.*, 1991). This kit was appealing to the forensic biology community because it was amenable to PCR, which helped to improve the sensitivity when screening crime scene samples that were often limited in quantity. The basis of the test was the binding of amplified target DNA to sequence-specific oligonucleotide probes that were immobilized on a nylon strip. The main disadvantage of the HLA-DQα kit was that it could only examine one locus of the DNA, thus limiting the statistical significance of a "match." Eventually this was remedied with the introduction of AmpliType® PolyMarker, which enabled simultaneous screening of five loci (Dimo-Simonin and Brandt-Casadevall, 1996).

In sum, sequence polymorphisms can be identified by sequencing genes, chromosomes, or even whole genomes. However, for the purposes of identification, sequencing such large portions of the genome could overwhelm the analyst with

superfluous information. Much smaller portions of the genome, called “markers”, can be examined for sequence polymorphisms. In microbial forensics, genetic techniques like DNA microarrays and bead-based technologies use markers to identify sequence polymorphisms. These techniques will be discussed in a later section.

4. Quantitative PCR

Quantitative PCR (sometimes referred to as real-time PCR) uses the real-time accumulation of PCR product to identify and quantify the target DNA. Quantitative PCR sometimes relies on intercalating fluorescent dyes that bind to the double-stranded PCR amplicons. Other protocols use special oligonucleotide probes to achieve greater specificity. These single-stranded probes have a fluorescent reporter at one end and a quencher at the other end. As long as the probe is intact, the quencher suppresses the fluorescence of the reporter. If the probe binds to the target DNA, the probe is cleaved by the *Taq* polymerase enzyme as it moves along the DNA template in the 5′ to 3′ direction during the extension phase of PCR. The cleavage of the probe separates the reporter from the quencher, thus permitting the reporter to emit its fluorescence.

Although quantitative PCR has a superior specificity and sensitivity (Pabbaraju *et al.*, 2011; Taniuchi *et al.*, 2011), its main drawback is its lack of multiplexing ability (Skillman *et al.*, 2006). Quantitative PCR is technically capable of multiplexing because of its use of the five-dye chemistry, but this does not nearly approach the multiplexing capabilities of other technologies such as DNA microarrays or bead-based assays.

5. DNA Microarrays

The DNA microarray is probably the most formidable competitor of the bead-based technologies. The most well-known and the most closely related predecessor of DNA microarrays is the Southern blot (Southern, 1975). The Southern blot relies on the hybridization of a labeled probe to target DNA that has been immobilized on an absorbent substrate. First proposed by Saiki *et al*. (1989), DNA microarrays differ from their blotting ancestors because microarrays use a non-porous substrate (typically glass) to immobilize the DNA. Microarray technology involves multiple reactions on a single glass plate, called a microarray or "DNA chip."

In DNA microarrays, a glass chip is first imprinted with covalently attached oligonucleotide probes. The analyte (*i.e.*, the unknown) usually consists of short segments of DNA that will hybridize easily with the immobilized probes, if any complementarity exists. The hybridization between probe and target is typically accomplished by using target DNA that has been radioactively or fluorescently labeled. In earlier applications of DNA microarrays, the analyte DNA was labeled with a radioactive marker, whose hybridization to the target could be visualized using autoradiography. Today it is more common to use safer methods of detection that rely on fluorescent dyes. The energy emitted by the tagged analyte upon hybridization is used as evidence that hybridization to the immobilized probe has occurred.

DNA microarrays achieved a great amount of success after their introduction because the use of a microarray essentially reduces the reagent and analyte requirements to a negligible amount. Many researchers were also drawn to microarrays because of the inherent multiplexing capabilities; a single microarray chip can be imprinted with

thousands of oligonucleotides. Laboratories that conduct many repetitive experiments also found their work amenable to microarray technology because the DNA chips can be manufactured reliably and reproducibly.

Although there are many advantages to microarray technology, there are also some drawbacks. Compared to other instrumental techniques, the start-up costs associated with microarray technology are fairly reasonable, yet for some laboratories the investment may not be feasible. Additionally, smaller laboratories may find it overwhelming to constantly produce the chips in-house for repetitive assays, thus requiring the allocation of funds for commercial production of the DNA chips. This continuous purchase of commercially prepared chips can represent an undesirable long-term investment in the microarray technology. Indeed, it is well-documented that many researchers are dissatisfied with the equipment requirements and the need to constantly regenerate new arrays (Spiro *et al.*, 2000; Grate *et al.*, 2003; Summerbell *et al.*, 2005).

There are also several problems associated with the immobilization of the probes on the glass chip. One problem is that the probes themselves may exhibit complementarity to their neighbors, so the position of the probes must be carefully chosen so as to avoid these undesirable interactions. The second problem is that the glass substrate often precludes a uniform tendency for hybridization throughout the probe molecule. For example, the bases closest to the end of the strand near the glass are more prone to steric hindrance, owing to the fact that they are sheltered by neighboring probes. The third, and most serious, problem is that DNA microarrays are based on solid-phase two-dimensional chemistry. This means that, because the probes are immobilized on a stationary glass chip, binding of the probe and target DNA depends solely on the mobility

of the target DNA. Ideally, both the probe and target DNA would be mobile, thus maximizing the opportunities for collision and hybridization.

There have been several proposed solutions to the problems with DNA microarrays, some of which have produced favorable results. However, these issues will never be totally resolved because they stem from characteristics that are intrinsic to solid-phase assays. In contrast to DNA microarrays, liquid-phase assays such as bead-based technologies are not as susceptible to the thermodynamics associated with the binding of the probe to its target, for reasons to be discussed later.

6. Luminex Technology

Several instances have made it abundantly clear that specific, sensitive, quantitative, and rapid identification of pathogenic microorganisms are necessary if fatalities are to be minimized in the event of a bioterror attack. It has become apparent that traditional methods of analysis are grossly inadequate in this regard. Even more modern genetic analyses are lacking in one or more areas. Luminex® is at the forefront of modern technologies poised to supplant traditional techniques that have outlived their usefulness.

Bead-based liquid array technologies like Luminex represent a convergence of several technologies, including DNA technology, microsphere technology, and flow cytometry. As a hybrid of these technologies, microsphere-based liquid arrays also represent the nexus of analytical chemistry and molecular biology (Chandler *et al.*, 2000). In the past, this has been a rare combination, but it is becoming an increasingly useful and popular trend in modern bioanalytical techniques.

At its core, the Luminex instrument is a flow cytometer. Flow cytometry is a relatively new technique that was patented in the early 1970s (Kettman *et al.*, 1998) and was introduced to clinical and research laboratories shortly thereafter (Horan and Wheeless, 1977). Flow cytometry is a method of sorting microscopic particles. It serves as an integral component of particle identification in the Luminex system. Thus, it is appropriate to briefly discuss flow cytometry by itself before its application to the Luminex technology is discussed.

As its name suggests, flow cytometry was originally designed to sort cells. As the technique evolved, so did its range of potential analytes. Today, flow cytometry is commonly used to sort and facilitate the detection of almost any microscopic particle. At the advent of its development, flow cytometry was generally considered a novelty with few practical applications. However, recent technological advancements have improved sample labeling, excitation, and detection methods, to the point that flow cytometry has established itself as a venerable necessity in many laboratories. Additionally, flow cytometry possesses the two hallmarks that any technique must have if it is to receive serious consideration and widespread use in today's scientific community: multiplexing capabilities and high sample throughput.

Flow cytometry consists of a pool of microscopic particles that are forced "single-file" through a narrow channel. The particles are usually immersed in some liquid medium and pressurized, which serves to guide the particles in their one-way movement. However, in order for the particle to be identified as it flows through the channel, it must be labeled in some way. It is usually not the particles themselves that are labeled; instead, the substance that binds to the particles (*i.e.*, the analyte, or unknown) is labeled.

Therefore, any particles that have not undergone a hybridization event pass through the channel undetected, whereas the particles that have hybridized to the analyte will have undergone some change that makes them detectable. Usually this detection is made possible by labeling the analyte with a fluorescent dye. The analyte could be something as diverse as an antigen, a peptide, a nucleotide, or another ligand.

The reporter dye would be useless unless there was a means of exciting it, thereby facilitating its detection. Fluorescence, by definition, is the excitation of some particle with high frequency energy in order to yield the low frequency fluorescent energy emission. The channel inside a flow cytometer has an illumination device that consists of a light, which is designed to excite the reporter dye. The light is usually a diode laser, but other light sources are also possible. When the light excites the reporter dye, the reporter dye fluoresces. This fluorescence is detected by a detection apparatus. The detection apparatus is capable of receiving fluorescent energy that has been scattered in the forward direction (*i.e.*, propagating in the same general direction as the light beam that originally bombarded the sample), but also light that has been "side-scattered." More complex experiments can seek detailed information about the particles by interpreting the relative quantities of forward- and side-scattered energy.

The main benefit of using flow cytometry stems from the rapidity with which this technique is capable of examining the analyte. Although the particles flow through the channel one at a time, they move so fast that anywhere from several hundred to several thousand of them may pass through the channel per second, depending on the capabilities of the particular instrument.

The Luminex instrumentation is essentially a basic flow cytometer with some modern adaptations. The Luminex instrument is capable of identifying not just one, but many analytes. This is accomplished via the use of two different dyes that are used to internally label small polystyrene microspheres (*i.e.*, beads) that are 5.6 μm in diameter. The two dyes are combined in five-hundred different proportions to create five-hundred different types of microspheres.

In bead-based assays like Luminex, a unique DNA probe is attached to each of the five-hundred different groups of microspheres (other ligands such as RNA or protein can also be attached to the microspheres, depending on the assay). Probes must be carefully selected so as to ensure minimum cross-reactivity. It is recommended that the probes be about 20 bases in length (Fulton, *et al.*, 1997; Dunbar and Jacobson, 2007). Some researchers have used modifiers that fit between the actual probe and the microsphere, in order to reduce steric hindrance (Wilson *et al.*, 2005). The specificity of the probes can be evaluated by comparing signal intensities for perfectly matched and single-base mismatched duplexes (Chandler and Jarrell, 2004).

The target DNA (*i.e.*, the unknown) is biotinylated. Biotin is a vitamin that has a strong affinity for streptavidin (a protein) which is complexed to phycoerythrin (a fluorophore). Thus, when the streptavidin-phycoerythrin (SAPE) reporter is introduced to the biotin on the target DNA, they bind and, in doing so, enable a quantitative estimate of the number of targets that have hybridized with the probe. Currently the most common reporter to be used is SAPE, but some researchers have experimented with other reporters (Wang *et al.*, 2004) in response to complaints that high reporter fluorescence may

interfere with the fluorescence of the internal dyes of the microspheres (Kettman *et al.*, 1998).

Target DNA is derived from highly polymorphic, thermodynamically stable, and non-hairpin-forming regions of the DNA (Summerbell *et al.*, 2005). It is recommended that the target DNA molecules be about 200 bp in length (Dunbar and Jacobson, 2007). Longer sequences are not preferred because this can create steric hindrance that precludes the target from binding its probe. However, the target sequences should be at least 200 bp in length because this minimizes the chance that slight sequence complementation between mixed targets would lead to the hybridization of two targets (instead of the intended hybridization between probe and target). Sequences shorter than 200 bp are not readily feasible because target DNA includes both a region that is complementary to the probe (~20 bases), plus the forward and reverse flanking regions that must be available for the PCR primers to bind (at least 20 bases in either direction). Furthermore, if one desires to use degenerate primers capable of binding multiple DNAs, regions of high similarity among the different DNAs (*i.e.*, the easiest place to design a degenerate primer) may be well outside the region of the target that is complementary to the probe.

Complementary probes and targets undergo hybridization (Figure 2.1). It is the fluorescence emitted by the reporter molecules (bound to the targets, which in turn, are hybridized to the probes) that is detected and quantified by the instrument. The fluorescence of a reporter can be attributed to a particular probe because the probe will have been attached to a microsphere containing a known proportion of the two internal dyes. Thus, the fluorescence of the reporter dye can be traced back to a particular DNA

probe solely because the unique internal dye combinations give each microsphere set a unique spectral signature.

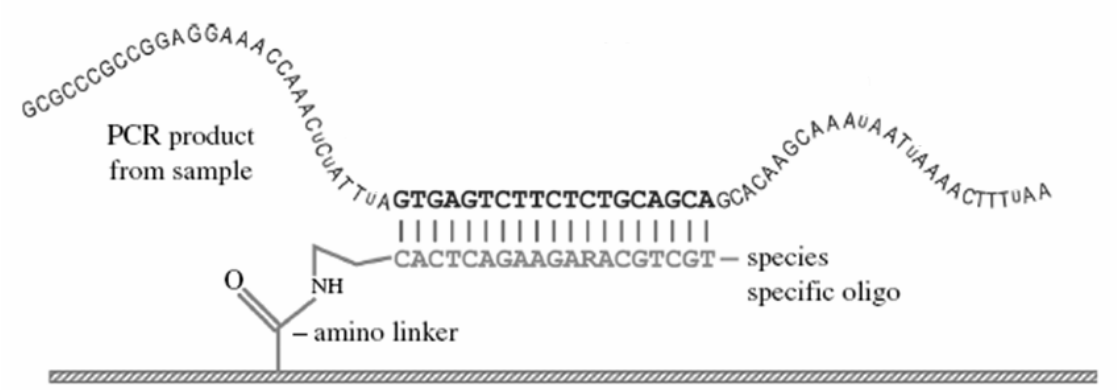


Figure 2.1. Hybridization of target DNA and probe. Probes are attached to microspheres via an amino linker. Originally published as Figure 1 on page 1900 in Microcoding: the second step in DNA barcoding (*Philososophical Transactions of the Royal Society B.* **360**: 1897-1903) by Summerbell, R.; Levesque, C.; Seifert, K.; Bovers, M.; Fell, J.; Diaz, M.; Boekhout, T.; de Hoog, G.; Stalpers, J.; and Crous, P. (2005). Reprinted with permission from Highwire Press and R.C. Summerbell.

a) Advantages of Luminex Technology

The multiplexing ability of the Luminex technology is perhaps its most impressive attribute. The Luminex system is able to achieve superior multiplexing ability because of its use of the dual dye system, which enables up to five-hundred different internal dye variations. Multiplexing is something that has become especially important in the face of time and resource limitations. Luminex is one of the few technologies that answers the need for multiplexing without sacrificing sensitivity or specificity. Luminex technology is amenable to multiplexing because the various internal dye combinations enable researchers to simultaneously screen for up to five-hundred different analytes.

Aside from its multiplexing capabilities, the greatest advantage of the Luminex technology is its high throughput. Not only does this enhance laboratory efficiency, but it

also poses the advantage of conducting numerous trials within a short time frame. Another advantage of this technology is that each assay occurs in its own well, translating into less contamination and fewer opportunities for cross-reactivity. Only very minute quantities of analyte are needed for each assay, which is especially important in forensic applications, when the sample may be very limited.

The sensitivity of the microsphere technology is quite impressive. One study demonstrated that detection was feasible when just 0.1% of the probes on the bead surface were occupied with target DNA (Spiro *et al.*, 2000). In another study (Biagini *et al.*, 2004), the limits of detection for high-performance liquid chromatography-mass spectrometry (HPLC-MS) were compared to those for the Luminex technology. The results indicated that HPLC-MS had a limit of detection of 0.4 ng/mL, whereas the Luminex technology had a limit of detection of 0.10 ng/mL. In yet another study, target sequences comprising as little as 0.3% of the total analyte solution were successfully detected and identified (Spiro and Lowe, 2002).

The simple but elegant design of the microsphere technology, coupled with its ability to produce real-time results, has led to the application of this technology to help solve numerous practical problems. For example, environmental monitoring for allergens, pesticides (Biagini *et al.*, 2002), and biowarfare agents (McBride *et al.*, 2003) has become increasingly prevalent in recent years. Luminex technology has even been combined with mobile formats, such as the autonomous pathogen detection system (APDS), which is designed for on-site, real-time monitoring for biowarfare agents (McBride *et al.*, 2003; Hindson *et al.*, 2004).

b) Disadvantages of Luminex Technology

One of the more general disadvantages is that the microsphere technology may require a considerable start-up investment. The retail price of the instrument itself is approximately $70,000, plus about $10,000 for the accompanying software. Operating costs may also be quite demanding, depending on the scope of the study. Each marker has its own probe which is attached to its own bead set. Probes, reverse probes, and target DNA must all be purchased from a commercial oligonucleotide vendor. Many researchers view the start-up and operating costs as a small price to pay for the capabilities of the Luminex instrumentation. Although these costs may be readily absorbed by a large research laboratory, smaller laboratories may find the financial commitment too demanding.

Luminex research also requires a substantial investment in background research that must precede the benchwork. The research must be carried out using probe and target sequences that have underdone extensive studies prior to their use. For example, it must be confirmed that the target DNA is derived from highly polymorphic, thermodynamically stable, and non-hairpin-forming regions of the DNA (Summerbell *et al.*, 2005). The research and preparation involved (especially if hundreds of analytes will be screened) can be quite laborious. Nevertheless, the data output mirrors the work that preceded it, so it is important to invest the appropriate amount of time in the early phases of the research. Once the initial background work is completed, many assays can be completed without the need for any further bioinformatics research.

Problems with cross-hybridization on the Luminex instrument have been encountered on several occasions. The impact of cross-reactivity may be more

pronounced when more analytes are tested. In order to resolve this issue, unpredictable probes (*i.e.*, those derived from regions of the genome that undergo high rates of mutation) should be avoided and/or removed from use (Chandler and Jarrell, 2004). Multiple probes can also be used to distinguish each microorganism. Doing so minimizes the chance that cross-reactivity with one probe will hamper identification using the other probes. In sum, cross-reactivity has been observed in Luminex assays, but it is often to a much lesser degree than the cross reactivity observed in alternative methods such as DNA microarrays (Spiro *et al.*, 2000).

7. Luminex Research at John Jay College

a) Overview

The Luminex instrument was purchased by John Jay College in August 2008. A pilot study was undertaken in order to explore the abilities of the Luminex platform. The findings of the pilot study were described in an unpublished Masters thesis (Battaglia, 2009) and in the literature (Battaglia *et al.*, 2011).

First, four distantly related microbes of minimum pathogenicity were selected: *Bacillus cereus*, *Escherichia coli*, *Salmonella enterica*, and *Staphylococcus aureus*. A unique twenty base probe (residing in the 23S rRNA gene *rrl*) was selected for each of the four microbes based on data available from prior research (Dunbar and Jacobson, 2007). The probes were each coupled to a unique microsphere set. Target DNA was amplified using genomic DNA of the four chosen microorganisms.

b) Specificity Study

In order to obtain some measure of this assay's specificity, each target DNA was separately combined with a mix of all four bead sets. The results indicated the specificity of the probe-coupled beads to their respective target DNAs (Table 2.1). Some false positives were observed in this and other assays (data not shown), indicating possible cross-reactivity between *E. coli* and *S. enterica*.

Sample	BC Probe	EC Probe	SE Probe	SA Probe
Bacillus cereus	1412	0	0	0
Escherichia coli	17	1130	0	0
Salmonella enterica	3	27	868	0
Staphylococcus aureus	1	0	10	579
Neg. control	180	116	104	100

Table 2.1. Raw MFIs for selected samples in specificity study. Highlighted values indicate designed complementarity. Originally published as Table 4 on page 763 in Pathogen Detection Using a Liquid Array Technology (*Journal of Forensic Sciences* **56**: 760-765) by Battaglia, A.; Schweighardt, A.; and Wallace, M. (2011). Reprinted with permission from John Wiley & Sons.

Further investigation into the cross-reactivity between *E. coli* and *S. enterica* revealed that these two microbes shared an 85% sequence homology in the region of the *rrl* gene which contained the probe. An additional source of the cross-reactivity observed between these two microbes may have been related to the location of the polymorphisms within these probes. Most of the polymorphism between the *E. coli* and *S. enterica* probe resided toward the end of the probe. Ideally, polymorphisms should reside in the center of the probe (Dunbar and Jacobson, 2005; Ishii *et al.*, 2008), because bases in the middle hybridize first and move outward to the 5′ and 3′ ends. Polymorphisms located toward the ends of the probe create a problem because, in the case of a mismatched probe and

target, most of the duplex may be irrevocably formed before the polymorphism is encountered at the termini.

c) Sensitivity Study

A sensitivity study was performed to establish the lower limits of detection (LLD). Target DNA sequences of various concentrations ranging from 0.25 to 25 ng were introduced to the probe-coupled beads. The goal was to find the point at which the target DNA could no longer be detected. The results indicated that the lower limit of detection was 0.5 ng for *B. cereus* and *E. coli*; 2 ng for *S. enterica*; and below 0.25 ng for *S. aureus* (Table 2.2). The LLD for *S. aureus* could not be precisely determined. This could have been due to the high sensitivity of the *S. aureus* probe. However, it was more likely due to slight errors in the quantitation process which led to *S. aureus* target DNA and/or beads being more concentrated than they should have been.

Sample	*B. cereus* probe	*E. coli* probe	*S. enterica* probe	*S. aureus* probe
B. cereus 0.25 ng	0	0	0	0
B. cereus 0.5 ng	81	0	0	0
B. cereus 1 ng	161	0	0	0
B. cereus 2 ng	364	0	0	0
B. cereus 5 ng	576	0	0	0
B. cereus 10 ng	754	0	0	0
B. cereus 25 ng	1119	0	0	0
E. coli 0.25 ng	0	0	0	0
E. coli 0.5 ng	0	101	0	0
E. coli 1 ng	0	211	0	0
E. coli 2 ng	0	332	0	0
E. coli 5 ng	0	616	0	0
E. coli 10 ng	0	965	0	8
E. coli 25 ng	0	1734	0	25
S. enterica 0.25 ng	0	0	0	0
S. enterica 0.5 ng	0	0	0	0
S. enterica 1 ng	0	0	0	0
S. enterica 2 ng	0	0	14	0
S. enterica 5 ng	0	0	74	0
S. enterica 10 ng	0	0	178	0
S. enterica 25 ng	0	0	453	0
S. aureus 0.25 ng	35	43	21	366
S. aureus 0.5 ng	5	9	0	614
S. aureus 1 ng	49	51	16	629
S. aureus 2 ng	75	60	44	1295
S. aureus 5 ng	224	218	205	1682
S. aureus 10 ng	179	170	169	2169
S. aureus 25 ng	234	203	161	2907
Negative control	18	17	29	15

Table 2.2. Net MFIs for selected samples in specificity study (*Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0"). Yellow highlighted cells indicate designed complementarity. Net MFIs in Blue font indicate a false positive. Positive controls were above threshold but are not shown.

d) Mixture Study

A mixture study was undertaken to determine the efficacy of detecting the components of several bacterial DNA sequences when present in a mixture. A 1:1 ratio (in terms of mass of sample, or amplicons) of all possible binary, ternary, and quaternary combinations was made.

In the mixture study, the Luminex never detected any bacterial DNA that was not present (*i.e.*, no false positives). However, there were several false negatives, all of which involved the *S. enterica* probe. This could have been another side effect of the cross-reactivity between *E. coli* and *S. enterica*. The results of the mixture study are summarized in Table 2.3.

Sample**	BC Probe	EC Probe	SE Probe	SA Probe
BC + EC	**486**	**599**	6	−2
BC + SE	**586**	−6	75	−9
BC + SA	**537**	−10	7	**1643**
EC + SE	−8	**665**	62	7
EC + SA	−7	**713**	−6	**1733**
SE + SA	−8	−1	76	**1750**
BC + EC + SE	**603**	**644**	**83**	−1
BC + EC + SA	**532**	**707**	10	**1672**
BC + SE + SA	**508**	10	79	**1642**
EC + SE + SA	11	**718**	78	**1362**
BC + EC + SE + SA	**502**	**519**	80	**722**

Table 2.3. Net MFIs for 1:1 mixtures. All samples contained 20 ng of each PCR amplicon. The combinations of designed complementarity are highlighted. Values two times the average background (81) indicate a positive identification and are denoted in bold. The net MFI values are equivalent to the raw MFI value less one background of the specific microsphere set. Positive controls were above threshold but are not shown. **BC = *B. cereus*, EC = *E. coli*, SE = *S. enterica*, SA = *S. aureus*. Originally published as Table 5 on page 764 in Pathogen Detection Using a Liquid Array Technology (*Journal of Forensic Sciences* **56**: 760-765) by Battaglia, A.; Schweighardt, A.; and Wallace, M. (2011). Reprinted with permission from John Wiley & Sons.

A second mixture study was conducted which involved studying only binary mixtures in which the bacterial DNA components were present in various ratios (1:1, 1:2, 1:5, and 1:10, plus all converse ratios). The sensitivity study indicated that the *highest* lower limit of detection was 2 ng (for *S. enterica*). Therefore, "1" was normalized to mean 2 ng and other values were adjusted relative to this.

In the second mixture study, the Luminex never detected any bacterial DNA that was not present (*i.e.*, no false positives). There were several false negatives, again involving the *S. enterica* probe. Considering the number of samples tested (forty-eight), there were not many false negatives (three). It is important to emphasize that the Luminex was able to detect both components of most binary mixtures, even when the minor component was overshadowed ten-fold by the major component. Although the net MFI observed in each mixture was generally proportional to the quantity of input DNA, the fluorescence did not always mirror the input DNA ratios. The results of this mixture study are summarized in Table 2.4.

Sample**	BC Probe	EC Probe	SE Probe	SA Probe	MFI Ratio***
BC:EC 10:1	**1035**	**254**	−10	−2	4:1
BC:EC 5:1	**817**	**313**	21	9	2.5:1
BC:EC 2:1	**533**	**283**	−2	−6	2:1
BC:EC 1:1	**341**	**303**	−2	2	1:1
BC:EC 1:2	**315**	**470**	1	5	1:1.5
BC:EC 1:5	**331**	**907**	8	7	1:2.5
BC:EC 1:10	**303**	**1276**	18	23	1:4
BC:SE 10:1	**1074**	−3	45	−8	24:1
BC:SE 5:1	**901**	−10	56	−12	16:1
BC:SE 2:1	**560**	−8	37	−6	15:1
BC:SE 1:1	**383**	−4	56	−11	7:1
BC:SE 1:2	**420**	1	**97**	7	4.5:1
BC:SE 1:5	**275**	−5	**189**	−9	1.5:1
BC:SE 1:10	**291**	11	**340**	−5	1:1
BC:SA 10:1	**1169**	−6	1	**462**	2.5:1
BC:SA 5:1	**795**	−9	−11	**297**	2.5:1
BC:SA 2:1	**619**	−7	−2	**822**	1:1.5
BC:SA 1:1	**424**	−4	−1	**925**	1:2
BC:SA 1:2	**408**	0	2	**1332**	1:3.5
BC:SA 1:5	**335**	5	−9	**1815**	1:5.5
BC:SA 1:10	**242**	−4	5	**2075**	1:8.5
EC:SE 10:1	2	**1648**	**115**	33	14.5:1
EC:SE 5:1	0	**1036**	**64**	26	16:1
EC:SE 2:1	−4	**571**	45	9	12.5:1
EC:SE 1:1	−6	**380**	49	3	8:1
EC:SE 1:2	0	**292**	**89**	3	3.5:1
EC:SE 1:5	−6	**249**	**207**	1	1:1
EC:SE 1:10	−2	**232**	**374**	−1	1:1.5
EC:SA 10:1	−4	**1592**	43	**1226**	1.5:1
EC:SA 5:1	5	**1109**	3	**1009**	1:1
EC:SA 2:1	−13	**672**	−4	**1002**	1:1.5
EC:SA 1:1	1	**442**	6	**1050**	1:2.5
EC:SA 1:2	0	**472**	9	**190**	2.5:1
EC:SA 1:5	−2	**469**	10	**1987**	1:4
EC:SA 1:10	−4	**499**	−3	**2261**	1:4.5
SE:SA 10:1	6	10	**436**	**1012**	1:2.5
SE:SA 5:1	8	4	**252**	**983**	1:4
SE:SA 2:1	−2	6	**124**	**986**	1:8
SE:SA 1:1	1	0	**84**	**1028**	1:12
SE:SA 1:2	−5	14	**76**	**1386**	1:18
SE:SA 1:5	4	14	**103**	**1938**	1:19
SE:SA 1:10	5	15	**87**	**2398**	1:27.5

Table 2.4. Net MFIs for ratio mixtures. The combinations of designed complementarity are highlighted. Values two times the average background (46) indicate a positive identification and are denoted in bold. False negative values are underlined. The net MFI values are equivalent to the raw MFI value less one background of the specific microsphere set. Positive controls were above threshold but are not shown. **BC = *B. cereus*, EC = *E. coli*, SE = *S. enterica*, SA = *S. aureus*. ***An approximate ratio of MFI values was determined between the two organisms present in the sample by dividing the larger MFI value by the smaller one and rounding to the nearest 0.5. Originally published as Table 6 on page 764 in Pathogen Detection Using a Liquid Array Technology (*Journal of Forensic Sciences* **56**: 760-765) by Battaglia, A.; Schweighardt, A.; and Wallace, M. (2011). Reprinted with permission from John Wiley & Sons.

CHAPTER III. RESEARCH DESIGN

A. Selecting Four Core Pathogens

There are many potential pathogens that could be used in a bioterrorist attack. Figure 3.1 summarizes just a few of the most common bacterial pathogens. The Centers for Disease Control and Prevention list *Bacillus anthracis*, *Clostridium botulinum*, *Francisella tularensis*, and *Yersinia pestis* as Category A high-priority pathogens that pose a risk to national security due to their ease of dispersal, high mortality rates, potential to cause panic, and the need for special emergency response procedures (CDC, 2011). For all these reasons, *B. anthracis*, *C. botulinum*, *F. tularensis*, and *Y. pestis* are highly likely to be used by a bioterrorist. The genomic DNA sequences of these bacteria are readily available (Read *et al.*, 2003; Smith *et al.*, 2007; Larsson *et al.*, 2005; Parkhill *et al.*, 2001). The Bacillus and Clostridium genera both reside in the Firmicutes phylum. The Francisella and Yersinia genera both reside in the Proteobacteria phylum, as shown in Figure 3.1.

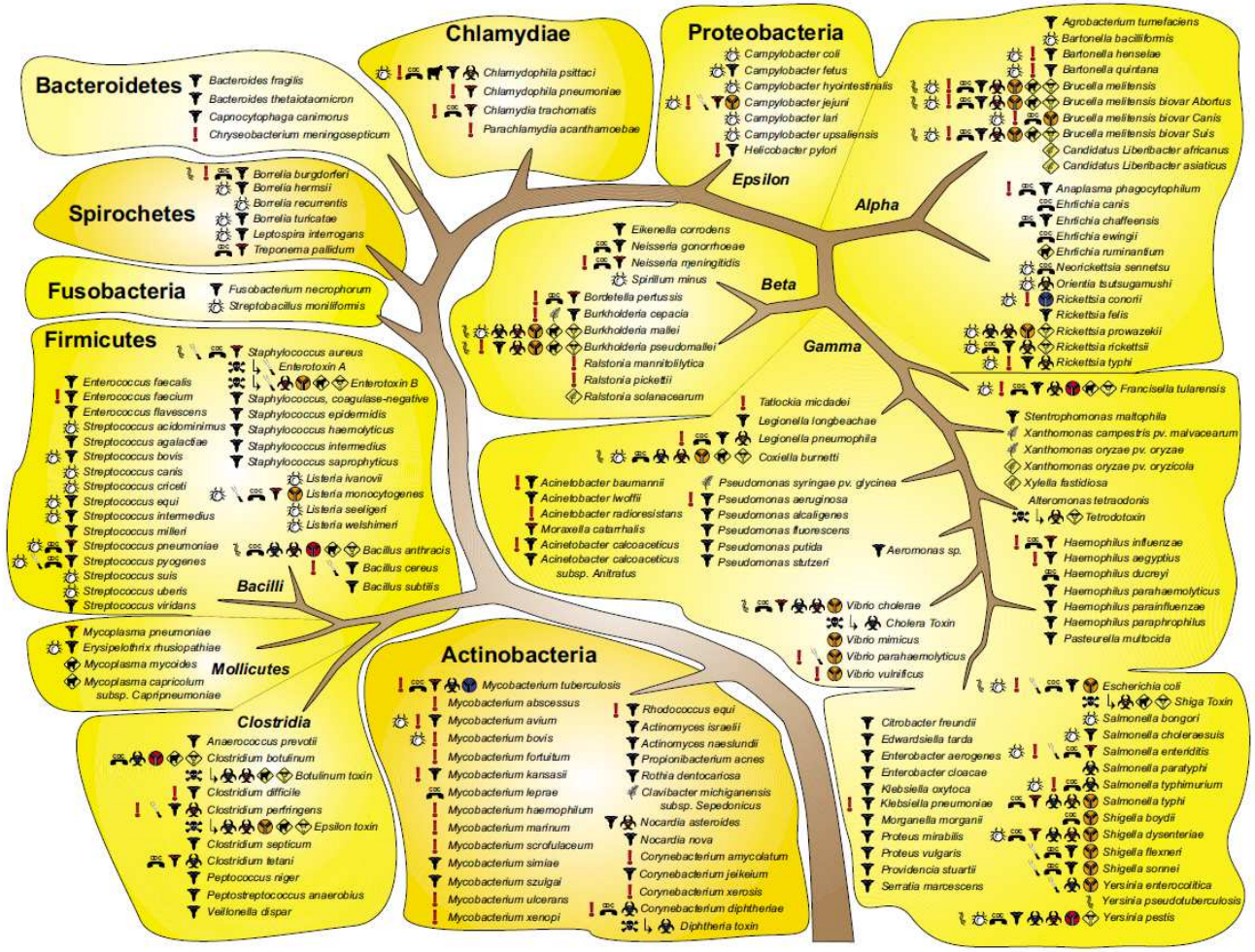


Figure 3.1. Bacterial phylogeny of some common bacterial pathogens. Originally published as Figure 2 on page 4 in The microbial rosetta stone database: a compilation of global and emerging infectious microorganisms and bioterrorist threat agents (*BMC Microbiology* **5**) by Ecker, D.; Sampath, R.; Willett, P.; Wyatt, J.; Samant, V.; Massire, C.; Hall, T.; Hari, K.; McNeil, J.; Buchen-Osmond, C.; and Budowle, B. (2005). Reprinted with permission from BioMed Central.

1. Anthrax

Bacillus anthracis was identified in 1876 by the now famous German physician and microbiologist Robert Koch (Koch, 1876). Aside from identifying the causative agent of anthrax and tuberculosis, Koch is also well known for his Postulates, which essentially state that microorganisms are the cause of disease. This was not a popular idea in Koch's time, when superstitions prevailed and popular culture dictated fallacious explanations for disease. It was the pioneering work done by researchers like Koch, Redi, Pasteur, Tyndall, and others that finally convinced skeptics of the "germ theory" of disease.

Today, our knowledge of the bacterium that once eluded physicians and microbiologists for centuries has grown quite impressive indeed. It is now known that *B. anthracis* is a spore-forming (one characteristic of its life cycle that complicated its discovery and classification), facultatively anaerobic, Gram-positive bacterium. It is also known that this bacterium can cause inhalational, cutaneous, or gastrointestinal infections, the former being the most lethal. Inhalational anthrax has been known to result in a mortality rate of up to 80% in the absence of timely and adequate antimicrobial therapy (Ciammaruconi *et al.*, 2008). Antimicrobial therapies have been used with some degree of success to treat victims of anthrax disease, provided that their illness is diagnosed in a timely fashion (Ghosh *et al.*, 2002). If diagnosis is delayed, the toxins secreted by the bacteria may have accumulated to the point where a neutralizing anti-toxin treatment is needed in addition to antibiotic treatment.

Research over time has steadily demystified *B. anthracis*, to the point that the entire genome of this bacterium is now known (Read *et al.*, 2003). However, there is still

much to be learned about *B. anthracis*. For example, this bacterium is difficult to identify by genetic means because of the high degree of homology that its genome has with other closely related microorganisms. Technologies designed for the identification of this pathogenic microorganism cannot advance until further genetic research provides a basis for genetically distinguishing *B. anthracis* from similar bacteria. Therefore, the discovery and proven effectiveness of polymorphic genetic markers for *B. anthracis* remains an area of great research interest.

2. Botulism

The name "botulism" comes from the Latin word for sausage, which is *botulus*. The illness was so named because it was often contracted from eating contaminated sausage (Black, 2005). Botulism is caused by a toxin secreted by the bacterium *Clostridium botulinum*, which is a gram-positive rod-shaped bacterium.

The unusually heat-resistant spores of *C. botulinum* can be found on edible plants or in the soil. The spores can also germinate in canned goods because they are obligate anaerobes. The toxin is secreted from the bacterium when contaminated food is ingested. Once ingested, the bacteria release an exotoxin that causes muscle paralysis by blocking the release of the neurotransmitter acetylcholine. This neurotransmitter is essential for carrying nerve impulses across the synaptic cleft that separates one nerve cell from the next. With acetylcholine disabled, neurons cannot transmit the electrical signals that cause muscles to contract, and therefore the muscles of infected individuals become flaccid and useless.

Initial muscle paralysis manifests itself in the head and neck but eventually progresses to all of the skeletal muscles. If the infection is not diagnosed or if the proper

antiserum is not administered in a timely fashion, death results from total respiratory arrest. Simultaneous infection of many individuals is a major concern because of the potential to easily overwhelm medical treatment facilities (Ghosh *et al.*, 2002). Even with the timely diagnosis and treatment, the recovery is a long process that often requires elaborate equipment such as a medical ventilator.

3. Tularemia

Francisella tularensis is the causative agent of tularemia. This bacterium is Gram-negative with a coccobacillus morphology. It was first isolated in Tulare County, California in 1911 and was named accordingly (Black, 2005).

Tularemia is normally transmitted by insect bites to various small mammals. Humans can also be infected by the bite of an insect, among other modes of transmission. The virulence of the disease is dependent on the method by which it has been transmitted, but the most destructive form of the disease comes from inhalation of the aerosolized bacteria, causing a severe form of the disease known as "pneumonic tularemia" (Ghosh *et al.*, 2002). The symptoms of the disease initially resemble pneumonia, but the condition of the victims gradually worsens as the bacteria infect the lymphatic system. Estimates of the mortality rate for tularemia range from 5% (Black, 2005) to 60% (Ghosh *et al.*, 2002). The high variability is due to the virulence of different isolates and/or the ability of a particular geographic area to effectively treat the symptoms. Increased virulence could also result from strains that have been genetically engineered for resistance to common antibiotics.

4. Plague

Plague is caused by *Yersinia pestis*, which is a Gram-negative bacterium with a short, thick rod-shaped morphology (Black, 2005). Bacteria are spread via fleas that live on and bite rodents. Once the rodent dies, the fleas seek a new rodent host. Historically, urban areas are at an especially high risk for the plague because when high numbers of rodents succumb to the disease, fleas seek alternative hosts which may include humans (Ghosh *et al.*, 2002).

Victims become symptomatic about one to eight days after the initial fleabite. Flu-like symptoms develop first but the lymphatic system is soon compromised which is evident from the characteristic "bubos," or swollen lymph nodes in the groin and underarm areas. The bacteria multiply in the lymphatic system and evade destruction because of their secretion of a protein that inhibits phagocytosis. As the disease progresses and the respiratory system becomes infected, the bacteria can be transmitted from human-to-human by coughing. Further progression of the disease causes the circulation of *Y. pestis* in the blood. This causes septicemia, or a generalized inflammatory response characterized by shock, organ failure, internal bleeding, and gangrenous extremities (Ghosh *et al.*, 2002). Plague is sometimes referred to as "Black Death" because the extremities turn black from prolonged gangrene. The mortality rate is close to 100% unless antibiotic treatments are administered.

B. Selecting Close Relatives of the Four Core Pathogens

Once the four core pathogens were selected, close relatives of each of the four pathogens were sought. The purpose of selecting close relatives was to test the ability to identify potential mimics and/or genetically modified relatives within the same genus.

The close relatives chosen for this research were *Bacillus cereus*, *Clostridium difficile*, *Francisella tularensis* subsp. *holarctica* and *Yersinia enterocolitica*. The genomes of these four bacteria have been fully sequenced (Ivanova *et al.*, 2003; Sebaihia *et al.*, 2006; Barabote *et al.*, 2009; and Thomson *et al.*, 2006).

An additional decision to be made before commencing the bioinformatics study was to choose an outlying microorganism. The outlier should theoretically have less DNA homology with the other eight microbes and thus its target DNA should not be very reactive with the probes used in this study. In order to select an appropriate outlier, the phyla of the four core pathogens and their respective close relatives were examined. The *Francisella* and *Yersinia* genera reside in the Proteobacteria phylum; the *Bacillus* and *Clostridium* genera reside in the Firmicutes phylum. The kingdom Bacteria has over twenty phyla (Garrity *et al.*, 2001) but it was important to select the outlier from a phylum other than Proteobacteria and Firmicutes.

Mycobacterium tuberculosis was ultimately selected as the outlier because it belongs to the Actinobacteria phylum which diverged before Firmicutes and Proteobacteria (Ecker *et al.*, 2005). The complete genome sequence of *M. tuberculosis* is available (Cole *et al.*, 1998), and the *rrl* gene sequence can also be found on the NCBI Gene database. *M. tuberculosis* is the causative agent of the respiratory disease tuberculosis and thus its genomic DNA can be readily acquired, along with the genomic DNA of other priority pathogens, from BEI Resources.

C. Identifying Genes for Probe Design

1. rrl Gene

Ribosomal RNA is an essential part of translation, the process by which the information contained in an mRNA sequence is converted into the protein language. Prokaryotic ribosomal RNA consists of a large (50S) and small (30S) subunit. The large subunit is made of 23S and 5S RNA while the small subunit is made of 16S RNA. The ribosomal RNAs are measured in Svedberg (S) units, which are a measure of sedimentation under standard conditions (Lodish *et al.*, 2004). Figure 3.2 illustrates the composition of prokaryotic ribosomal RNA.

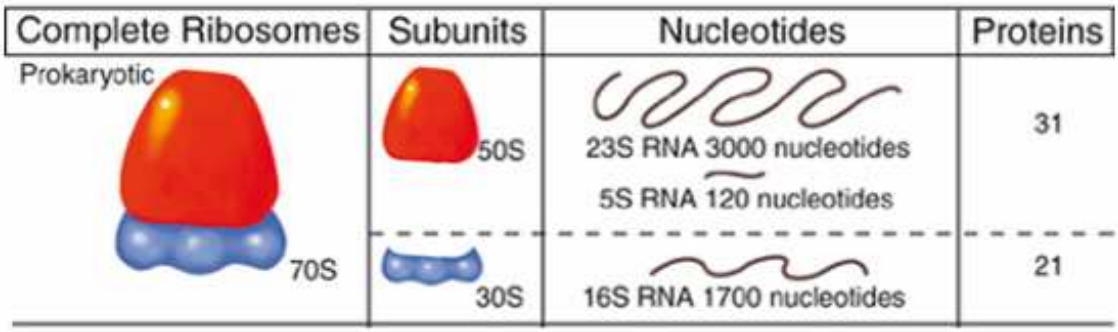


Figure 3.2. Composition of prokaryotic ribosomal RNA. Originally published as Figure 8.23 on page 278 in *Genetics: from genes to genomes*, 3rd edition (McGraw-Hill Higher Education) by Hartwell, L.; Hood, L.; Goldberg, M.; Reynolds, A.; Silver, L.; and Veres, R. (2008). Reprinted with permission from The McGraw-Hill Companies.

This study relied, in part, upon probes derived from DNA sequences that code for 23S rRNA. The 23S rRNA is coded for by DNA sequences that collectively comprise the *rrl* gene. The *rrl* gene is usually part of a larger gene cluster that includes the upstream *rrs* gene (which codes for the small, 16S rRNA subunit) and the downstream *rrf* gene (which codes for the 5S component of the large rRNA subunit; Srivastava and Schlessinger, 1990). Genetic analyses of bacteria can sometimes be complicated by the

propensity for prokaryotic genomes to undergo high rates of mutation. However, such complications are often minimized by studying ribosomal RNA genes, which are highly conserved because they code for structural RNA molecules involved in processes, such as translation, that are essential for life (Nomura *et al.*, 1984). Genes that are essential for life processes are ideal for probe design because they are under high selective pressure to minimize mutations. Probing for ribosomal RNA gene sequences may also help to improve sensitivity, as these genes are often present in multiple copies in order to accommodate the cell during times when many ribosomes are needed for translation.

The nine pathogens selected for use in this study had a fully sequenced *rrl* gene available from the NCBI Gene database (Maglott *et al*., 2005).

2. TOX Genes

a) Overview

Many pathogenic microorganisms have toxin-coding sections of their genome that separate them from their non-pathogenic cousins. These unique sections of the genome, from which the organisms derive their pathogenicity, were thought to be good candidates for confirmatory genetic identification. Each pathogenic species has its own special toxin-coding genes, and therefore these probes were designed using a different gene for each microbe, unlike those for the presumptive identification which targeted the same gene (*rrl*) in all organisms. Accordingly, probe design for the toxin-coding genes began by searching the literature to glean information about the toxin-coding genes for each of the four core pathogens.

b) Bacillus anthracis

The *Sterne* strain of *B. anthracis* was selected for this research because this was the only strain available from BEI Resources, the genomic DNA provider. The genome of this strain was completely sequenced in 2004 by the Joint Genome Institute of the U.S. Department of Energy. A paper formally detailing the sequencing effort was not published, but a paper on the plasmids from which *B. anthracis* str. *Sterne* mainly derives its toxicity has been published (Okinaka *et al.*, 1999). The pathogenicity of *B. anthracis* stems mainly from the genes on the pXO2 plasmid that code for capsule synthesis (*capA*, *capB*, and *capC*) and from the genes on the pXO1 plasmid that code for the toxins *pagA* (protective antigen), *lef* (lethal factor), and *cya* (edema factor).

Even though *B. anthracis* derives its toxicity from genes on both the pXO1 and pXO2 plasmids, probe design for this study was restricted to genes on pXO1 because the *Sterne* strain is acapsulate (*i.e.*, lacking the pXO2 plasmid). The three toxin-coding genes (*pagA*, *lef*, and *cya*) on pXO1 were excellent candidates for probe design because prior research demonstrated that even closely related microbes such as *B. cereus* and *B. thuringiensis* lacked these genes (Read *et al.*, 2003). Furthermore, there appeared to be high homology in the toxin-coding gene sequences for all *B. anthracis* isolates (Leppla, 1999). This meant that a probe designed within one of these three toxin-coding genes would most likely be highly specific for *B. anthracis*.

The *pagA* and *cya* genes are essential in the early infective stages of anthrax (Keppie *et al.*, 1963), whereas the *lef* gene is important later on in the infection, just before death occurs (Leppla, 1995). The toxin coded by the *cya* gene is an adenylate cyclase that causes abnormally high levels of cyclic AMP, thereby leading to significant incapacitation of neutrophils (O'Brien *et al.*, 1985). The toxin coded by the *lef* gene is a

metalloprotease (Klimpel *et al.*, 1994) that was first discovered in 1961 (Stanley and Smith). It works in tandem with *pagA* to create what is known as the lethal toxin. The gene *pagA* codes for a protein that binds to cell surface receptors, subsequently creating a channel that allows the protein coded by *lef* to penetrate the cell and enter the cytosol (Milne *et al.*, 1994). The metalloprotease coded by *lef* may exert its pathogenesis by inhibiting the mitogen-activated protein kinase (MAPK) signal transduction pathway (Duesbery *et al.*, 1998). This pathway, however, is not known to control short-term homeostasis – a process that is severely compromised in the presence of the lethal toxin (Leppla, 1999). Accordingly, it is believed that the protein coded by *lef* may act on other cytosolic components, perhaps even some that cause macrophages to secrete abnormally high levels of cytokines, thus leading to systemic shock (Hanna *et al.*, 1993). The *lef* gene was ultimately chosen for probe design because it is essential for pathogenicity – studies have shown that virulence is more than 1000 times less when *lef* is inactivated by insertional mutagenesis (Pezard *et al.*, 1991). Indeed, the *lef* gene is regarded as the most important virulence factor (Leppla, 1999) in *B. anthracis*.

c) Clostridium botulinum

Neuroexocytosis is the process by which synaptic vesicles fuse with the presynaptic membrane in a neuron and release neurotransmitters into the synaptic cleft. It is a key step in the transmission of an action potential from one neuron to a neighboring neuron. Botulinum neurotoxin (BoNT) is a metalloprotease that interferes with this process by cleaving proteins called SNAREs that are essential for fusion of the synaptic vesicle to the presynaptic membrane (Simpson, 2004; Poulain *et al.*, 2006).

C. botulinum is divided into four phenotypic groups, numbered I through IV, with groups I and II being responsible for most cases of botulism in humans (Lindstrom and Korkeala, 2006). Group I is designated as "proteolytic" because the neurotoxin achieves maximum potency after the botulinum neurotoxin is cleaved by an endogenous protease, whereas Group II organisms are designated as "non-proteolytic" because the neurotoxin achieves maximum potency after the BoNT is cleaved by an exogenous protease (Lindstrom and Korkeala, 2006). Each group can be further classified into serotypes A through G, with specific serotypes being found only in certain groups (East *et al.*, 1998). The serotype letter corresponds to the BoNT produced by that particular organism.

The only *C. botulinum* genomic DNA available (from BEI Resources) was *C. botulinum* Strain Walls 8G. According to the product information sheet, this is a group I (proteolytic) BoNT type F toxin-producing strain. The BoNT/F gene is usually part of a gene cluster with NTNH, BotR, and p47, as well as polycistronic open reading frames orfX3, orfX2, and orfX1 (East *et al.*, 1998; Smith *et al.*, 2007).

d) Francisella tularensis

Upon entering the host organism, *Francisella tularensis* uses components of its outer surface (such as its pili) to adhere to host macrophages (Gil *et al.*, 2004). The microbe is engulfed by the macrophage and becomes sequestered in a phagosome, where it evades lysosomal degradation by creating a basic environment due to ammonia generated by enzymes such as citrulline ureidase (Fleming and Foshay, 1955; Gordon *et al*., 1980). Several hours after being engulfed by the macrophage, the bacteria degrade the phagosomal membrane and enter the cytosol where they replicate (Golovliov *et al.*,

2003). Severe cytopathogenicity causes apoptosis of the macrophage within twenty-four hours of the initial infection (Lai *et al.*, 2001).

A great deal is known about the mechanism by which *F. tularensis* enters host cell macrophages and eventually causes apoptosis. However, very little is known about the genes within *F. tularensis* that account for its destructiveness. This microbe does not derive its virulence from toxin-coding genes. Instead, its virulence ostensibly stems from enzymes that enable it to enter, replicate within, and destroy host macrophages. An overview of some of the genes thought to be involved in these processes can be found in the literature (Barker and Klose, 2007). Most efforts to identify *F. tularensis* have focused on the genes that are responsible for enabling the bacteria to grow within, and eventually be released from, the host macrophages. The genes *mglA* and *mglB* (macrophage growth loci A and B), for example, are critical for bacterial growth within the macrophage (Baron and Nano, 1998). Also of interest are genes on the Francisella pathogenicity island (FPI) such as *iglA*, *iglB*, and *iglC*, which are important for sustaining growth of the bacteria within the macrophage, as well as orchestrating the release of the bacteria from the phagosome (Bruin *et al.*, 2007 and Santic *et al.*, 2005).

Although sufficient for identifying organisms belonging to the Francisella genus, many of the aforementioned genes lack the uniqueness necessary to specifically identify *F. tularensis* subsp. *tularensis* (the most virulent subspecies). One gene that may allow specific identification of the *tularensis* subspecies is FTT0435 (also known as *ctu*), which codes for citrulline ureidase, the enzyme whose generation of ammonia contributes to the bacteria's ability to evade lysosomal degradation (Larsson *et al.*, 2005). Citrulline

ureidase is thought to be unique to the *tularensis* subspecies (Sjostedt, 2004), thus making FTT0435 a good candidate for identification of *F. tularensis* subsp. *tularensis*.

e) Yersinia pestis

Humans can acquire plague through the bite of an infected flea or mammal, or from inhaling respiratory droplets containing the bacteria (Perry and Fetherston, 1997). There are eleven species within the Yersinia genus, but only *Y. enterocolitica*, *Y. pseudotuberculosis*, and *Y. pestis* are pathogenic to humans. The bacteria cause illness in humans not by producing toxins, but instead by evading immune system defenses. Some genes code for proteins that facilitate entry into host cells (Forman *et al.*, 2008), whereas others code for proteins that help the bacteria to avoid phagocytosis (Viboud and Bliska, 2005), leading eventually to apoptosis. Some of the virulence genes are located on the bacterial chromosome, but some also reside in the plasmids. Genes on the plasmids were preferentially sought for this research because these genes are more susceptible to malicious genetic manipulations that could confer toxicity on an otherwise benign strain.

All three pathogenic strains possess the pCD1 (calcium dependence) plasmid. However, the plasmids pMT1 (murine toxin) and pPCP1 (pesticin, coagulase, plasminogen activator) are unique to *Y. pestis*. Genes on plasmid pMT1 were not pursued for this study because of complications anticipated from the genetic relatedness of pMT1 to a plasmid in *Salmonella enterica* (Prentice *et al.*, 2001). This left genes on plasmid pPCP1 for probe design. One gene that has been used to successfully identify *Y. pestis* is the plasminogen activator, *pla* (Leal and Almeida, 1999; Lathem *et al.*, 2007). This gene was mapped in 1988 (Sodeinde and Goguen) and is thought to activate a protease that inhibits coagulation and causes the lysis of fibrin clots (Sodeinde *et al.*, 1992).

CHAPTER IV. METHODOLOGY

A. Bioinformatics

1. Searching for Gene Sequences

Sequences for the *rrl* gene of the nine bacteria were accessed on the NCBI website using their GenBank accession number as indicated in Table 4.1.

Organism	GenBank Number
Bacillus anthracis str. Sterne	AE017225
Bacillus cereus	CP000001
Clostridium botulinum	CP000728
Clostridium difficile	AM180355
Francisella tularensis tularensis SCHU S4	AJ749949
Francisella tularensis holarctica	AM233362
Yersinia pestis CO92	AL590842
Yersinia enterocolitica enterocolitica	AM286415
Mycobacterium tuberculosis H37rv	AL123456

Table 4.1. Names and GenBank accession numbers for nine bacteria.

A literature search for the toxin-coding genes returned several candidate genes for each of the four core microbes. However, after closer analysis (see section on identifying genes for probe design in Chapter III: Research Design), one gene for each core microbe was ultimately selected. The outcome of the literature search for toxin-coding genes is summarized in the Table 4.2.

Organism	Toxin-Coding Gene	Cellular Location	Base Location	GenBank Number
B. anthracis str. Sterne	lef	pXO1	127442 to 129871	AF065404
C. botulinum F str. Langeland	BoNT/F	chrom.*	880105 to 883716	X99064
F. tularensis tularensis str. SCHU S4	FTT0435	chrom.*	448883 to 449743	AJ749949
Y. pestis CO92	pla	pPCP1	6665 to 7603	AL109969

Table 4.2. Toxin-coding genes selected for probe design. Abbreviation "chrom" = chromosome.

2. Probe and Primer Design

RRL probe and primer sequences were designed to be complementary to sequences within the 23S ribosomal RNA gene *rrl*. Initial attempts to design RRL probes and primers at separate loci were not successful due to the high degree of homology in the *rrl* gene. The locus ultimately selected for probe and primer design was based on one identified in prior research (Dunbar and Jacobson, 2007). This locus was optimal because there existed a sufficient number of polymorphisms to permit the design of unique probes, yet a substantial degree of homology in the flanking regions made these target sequences amenable to amplification using degenerate PCR primers.

To facilitate degenerate primer design, the locus was aligned for all nine bacteria using the ClustalW multiple sequence alignment tool (ClustalW, 2010). Probe and reverse probe sequences are shown in Tables 4.3 and 4.4. RRL primer sequences are shown in Table 4.5. The RRL primers and probes were checked for general uniformity in terms of melting temperature and GC content using a DNA calculator available on the World Wide Web (Sigma, 2010). Primer and probe specificity (see Section 3, below) and propensity for secondary structure formation (see Section 4, below) were also evaluated.

Ideally, the range of melting temperatures for probes and primers would not exceed 5°C, but this goal had to be somewhat flexible to accommodate other important parameters in sequence design such as maximizing specificity and minimizing the potential for self- or cross-dimerization.

Probe	RRL Probe Sequence	T_M (°C)	%GC
Bac/RRL	CGTAATGGTATGGTATCCTT	56.3	40
Cdif/RRL	AGTAATGTCTAAGTATCATT	50.2	25
Cbot/RRL	GGTAACCCCATGTATCATG	58.0	47.4
Fran/RRL	ACACAAGTTGGGTACTCACT	58.4	45
Yent/RRL	CAATTCGTTGCACTATTGCA	56.3	40
Ypes/RRL	CAATACGTTGCACTATCGTT	56.3	40
Myc/RRL	AGTGATGTCGTGCTACCCGC	64.5	60

Table 4.3. RRL probe sequences (all shown in 5′ to 3′ orientation).

Rev. Probe	RRL Reverse Probe Sequence	T_M (°C)	%GC
Bac/RRL	AAGGATACCATACCATTACG	56.3	40
Cdif/RRL	AATGATACTTAGACATTACT	50.2	25
Cbot/RRL	CATGATACATGGGGTTACC	58.0	47.4
Fran/RRL	AGTGAGTACCCAACTTGTGT	58.4	45
Yent/RRL	TGCAATAGTGCAACGAATTG	56.3	40
Ypes/RRL	AACGATAGTGCAACGTATTG	56.3	40
Myc/RRL	GCGGGTAGCACGACATCACT	64.5	60

Table 4.4. RRL reverse probe sequences (all shown in 5′ to 3′ orientation).

Primer Name	RRL Primer Sequence	T_M (°C)	%GC
RRL Fwd. Primer #1	A[00-30-00-70]CC[20-00-80-00][50-00-50-00][70-00-30-00]GATTTCCGAATG[30-00-70-00]G	61.8	50.5
RRL Fwd. Primer #2	ACCGACGATACCCGAATGGG	64.5	60
RRL Rev. Primer	[30-00-70-00]GGTACT[00-00-50-50]AGATGTTTCA[00-30-70-00]TTC	57.6	37.7

Table 4.5. RRL primer sequences (all shown in 5′ to 3′ orientation). Brackets denote proportions of degenerate bases at these locations. Within each set of brackets, base symbols are omitted and proportions of each base (in alphabetical order) are listed in their place.

TOX gene sequences were opened in FASTA format and aligned (ClustalW, 2010) to identify regions in which there was no homology with other genes used in this study. Unique sections of each gene were pasted into the source sequence box in Primer3 software (Rozen and Skaletsky, 2000), which was used to design TOX probe and primer sequences. Default parameters were used in the analysis. Briefly, the optimum target DNA size was between 150 and 250 bases. Primer specifications were set to an optimum size of 20 bases and an optimum melting temperature of 60ºC. The same length and melting temperature specifications were used for the probe. Primer, probe, and reverse probe sequences for the TOX targets are shown in Tables 4.6, 4.7, and 4.8, respectively.

Primer Name	TOX Primer Sequence	T_M (°C)	%GC
Bant/TOX Fwd	TGGACATGCTGTGGATGATT	58.4	45
Bant/TOX Rev	ATTGGAAAGTTTTCGGAGCA	56.3	40
Cbot/TOX Fwd	AAAAGGCATATTTATGGACATTGAA	56.4	28
Cbot/TOX Rev	TTGTATCATCATTAACAGGGTCA	57.4	34.8
Ftul/TOX Fwd	AATGATTGATGCGGATGGTT	56.3	40
Ftul/TOX Rev	TCCCAGCAAATACCAACTCC	60.4	50
Ypes/TOX Fwd	GACATCCGGCTCACGTTATT	60.4	50
Ypes/TOX Rev	AGCATCTCCGCCAATAGAGA	60.4	50

Table 4.6. TOX primer sequences (all shown in 5′ to 3′ orientation).

Probe	TOX Probe Sequence	T_M (°C)	%GC
Bant/TOX	ATGGGAGAACAAATGAAGCG	58.4	45
Cbot/TOX	GGGGATTTTATGCCAGTTGT	58.4	45
Ftul/TOX	TGCTGGTTTTAAGGTTTGGG	58.4	45
Ypes/TOX	TATGATGAGGGCAAAGGAGG	60.4	50

Table 4.7. TOX probe sequences (all shown in 5′ to 3′ orientation).

Rev. Probe	TOX Reverse Probe	T_M (°C)	%GC
Bant/TOX	CGCTTCATTTGTTCTCCCAT	58.4	45
Cbot/TOX	ACAACTGGCATAAAATCCCC	58.4	45
Ftul/TOX	CCCAAACCTTAAAACCAGCA	58.4	45
Ypes/TOX	CCTCCTTTGCCCTCATCATA	60.4	50

Table 4.8. TOX reverse probe sequences (all shown in 5′ to 3′ orientation).

3. Search of the BLASTN Database for Specificity

The RRL and TOX primer and probe sequences were uploaded in the FASTA format to the query sequence box for the blastn algorithm on the website of the U.S. National Center for Biotechnology Information (NCBI, 2010). The query sequence was searched (using default algorithm parameters) against sequences found in all genomes (approximately 1700) in the nucleotide collection (nr/nt) database to confirm the specificity of each primer and probe.

4. AutoDimer Analysis

All probe, primer, and target sequences were analyzed by an AutoDimer (Vallone and Butler, 2004) hairpin screen to assess the likelihood of self-dimerization. Additionally, all primer, probe, and target sequences were analyzed by an AutoDimer primer dimer screen for cross-dimerization. The software analyzes each potential interaction for basepair matches and mismatches; the sum of the mismatches is then subtracted from the sum of the matches. Any interaction that exceeds the threshold score value (default is seven) is reported. Scores that are high above the threshold indicate a strong propensity for dimerization, while scores closer to (or below) the threshold indicate the opposite and are thus more favorable. If an interaction scores below the threshold, it is not reported.

5. Determination of Copy Number

The whole genome of each microbe was searched to determine the copy number of its target sequence. The purpose of this was to determine the number of times the designated primers would amplify the target sequence. Target sequences were

individually uploaded to the query sequence box of the blastn algorithm on the NCBI BLAST search engine (Figure 4.1). The microbe's GenBank number was then entered into the Entrez Query box under the Search Set parameters. This limited the search for the target sequence to the selected microbe.

Figure 4.1. Screenshot of the blastn algorithm for determining target sequence copy number.

When the BLAST search is complete, a dialogue box at the top of the results page reminds the viewer that the search has been limited to records matching the input entrez query (Figure 4.2). In other words, results are not returned unless they belong to sequences present in the organism whose GenBank accession number matches the number typed into the Entrez Query box at the beginning of the search.

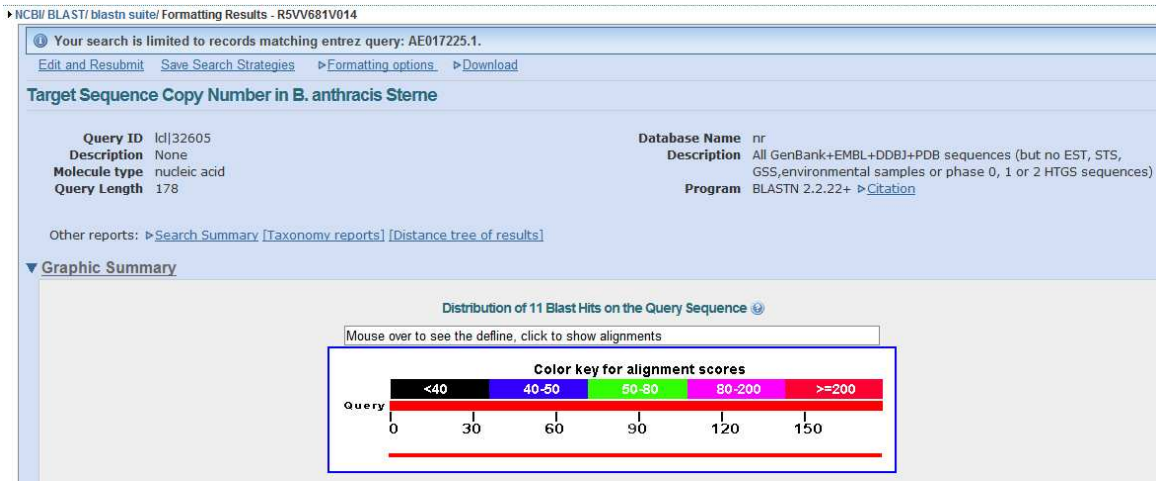


Figure 4.2. Screenshot of the results page for blastn algorithm for determining target sequence copy number.

B. Probe/Primer Synthesis

The synthesis of all primers, probes, and reverse probes was performed by Eurofins MWG Operon (Huntsville, AL). The probes (50 nmol; HPLC purity) were synthesized with a 5′ amino group and C12 spacer to facilitate attachment (via a carbodiimide bond) to the carboxyl groups on the microspheres. All forward primers (1.0 umol; salt-free purity) were synthesized with a 5′ phosphate to facilitate enzymatic degradation of the non-complementary target strand. Reverse primers (1.0 umol; salt-free purity) were synthesized with a 5′ biotin moiety to facilitate fluorescent detection. Reverse probes (0.2 umol; salt-free purity) were biontinylated on the 5′ end and synthesized for use in the titration study, as well as for positive controls throughout the study.

C. PCR Preparation

1. Acquiring Genomic DNA

Lyophilized genomic DNA of nine bacterial species was acquired from BEI Resources (Manassas, VA) as described in Table A1.1 in Appendix I.

2. Rehydrating the Genomic DNA

Lyophilized genomic DNA was rehydrated in molecular grade water as recommended by the certificate of analysis provided with each vial. The necessary volume of water was added to each genomic DNA to achieve a final concentration of 20 ng/uL. To ensure thorough rehydration, all tubes were shaken for twenty minutes at 300 rpm on an Eppendorf Thermomixer® R (Eppendorf, Hamburg, Germany) and stored at 4°C until use.

3. Rehydrating the PCR Primers

PCR primers were rehydrated in 1X Tris-EDTA (TE) buffer solution as recommended by the oligonucleotide data sheet. The necessary volume of 1X TE was added to achieve a stock concentration of 100 uM.

D. Monoplex PCR Specificity Check

1. PCR Parameters

Templates needed for PCR were rehydrated as previously described. The 100 uM stock primer solutions were diluted to 5 uM working solutions. Other reagents were acquired (Qiagen, Hilden, Germany) and used to prepare each sample by combining the reagents as indicated in Table 4.9.

Reagent	Volume
Fwd. primer (5 uM)	5 uL
Rev. primer (5 uM)	5 uL
$MgCl_2$ (25 mM)	1 uL
dNTP Mix (10 mM of each dNTP)	1 uL
10X Buffer	5 uL
Template (20 ng/uL)	5 uL
HotStarTaq Plus DNA polymerase (5 units/uL)	0.7 uL

Table 4.9. PCR reagents and volumes used in the monoplex study. Reagents were combined and enough molecular grade water was added to bring total reaction volume to 50 uL.

All templates were amplified in six separate batches (Table 4.10). The first batch of samples was amplified with the RRLfwd#1/RRLrev degenerate primer set. The second batch was amplified with the RRLfwd#2/RRLrev degenerate primer set. The remaining four batches consisted of all templates amplified by each individual TOX primer set.

PCR Batch	Primer Set
1	RRLfwd#1/RRLrev
2	RRLfwd#2/RRLrev
3	BantTOXfwd/BantTOXrev
4	CbotTOXfwd/CbotTOXrev
5	FtulTOXfwd/FtulTOXrev
6	YpesTOXfwd/YpesTOXrev

Table 4.10. Summary of sample batches in monoplex PCR specificity study.

All sample tubes were loaded into the GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA) and amplified according to the touchdown PCR (TD-PCR) parameters outlined in Table 4.11. Samples were stored at 4°C until quantitation.

Stage	Temp. (°C)	Cycle Number	Time
Initial Incubation	95	1	5 min
Denaturation	94	12	30 sec.
Annealing	67*		30 sec.
Extension	72		30 sec.
Denaturation	94	13	30 sec.
Annealing	55**		30 sec.
Extension	72		30 sec.
Final Extension	72	1	3 min.
Hold	4	1	∞

Table 4.11. TD-PCR parameters. *Starting annealing temperature was 67°C and was decreased by one degree over the course of twelve cycles. **Annealing temperature was held at 55°C for remaining thirteen cycles.

2. Quantitation

Quantitation of the PCR amplicons was performed using the Agilent 2100 Bioanalyzer and accompanying Agilent DNA 1000 reagent kit and 2100 Expert software (Agilent Technologies, Santa Clara, CA). The gel-dye mix was prepared as per the manufacturer's instructions and loaded in the appropriate wells of the chip secured within the priming station. DNA marker (5 uL) was added to each of the twelve wells in the chip plus the ladder well. Ladder (1 uL) was added to the ladder well. PCR amplicons (1 uL) were added to their respective wells in the chip. Additional DNA marker (1 uL) was added to any unused sample wells. The chip was vortexed at 24,000 rpm for one minute using the IKA® MS 3 vortex equipped with a chip adaptor (IKA, Wilmington, NC). The chip was loaded into the Agilent 2100 Bioanalyzer and sample acquisition was initiated.

E. Multiplex PCR Specificity Check and PCR Optimization

1. Optimization of Primer Concentration

Several experiments were performed to optimize the multiplex. The first trial consisted of separate forward and reverse primer mixes in which each primer had a

concentration of 5 uM, except the degenerate RRL primers, which had a higher concentration because they were designed to amplify multiple templates, as shown in Table 4.12. The samples were amplified according to the TD-PCR protocol described above. All PCR amplicons were quantified using the Agilent 2100 Bioanalyzer.

Fwd Primer Name	Final Conc.		Rev Primer Name	Final Conc.
BanthTOXfwd	5 uM		BanthTOXrev	5 uM
CbotTOXfwd	5 uM		CbotTOXrev	5 uM
FtulTOXfwd	5 uM		FtulTOXrev	5 uM
YpestTOXfwd	5 uM		YpestTOXrev	5 uM
RRLfwd#1	35 uM		RRLrev	45 uM
RRLfwd#2	10 uM			

Table 4.12. Summary of primers in the original Fwd. and Rev. primer mixes.

The primer concentrations described in Table 4.12 were modified with the intent of improving uniformity in terms of the amount of amplicons produced for each template. Samples were prepared using the modified primer concentrations described in Table 4.13 and re-amplified according to the TD-PCR protocol described above. All PCR amplicons were quantified using the Agilent 2100 Bioanalyzer.

Fwd Primer Name	Final Conc.		Rev Primer Name	Final Conc.
BanthTOXfwd	5 uM		BanthTOXrev	5 uM
CbotTOXfwd	5 uM		CbotTOXrev	5 uM
FtulTOXfwd	5 uM		FtulTOXrev	5 uM
YpestTOXfwd	5 uM		YpestTOXrev	5 uM
RRLfwd#1	15 uM		RRLrev	20 uM
RRLfwd#2	10 uM			

Table 4.13. Summary of primers in the modified Fwd. and Rev. primer mixes.

The primer concentrations described in Table 4.13 were modified once more for the purpose of diminishing the occurrence of primer dimers among the amplicons. The modified primer mixes were diluted 1:2 to yield diluted primer mixes with each primer having the concentrations summarized in Table 4.14. The samples were re-amplified

according to the TD-PCR protocol described above and quantified using the Agilent 2100 Bioanalyzer.

Fwd Primer Name	Final Conc.		Rev Primer Name	Final Conc.
BanthTOXfwd	2.5 uM		BanthTOXrev	2.5 uM
CbotTOXfwd	2.5 uM		CbotTOXrev	2.5 uM
FtulTOXfwd	2.5 uM		FtulTOXrev	2.5 uM
YpestTOXfwd	2.5 uM		YpestTOXrev	2.5 uM
RRLfwd#1	7.5 uM		RRLrev	10 uM
RRLfwd#2	5 uM			

Table 4.14. Summary of primers in the diluted Fwd. and Rev. primer mixes.

2. Optimization of Template Concentration

PCR reactions in the primer optimization study each used 100 ng of template (20 ng/uL). The diluted primer mix was used to amplify just 10 ng of each individual template. Additionally, each primer set was used to amplify mixtures containing 10 ng of each template. Samples were amplified according to the TD-PCR protocol and quantified using the Agilent 2100 Bioanalyzer.

F. Preparation for Attachment of Probes to Microspheres

1. Rehydrating the Probes

Lyophilized probes were rehydrated in molecular grade water to a stock concentration of 250 uM. Working solutions (100 uM) of each probe were also prepared.

2. Bead Preparation

Eleven bead sets were acquired from Luminex® Corporation (Austin, TX). The vials of uncoupled beads were vortexed for twenty seconds and sonicated for twenty seconds. Stock beads (5.0×10^6) from each set were transferred to respective microcentrifuge tubes and centrifuged at 8000 × g for two minutes. The supernatant was

discarded and the pelleted microspheres were resuspended in 50 uL of 0.1M MES (2-[*N*-morpholino] ethanesulfonic acid) adjusted to pH 4.5. Each tube was vortexed for twenty seconds and sonicated for twenty seconds.

G. Coupling Probes and Beads

Each of the amino-labeled probes was paired with its own set of beads according to Table 4.15.

Probe	Bead Set
Bac/RRL	112
Bant/TOX	113
Cdif/RRL	114
Cbot/RRL	116
Cbot/TOX	115
Fran/RRL	118
Ftul/TOX	117
Yent/RRL	120
Ypes/RRL	121
Ypes/TOX	119
Mtub/RRL	122

Table 4.15. Probe-Bead Coupling.

Working solution of each rehydrated probe (100 uM; 2 uL) was combined with its respective bead set and vortexed for twenty seconds. Fresh EDC (1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride) solution (10 mg EDC powder in 1 mL distilled water) was prepared and added (2.5 uL) to each reaction tube. The reaction tubes were vortexed and then incubated in the dark at room temperature for thirty minutes. Another fresh EDC solution was prepared and added (2.5 uL) to each reaction tube, followed by a pulse vortex and a thirty minute incubation in the dark at room temperature. Tween-20 (0.02%; 1.0 mL) was added to each reaction tube and the tubes were inverted to ensure mixing. The reaction tubes were centrifuged at 8000 × g for two minutes. The supernatant was discarded and the pellet was resuspended in 1X TE (100

uL; pH 8.0). The probe-coupled beads were vortexed for twenty seconds, sonicated for twenty seconds, and stored at 4°C in the dark.

H. Counting the Probe-Coupled Beads

A 1 uL aliquot of each coupled bead set was withdrawn and diluted 1:100 in distilled water. Each diluted probe-coupled bead set was individually and sequentially treated as follows: The diluted beads (10 uL) were transferred to a hemacytometer and covered with a coverslip. The number of beads in each of the four corners (all four 1mm × 1mm corner squares) was counted under 100X magnification. The number of beads per microliter was tabulated by multiplying the sum of beads counted in all four corners by 2.5 (volume conversion) and 100 (dilution factor). The necessary volume of 1.5X TMAC (tetramethylammonium chloride) was added to each probe-coupled bead set such that a uniform stock concentration of 1650 beads/uL was achieved for all bead sets.

I. Verification of Probe-Bead Coupling and Titration Study

1. Rehydrating the Reverse Probes

Lyophilized reverse probes were rehydrated in 1X TE to a stock concentration of 5×10^{-5} umol/uL. A serial dilution was performed until the stock had been diluted 1:100,000.

2. Preparing the Beads

The necessary amount of 1.5X TMAC was added to each stock bead set to achieve a working solution (approximately 500 uL) of each bead set with a concentration of 150 beads/uL.

3. Titration Study

Various increments of reverse probe from the appropriate working solution were combined with the corresponding bead set (1) to verify probe-bead coupling, and (2) to determine the approximate number of probes per bead. Table 4.16 represents a template on which each titration study was based. Included in the table are the various increments of reverse probe tested, the reverse probe dilution from which the reverse probe originated, the volume of reverse probe needed to achieve the indicated amount, and the volume of 1X TE needed to bring the total amount of sample to 17 uL in the Luminex assay.

Sample	Rev. Probe Dilution	Rev. Probe Concentration	Vol. Rev. Probe Needed	Vol. 1X TE Needed
0 fmol	n/a	n/a	0 uL	17 uL
0.25 fmol	1:100,000	0.5 fmol/uL	0.5 uL	16.5 uL
0.5 fmol	1:100,000	0.5 fmol/uL	1 uL	16 uL
1 fmol	1:100,000	0.5 fmol/uL	2 uL	15 uL
10 fmol	1:10,000	5 fmol/uL	2 uL	15 uL
50 fmol	1:1000	50 fmol/uL	1 uL	16 uL
100 fmol	1:1000	50 fmol/uL	2 uL	15 uL
150 fmol	1:1000	50 fmol/uL	3 uL	14 uL
200 fmol	1:1000	50 fmol/uL	4 uL	13 uL
250 fmol	1:100	500 fmol/uL	0.5 uL	16.5 uL

Table 4.16. Titration study sample template.

Following the Luminex analysis of each bead set, a chart of MFI versus Reverse Probe (fmol) was plotted to determine the point at which the fluorescence plateaued. The amount of reverse probe present at the plateau was converted, through a series of calculations (see Results, Section D), to an approximate number of probes per bead.

J. Instrument Maintenance

1. Daily Start-up and Shut-down Procedures

The Luminex 100, sheath delivery system, and XY platform were all powered on sequentially. The laser was warmed up for thirty minutes, followed by a prime to remove air bubbles from the fluidics subsystem. A 70% alcohol flush was performed to remove air bubbles from the cuvette. The start-up procedure concluded with a sheath fluid wash to clear dust and other debris from the system.

After each assay, the system was sanitized using a 10% bleach solution to decontaminate the sample lines and remove any trace amounts of biohazardous materials. A soak step was performed using distilled water to remove any remaining sheath fluid that could otherwise dry out and crystallize.

2. Monthly Maintenance

The system was calibrated on a monthly basis with control (CON1 and CON2) and calibration (CAL1 and CAL2) microspheres. Calibration microspheres have a known fluorescent light intensity and are designed to calibrate settings for both classification channels and the doublet discrimination channel. Control microspheres verify the calibration and optical integrity for both classification channels and the doublet discrimination channel. A calibration log was maintained to document the results.

3. Semi-Annual Maintenance

Heavy use of the Luminex instrument sometimes caused obstructions to form in the instrument probe despite the start-up and shut-down procedures designed to prevent such blockages. Minor obstructions were removed by running two backflushes, three

drains, two alcohol flushes, and five washes with distilled water. Severe clogs required the removal of the probe, which was sonicated for five minutes to loosen the obstruction. A syringe was used to flush water through the probe in the direction opposite of normal flow. The instrument probe was then replaced in the instrument and probe height was adjusted.

At the time of purchase, the instrument was certified by a Luminex field engineer. Periodic maintenance of the instrument was performed by Luminex personnel as needed according to the benefits allowed under the service contract.

K. Luminex Assay

Each Luminex assay was performed using a Luminex 100 (Luminex, Austin TX) and MiraiBio MasterPlex® CT software version 1.2 (Hitachi Solutions America, San Francisco, CA). Reverse probes were used as positive controls where indicated. Samples and controls were loaded into a 96-well flat-bottom microtiter plate outfitted with a polycarbonate PCR microplate with V-bottom wells (Corning Inc., Corning, NY). Three 1X TE washes (75 uL each) were included between each sample unless otherwise indicated. The remainder of the assay was performed under the illumination of only a darkroom light to protect the beads from exposure to light.

The appropriate bead set (titration study and enzymatic degradation study) or bead mix (all other assays) was added (33 uL) to each sample well and mixed by pipeting up and down. Any given bead set was added to the sample wells such that its concentration was 150 beads/uL, whether one or more bead sets were present. The reaction plate was covered with adhesive aluminum foil (Corning Inc., Corning, NY) and incubated in a 95°C water bath. After five minutes, the reaction plate was transferred to a

48°C water bath for fifteen minutes. This hybridization temperature was selected because it was just below the lowest melting temperature of all probes used in this study (see Tables 4.3 and 4.7).

The reaction plate was then centrifuged at 2250 × g for three minutes in an Eppendorf 5430 microcentrifuge equipped with a swingbucket rotor for microtiter plates (Eppendorf, Hamburg, Germany). A streptavidin-R-phycoerythrin (SAPE; Invitrogen, Carlsbad, CA) reporter mix was prepared by combining 4 uL of stock SAPE (1 mg/mL) with 996 uL of 1X TMAC. After centrifugation the supernatant was discarded and 75 uL of the reporter mix was added to each sample well. The reaction plate was loaded in the Luminex 100 and incubated at 48°C for five minutes. Samples were sequentially analyzed at 48°C.

L. Options for Improving Sensitivity

1. Salmon Sperm DNA

All target sequences were single-stranded, but because they were amplified by PCR, the complement of the target sequence was always amplified simultaneously. The complement of the target sequence competes with the probe for binding to the target sequence. This competition may be detrimental to the sensitivity of the assay. Although the LLDs are heavily dependent on the chosen quantitation method, it is fair to say that the LLDs in prior studies might have been improved if the target sequence was present without its complement.

The use of salmon sperm DNA was explored as a method for removing unwanted DNA. Salmon sperm DNA contains very heterologous sequences and has traditionally been used as a blocking agent in blotting for the purpose of increasing specificity and

sensitivity. It has been used more recently in DNA microarray protocols – applying salmon sperm DNA during the hybridization step may reduce background fluorescence signals. In the current research, salmon sperm DNA could have been used to bind non-specifically to the unwanted complement of the target DNA. However, this idea did not materialize because there was no apparent way to prevent the salmon sperm DNA from binding the target sequence, which must be kept free for binding with the probe. The use of salmon sperm DNA was also discouraged by Luminex technical support because it was believed that this would only slow (and not entirely prevent) the re-association of the PCR amplicon duplex (I. Rupani, personal communication, 19 February 2010).

2. Streptavidin-coated Microtiter Plates

Streptavidin-coated microtiter plates were also considered for improving sensitivity. Recall that the strand of target DNA to which the probe binds is biotinylated. The unwanted strand of the target DNA is not biotinylated. Therefore, a post-PCR purification step could be performed in which the target DNA amplicon is introduced to a streptavidin-coated plate. The biotinylated strand should bind the streptavidin, allowing the non-biotinylated strand to be removed. The biotin-streptavidin bond would then have to be broken, allowing elution of the biotinylated strand of the PCR product. This would make the purified, biotinylated strand of the target DNA free for binding with the probe-coupled beads. This idea was also not feasible because there was no apparent way to break the biotin-streptavidin bond without compromising the biotinylated target DNA. Biotin and streptavidin have a very strong association constant ($K_a \sim 10^{15}$ M^{-1}; Morag *et al.*, 1996) which can be a very useful characteristic at times, and yet very disadvantageous at others. Researchers have modified their methods in an effort to

duplicate the strength of the biotin-streptavidin bond, yet with more options concerning reversibility. These protocols have involved modifiying either the biotin (Olejnik *et al*., 1995; Hirsch *et al*., 2002) or the streptavidin (Morag *et al*., 1996) moieties. Although biotin and streptavidin may still bind to each other with these modifications, the association is often weaker. Still other researchers have attempted to solve the problem by experimenting with reagents such as EDTA/formamide mixtures in order to elute the biotinylated analyte (Tong and Smith, 1992). Although these attempts have been met with some success, there is speculation that the use of harsh reagents may have a deleterious impact on the DNA analyte. Thus, the analyte may be purified, but its quantity and quality could be seriously compromised in the process. The idea of using streptavidin-coated plates in this research was abandoned because of the difficulties anticipated in breaking the biotin-streptavidin bond without putting precious target DNA at risk for destruction.

3. Allele-specific Primer Extension

One option for improving sensitivity was suggested by a technical support specialist at Luminex (I. Rupani, personal communication, 12 February 2010). Allele-specific primer extension (ASPE) is a technique that could potentially improve sensitivity and minimize cross reactivity. This technique works by first amplifying the target DNA by PCR. The DNA is then reduced to its single-stranded state by heating and an allele-specific primer (with a 5′ TAG sequence attached) anneals to the target DNA wherever a complementary sequence is recognized. A polymerase then extends the primer, adding biotinylated nucleotides that will facilitate detection inside the Luminex instrument. When the extended primers are introduced to the beads, the 5′ TAG sequence on the

primers should hybridize with a complementary anti-TAG sequence, which is covalently linked to the beads by the manufacturer. Figure 4.3 shows a schematic of the overall ASPE technique and Figure 4.4 shows a close-up view of the TAG sequence hybridizing with the Anti-TAG sequence attached to the bead surface.

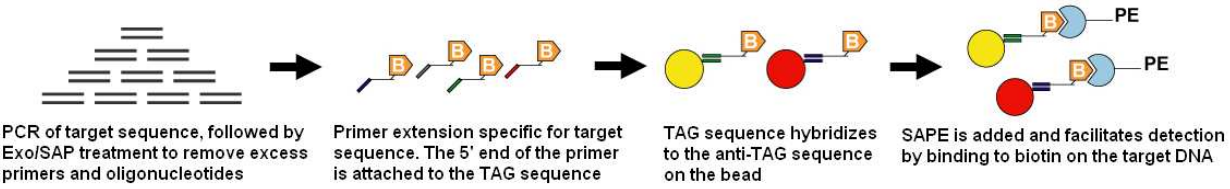


Figure 4.3. Schematic of ASPE and detection using Luminex MicroPlex xTAG microspheres. Target sequence is first amplified and then a biotinylated complement of the target sequence is extended using a primer with TAG sequence attached. Microspheres with anti-TAG sequences attached are introduced and bind to TAG sequences. Hybridization is detected by fluorescent emission when SAPE binds the biontinylated target. Schematic adapted from original image found at >http://www.luminexcorp.com/TechnologiesScience/xTAGTechnology/index.htm<. Original image accessed on the World Wide Web on 09Aug2011. Reprinted with permission from Luminex Corp.

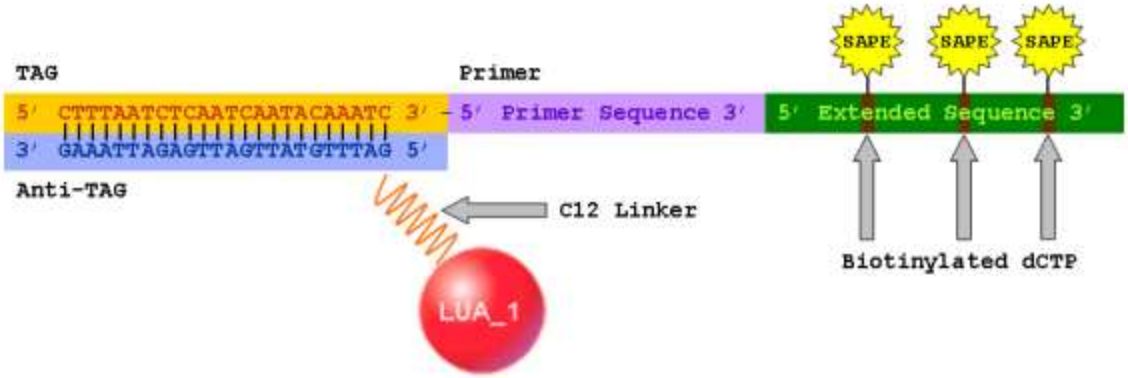


Figure 4.4. Close-up view of the TAG sequence hybridizing with the anti-TAG sequence attached to the microsphere. Image accessed on the World Wide Web on 09Aug2011 at >http://www.miraibio.com/microspheres/luminex-microplex-xtag-microspheres-flexmap-overview.html?Itemid=0<. Reprinted with permission of MiraiBio Group of Hitachi Solutions America, Ltd.

The main advantage of using ASPE is that it eliminates the need for researchers to couple probes and beads. MicroPlex-xTAG™ beads are sold with the anti-TAG sequences already coupled to the beads by the manufacturer. Once the template has been

amplified, the researcher has only to combine the TAG-labeled amplicon with the anti-TAG-labeled beads in order for the reaction to occur. Use of the ASPE technique also obviates the need for probe design; however, primers specific for the target sequence must be designed instead. Special software is available to help researchers design appropriate primers that will not cross-react with the TAG sequences. An ancillary advantage of ASPE is that it has the potential to enhance specificity enough to permit differentiation of targets whose sequences may differ by one SNP (Song *et al.*, 2010). Researchers have had greater success distinguishing sequences with SNPs because the TAG sequence (not the region of the target with the SNP) is what actually binds the probe (anti-TAG). Sensitivity is also improved with ASPE because there is nothing (other than the anti-TAG) competing to bind with the TAG sequence. This contrasts with traditional detections in which the probe is competing with single-stranded complementary PCR product for binding with the target DNA.

Unfortunately, the option to perform ASPE did not become apparent until the research was well underway – primers, probes, and target DNA had all been designed and/or ordered from the vendor. Additionally, ASPE requires the use of special MicroPlex-xTAG™ beads offered by Luminex. These beads are different from the conventional MicroPlex beads because xTAG beads are sold with the anti-TAG sequence already attached. Applying the ASPE technique would have essentially rendered the beads already purchased useless, and purchasing new beads was not feasible within the limits of the project budget. For these reasons, the ASPE technique was not applied, but future researchers may want to consider it for improving sensitivity as well as expediting benchwork.

4. Enzymatic Digestion

Although there was at least one flaw with each of the proposed methods for improving sensitivity, there was still one viable alternative. There remained some appeal to the concept of somehow tagging for destruction the unwanted complement of the target strand. Recall that the PCR primers were designed such that the reverse primer was biotinylated, thus imparting a biotin moiety which enabled detection of the target strand. Using similar logic, the *forward* primer could be phosphorylated on the 5′ end. This would enable post-PCR enzymatic degradation of the unwanted strand. A method has been described for adding phosphate groups to unwanted DNA and subsequently degrading this DNA using an enzyme called lambda exonuclease (Schwieger and Tebbe, 1998; Mitsis and Kwagh, 1999). The PCR amplicon would be almost exclusively composed of the target DNA, thus relieving the probe of competition that it faces for the target DNA.

Exonucleases are enzymes used by cells *in vivo*, usually for degrading DNA. These enzymes vary slightly in terms of substrate and product, but in general they begin at the terminus of a DNA sequence and work inwards, cleaving one base at a time until the entire sequence is degraded. In order for lambda exonuclease to function, double-stranded DNA sequences with a 5′ phosphate on the strand(s) targeted for degradation must be available as the substrate. Figure 4.5 illustrates how the phosphate group, when attached to the 5′ end of the primer which amplifies the unwanted strand of the PCR product, can be incorporated into all PCR products. When the PCR amplicons are subjected to lambda exonculease under specific reaction conditions, the unwanted PCR strands with the 5′ phosphate are degraded and the target sequence is left untouched.

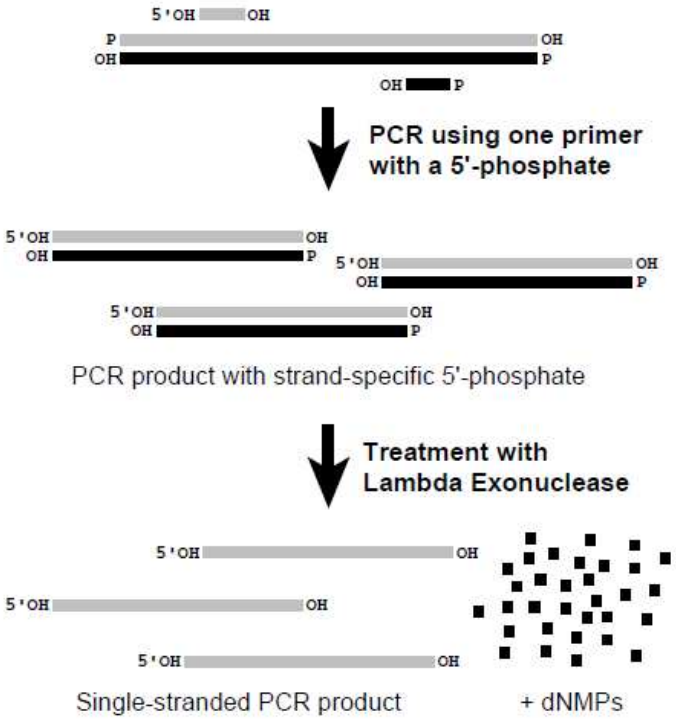


Figure 4.5. Lambda exonuclease selectively digests the strand of a PCR product produced using a PCR primer with a 5´-phosphate. Copyright 2011 Epicentre (an Illumina company), Madison, WI. Used with permission.

It has already been mentioned that the PCR primers were designed such that the reverse primer was biotinylated, thus imparting a biotin moiety which enabled detection of the target strand. By phosphorylating the forward primer on the 5´ end, the non-complementary PCR strand can be degraded (post-PCR) using lambda exonuclease. The PCR amplicon would be almost exclusively composed of the target DNA, thus relieving the probe of competition that it faces for the target DNA.

A study was undertaken to examine the effectiveness of enzymatically degrading the non-complementary PCR strand. It was thought that the enzymatic degradation of competing sequences in the reaction mixture might enhance the hybridization efficiency of probe and target DNA. Lambda exonuclease (New England BioLabs, Ipswich, MA) was acquired for this purpose. This enzyme is a highly processive 5′→3′ exonuclease

isolated from a recombinant source. Its preferred substrate is 5′-phosphorylated duplex DNA. If one strand of the duplex contains a 5′-phosphate group and the other does not, only the phosphorylated strand will be digested, leaving a single-stranded product.

Multiple samples of *B. cereus* were amplified using the diluted primer mix and 10 ng of template (as determined in the PCR optimization study) according to the TD-PCR protocol. PCR amplicons were quantified using the Agilent 2100 Bioanalyzer. Lambda exonuclease 10X reaction buffer (5 uL) was added to the control and experimental tubes. Lambda exonuclease enzyme (0.5 uL; 5000 U/mL) was added to the experimental tubes. Enough water to bring the total reaction volume to 50 uL was added to both control and experimental tubes. Samples were incubated at 37°C for 30 minutes followed by a 10 minute heat inactivation at 75°C. Post-digestion quantitation of control and experimental tubes was performed using the Agilent 2100 Bioanalyzer. Select control and experimental tubes were analyzed on the Luminex 100 as previously described. The only modification was that Bac/RRL beads (no other bead sets) were added to each sample well.

It should be noted that even though the contents of the digested experimental tubes were loaded on to the Luminex plate, the post-digest concentration of these tubes was not used to determine the quantity of DNA added to each Luminex well. Although the post-digest quantitation of experimental tubes helped to establish the presence of digested DNA, the concentrations reported were not regarded as accurate because the Agilent kit is designed for ds-DNA, and digested DNA is mostly ss-DNA. Furthermore, loading equal amounts of control and experimental samples provided a better basis for comparison. Accordingly, digested amplicons from the experimental tubes were loaded into the Luminex sample wells based on the post-digest concentration reported for the

RRL band in the corresponding control tubes. For example, if the RRL band in the *B. cereus* post-digest control tube had a concentration of 5 ng/uL, then this (for the purposes of loading the Luminex plate) was also assumed to be the concentration of the *B. cereus* post-digest experimental tube. A graphical comparison of MFI values for control and experimental samples was prepared to assess the effectiveness of enzymatic degradation of the non-complementary PCR strand.

The enzymatic degradation study was repeated as described above for *B. anthracis* (only Bac/RRL and Bant/TOX beads were used in the Luminex assay) and *Y. pestis* (only Ypes/RRL and Ypes/TOX beads were used in the Luminex assay).

M. Specificity/Sensitivity Study

Templates from previous amplifications were pooled and the contents of each PCR tube were split equally between a control and experimental tube. The experimental tubes were digested with lambda exonuclease enzyme. Post-digest samples were quantified using the Agilent 2100 Bioanalyzer. Samples of each template were loaded into individual wells on a Luminex plate in the following increments: 0.1 ng, 0.5 ng, 1.0 ng, 5 ng, 10 ng, 20 ng, 40 ng, and 80 ng (unless otherwise indicated). A bead mix consisting of all bead sets was prepared in which each bead set had a concentration of 150 beads/uL. The samples were analyzed on the Luminex 100.

N. Mixture Study (1:1)

A mixture study was performed in order to evaluate the ability of the Luminex platform to identify all components in binary, ternary, and quaternary mixtures made up of the four core bacteria and their close relatives. There were many possible mixtures that

could have been created (excluding the outlier). The number of possible mixtures was calculated using the formula $C(n,r) = n!/r!(n-r)!$ where n = total number of bacterial species and r = number of bacteria in each mixture. Mixtures in which both bacteria belonged to the same genus were omitted from the calculations because intra-genus mixtures were examined in a later study. The calculations indicated that there were 24 possible binary mixtures, 32 possible ternary mixtures, and 16 possible quaternary mixtures. Studying all of these mixtures would have been impractical and wasteful of limited reagents. Mixtures in which one or more components consisted of a core bacterium (*B. anthracis*, *C. botulinum*, *F. tularensis tularensis*, or *Y. pestis*) were deemed most important and thus received preference when the mixtures were selected for Luminex analysis (Table 4.17).

Select samples were split into three batches and amplified using the diluted primer mix and 10 ng of template (as determined in the PCR optimization study). Samples were amplified according to the TD-PCR protocol and quantified using the Agilent 2100 Bioanalyzer.

	No Tox	One Tox			Two Tox			Three Tox	Four Tox
Binary	by	By		bY	BY				
	bf	Bf		bF	BF				
	bc	Bc		bC	BC				
	yf	Yf		yF	YF				
	yc	Yc		yC	YC				
	fc	Fc		fC	FC				
Ternary	byf	byF	bYf	Byf	BYf	ByF	bYF	BYF	
	byc	byC	bYc	Byc	BYc	ByC	bYC	BYC	
	bfc	bfC	bFc	Bfc	BFc	BfC	bFC	BFC	
	yfc	yfC	yFc	Yfc	YFc	YfC	yFC	YFC	
Quat.	byfc	Byfc			BYfc		bYFc	BYFc	BYFC
		bYfc			ByFc		bYfC	BYfC	
		byFc			ByfC		byFC	ByFC	
		byfC						bYFC	

Table 4.17. Summary of all possible 1:1 mixtures (excluding *M. tuberculosis*). B = *B. anthracis*; b = *B. cereus*; C = *C. botulinum*; c = *C. difficile*; F = *F. tularensis tularensis*; f = *F. tularensis holarctica*; Y = *Y. pestis*; y = *Y. enterocolitica*. Cells highlighted in yellow indicate samples in batch #1; cells highlighted in blue indicate samples in batch #2; cells highlighted in orange indicate samples in batch #3.

Digested DNA was added to each sample well in the Luminex plate so as to achieve a total target amount of 10 ng (the highest LLD observed in the specificity/sensitivity study). When loading samples on the Luminex plate, the amount of digested template needed to satisfy this threshold was determined using whichever one of the following strategies applied:

- if only one RRL band was detected in a given control tube (for example, see Appendix VII, Table A7.1, Lane 1), the concentration of digested template in the corresponding experimental tube was estimated by dividing the concentration of the sole RRL band by the number of components in the mixture;
- if one RRL band was detected for <u>*each*</u> component of the mixture (for example, see Appendix VII, Table A7.1, Lane 2), the concentration of digested template in

the experimental tube was estimated by taking the average concentration of all RRL bands observed in the control sample.

The rationale behind the aforementioned calculations was to establish the amount of DNA contributed by each component in a mixture, thereby ensuring that components were present in roughly equal proportions. The concentration of the RRL band(s) were selected as the basis for Luminex input amounts because concentration data were available whether the bacterium was one of the four core microbes or a close relative. Basing Luminex input amounts on TOX bands would not have been prudent because only the four core microbes (not their close relatives) would display a TOX band.

O. Mixture Study (Intra-Genus; 1:1)

A binary mixture (1:1) of *C. difficile* and *C. botulinum*, as well as a binary mixture (1:1) of *Y. enterocolitica* and *Y. pestis* were selected for analysis in the intra-genus mixture study. Each sample was amplified using the diluted primer mix and 10 ng of each template (as per the PCR optimization study). The samples were quantified on the Agilent 2100 Bioanalyzer, digested, and re-quantified post-digestion. The amount of digested DNA to be loaded on the Luminex plate was determined according to the method described above. Samples were analyzed on the Luminex as previously described.

P. Mixture Study (Ratios)

Four binary mixtures were selected for further analysis as mixtures with varying ratios. These were *B. anthracis* and *F. tularensis tularensis*, *B. anthracis* and *Y. pestis*, *Y. pestis* and *F. tularensis tularensis*, and *B. anthracis* and *C. botulinum*. These binary

mixtures were selected based on favorable results observed in the 1:1 mixture study (*e.g.*, adequate fluorescent response to both components, few false positives, and few false negatives). Each binary mixture was amplified such that the template amounts were in the following ratios: 10:1, 5:1, 2:1, 1:1, 1:2, 1:5, and 1:10. The PCR optimization study determined that the optimum template amount was 10 ng, so this amount was added for any component that made up one part of a mixture; 20 ng of template was added for any component that made up two parts of a mixture, etc. Samples were then amplified using the TD-PCR protocol. The PCR amplicons were quantified using the Agilent 2100 Bioanalyzer, digested using lambda exonuclease, and re-quantified post-digestion.

For the Luminex assay, the amount of digested DNA from the experimental tubes to be loaded on the Luminex plate was determined using whichever one of the following strategies applied:

- If separate RRL bands were observed for each component in the mixture (except 1:1 mixtures), the concentration of digested DNA in the experimental tube was taken to equal the concentration of the RRL band of the minor contributor (according to the ratio of template) in the corresponding control tube (for example, see Appendix IX, Table A9.1, Lane 1);
- If separate RRL bands were observed for each component in a 1:1 mixture, the concentration of digested DNA in the experimental tube was estimated by taking the average of the two RRL bands (for example, see Appendix IX, Table A9.1, Lane 4);
- If only one combined RRL band was detected for both components in the control tube, the amount of digested DNA in the corresponding experimental tube was

estimated by dividing the concentration of the RRL band (from the control tube) by $n+1$, where n = parts of the mixture contributed by the major component (for example, see Appendix IX, Table A9.3, Lane 1). The rationale behind this calculation was to establish the amount of DNA contributed by the minor component only, thus ensuring that the amount of input DNA for both components would meet or exceed the Luminex LLD.

The samples were analyzed on the Luminex as previously described.

Q. Data Analysis

The uniformity of amplicons produced in the monoplex PCR study was assessed by computing the average, standard deviation, and range of all reported concentrations. Selection of the best primer mix for use in the multiplex PCR study was accomplished by assessing the average and standard deviation for all concentrations reported when (1) the original primer mix was used, (2) when the modified primer mix was used, (3) when the modified, diluted primer mix was used, and (4) when the modified, diluted primer mix was used with reduced template. The abundance of primer dimers was also a factor considered when choosing the best primer mix.

The Luminex instrument has a red laser (635 nm) that classifies each bead type. Light scatter from the red laser is interpreted by a series of avalanche photo diodes (APDs). Two tandem APDs classify each bead by interpreting light scattered when the internal red and infrared dyes are excited. The third APD acts as a doublet discriminator and ensures that fluorescence is reported only for beads that pass through the instrument channel as singlets and not doublets or triplets.

The instrument also has a green laser (532 nm) that excites the reporter fluorophore (SAPE) attached to each target sequence that has hybridized to a probe. A photo multiplier tube (PMT) works with the green laser to interpret the light scatter produced by the probe-target duplexes as each bead passes through the channel. The reporter fluorescences emitted by all probe-target duplexes on a single bead in a given bead set for a given sample were pooled and the median fluorescence intensity (MFI) was calculated by the analysis software. The median statistic is preferred over the mean because it discounts most (if not all) statistical outliers and helps to minimize a skew in the data if one exists (J. Callaway, personal communication, 22 June 2011). Background noise (three times the MFI for the negative control) was subtracted from each raw MFI value to yield net MFI values.

In the titration study, the enzymatic degradation study, and the specificity/sensitivity study, net MFIs as a function of target sequence quantity were graphed as XY scatter plots in Microsoft® Excel. The plots created for the enzymatic degradation study and the specificity/sensitivity study were further analyzed by displaying the line equation and coefficient of determination (R^2) for each assay.

For mixtures analyzed in the 1:1 inter- and intra-genus mixture studies, net MFIs for each probe were plotted in bar graph format using Microsoft Excel. Average net MFI values along with corresponding standard deviations for each bead set in the 1:1 inter-genus mixture study were computed (but only for cases in which there was designed complementarity between the probe and target sequence). Coefficients of variation (CV) were also calculated for each bead set because CV facilitates an easier comparison of variation across several data sets that have a wide range of means.

Net MFIs for each of the four mixtures analyzed in various ratios were plotted in Microsoft Excel as a three-dimensional column chart which enabled comparison of data across multiple categories (ratios) and series (probes). Plotting the data in this format permitted the visualization of trends in the net MFI values as the input ratios were adjusted. Mean, standard deviation, and coefficient of variation were calculated for net MFI values of the minor component in each mixture. Microsoft Excel was used to perform an Analysis of Variance (ANOVA) at a 95% confidence level on net MFI data from the Bac/RRL and Bant/TOX probes for those mixtures in which *B. anthracis* was the minor contributor. The purpose of this analysis was to determine if the variation in net MFIs observed with different major components was dependent on the identity of the major component.

CHAPTER V. RESULTS

A. Bioinformatics

1. RRL Probe and Primer Design

RRL probe and primer sequences were designed to be complementary to sequences within the 23S rRNA gene *rrl* that had been previously studied using other bacteria (Dunbar and Jacobson, 2007). The high degree of homology in the primer binding regions (see yellow highlighted bases in Figure 5.1) prompted the design of degenerate primers. Unfortunately a separate forward primer had to be developed for *Y. enterocolitica* and *Y. pestis*. Even though the forward primer binding region of the two Yersinia species shared a great degree of homology (up to 16 out of 20 bases) with the other sequences, the Yersinia forward primer was offset by one base, thus resulting in a "frameshift" that made it difficult to align with the other sequences. Accordingly, RRL fwd. primer #1 was designed to amplify the RRL target in *B. cereus*, *B. anthracis*, *C. difficile*, *C. botulinum*, *F. tularensis holarctica*, *F. tularensis tularensis*, and *M. tuberculosis*. RRL fwd. primer #2 was designed to amplify the RRL target in *Y. enterocolitica* and *Y. pestis*. One degenerate reverse primer was designed to amplify the reverse strand in all nine templates.

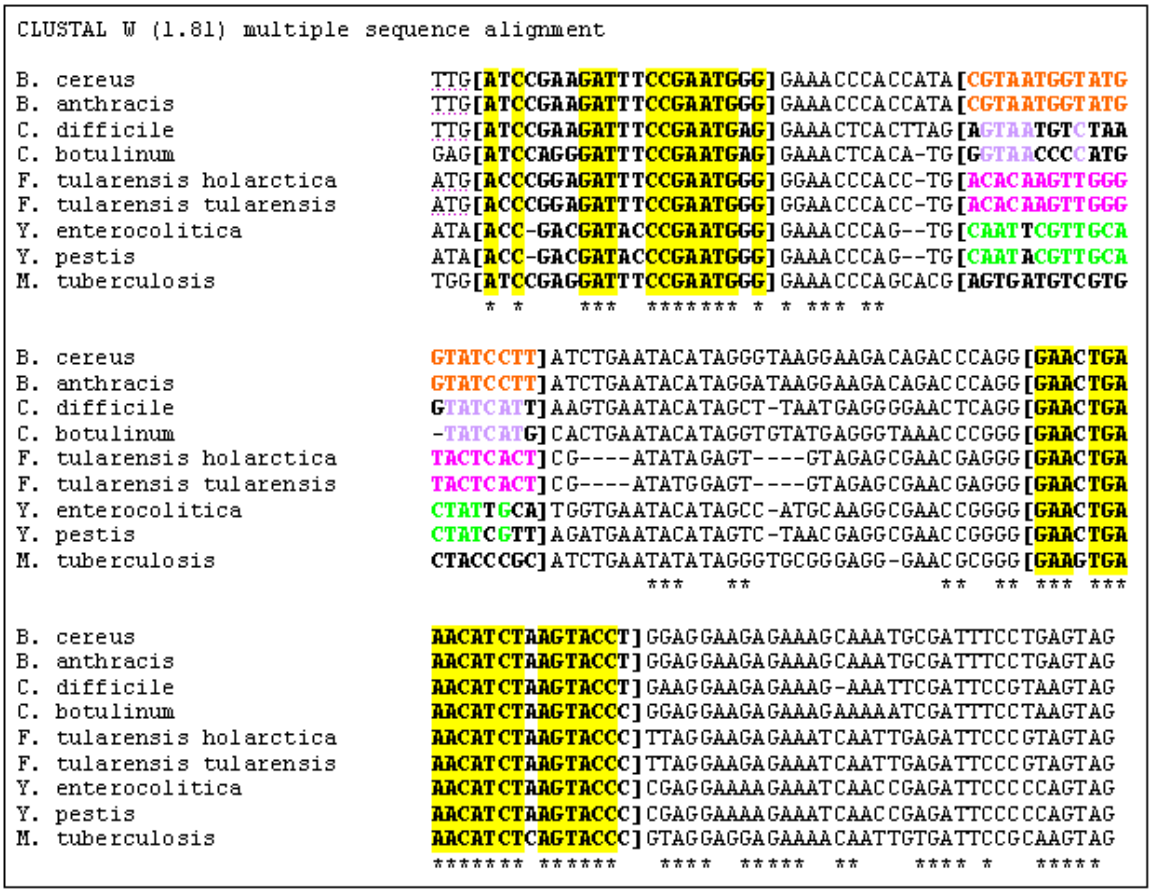


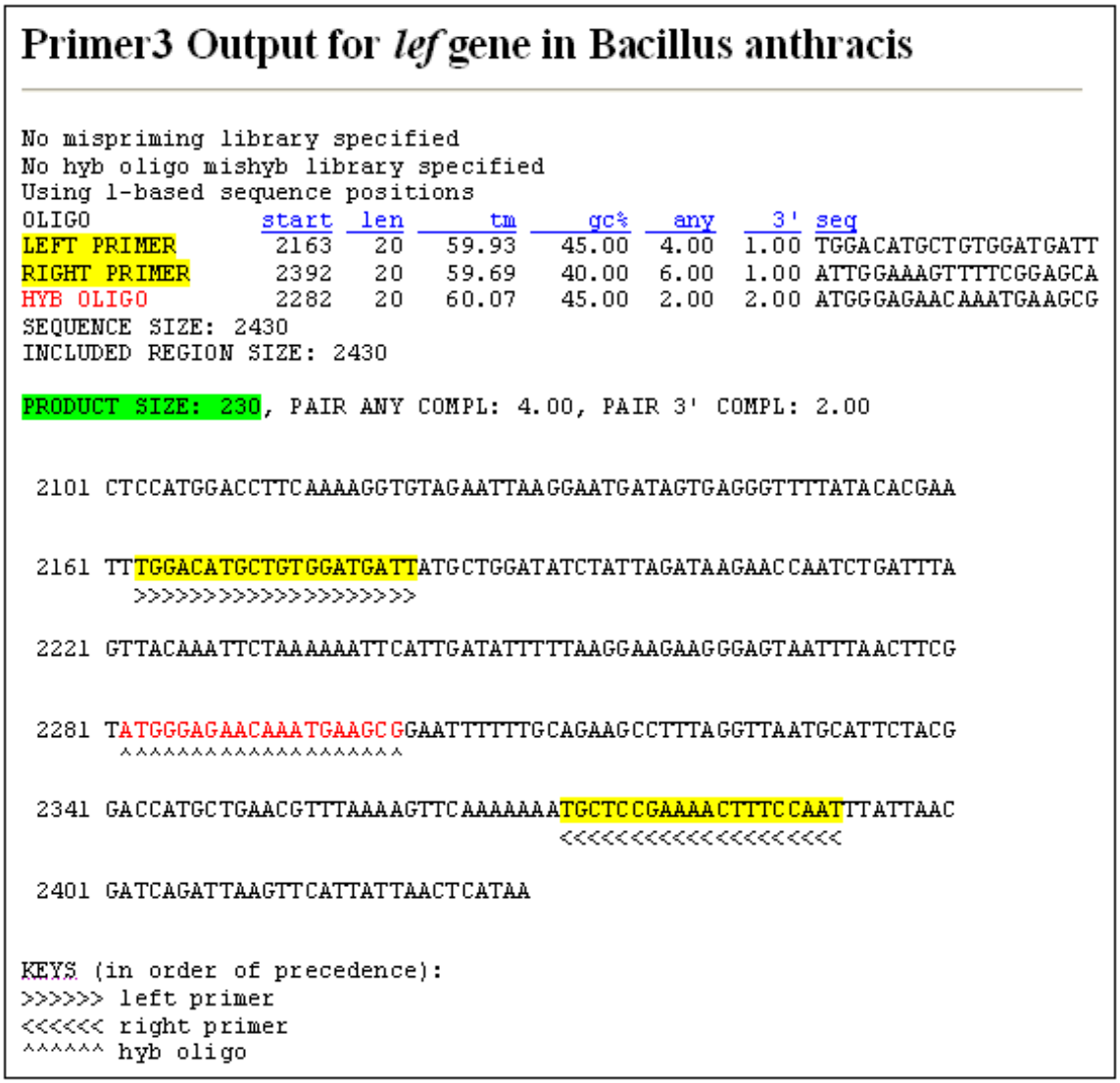
Figure 5.1. ClustalW alignment of target regions in the *rrl* gene for the nine bacteria. The forward primer-binding region begins at the fourth base and continues for the next twenty bases (see bracketed region). The probe sequences are designated by the next bracketed region (multicolors). A unique probe-binding region is identified by a different font color for each genus. Polymorphisms in the probe between two bacteria of the same genus were left in black font. The reverse primer-binding region is the bracketed region furthest downstream. Bases in the primer-binding region that are homologous across all nine bacteria are highlighted in yellow.

Initially a presumptive probe was desired for each of the nine bacteria, but unfortunately this was not possible due to homologies across species and genera. The goal was modified such that a presumptive probe for each genus was sought. Total homology between the probe binding regions of both species in some genera (Bacillus and Francisella) enabled the design of one RRL probe for both species within that genus. However, other bacteria (Clostridium and Yersinia) possessed one or more

polymorphisms between two species of the same genus, and thus a unique probe for each species had to be designed.

2. TOX Probe and Primer Design

Toxin-coding genes among the four pathogenic bacteria were selected. Primer3 software was used to identify TOX target sequences (including the flanking PCR primers and the probe sequence within). Figure 5.2 shows an example of output data when Primer3 was used to design the primer and probes for the TOX target sequence of *B. anthracis*.



Primer3 Output for *lef* gene in Bacillus anthracis

```
No mispriming library specified
No hyb oligo mishyb library specified
Using 1-based sequence positions
OLIGO            start  len      tm      gc%   any   3'  seq
LEFT PRIMER       2163   20   59.93    45.00  4.00  1.00 TGGACATGCTGTGGATGATT
RIGHT PRIMER      2392   20   59.69    40.00  6.00  1.00 ATTGGAAAGTTTTCGGAGCA
HYB OLIGO         2282   20   60.07    45.00  2.00  2.00 ATGGGAGAACAAATGAAGCG
SEQUENCE SIZE: 2430
INCLUDED REGION SIZE: 2430

PRODUCT SIZE: 230, PAIR ANY COMPL: 4.00, PAIR 3' COMPL: 2.00


 2101 CTCCATGGACCTTCAAAAGGTGTAGAATTAAGGAATGATAGTGAGGGTTTTATACACGAA


 2161 TTTGGACATGCTGTGGATGATTATGCTGGATATCTATTAGATAAGAACCAATCTGATTTA
        >>>>>>>>>>>>>>>>>>>>


 2221 GTTACAAATTCTAAAAAATTCATTGATATTTTTAAGGAAGAAGGGAGTAATTTAACTTCG


 2281 TATGGGAGAACAAATGAAGCGGAATTTTTTGCAGAAGCCTTTAGGTTAATGCATTCTACG
       ^^^^^^^^^^^^^^^^^^^^


 2341 GACCATGCTGAACGTTTAAAAGTTCAAAAAAATGCTCCGAAAACTTTCCAATTTATTAAC
                                      <<<<<<<<<<<<<<<<<<<<


 2401 GATCAGATTAAGTTCATTATTAACTCATAA


KEYS (in order of precedence):
>>>>>> left primer
<<<<<< right primer
^^^^^^ hyb oligo
```

Figure 5.2. Screenshot of Primer3 output for TOX target sequence design in *B. anthracis*. The highlighted region beginning at base 2163 indicates the forward primer. The highlighted region beginning at base 2392 indicates the complement of the reverse primer. The bases in red font indicate the probe sequence.

3. Search of the BLASTN Database for RRL and TOX Primer and Probe Sequences

All RRL and TOX primer and probe sequences were uploaded in the FASTA format to the blastn algorithm of the NCBI search engine. No significant alignments with other sequences in the nucleotide collection (nr/nt) were reported.

4. AutoDimer Analysis of RRL and TOX Primers/Probes

The AutoDimer hairpin and primer dimer screens (threshold of seven) of all primers, probes, and target sequences predicted no hairpin-forming or dimerizing sequences.

5. Determination of Copy Number for RRL and TOX Target Sequences

Table 5.1 provides a summary of RRL and TOX target sequence copy numbers found in each of the nine bacteria. Primer and probe sequences targeted all copies of the *rrl* gene in each species.

Organism	***rrl* Copy Number**	**Tox Gene Copy Number**	**RRL Target Seq. Copy Number**	**TOX Target Seq. Copy Number**
Bacillus cereus	13	n/a	13	n/a
Bacillus anthracis str. Sterne	11	1	11	1
Clostridium difficile	11	n/a	11	n/a
Clostridium botulinum	9	1	9	1
Francisella tularensis subsp. *holarctica*	3	n/a	3	n/a
Francisella tularensis subsp. *tularensis*	3	1	3	1
Yersinia enterocolitica	7	n/a	7	n/a
Yersinia pestis	6	1	6	1
*Mycobacterium tuberculosis**	1	n/a	1	n/a

Table 5.1. Summary of RRL and TOX target sequence copy numbers. *Notice that *M. tuberculosis* has only one copy of the RRL gene, thus causing this bacterium to have a slow growth rate (Cox, 2004). A "not applicable" (n/a) sign indicates that no TOX gene data was available for the non-pathogenic species.

In order to provide a better picture of the effective concentration of each target region, the amplicon concentrations shown in Table A2.4 (Appendix II) were converted to gene copies per microliter. The values in Table A2.4 were chosen for these calculations because they came from amplicons that were produced from the primer

dilution and template amount that was ultimately selected after the PCR optimization. Briefly, the number of basepairs in the amplicon was multiplied by the mass in one mole of basepairs, divided by the number of basepairs in one mole, and then multiplied by a conversion factor. Next, concentration of the PCR product was multiplied by the number of gene copies to yield the number of gene copies per microliter. These calculations are summarized in Figure 5.3. The source data and final answer of these calculations for each target sequence are summarized in Table 5.2.

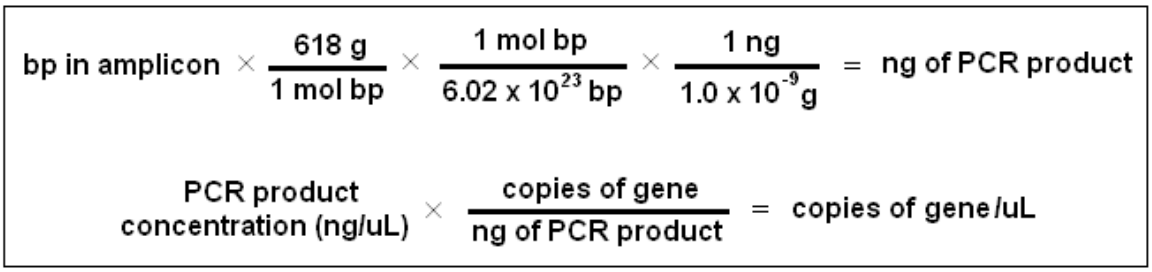


Figure 5.3. Sample calculations for determining gene copies per microliter.

Organism	Target sequence	Length (bp)	Copy number*	Concentration (ng/uL)†	Copies per microliter
B. cereus	Bac/RRL	111	13	12.90	1.47×10^{12}
B. anthracis	Bac/RRL	111	11	9.35	9.03×10^{11}
B. anthracis	Bant/TOX	230	1	6.35	2.68×10^{10}
C. difficile	Cdif/RRL	110	11	13.20	1.29×10^{12}
C. botulinum	Cbot/RRL	109	9	2.54	2.04×10^{11}
C. botulinum	Cbot/TOX	100	1	1.01	9.83×10^{9}
F. tularensis holarctica	Fran/RRL	102	3	11.57	3.31×10^{11}
F. tularensis tularensis	Fran/RRL	102	3	14.88	4.26×10^{11}
F. tularensis tularensis	Ftul/TOX	159	1	8.46	5.18×10^{10}
Y. enterocolitica	Yent/RRL	107	7	19.29	1.23×10^{12}
Y. pestis	Ypes/RRL	107	6	14.50	7.92×10^{11}
Y. pestis	Ypes/TOX	169	1	6.99	4.03×10^{10}
M. tuberculosis	Myc/RRL	110	1	1.54	1.36×10^{10}

Table 5.2. Source data and final answers for calculating gene copies per microliter. *From Table 5.1. †From Table A2.4; if more than one band was detected for a single target sequence, the sum of the both concentrations was used for the calculations.

The calculations in Table 5.2 were based on amplicon concentrations from a single quantitation and although the number of copies per microliter would undoubtedly be different for amplicons from another PCR batch, the overall trends should not deviate much from the values reported in Table 5.2.

Note that the number of gene copies per microliter reported in Table 5.2 are specific to those gene copies that contain the probe sequence, as these are the only sites that were targeted in the PCR reaction. The results indicate that the number of gene copies ranges from approximately 10^9 to 10^{12}. The lack of uniformity among gene copy number may have contributed to the variation in fluorescent response that was observed for different probes later on in this study.

B. Monoplex PCR Specificity Check

1. Nine Templates Amplified with RRL Fwd. Primer #1 and RRL Rev. Primer

Figure 5.4 and Table 5.3 provide a summary of the quantitation results of all nine templates when amplified using the RRL fwd. primer #1 and RRL rev. primer. This primer set was designed to amplify the RRL target in all templates except *Y. enterocolitica* and *Y. pestis*, for which a different forward primer had to be designed.

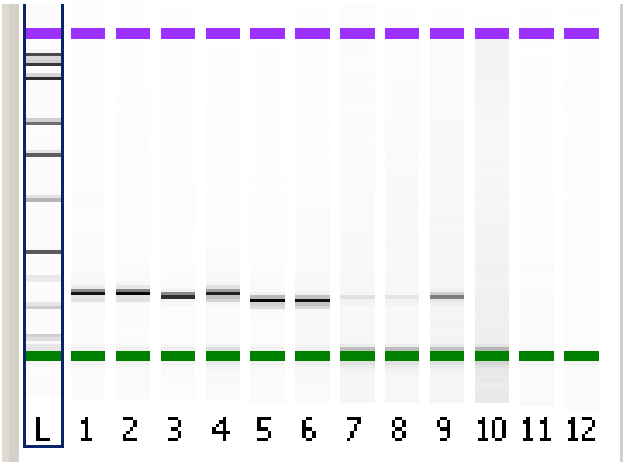


Figure 5.4. Gel image of nine templates amplified with RRL Fwd. primer #1 and RRL Rev. primer. Key: L = ladder, 1 = *B. cereus* RRL target, 2 = *B. anthracis* RRL target, 3 = *C. difficile* RRL target, 4 = *C. botulinum* RRL target, 5 = *F. tularensis holarctica* RRL target, 6 = *F. tularensis tularensis* RRL target, 7 = faint band observed with *Y. enterocolitica* template but concentration too low to be reported, 8 = faint band observed with *Y. pestis* template but concentration too low to be reported, 9 = *M. tuberculosis* RRL target, 10 = negative control, 11 = blank, 12 = blank.

	Lane								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Expected Length (bp)	111	111	110	109	102	102	n/a	n/a	110
Reported Length (bp)	124	123	118	121	109	109	*	*	117
Concentration (ng/uL)	12.04	8.58	10.29	7.96	7.41	6.63	n/a	n/a	2.43

Table 5.3. Summary of quantitation data for amplification of all nine templates using RRL Fwd. primer #1 and RRL Rev. primer. An asterisk (*) symbol indicates that a faint band was visible but its concentration was too low to be reported.

The Bac/RRL target sequence was detected when both *B. cereus* and *B. anthracis* served as template (Lanes 1 and 2). The actual length of the Bac/RRL target sequence was 111 bp, which was in close proximity to the band lengths reported in lanes 1 and 2 (124 and 123 bp, respectively). A sufficient amount of Bac/RRL target was present for both species (12.04 and 8.58 ng/uL, respectively).

The Cdif/RRL target sequence was detected when *C. difficile* served as template (lane 3). The actual length of the Cdif/RRL target sequence was 110 bp, which was in

close proximity to the 118 bp band detected in lane 3. The concentration of the Cdif/RRL target in lane 3 was reported as 10.29 ng/uL, indicating a sufficient amount of target DNA.

The Cbot/RRL target sequence was detected when *C. botulinum* served as template (lane 4). The actual length of the Cbot/RRL target sequence was 109 bp, which was in close proximity to the 121 bp band detected in lane 3. The concentration of the Cbot/RRL target in lane 4 was reported as 7.96 ng/uL, indicating a sufficient amount of target DNA.

The Fran/RRL target sequence was detected when both *F. tularensis holarctica* and *F. tularensis tularensis* served as template (lanes 5 and 6). The actual length of the Fran/RRL target sequence was 102 bp, which was in close proximity to the band lengths reported in lanes 5 and 6 (109 and 109 bp, respectively). A sufficient amount of Fran/RRL target was present in roughly equal concentrations for both species (7.41 and 6.63 ng/uL, respectively).

Lanes 7 and 8 contained *Y. enterocolitica* and *Y. pestis* template, respectively. A faint band was observed in each of these lanes, although not officially detected by the software because of insufficient concentration (the Agilent LLD is 0.02 ng/uL). The RRL fwd. primer #1 was not designed to amplify *Y. enterocolitica* and *Y. pestis*, but the sequence of this primer differed by only a few bases from RRL fwd. primer #2, which was designed to amplify these two species. The non-specific amplicons observed in lanes 6 and 7 were likely the consequence of a minor deficiency in primer specificity.

The Myc/RRL target sequence was detected when *M. tuberculosis* served as template (lane 9). The actual length of the Myc/RRL target sequence was 110 bp, which

was in close proximity to the 117 bp band detected in lane 9. The concentration of the Myc/RRL target in lane 9 was reported as 2.43 ng/uL, indicating a sufficient amount of target DNA. The concentration of the Myc/RRL target was expected to be the lowest of all target sequences because there is only one copy of the target sequence in *M. tuberculosis*. The RRL target copy number is much higher for the other bacteria, ranging from three to thirteen copies. Having fewer copies of the RRL target sequence means that the concentration of the amplicons was expected to be lower because there was a lower quantity of starting template (see Table 5.2; the Myc/RRL target sequence has a concentration of 1.36×10^{10} copies/uL, whereas most other target sequences have more copies per microliter).

2. Nine Templates Amplified with RRL Fwd. Primer #2 and RRL Rev. Primer

Figure 5.5 and Table 5.4 provide a summary of the quantitation results of all nine templates when amplified using the RRL fwd. primer #2 and RRL rev. primer. This primer set was designed to amplify the RRL target in *Y. enterocolitica* and *Y. pestis*. These two species had a sequence that differed enough from the other bacteria at the RRL target sequence locus to warrant the design of a unique forward primer.

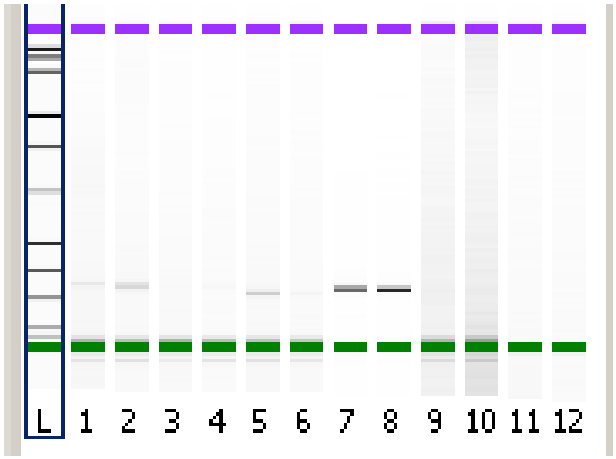


Figure 5.5. Gel image of nine templates amplified with RRL Fwd. primer #2 and RRL Rev. primer. Key: L = ladder, 1 = faint band observed with *B. cereus* template but concentration too low to be reported, 2 = band observed with *B. anthracis* template, 3 = no band observed with *C. difficile* template, 4 = no band observed with *C. botulinum* template, 5 = faint band observed with *F. tularensis holarctica* template but concentration too low to be reported, 6 = no band observed with *F. tularensis tularensis* template, 7 = *Y. enterocolitica* RRL target, 8 = *Y. pestis* RRL target, 9 = no band observed with *M. tuberculosis* template, 10 = negative control, 11 = blank, 12 = blank.

	Lane								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Expected Length (bp)	n/a	n/a	n/a	n/a	n/a	n/a	107	107	n/a
Reported Length (bp)	*	121	n/a	n/a	*	n/a	115	114	n/a
Concentration (ng/uL)	n/a	1.03	n/a	n/a	n/a	n/a	18.58	20.69	n/a

Table 5.4. Summary of quantitation data for amplification of all nine templates using RRL Fwd. primer #2 and RRL Rev. primer. An asterisk (*) symbol indicates that a faint band was visible but its concentration was too low to be reported. A "not applicable" (n/a) sign in either "length" row indicates that a PCR product was neither expected nor reported because the primers were not designed to amplify the template in that sample. A "n/a" sign in the concentration row indicates that the concentration of the visualized band was too weak to be reported, or a band was absent altogether because the primers were not designed to amplify the template in that sample.

As expected, no bands were detected in lanes 3, 4, 6, or 9 because RRL fwd. primer #2 was not designed to amplify the template used in the PCR reactions whose amplicons were loaded into these wells. Lanes 1, 2, and 5 contained *B. cereus*, *B. anthracis*, and *F. tularensis holarctica* template, respectively. A faint band was observed

in lanes 1 and 5, although not officially detected by the software because of insufficient concentration. A faint band was also observed in lane 2, whose concentration (1.03 ng/uL) was low but nevertheless adequate enough to exceed the LLD of the Agilent Bionanalyzer (0.02 ng/uL). The RRL fwd. primer #2 was not designed to amplify *B. cereus*, *B. anthracis*, or *F. tularensis holarctica* target sequences, but the sequence of this primer differed by only four bases from RRL fwd primer #1, which was designed to amplify these three templates. The non-specific amplicons observed in lanes 1, 2, and 5 were likely the consequence of a minor deficiency in primer specificity.

The Yent/RRL target sequence was detected when *Y. enterocolitica* served as template (lane 7). The actual length of the Yent/RRL target sequence was 107 bp, which was in close proximity to the 115 bp band detected in lane 7. The concentration of the Yent/RRL target in lane 7 was reported as 18.58 ng/uL, indicating a sufficient amount of target DNA.

The Ypes/RRL target sequence was detected when *Y. pestis* served as template (lane 8). The actual length of the Ypes/RRL target sequence was 107 bp, which was in close proximity to the 114 bp band detected in lane 8. The concentration of the Ypes/RRL target in lane 8 was reported as 20.69 ng/uL, indicating a sufficient amount of target DNA.

3. Nine Templates Amplified with TOX Primers

Figure 5.6 illustrates the quantitation results of all nine templates when amplified using the Bant/TOX primer set. The *B. anthracis* template in lane 2 was the only sample that yielded a detectable amplicon with the Bant/TOX primer set. This band was detected as a 231 bp fragment, which was in close proximity to the actual length (230 bp) of the

Bant/TOX target sequence. The concentration of the amplicon in lane 2 was reported as 17.16 ng/uL, indicating a sufficient amount of target DNA.

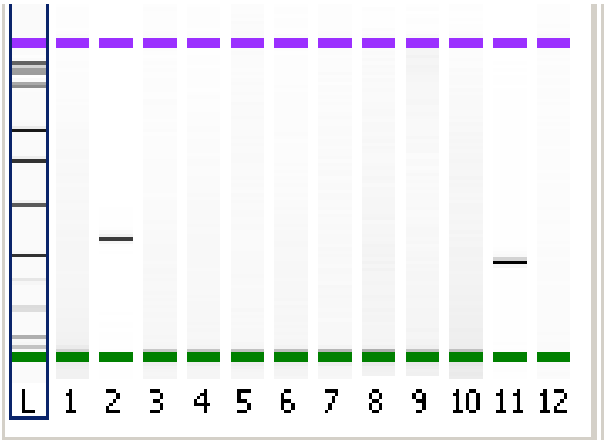


Figure 5.6. Gel image of nine templates amplified with Bant/TOX primer set. Key: L = ladder, 1 = *B. cereus*, 2 = *B. anthracis* TOX target, 3 = *C. difficile*, 4 = *C. botulinum*, 5 = *F. tularensis holarctica*, 6 = *F. tularensis tularensis*, 7 = *Y. enterocolitica*, 8 = *Y. pestis*, 9 = *M. tuberculosis*, 10 = PCR negative control, 11 = sample unrelated to this study, 12 = blank.

Figure 5.7 illustrates the quantitation results of all nine templates when amplified using the Cbot/TOX primer set. The *C. botulinum* template in lane 4 was the only sample that yielded a detectable amplicon with the Cbot/TOX primer set. This band was detected as a 112 bp fragment, which was in close proximity to the actual length (100 bp) of the Cbot/TOX target sequence. The concentration of the amplicon in lane 4 was reported as 13.83 ng/uL, indicating a sufficient amount of target DNA.

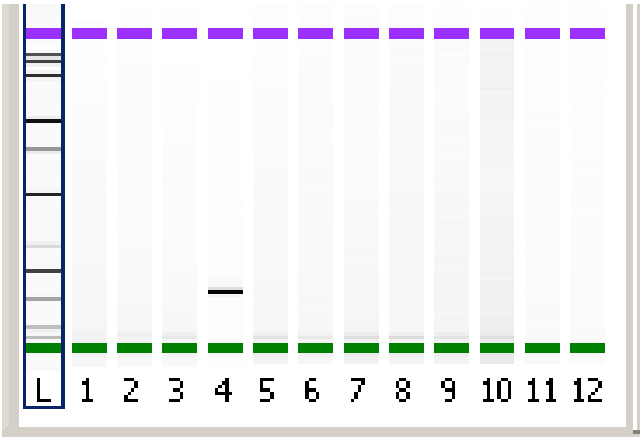


Figure 5.7. Gel image of nine templates amplified with Cbot/TOX primer set. Key: L = ladder, 1 = *B. cereus*, 2 = *B. anthracis*, 3 = *C. difficile*, 4 = *C. botulinum* TOX target, 5 = *F. tularensis holarctica*, 6 = *F. tularensis tularensis*, 7 = *Y. enterocolitica*, 8 = *Y. pestis*, 9 = *M. tuberculosis*, 10 = PCR negative control, 11 = blank, 12 = blank.

Figure 5.8 illustrates the quantitation results of all nine templates when amplified using the Ftul/TOX primer set. Amplicons were detected in lane 5 (*F. tularensis holarctica* template) and lane 6 (*F. tularensis tularensis* template). The *F. tularensis tularensis* template in lane 6 was expected to yield a detectable amplicon when amplified with the Ftul/TOX primer set. This band was detected as a 166 bp fragment, which was in close proximity to the actual length (159 bp) of the Ftul/TOX target sequence. The concentration of the amplicon in lane 6 was reported as 20.64 ng/uL, indicating a sufficient amount of target DNA.

The amplicon in lane 5 was not expected because this sample contained *F. tularensis holarctica* template and the Ftul/TOX primer set was designed to have specificity for the Ftul/TOX target sequence found only in *F. tularensis tularensis*. This non-specific amplicon was detected as a 166 bp fragment, which was in close proximity to the actual length (159 bp) of the Ftul/TOX target sequence. The concentration of the non-specific amplicon in lane 5 was reported as 17.34 ng/uL, which was similar to the concentration of the amplicon in lane 6. This observation, combined with the fact that the

non-specific amplicon in lane 5 and the Ftul/TOX sequence were similar in length, provided evidence for the possibility that the Ftul/TOX primers were not just amplifying a random section of the *F. tularensis holarctica* genome, but were instead amplifying the Ftul/TOX target sequence in *F. tularensis holarctica*. This possibility was further explored (see Discussion) to (1) determine what section of the *F. tularensis holarctica* genome was being amplified by the Ftul/TOX primers, and (2) to determine if the Ftul/TOX probe sequence was included in the non-specific amplicon.

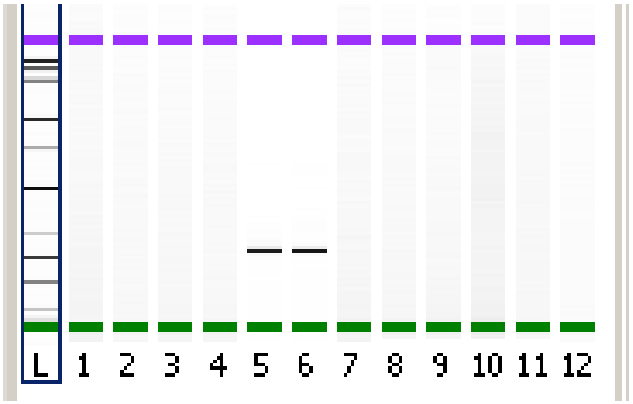


Figure 5.8. Gel image of nine templates amplified with Ftul/TOX primer set. Key: L = ladder, 1 = *B. cereus*, 2 = *B. anthracis*, 3 = *C. difficile*, 4 = *C. botulinum*, 5 = *F. tularensis holarctica* (non-specific amplicon), 6 = *F. tularensis tularensis* TOX target, 7 = *Y. enterocolitica*, 8 = *Y. pestis*, 9 = *M. tuberculosis*, 10 = PCR negative control, 11 = blank, 12 = blank.

Figure 5.9 illustrates the quantitation results of all nine templates when amplified using the Ypes/TOX primer set. The *Y. pestis* template in lane 8 was the only sample that yielded a detectable amplicon with the Ypes/TOX primer set. This band was detected as a 180 bp fragment, which was in close proximity to the actual length (169 bp) of the Ypes/TOX target sequence. The concentration of the amplicon in lane 8 was reported as 20.24 ng/uL, indicating a sufficient amount of target DNA.

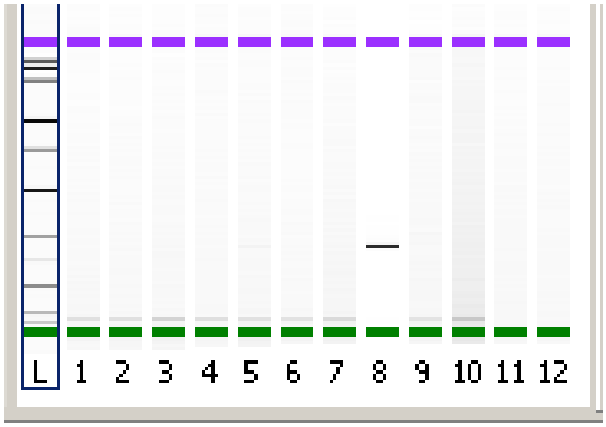


Figure 5.9. Gel image of nine templates amplified with Ypes/TOX primer set. Key: L = ladder, 1 = *B. cereus*, 2 = *B. anthracis*, 3 = *C. difficile*, 4 = *C. botulinum*, 5 = *F. tularensis holarctica*, 6 = *F. tularensis tularensis*, 7 = *Y. enterocolitica*, 8 = *Y. pestis* TOX target, 9 = *M. tuberculosis*, 10 = PCR negative control, 11 = blank, 12 = blank.

C. Multiplex PCR Study & Optimization Using Primer Mix

1. Multiplex TD-PCR Using Original Primer Concentrations

All eleven primer sets were combined into a single PCR primer mix such that each individual primer set had a final concentration of 5 uM, except the RRL primers, which had a higher concentration. Whereas each TOX primer set was designed to amplify its respective target sequence, the degeneracy of the RRL primers meant that they would have multiple targets. Accordingly, the concentration of the RRL primers was increased so that the primer mix contained RRL fwd. #1 (35 uM), RRL fwd. #2 (10 uM), and RRL rev. (45 uM). The forward primer with the highest concentration was RRL fwd. #1 because it was designed to amplify the RRL target sequence in *B. cereus*, *B. anthracis*, *C. difficile*, *C. botulinum*, *F. tularensis holarctica*, *F. tularensis tularensis*, and *M. tuberculosis*. RRL fwd. #2 was present in a lower concentration (10 uM) because it was designed to amplify the RRL target sequence only in *Y. enterocolitica* and *Y. pestis*. The RRL rev. primer was the most concentrated of all primers because it was designed to

amplify the reverse strand in all RRL target sequences. Figure 5.10 below and Table A2.1 in Appendix II illustrate the amplicons produced with the original PCR primer mix.

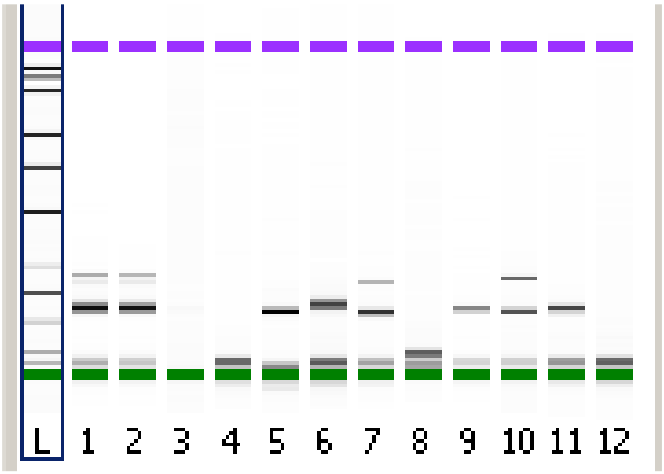


Figure 5.10. Gel image of nine templates amplified using original primer mix. L = ladder; 1 = all nine templates; 2 = all nine templates; 3 = *B. cereus* template; 4 = *B. anthracis* template; 5 = *C. difficile* template; 6 = *C. botulinum* template; 7 = *F. tularensis holarctica* template; 8 = *F. tularensis tularensis* template; 9 = *Y. enterocolitica* template; 10 = *Y. pestis* template; 11 = *M. tuberculosis* template; 12 = negative control.

In addition to amplifying each individual template with the primer mix, two samples were also prepared in which all nine genomes were combined. In Figure 5.10, the bands observed in lanes 1 and 2 represent samples in which all genomic DNA sequences were added as template. It was difficult to establish the specificity of the primer mix when all template sequences were present simultaneously. The reason for this was that many of the target sequences had overlapping size ranges, and oftentimes a single band detected by the Agilent Bioanalyzer represented multiple target sequences.

In lane 3, the 111 bp Bac/RRL target sequence was expected. Although not concentrated enough to be officially reported by the software, a faint band was observed in lane 3 that fell within the general size range of the Bac/RRL target.

In the lane 4 sample, *B. anthracis* was used as the template, so the 111 bp Bac/RRL and the 230 bp Bant/TOX target sequences were expected. Neither of these

were detected. The only bands detected were primer dimers (19 and 29 bp bands), indicating that the primer concentration may have been too high in the PCR primer mix.

The Cdif/RRL target sequence (110 bp) was most likely detected as the 117 bp band present in lane 5 (*C. difficile* template). In lane 6 (*C. botulinum* template), there was a noticeable discrepancy in size between the observed band (126 bp) and actual length of the Cbot/RRL target sequence (109 bp), but prior data from the monoplex PCR study (see Lane 4 in Figure 5.4 and Table 5.3) suggested that the observed band was indeed the Cbot/RRL target sequence. The Cbot/TOX target sequence (100 bp) was not detected in lane 6. As in lane 4, the primer dimers in lanes 5 (23 bp band) and 6 (19 and 29 bp bands) were particularly abundant.

The 114 bp band detected in lane 7 (*F. tularensis holarctica* template) represents the 102 bp Fran/RRL target sequence. In addition to some primer dimer bands (18 and 29 bp bands), a 170 bp band was also detected in lane 7, which was probably the non-specific amplicon produced by the Ftul/TOX primers in the primer mix. This non-specific amplicon was also observed in the monplex PCR study.

The primer dimers observed in lane 8 (*F. tularensis tularensis* template) were especially abundant (22, 26, and 46 bp bands). However, neither the 102 bp Fran/RRL target sequence nor the 159 bp Ftul/TOX target sequence were detected.

In lane 9 (*Y. enterocolitica* template), the Yent/RRL target sequence (107 bp) was detected as the 120 bp amplicon. Similarly, the Ypes/RRL target sequence (107 bp) was detected as the 119 bp fragment in lane 10 (*Y. pestis* template). The Ypes/TOX target sequence (169 bp) was also detected in lane 10 as the 177 bp fragment.

Finally, in lane 11 (*M. tuberculosis* template), the 110 bp Myc/RRL target sequence was detected as the 121 bp fragment. Primer dimers were observed in lane 11 (19 and 29 bp bands) as in other samples.

A primer dimer band just above the green lower marker was observed in many samples (see lanes 1-12 in Figure 5.10). The consistency and relatively high concentration of the primer dimer band in most samples was indicative of the need to reduce the primer concentration.

2. Multiplex TD-PCR Using Modified Primer Mix

The original primer mix was modified to reduce the abundance of primer dimers and prevent the occurrence of amplification drop-out. The modified primer mix consisted of TOX primers with the same concentration (5 uM each), but the concentration of the RRL primers was reduced (RRL fwd. #1 – 15 uM; RRL fwd. #2 – 10 uM; RRL rev. – 20 uM). The results of the amplification using this modified primer mix are summarized in Figure 5.11 below and Table A2.2 in Appendix II.

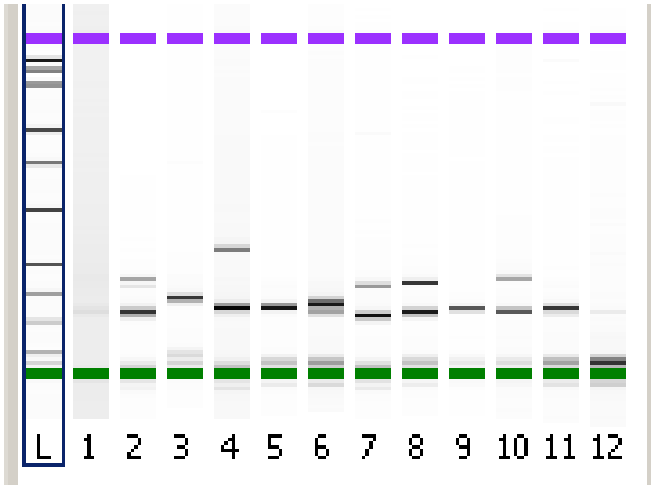


Figure 5.11. Gel image of nine templates amplified using modified primer mix. L = ladder; 1 = all nine templates; 2 = all nine templates; 3 = *B. cereus* template; 4 = *B. anthracis* template; 5 = *C. difficile* template; 6 = *C. botulinum* template; 7 = *F. tularensis holarctica* template; 8 = *F. tularensis tularensis* template; 9 = *Y. enterocolitica* template; 10 = *Y. pestis* template; 11 = *M. tuberculosis* template; 12 = negative control.

As in the amplification using the original primer mix, the modified primer mix was used to simultaneously amplify all eleven target sequences (see lanes 1 and 2). There was a universal sample drop-out observed in lane 1, and as expected, many of the amplicons in lane 2 had overlapping size ranges that prevented the resolution of each individual target sequence.

The Bac/RRL target sequence (111 bp) was probably detected as a combination of the 129 bp and 140 bp fragments observed in lane 3 (*B. cereus* template). In lane 4 (*B. anthracis* template), the Bac/RRL target sequence was detected as the 123 bp fragment. Also in lane 4, the Bant/TOX target sequence was detected as the 230 bp fragment.

In lane 5 (*C. difficile* template), the Cdif/RRL target sequence (110 bp) was detected as the 124 bp fragment. In lane 6 (*C. botulinum* template), the Cbot/RRL target sequence (109 bp) was detected as the 127 bp fragment. The Cbot/TOX target sequence (100 bp) was also detected in lane 6 as the 114 bp fragment.

The Fran/RRL target sequence (102 bp) was detected in lane 7 (*F. tularensis holarctica* template) as the 109 bp fragment. A 164 bp fragment was also detected in lane 7. This fragment was consistent with results observed for samples amplified in the monoplex study, and in the multiplex study using the original primer mix. It was most likely a non-specific amplicon produced when *F. tularensis holarctica* template was amplified with the Ftul/TOX primers.

The Fran/RRL target sequence was also detected in lane 8 when template DNA originating from *F. tularensis tularensis* was used. The Ftul/TOX target sequence (159 bp) was detected in lane 8 as the 169 bp fragment.

In lane 9 (*Y. enterocolitica* template), the Yent/RRL target sequence (107 bp) was detected as the 121 bp fragment. The Ypes/RRL target sequence (107 bp) was detected as the 118 bp fragment in lane 10 (*Y. pestis* template). Also in lane 10, the Ypes/TOX target sequence (169 bp) was detected as the 177 bp fragment.

The Myc/RRL target sequence (110 bp) was detected in lane 11 (*M. tuberculosis* template) as the 121 bp fragment.

The results of the amplification using the modified primer mix demonstrated an improvement over the original primer mix. Primer dimers were still consistently observed in most samples (see bands just above green lower marker in lanes 1-12, Figure 5.11), although their concentrations were usually much lower than observed with the original primer mix. A subsequent amplification was planned to determine whether diluting the modified primer mix two-fold would maintain amplification efficiency while reducing the occurrence of primer dimers.

3. Multiplex TD-PCR Using Modified Primer Mix Diluted 1:2

In order to determine if the concentration of primer dimers could be further reduced, the modified primer mix was diluted 1:2 and used to re-amplify the samples. The diluted primer mix concentrations were half of those in the modified primer mix (Each TOX primer – 2.5 uM; RRL fwd. #1 – 7.5 uM; RRL fwd. #2 – 5 uM; RRL rev. – 10 um). The results of the amplification using the diluted primer mix are summarized in Figure 5.12 below and Table A2.3 in Appendix II.

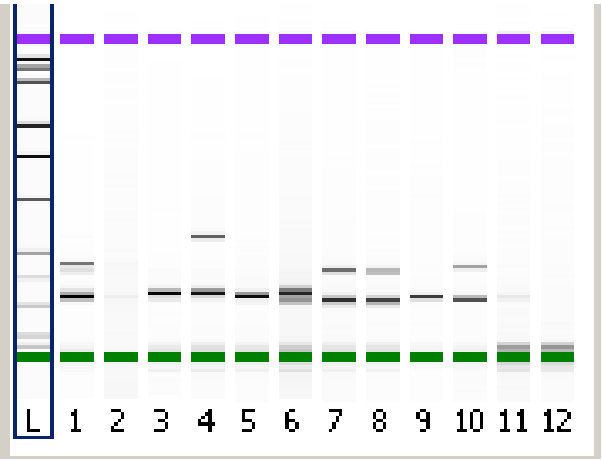


Figure 5.12. Gel image of nine templates amplified using modified primer mix diluted 1:2. L = ladder; 1 = all nine templates; 2 = all nine templates; 3 = *B. cereus* template; 4 = *B. anthracis* template; 5 = *C. difficile* template; 6 = *C. botulinum* template; 7 = *F. tularensis holarctica* template; 8 = *F. tularensis tularensis* template; 9 = *Y. enterocolitica* template; 10 = *Y. pestis* template; 11 = *M. tuberculosis* template; 12 = negative control.

As in the other assays performed in the primer optimization study, the diluted primer mix was used to simultaneously amplify all eleven target sequences (see lanes 1 and 2). There was a universal sample drop-out observed in lane 2, and as expected, many of the amplicons in lane 1 had overlapping size ranges that prevented the resolution of each individual target sequence.

In lane 3 (*B. cereus* template), the Bac/RRL target sequence (111 bp) was detected as the 122 bp fragment. The Bac/RRL target sequence was also detected in lane

4 (*B. anthracis* template) as the 122 bp fragment. The Bant/TOX target sequence (230 bp) was detected in lane 4 as the 228 bp fragment.

The Cdif/RRL target sequence (110 bp) was detected in lane 5 (*C. difficile* template) as the 117 bp fragment. The Cbot/RRL target sequence (109 bp) was detected in lane 6 (*C. botulinum* template) as the 120 bp fragment. Also detected in lane 6 was a 107 bp fragment corresponding to the 100 bp Cbot/TOX target.

In lane 7 (*F. tularensis holarctica* template), the Fran/RRL target sequence (102 bp) was detected as the 109 bp fragment. A 164 bp fragment was also observed in lane 7, which most likely corresponded to the non-specific amplification of the Ftul/TOX target sequence when *F. tularensis holarctica* DNA was used as template. This observation was consistent with other results observed in the monoplex and multiplex PCR studies.

The Fran/RRL target sequence was detected as the 108 bp fragment in lane 8 (*F. tularensis tularensis* template). Also detected in this lane was a 162 bp fragment corresponding to the 159 bp Ftul/TOX target sequence.

In lane 9 (*Y. enterocolitica* template), the Yent/RRL target sequence (107 bp) was detected as the 114 bp fragment. The Ypes/RRL target sequence (107 bp) was detected in lane 10 (*Y. pestis* template) as the 170 bp fragment. Also detected in lane 10 was a 170 bp fragment corresponding to the Ypes/TOX target sequence (169 bp).

In lane 11 (*M. tuberculosis* template), the Myc/RRL target sequence (110 bp) was not officially detected, although a faint band was visible on the gel image.

The results demonstrated that using a diluted version of the modified primer mix reduced the frequency and concentration of primer dimers (see bands just above the green lower marker in lanes 1-12, Figure 5.12). Although the concentration of the target

sequences was also reduced, the concentrations were still within an acceptable range for input into the Luminex assay.

4. Multiplex TD-PCR – Template Concentration Study

The multiplex PCR primer optimization study indicated that the diluted, modified primer mix was the most capable of amplifying the targets within an acceptable concentration range while still minimizing primer dimers. Although having the appropriate concentration and ratio of primers is important for amplification efficiency, templates that are too plentiful can detract from optimum performance. The PCR reactions performed in the primer optimization study all used 100 ng of template. The amplification using the diluted primer mix was repeated, using only 10 ng of template. A summary of the results is provided in Figure 5.13 below and Table A2.4 in Appendix II.

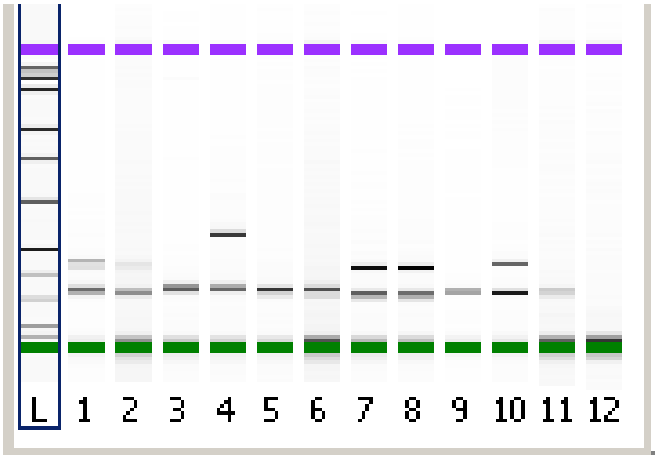


Figure 5.13. Gel image of amplicons produced using 10 ng of template and modified primer mix diluted 1:2. L = ladder; 1 = all nine templates; 2 = all nine templates; 3 = *B. cereus* template; 4 = *B. anthracis* template; 5 = *C. difficile* template; 6 = *C. botulinum* template; 7 = *F. tularensis holarctica* template; 8 = *F. tularensis tularensis* template; 9 = *Y. enterocolitica* template; 10 = *Y. pestis* template; 11 = *M. tuberculosis* template; 12 = negative control.

As in the assays performed in the primer optimization study, the diluted primer mix was used to simultaneously amplify all eleven target sequences (see lanes 1 and 2).

As expected, many of the amplicons had overlapping size ranges that prevented the resolution of each individual target sequence.

In lane 3 (*B. cereus* template), the Bac/RRL target sequence (111 bp) was detected as a combination of the 109 and 120 bp fragments. The Bac/RRL target sequence was also detected in lane 4 (*B. anthracis* template) as the 120 bp fragment. The Bant/TOX target sequence (230 bp) was detected as the 231 bp fragment in lane 4.

The Cdif/RRL target sequence (110 bp) was detected as the 116 bp fragment in lane 5 (*C. difficile* template). The Cbot/RRL target sequence (109 bp) was detected in lane 6 (*C. botulinum* template) as the 119 bp fragment. Also in lane 6, the Cbot/TOX target sequence (100 bp) was detected as the 107 bp fragment.

In lane 7 (*F. tularensis holarctica* template), the Fran/RRL target sequence (102 bp) was detected as the 107 bp fragment. A 163 bp fragment was also observed in lane 7, which most likely corresponded to the non-specific amplification of the Ftul/TOX target sequence when *F. tularensis holarctica* DNA was used as template. This observation was consistent with other results observed in the monoplex and multiplex PCR studies.

The Fran/RRL target sequence (102 bp) was detected in lane 8 (*F. tularensis tularensis* template) as a combination of the 97 and 107 bp fragments. Also in lane 8, the Ftul/TOX target sequence (159 bp) was detected as the 162 bp fragment.

In lane 9 (*Y. enterocolitica* template), the Yent/RRL target sequence (107 bp) was detected as the 114 bp fragment. The Ypes/RRL target sequence (107 bp) was detected in lane 10 (*Y. pestis* template) as the 112 bp fragment. Also detected in lane 10 was a 170 bp fragment corresponding to the Ypes/TOX target sequence (169 bp).

The Myc/RRL target sequence (110 bp) was detected in lane 11 (*M. tuberculosis* template) as the 114 bp fragment.

Amplifying the reduced template (10 ng) with the 1:2 diluted primer mix yielded results that were similar to those that were observed when 100 ng of template was amplified with the same diluted primer mix. The parameters of all subsequent PCR reactions were modified with the conclusions of the template concentration study in mind.

D. Titration Study

Before the target DNA was introduced to the beads, the attachment of the probes to the beads was first confirmed. This was accomplished by designing and purchasing biotinylated reverse probes which were complementary to each probe sequence. If hybridization between the probe and reverse probe was detected, this was indicative of successful probe-bead coupling. Various increments of reverse probe were used not only to verify successful probe-bead coupling, but also to enable an extrapolation of the approximate number of probes per bead. Raw data and plots of MFI vs. Reverse Probe (fmol) for each bead set are shown in Figures A4.1-4.11 in Appendix IV.

The results of the titration study indicated that the curve platueaued at approximately 50 fmol of reverse probe for each of the eleven bead sets studied. This value was first converted to moles of reverse probe and then to molecules of reverse probe, using Avogadro's number. The molecules of reverse probe were divided by ten to account for the ten-fold excess of reverse probe added to each sample to drive the reaction to completion (C. Carrasco, personal communication, 12 October 2009). This yielded an estimate of the molecules of probe in each reaction well, based on the 1:1 ratio

of probe bound to reverse probe. However, each reaction well had approximately 5000 beads, so the molecules of probe was divided by 5000 to ultimately provide the number of probes attached to each bead. Assuming that each titration curve pleateued around 50 fmol, it was concluded that each bead had approximately 608,282 probes attached. The conversion of reverse probe to an estimation of the number of probes per bead is summarized in Figure 5.14. A better baseline for comparing MFI values arising from the eleven bead sets was established after it was confirmed that each bead set consisted of beads with roughly the same number of probes attached.

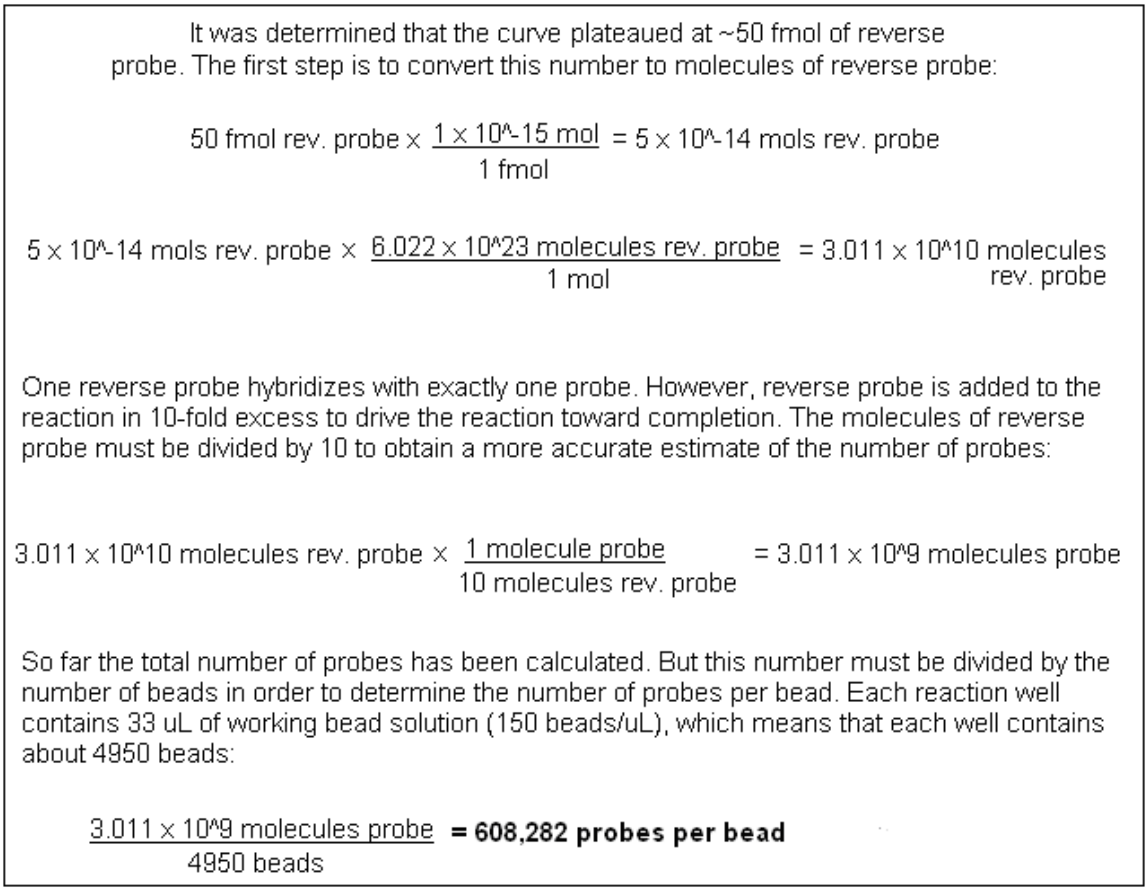


Figure 5.14. Sample calculations for determining number of probes per bead.

E. Enzymatic Degradation of the Competing PCR Strand

1. Enzymatic Degradation of B. cereus

Before the enzymatic degradation protocol was universally applied, its effectiveness was first studied using just one bacterium: *B. cereus*. This bacterium was chosen because prior assays demonstrated that amplification of the Bac/RRL target sequence yielded PCR products that were generally free of non-specific amplicons and of an ideal concentration.

B. cereus control and experimental samples were amplified in triplicate. The pre-digest quantitation results (Figure A5.1 and Table A5.1 in Appendix V) for all six samples displayed a band having a length of approximately 120 bp, which was in close proximity to the length of the Bac/RRL target sequence (111 bp). The quantitation results also indicated a concentration of amplicons in all samples that was sufficient to proceed with the enzymatic digestion. The average concentration of the Bac/RRL target was 11.2 ng/uL.

The enzymatic digestion on the samples described above was carried out once a sufficient amount of pre-digest DNA was verified. Figure A5.2 and Table A5.2 show the post-digest quantitation results of all samples. Once again a band of approximately 120 bp was observed in all samples. The average concentration among the control samples was 6.49 ng/uL, which was approximately half the concentration observed in the same samples pre-digestion (11.2 ng/uL). This reduction in concentration was expected because of the 1:2 dilution inherent to the digestion protocol, as described previously. With this dilution in mind, the concentration of the original PCR products and the pre-digest control samples were approximately equal.

Although a band having an approximate length of 120 bp was observed in all experimental samples, the concentration was too low to be reported. As discussed previously, the Agilent DNA 1000 kit was not designed to quantify single-stranded DNA. The post-digest quantitation of experimental samples was conducted for qualitative verification of digested DNA, not to provide a quantitative estimate of the amount of digested DNA. Therefore, the lack of quantitative data for post-digestion experimental samples did not preclude the Luminex analysis.

One control (7.21 ng/uL) and one experimental (concentration too low to be reported) sample from those in Table A5.2 (lanes 2 and 6; highlighted samples) were chosen for Luminex analysis. The basis for choosing the control sample in lane 2 was that its band had a concentration higher than the band in either of the other two control samples in lanes 1 and 3. The basis for choosing the experimental sample in lane 6 was that the experimental samples in lanes 4 and 5 had been exhausted in preceding trial runs (data not shown). Also analyzed on the Luminex were original PCR products from which the control and experimental samples originated (data not shown). The MFI values observed for these samples are summarized in Table 5.5. Figures A5.5 and A5.6 in Appendix V compare the instrument's sensitivity when double-stranded (undigested) and single-stranded (digested) target DNA were analyzed.

Sample	Bac/RRL
NC	71
NC × 3	213
Net MFI*	
BC/con/5ng	128
BC/exp/5ng	388
BC/con/10ng	239
BC/exp/10ng	507
BC/con/20ng	348
BC/exp/20ng	1010
BC/con/40ng	501
BC/exp/40ng	1493
BC/con/80ng	748
BC/exp/80ng	2039
BC – PC	8911

Table 5.5. Summary of net MFI values produced for digested *B. cereus* DNA. *Net MFI = Raw MFI – 3X NC for selected bead set.

2. Enzymatic Degradation of B. anthracis and Y. pestis

Enzymatic digestion of amplicons from the non-pathogenic *B. cereus* produced favorable results. Two pathogenic bacteria (*B. anthracis* and *Y. pestis*, respectively) were studied next in order to determine if the same success observed with the enzymatic degradation of RRL target sequences would be observed with TOX target sequences.

As shown in Figure A5.3 and Table A5.3 in Appendix V, all samples for *B. anthracis* and *Y. pestis* were amplified in duplicate. The pre-digest quantitation results for the four *B. anthracis* samples all displayed a band having an approximate length of 118 bp, which was similar to the length of the Bac/RRL target sequence (111 bp). The four *B. anthracis* samples also displayed a band having an approximate length of 237 bp, which was probably from the 230 bp Bant/TOX target sequence. The pre-digest quantitation results for the four *Y. pestis* samples also displayed two bands each. The first was a band having an approximate length of 111 bp, which corresponded to the 107 bp Ypes/RRL target sequence. The second was a band having an approximate length of 171 bp, which

corresponded to the 169 bp Ypes/TOX target sequence. The quantitation results also indicated a concentration of amplicons in all samples that was sufficient to proceed with the enzymatic digestion. The average concentration of the Bac/RRL and Bant/TOX targets were 11.3 ng/uL and 8.3 ng/uL, respectively. The average concentration of the Ypes/RRL and Ypes/TOX targets were 16.3 ng/uL and 9.7 ng/uL, respectively.

The enzymatic digestion on the samples above was carried out once a sufficient amount of pre-digest DNA was verified. Figure A5.4 and Table A5.4 show the post-digest quantitation results of all samples. Two bands (approximately 120 and 230 bp in length) were again observed in all *B. anthracis* samples. The average concentrations of the Bac/RRL target and Bant/TOX target in the two *B. anthracis* control samples were 6.2 ng/uL and 5.3 ng/uL, respectively. The average concentrations of the Ypes/RRL target and Ypes/TOX target in the two *Y. pestis* control samples were 10.0 ng/uL and 5.3 ng/uL, respectively. All post-digest concentrations in the control tubes were roughly half the concentration observed in pre-digest samples, because of the 1:2 dilution involved in the enzymatic degradation protocol. With this dilution in mind, the concentration of the original PCR products and the pre-digest control samples were approximately equal.

Both target sequences were detected in all four experimental samples, except one. Although faintly visible on the gel, the Bac/RRL band was not concentrated enough to be reported (lane 3). Similar observations were made for experimental samples in other assays, but this did not eliminate these samples from Luminex analysis. The faint band observed on the gel confirmed that some DNA was present. Furthermore, the fact that concentrations were not reported for experimental samples was somewhat expected

because the Agilent DNA 1000 kit was not optimized for quantifying single-stranded DNA.

One control (7.81 ng/uL) and one experimental (2.95 ng/uL) sample from the *B. anthracis* samples in Table A5.4 (highlighted samples) were chosen for Luminex analysis. The MFI values observed for these samples are summarized in Table 5.6. Figures A5.7 and A5.8 in Appendix V compare the fluorescent response of double-stranded and single-stranded target.

Sample	Bac/RRL	Bant/TOX
NC	59	57
NC × 3	177	171
Net MFI*		
BA/con/5ng	149	825
BA/exp/5ng	410	1843
BA/con/10ng	241	1528
BA/exp/10ng	727	2493
BA/con/20ng	370	2031
BA/exp/20ng	1095	3121
BA/con/40ng	509	2254
BA/exp/40ng	1478	3487
BA/con/80ng	569	2255
BA/exp/80ng	2216	4174
BA – PC	2209	3192

Table 5.6. Summary of net MFI values produced for digested *B. anthracis* DNA. *Net MFI = Raw MFI – 3X NC for selected bead set.

One control (10.47 ng/uL) and one experimental (4.02 ng/uL) sample from the post-digest *Y. pestis* samples in Table A5.4 (highlighted samples) were chosen for Luminex analysis. An internal obstruction in the Luminex instrument caused aberrant MFI values when these samples were first analyzed (data not shown). After the blockage was cleared, the analysis was successfully repeated. The MFI values observed for these samples are summarized in Table 5.7. Figures A5.9 and A5.10 in Appendix V compare the fluorescent response of double-stranded and single-stranded target.

Sample	Ypes/RRL	Ypes/TOX
NC	96	348
NC × 3	288	1044
Net MFI*		
YP/con/5ng	0	1110
YP/exp/5ng	0	1037
YP/con/10ng	0	1139
YP/exp/10ng	0	1885
YP/con/20ng	0	1528
YP/exp/20ng	0	2676
YP/con/40ng	0	2167
YP/exp/40ng	142	3204
YP/con/80ng	0	3147
YP/exp/80ng	325	3522
YP – PC	4387	6292

Table 5.7. Summary of net MFI values produced for digested *Y. pestis* DNA. *Net MFI = Raw MFI – 3X NC for selected bead set. Net MFIs ≤ 0 reported as "0".

F. Specificity/Sensitivity Study

1. Sample Preparation

The results of the enzymatic degradation study demonstrated that single-stranded target DNA yielded a stronger fluorescent response than double-stranded DNA. With this conclusion in mind, the samples used in the specificity/sensitivity study were enzymatically digested prior to analysis on the Luminex instrument.

Pre-digested samples for the specificity/sensitivity study did not all come from the same amplification batch. Many amplifications were performed and pre-digested samples from the most successful amplifications (in which there was no evidence of superfluous bands, contamination, drop-out, etc.) were pooled and quantified as a group. The contents of each sample tube were split equally between a control and experimental tube. The appropriate treatment was administered and the samples were quantified in separate

control (Figure A6.1 and Table A6.1 in Appendix VI) and experimental (Figure A6.2 and Table A6.2 in Appendix VI) batches.

As described in the Methodology section, the concentration of the RRL band in all control tubes was used to determine the amount of DNA from the corresponding experimental tube to be loaded on to the Luminex plate. A concentration of 0.05 ng/uL was assumed for the RRL band of any control sample (*e.g.*, *M. tuberculosis*) whose target DNA was visible on the gel image but had a concentration too low to be reported. The basis of this assumption was that the LLD on the Agilent is ~0.02 ng/uL and the presence of a faint band was indicative of a concentration at or slightly above the LLD.

Net MFI values for the Specificity/Sensitivity Study can be found in the Figures A6.3-6.11 and Tables A6.3-6.12 in Appendix VI. Representative charts of excellent, satisfactory, and poor results are shown below in Figures 5.15-5.17 (see Appendix VI for all figures pertaining to this study).

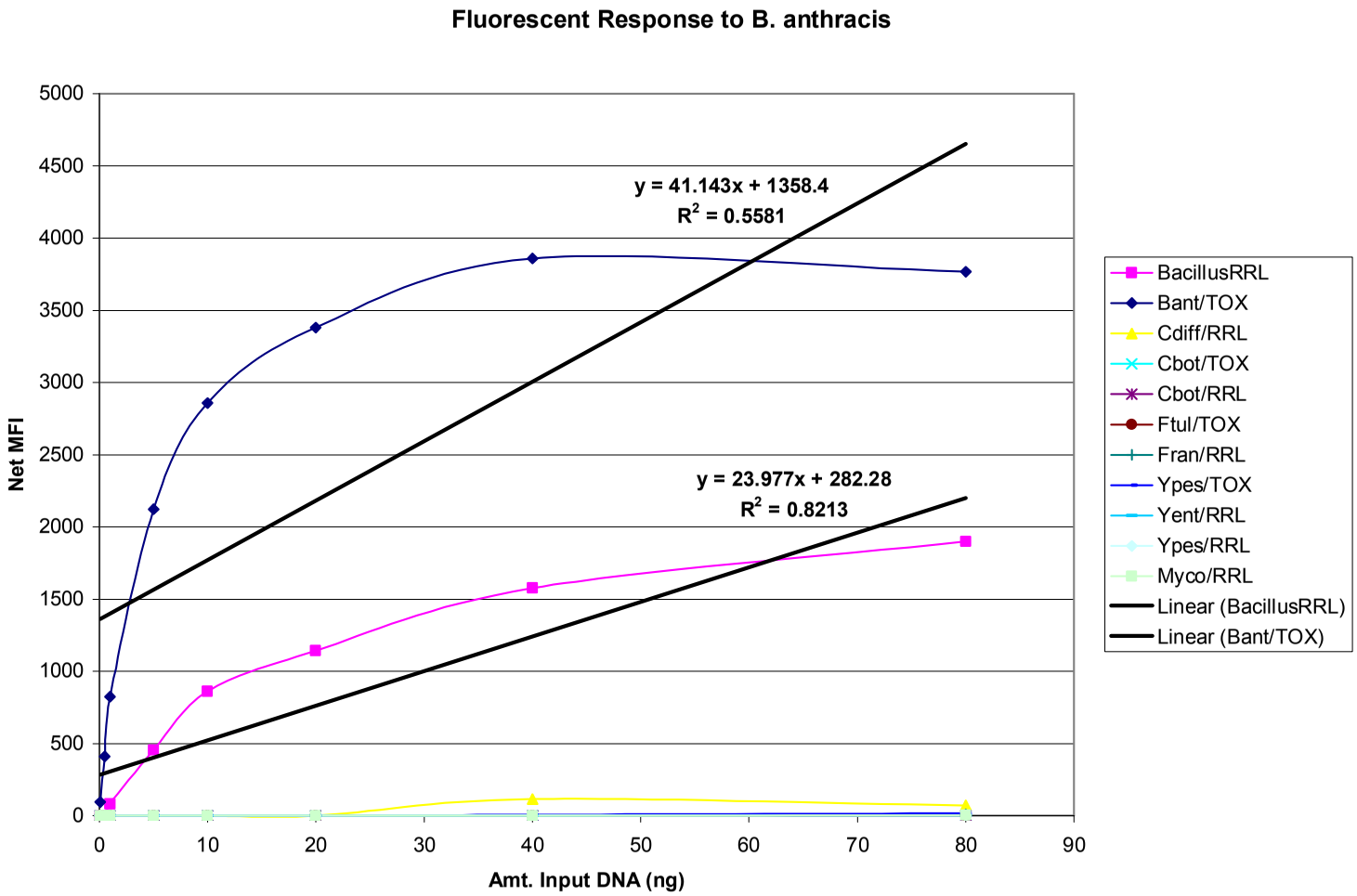


Figure 5.15. Fluorescent response of eleven probes to *B. anthracis* target sequence. This is an example of excellent results for this study. Notice the superior specificity and sensitivity for the Bac/RRL and Bant/TOX probes. Fluorescent response of other probes (bottom of graph) is low as expected.

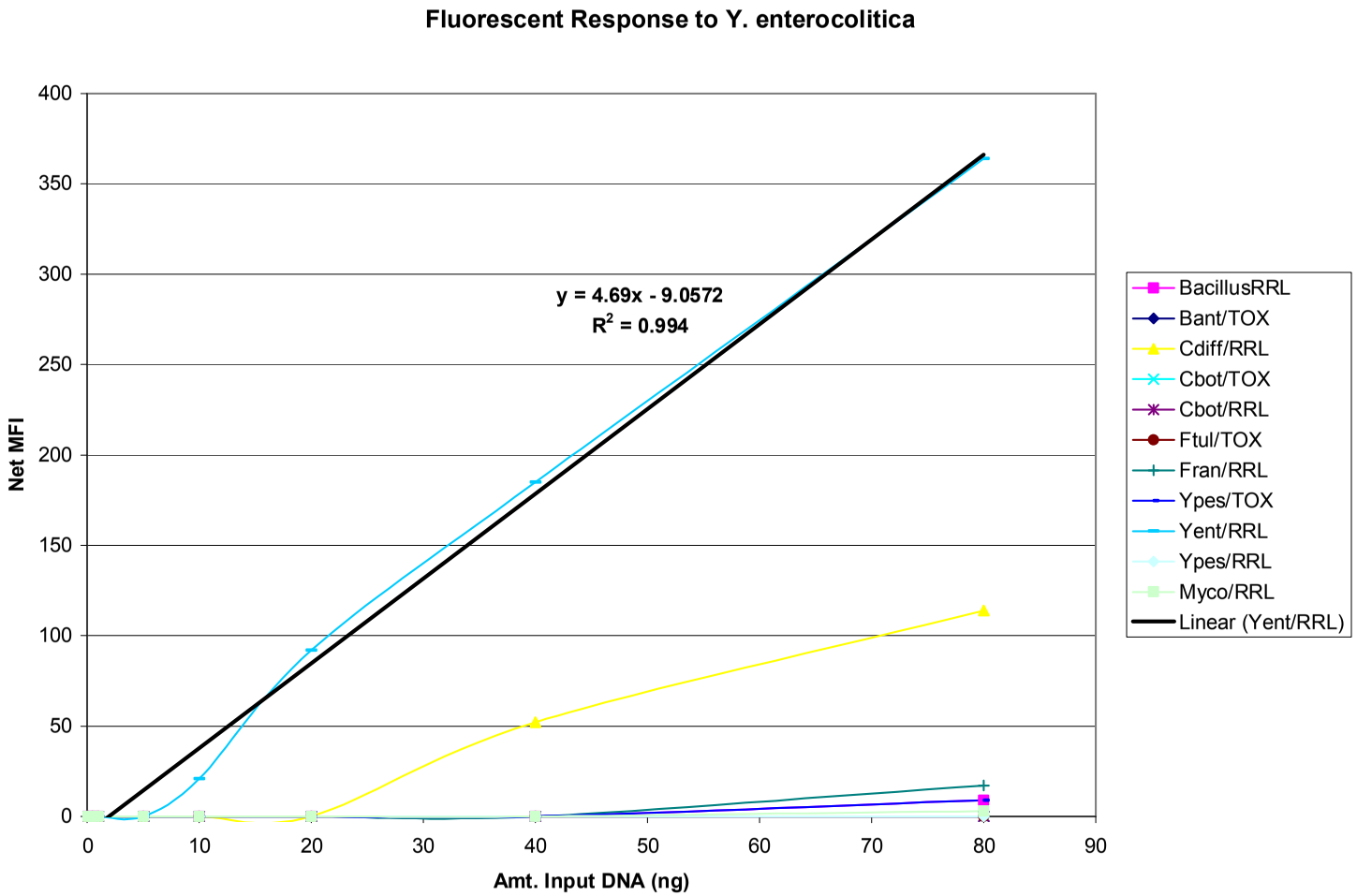


Figure 5.16. Fluorescent response of eleven probes to *Y. enterocolitica* target sequence. This is an example of satisfactory results for this study. Notice the superior specificity and sensitivity of the Yent/RRL probe. The fluorescent response of most other probes was weak, except for the Cdif/RRL probe, which yielded MFI values that could be difficult to distinguish from a genuine positive result.

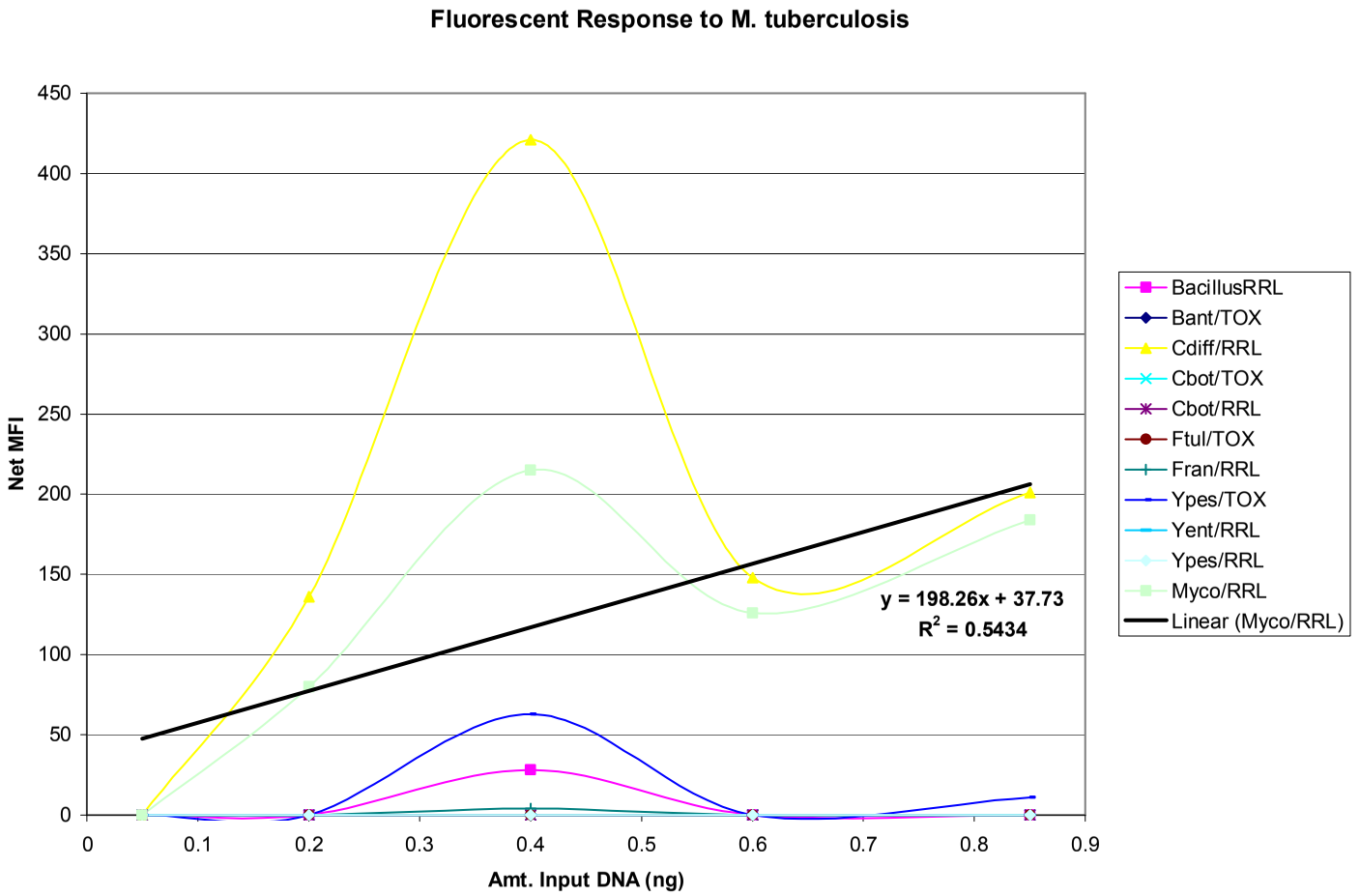


Figure 5.17. Fluorescent response of eleven probes to *M. tuberculosis* target sequence. This is an example of poor results for this study. The fluorescent response of the Myc/RRL probe is neither linear nor logarithmic. The MFI values for the Cdif/RRL probe surpass those for the Myc/RRL probe at almost every amount of input DNA. The MFI values for the remaining probes could also be very difficult to distinguish from a genuine positive result.

G. Mixture Study – 1:1

Three batches (see Table 4.17) of 1:1 mixtures were amplified, quantified, digested, and re-quantified. The post-digest quantitation results for all control samples in each batch are shown in Tables A7.1, A7.3, and A7.5 in Appendix VII. Guidelines for determining Luminex input amounts were followed as described in the Methodology section. The corresponding Luminex data is summarized in Tables A7.2, A7.4, and A7.6. The fluorescent response of all probes to each mixture is depicted graphically in Figures 5.18-5.50 below.

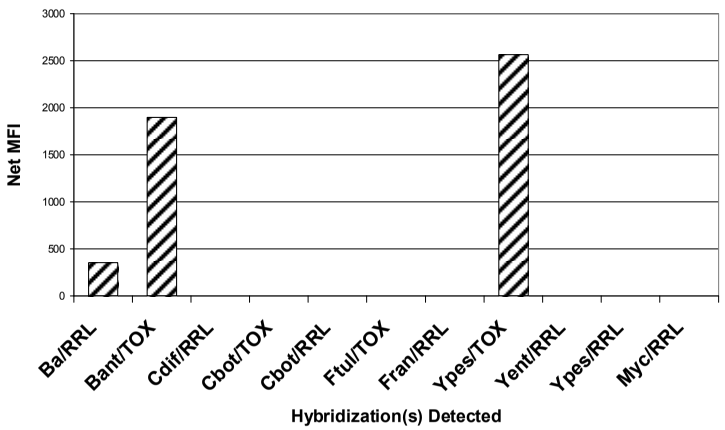


Figure 5.18. Fluorescent response to BA/YP mixture.

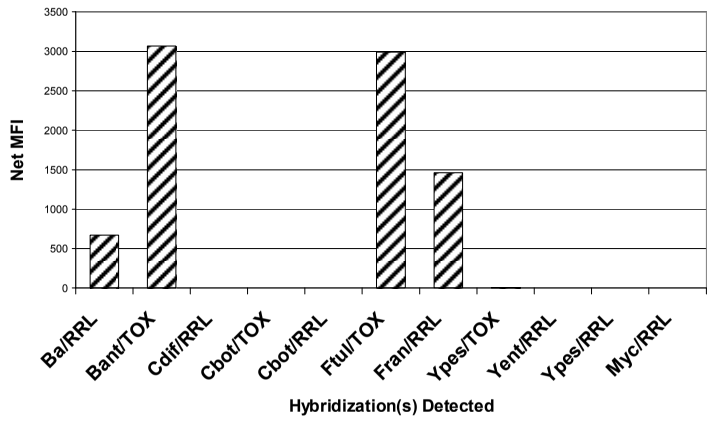


Figure 5.19. Fluorescent Response to BA/FT mixture.

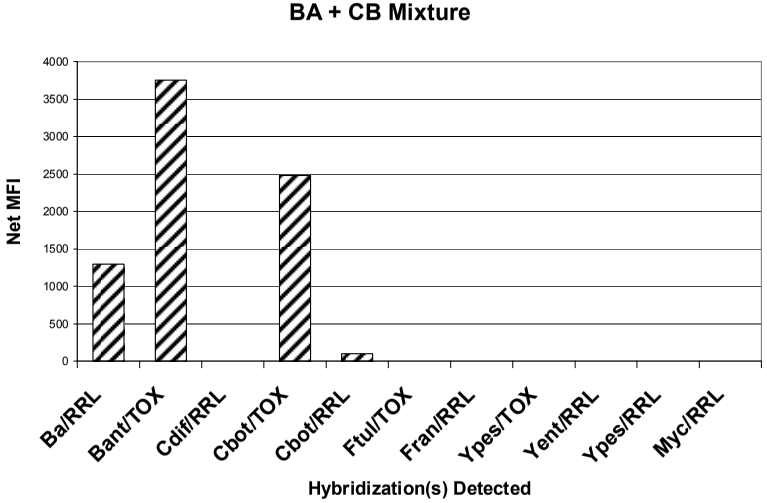


Figure 5.20. Fluorescent response to BA/CB mixture.

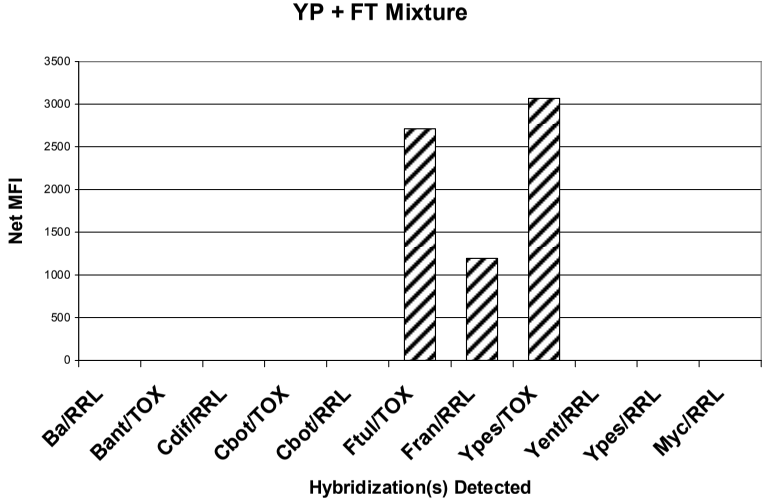


Figure 5.21. Fluorescent response to YP/FT mixture.

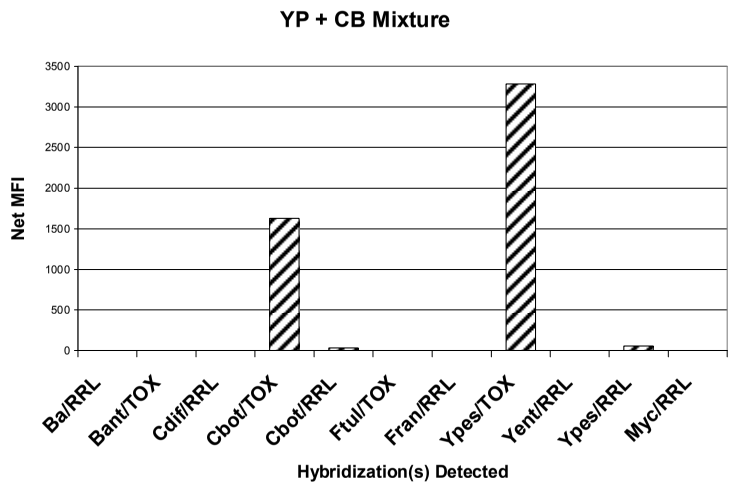


Figure 5.22. Fluorescent response to YP/CB mixture.

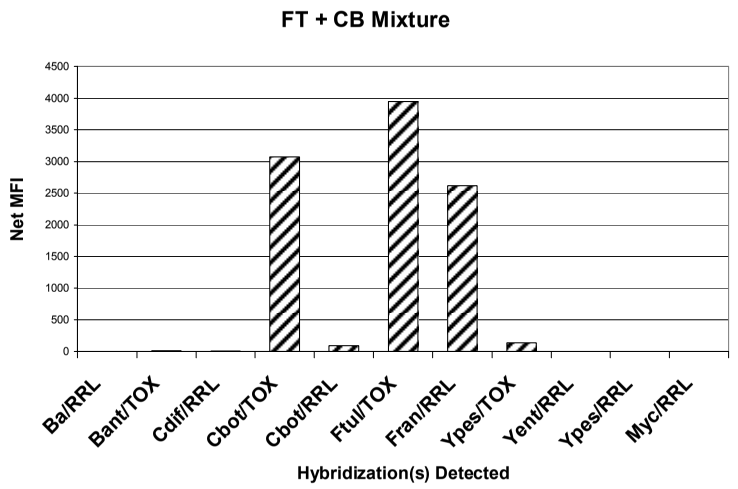


Figure 5.23. Fluorescent response to FT/CB mixture.

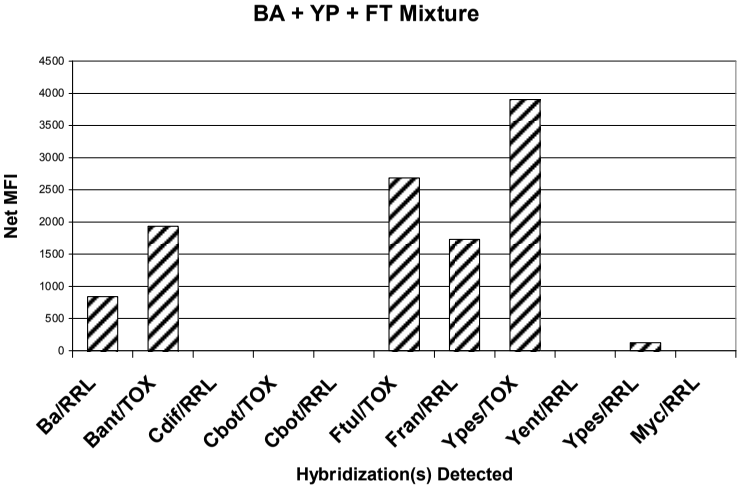


Figure 5.24. Fluorescent response to BA/YP/FT mixture.

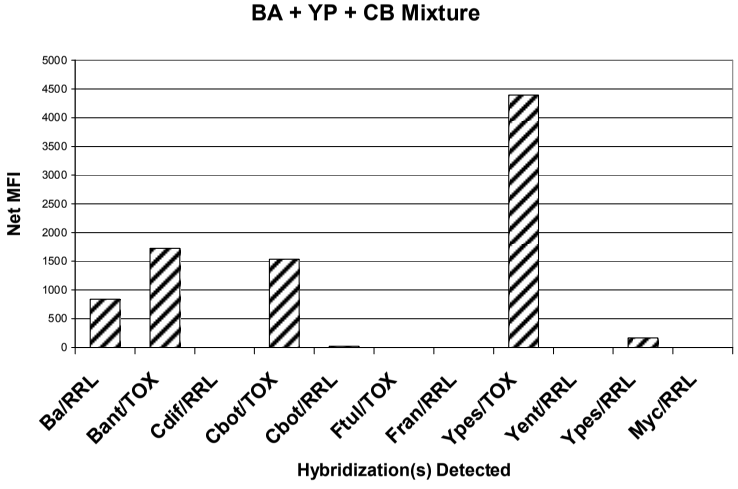


Figure 5.25. Fluorescent response to BA/YP/CB mixture.

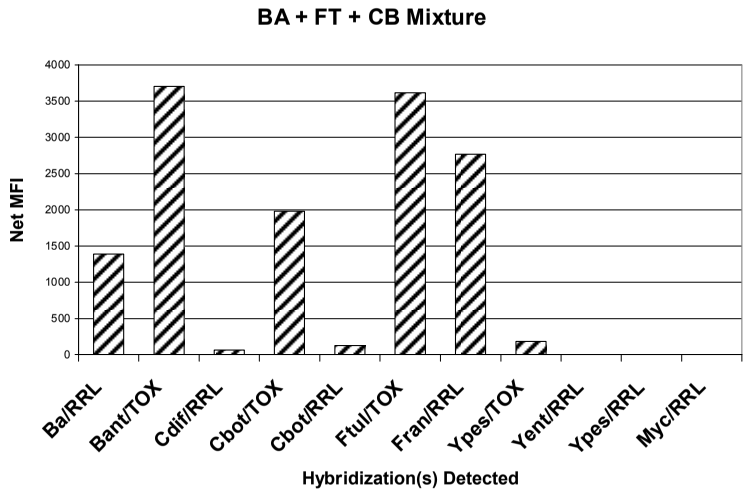


Figure 5.26. Fluorescent response to BA/FT/CB mixture.

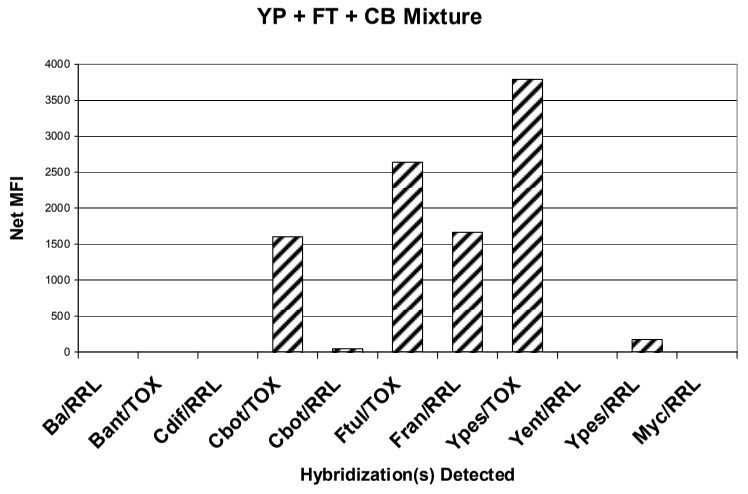


Figure 5.27. Fluorescent response to YP/FT/CB mixture.

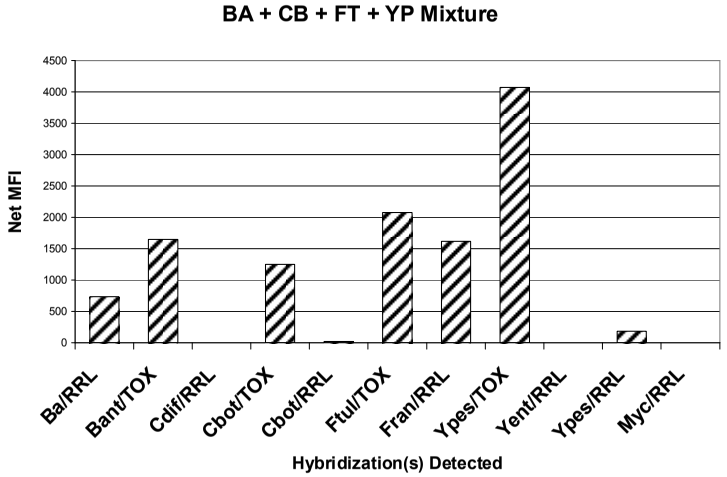


Figure 5.28. Fluorescent response to BA/CB/FT/YP mixture.

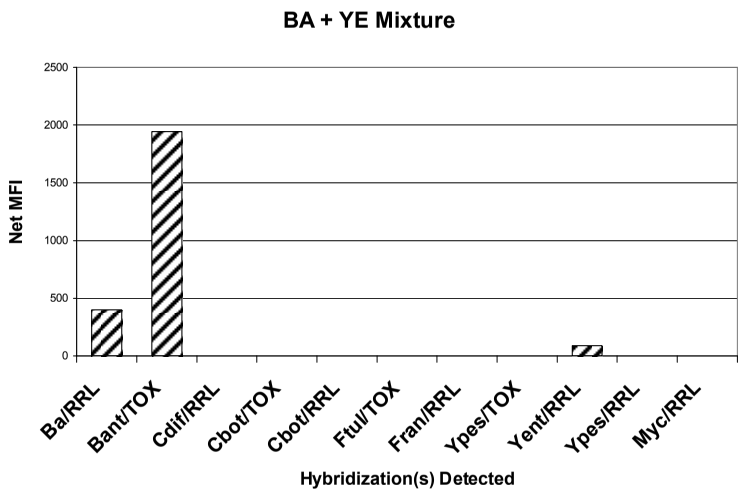


Figure 5.29. Fluorescent response to BA/YE mixture.

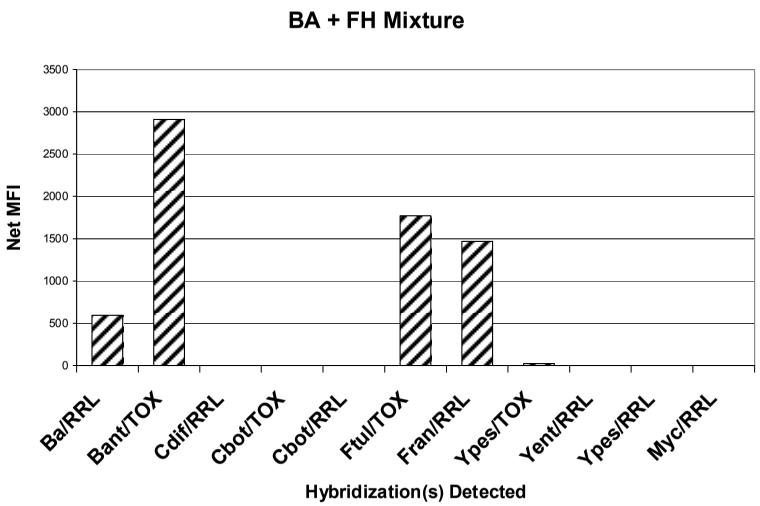


Figure 5.30. Fluorescent response to BA/FH mixture.

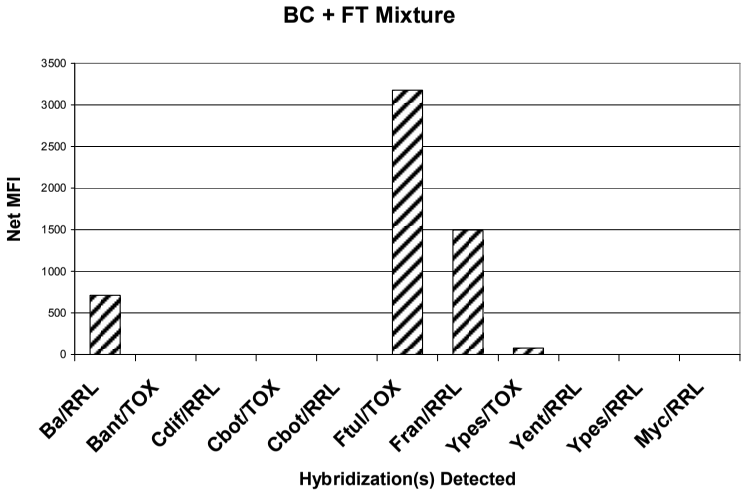


Figure 5.31. Fluorescent response to BC/FT mixture.

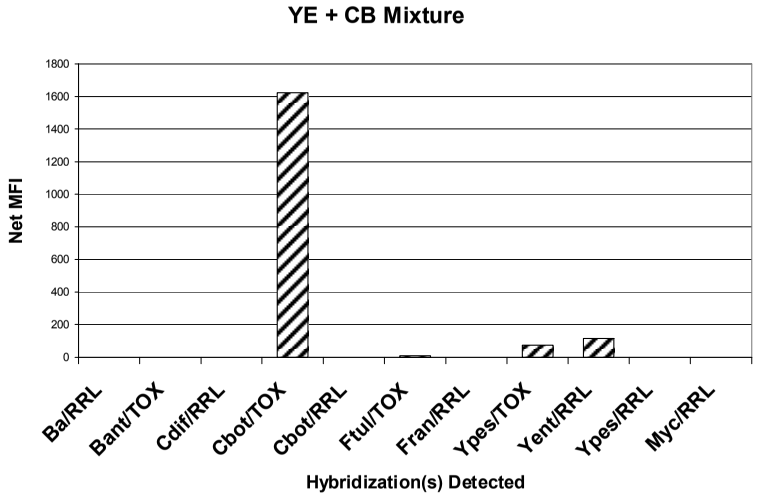


Figure 5.32. Fluorescent response to YE/CB mixture.

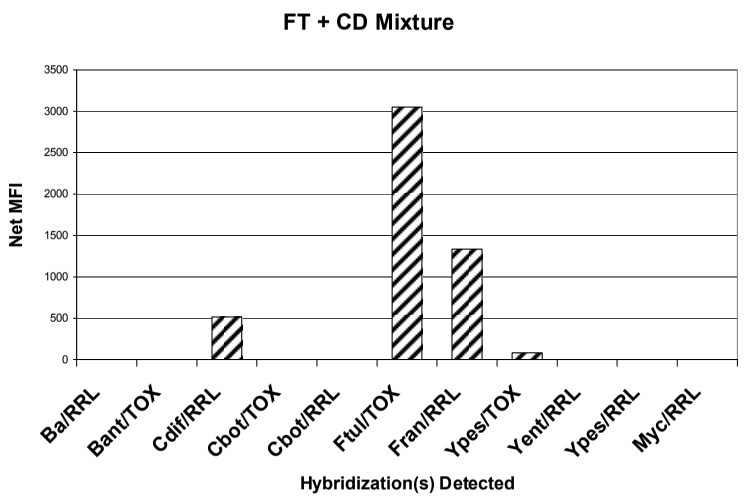


Figure 5.33. Fluorescent response to FT/CD mixture.

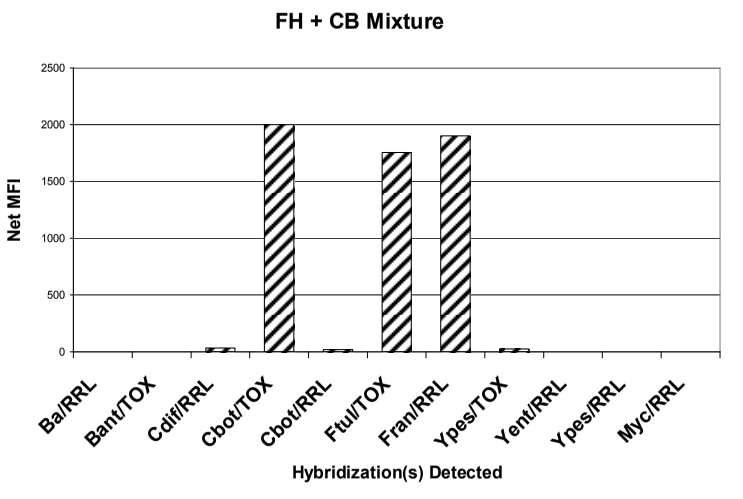


Figure 5.34. Fluorescent response to FH/CB mixture.

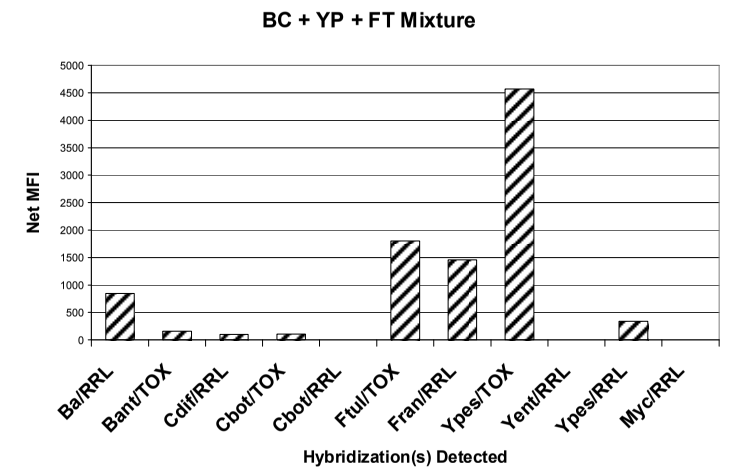


Figure 5.35. Fluorescent response to BC/YP/FT mixture.

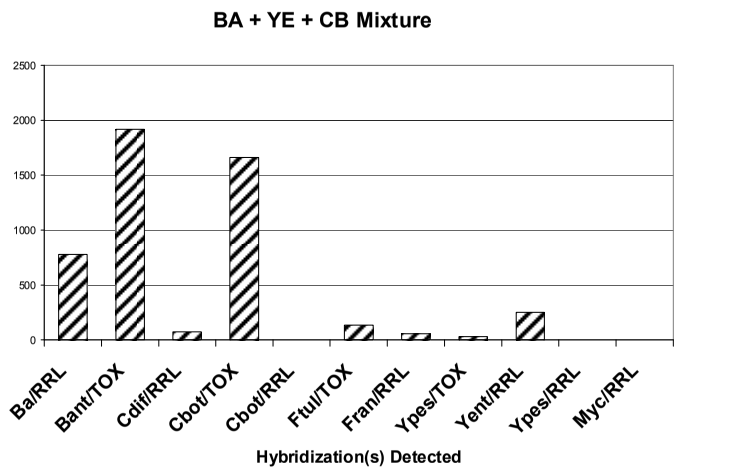


Figure 5.36. Fluorescent response to BA/YE/CB mixture.

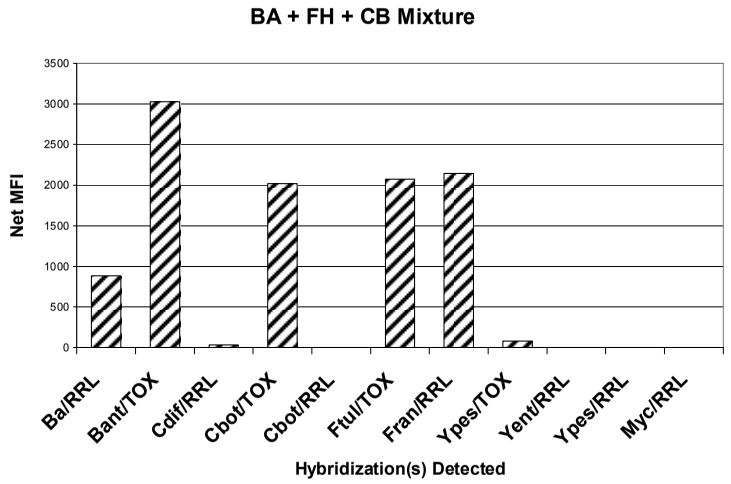


Figure 5.37. Fluorescent response to BA/FH/CB mixture.

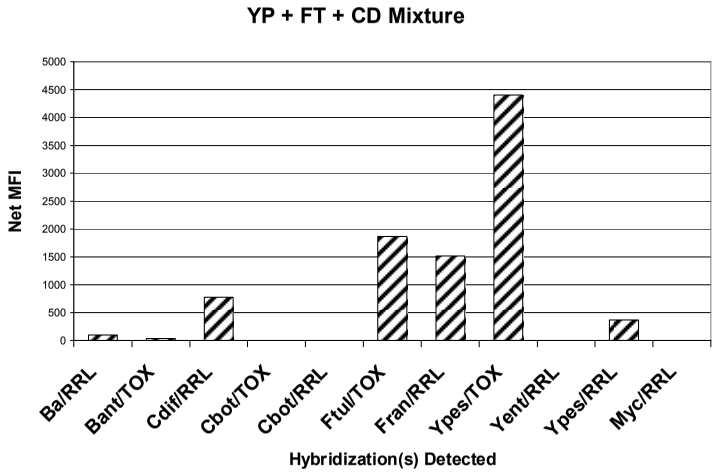


Figure 5.38. Fluorescent response to YP/FT/CD mixture.

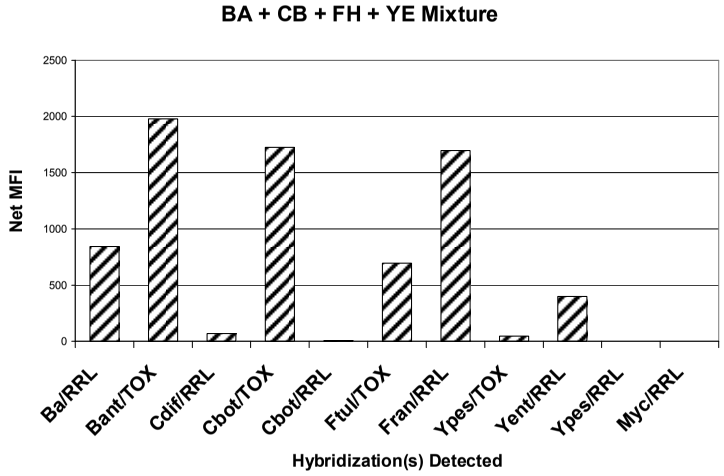


Figure 5.39. Fluorescent response to BA/CB/FH/YE mixture.

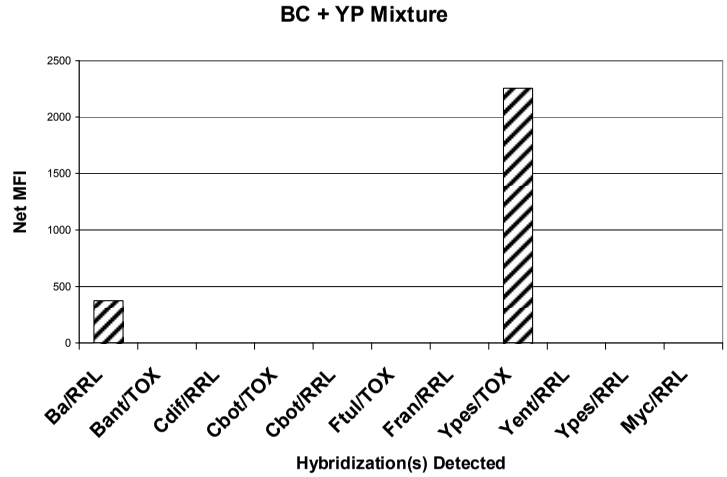


Figure 5.40. Fluorescent response to BC/YP mixture.

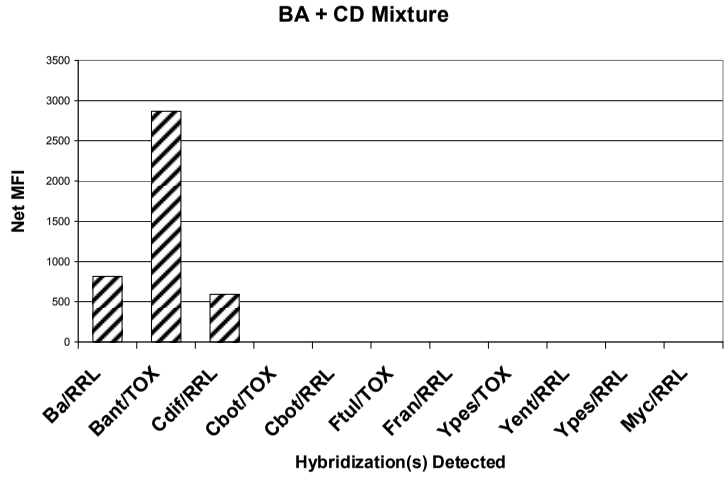


Figure 5.41. Fluorescent response to BA/CD mixture.

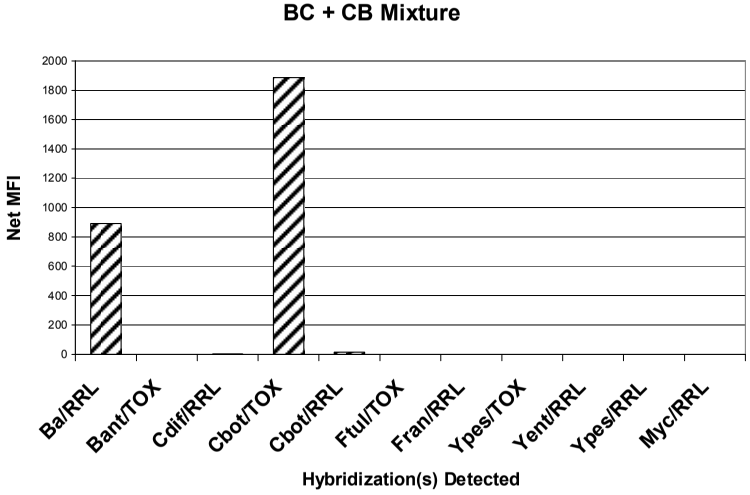


Figure 5.42. Fluorescent response to BC/CB mixture.

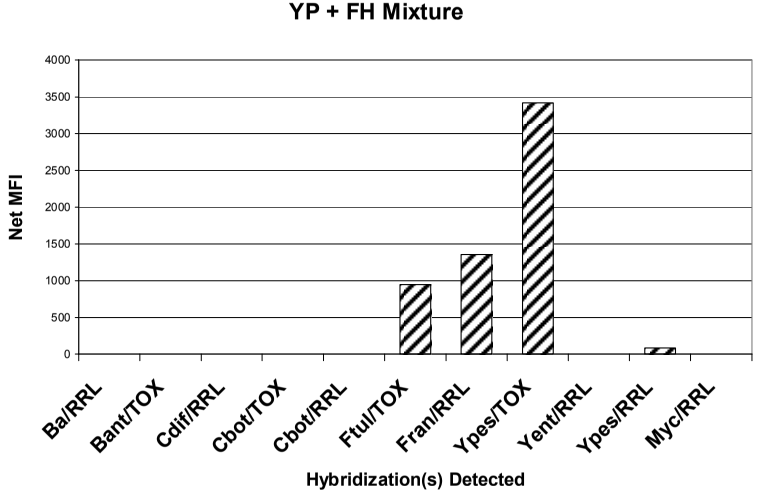


Figure 5.43. Fluorescent response to YP/FH mixture.

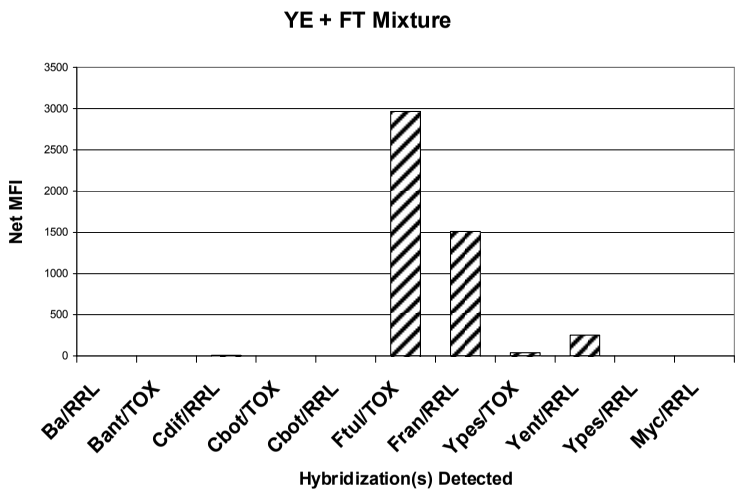


Figure 5.44. Fluorescent response to YE/FT mixture.

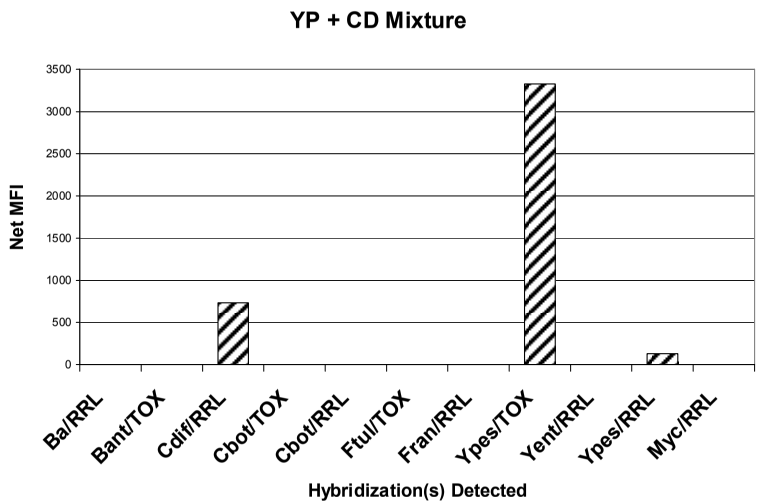


Figure 5.45. Fluorescent response to YP/CD mixture.

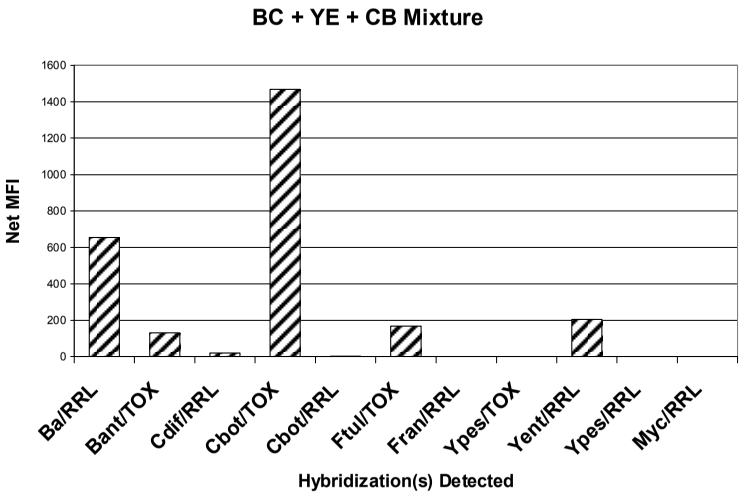


Figure 5.46. Fluorescent response to BC/YE/CB mixture.

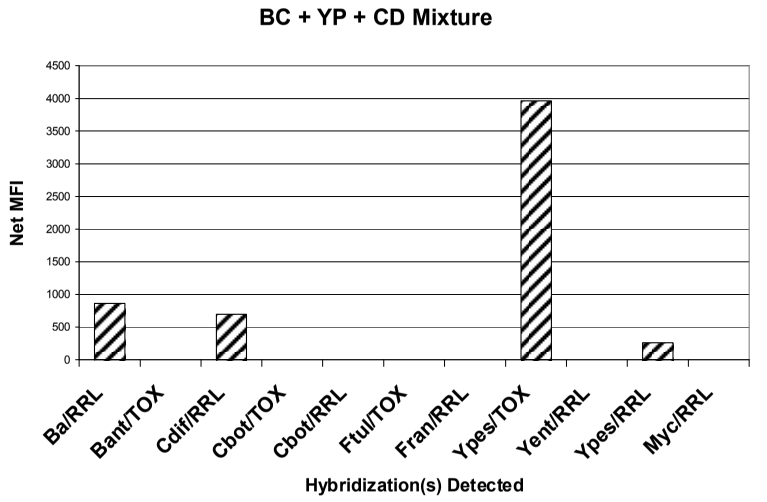


Figure 5.47. Fluorescent response to BC/YP/CD mixture.

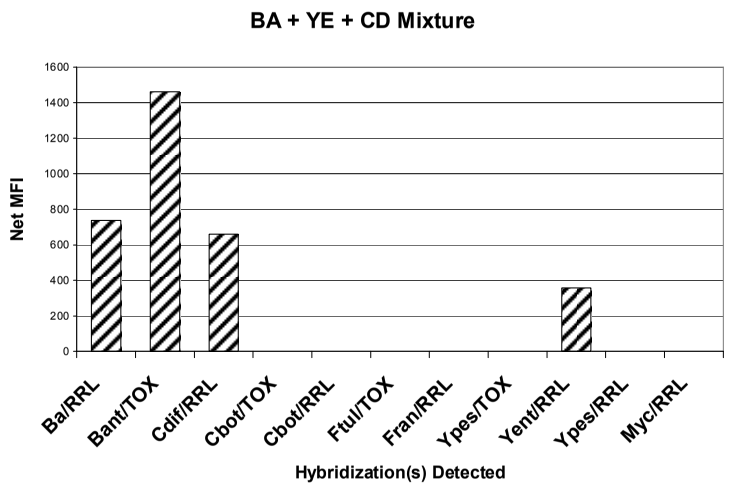


Figure 5.48. Fluorescent response to BA/YE/CD mixture.

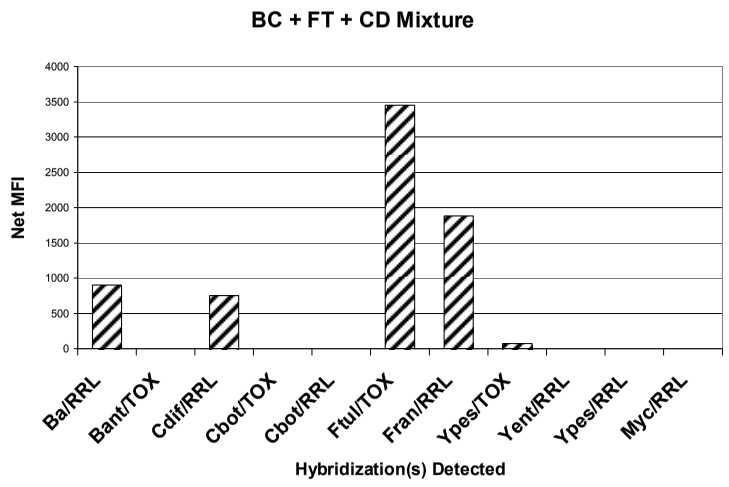


Figure 5.49. Fluorescent response to BC/FT/CD mixture.

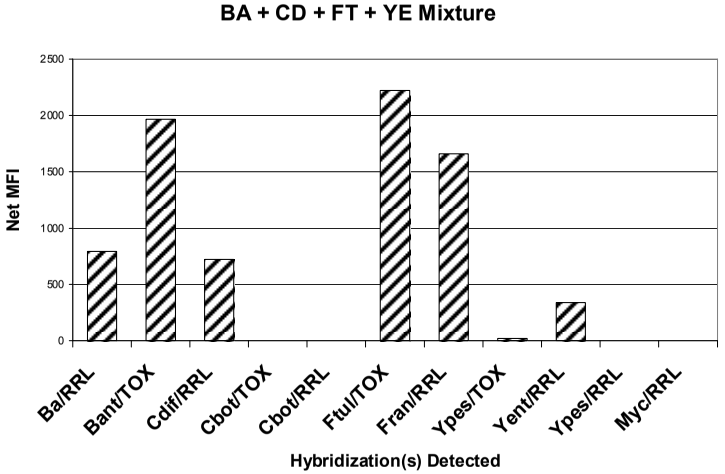


Figure 5.50. Fluorescent response to BA/CD/FT/YE mixture.

H. Intra-genus Mixture Study (1:1)

Mixtures studied in the 1:1 and ratio mixture studies (discussed below) were limited to mixtures in which each component belonged to a different genus. The specificity of the Luminex platform was further challenged by analyzing 1:1 binary mixtures in which both components belonged to the same genus.

One reaction tube containing template from *C. difficile* and *C. botulinum*, and another containing template from *Y. enterocolitica* and *Y. pestis*, were prepared and amplified according to the TD-PCR (mixtures of bacteria in the Bacillus or Francisella genera were not studied because they would not have been distinguishable due to the absence of a unique RRL probe for these bacteria). The PCR amplicons were split into a control group (undigested) and experimental group, with the experimental group being digested with lambda exonuclease as per the usual protocol. The post-digest quantitation results are shown in Table A8.1 in Appendix VIII. The standard rules for determining Luminex input amounts were applied. Table A8.2 provides a summary of net MFI values for each mixture analyzed in this assay. Figures 5.51 and 5.52 below illustrate the MFI results for this assay in bar graph format.

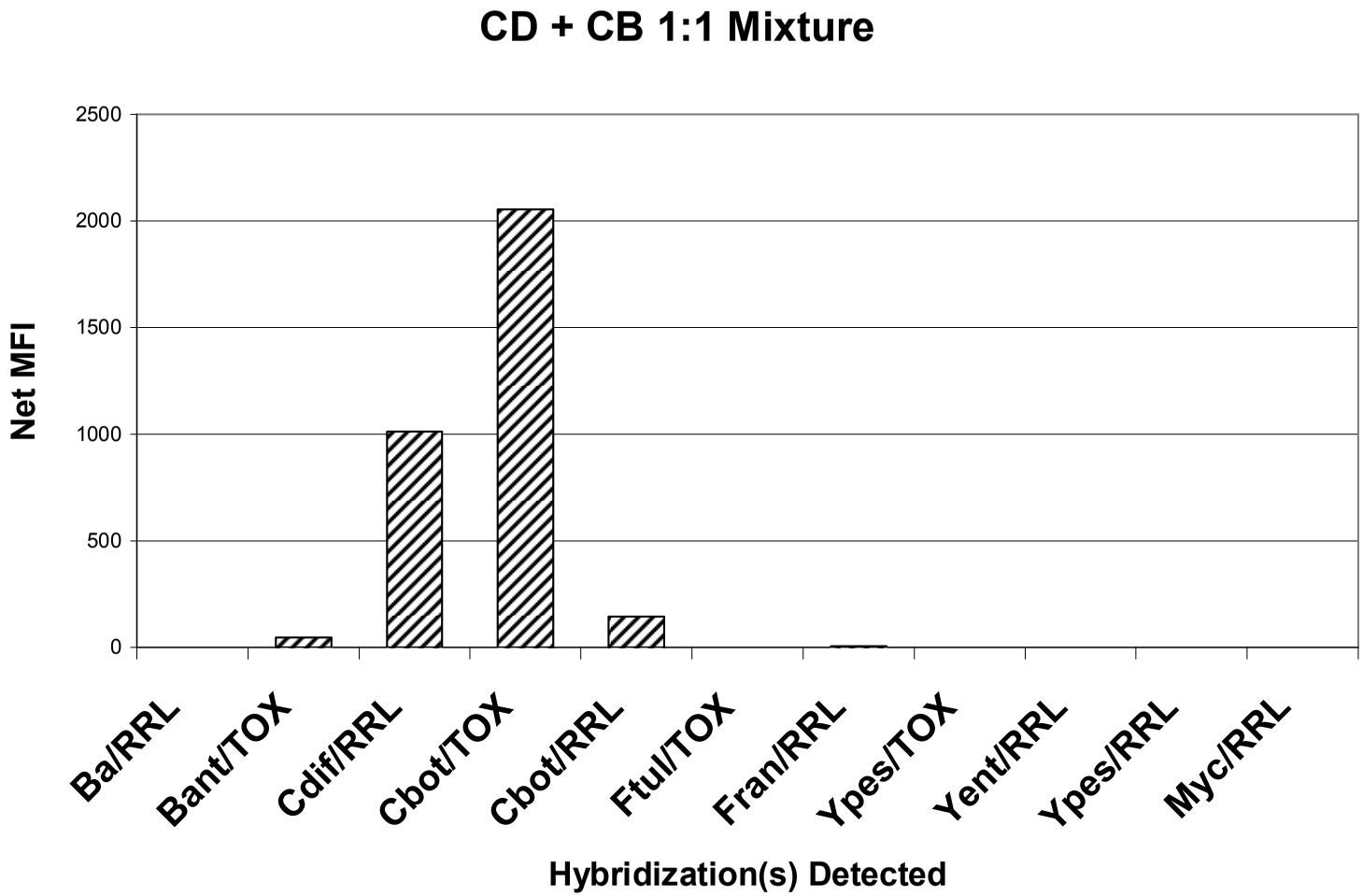


Figure 5.51. Fluorescent response of the eleven probes to a 1:1 mixture of *C. difficile* and *C. botulinum.*

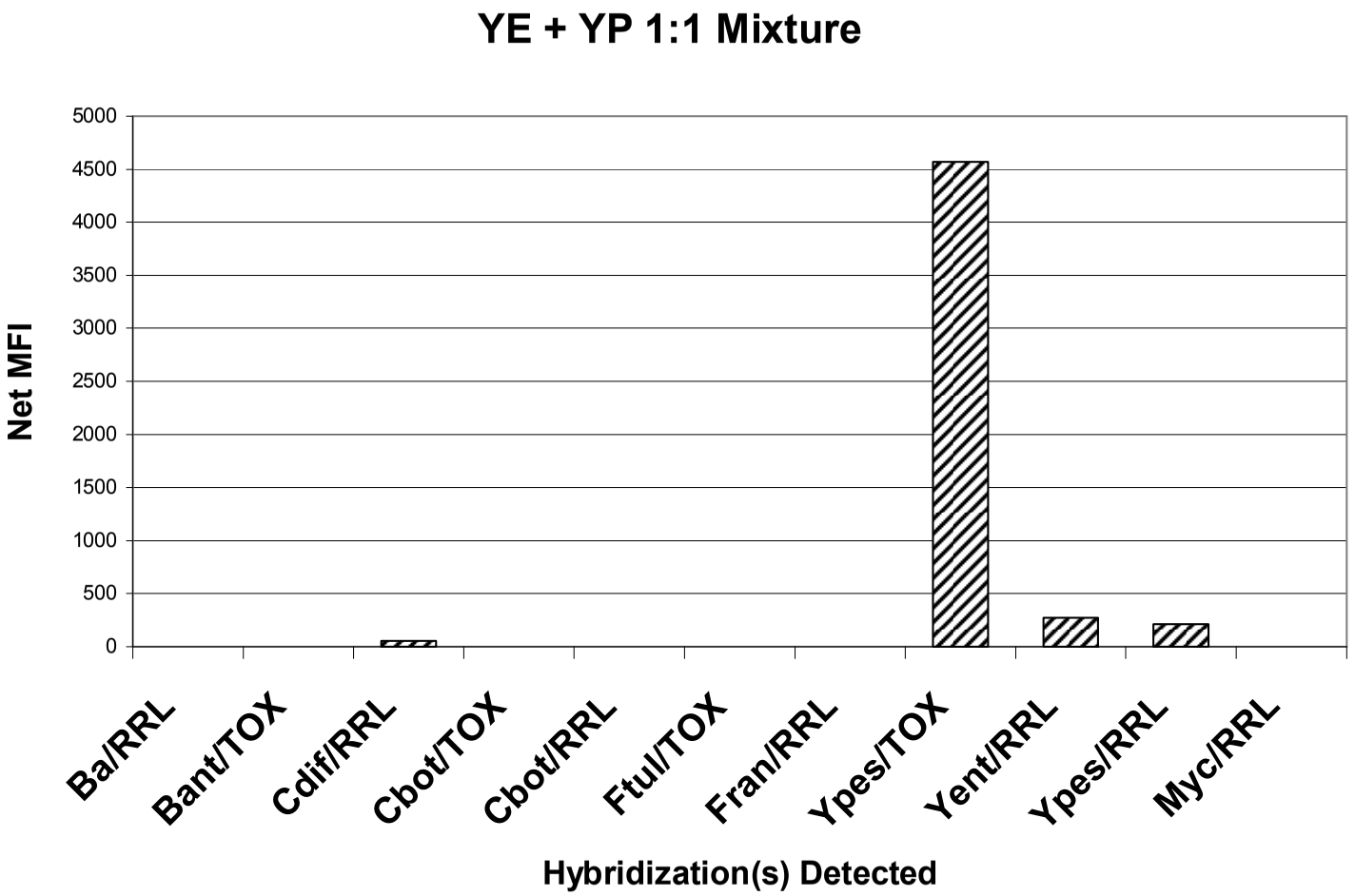


Figure 5.52. Fluorescent response of the eleven probes to a 1:1 mixture of *Y. enterocolitica* and *Y. pestis.*

I. Mixture Study – Ratios

1. BA/FT Ratio Mixtures

Mixtures of *B. anthracis* and *F. tularensis tularensis* genomic DNA were amplified and the PCR amplicons were split into a control group (undigested) and experimental group, with the experimental group being digested with lambda exonuclease as per the usual protocol. The post-digest quantitation results for all control samples in the first batch of ratio mixtures is shown in Table A9.1 in Appendix IX. The quantitation results indicated that, for most target sequences, concentration values generally increased when more template was present, and vice versa. Table A9.2 provides a summary of net MFI values for each mixture analyzed in this assay. The MFI data is depicted graphically in Figure 5.53 below.

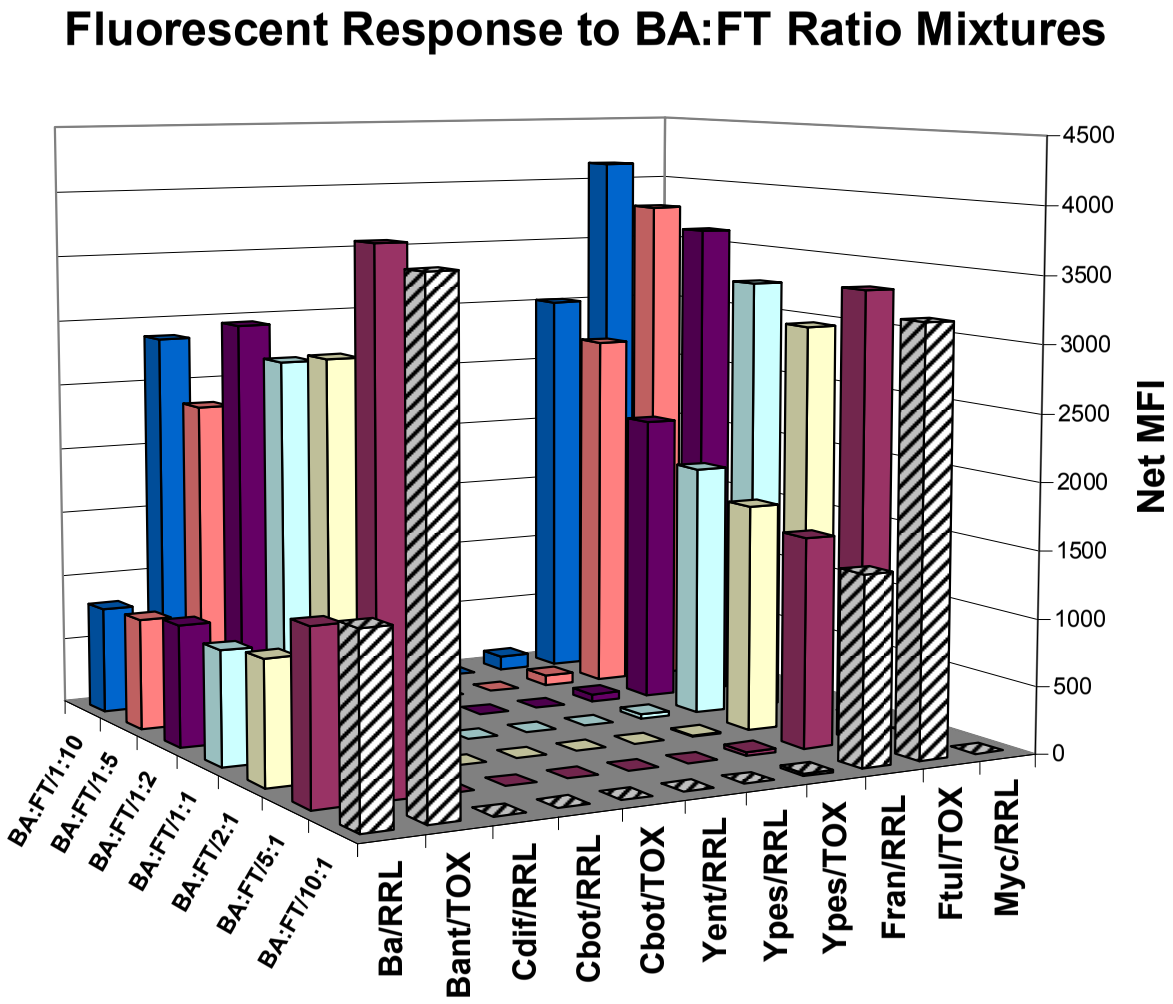


Figure 5.53. Fluorescent response of the eleven probes to various ratios of *B. anthracis* and *F. tularensis tularensis*.

2. BA/YP Ratio Mixtures

Mixtures of *B. anthracis* and *Y. pestis* genomic DNA were amplified and the PCR amplicons were split into a control group (undigested) and experimental group, with the experimental group being digested with lambda exonuclease as per the usual protocol. The post-digest quantitation results for all control samples in the second batch of ratio mixtures is shown in Table A9.3. Table A9.4 provides a summary of net MFI values for each mixture analyzed in this assay. The MFI data is depicted graphically in Figure 5.54 below.

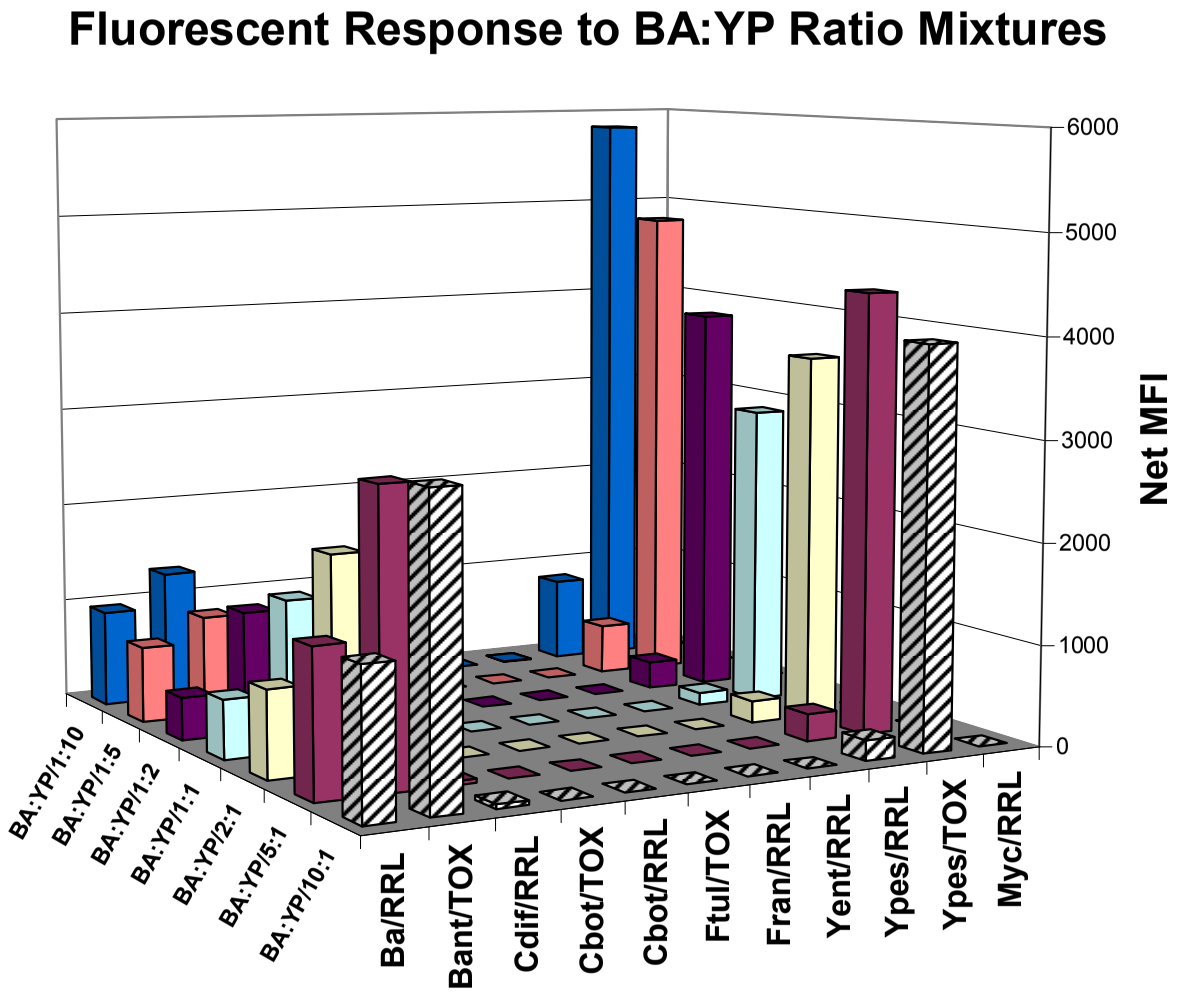


Figure 5.54. Fluorescent response of the eleven probes to various ratios of *B. anthracis* and *Y. pestis.*

3. YP/FT Ratio Mixtures

Mixtures of *Y. pestis* and *F. tularensis tularensis* genomic DNA were amplified and the PCR amplicons were split into a control group (undigested) and experimental group, with the experimental group being digested with lambda exonuclease as per the usual protocol. The post-digest quantitation results for all control samples in the third batch of ratio mixtures is shown in Table A9.5.

The quantitation results indicated that, for most target sequences, concentration values generally increased when more template was present, and vice versa. The main exception was the Ypes/TOX target sequence, which had a consistently high concentration whether it was one or ten parts of the mixture. Table A9.6 provides a summary of net MFI values for each mixture analyzed in this assay. The MFI data is depicted graphically in Figure 5.55 below.

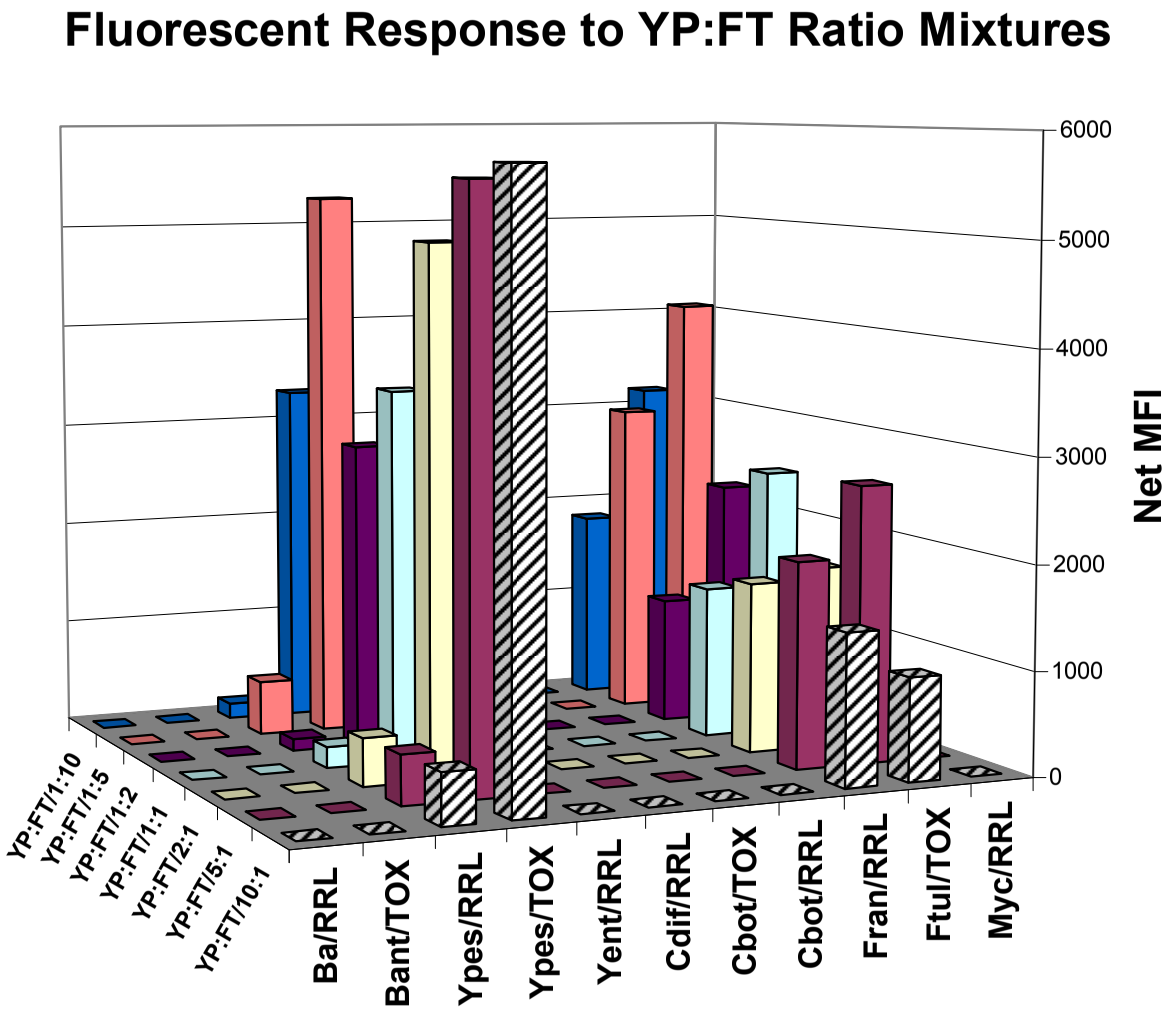


Figure 5.55. Fluorescent response of the eleven probes to various ratios of *Y. pestis* and *F. tularensis tularensis.*

4. BA/CB Ratio Mixtures

Mixtures of *B. anthracis* and *C. botulinum* genomic DNA were amplified and the PCR amplicons were split into a control group (undigested) and experimental group, with the experimental group being digested with lambda exonuclease as per the usual protocol. The post-digest quantitation results for all control samples in the fourth batch of ratio mixtures is shown in Table A9.7.

The standard rules for determining Luminex input amounts were applied. There were two exceptions preventing these rules from being applied as usual. The concentration of the RRL band reported for the BA:CB/1:5 and the BA:CB/1:10 samples was so low that the minimum amount of minor contributor needed to satisfy the Luminex LLD (10 ng) could not be loaded into the Luminex well. Each Luminex well has a capacity of 17 uL of DNA and the concentration of these samples was so low that >17 uL would have been necessary to satisfy the Luminex LLD. Thus, it was decided to load the maximum volume of target DNA (17 uL) into each of these wells. Table A9.8 provides a summary of net MFI values for each mixture analyzed in this assay. The MFI data is depicted graphically in Figure 5.56 below.

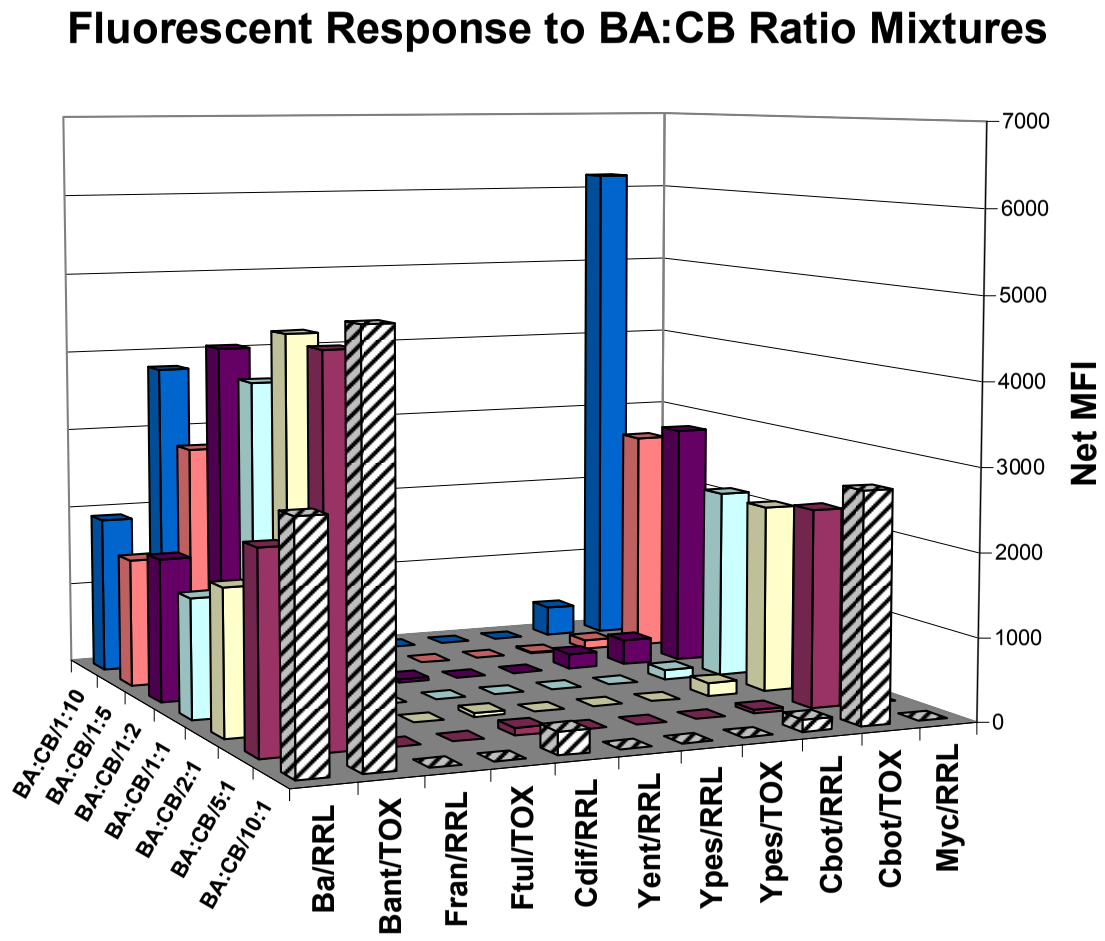


Figure 5.56. Fluorescent response of the eleven probes to various ratios of *B. anthracis* and *C. botulinum*.

CHAPTER VI. DISCUSSION

A. Probe Design for the *rrl* Gene

The first obstacle in this research was encountered when searching for the *rrl* gene sequence of the nine microbes. Some *rrl* gene sequences were easy to locate in the NCBI Gene database, while others were more difficult to find. After some experimentation, it became apparent that these difficulties stemmed from the database nomenclature – there was no standardization for the manner in which uploaded gene sequences were named and/or abbreviated. When the nucleotide sequence of a particular gene was discerned, some researchers uploaded the sequence to the Gene database and named the file according to the gene name, while other researchers uploaded the file and named it according to the gene product. Therefore, *rrl* gene sequences were found when searching both *rrl* (the gene name) and 23S rRNA (the gene product). Once this subtlety of the system was realized, the search for *rrl* gene sequences progressed with greater ease.

Initially a separate locus was sought for all RRL target sequences. However, a high degree of homology in the *rrl* gene made it difficult to establish primers that possessed an adequate degree of specificity. In preliminary experimentation, an attempt was made to improve primer specificity by manipulating template amounts, PCR cycling parameters, and primer concentrations (data not shown). The results were still unsatisfactory and ultimately a locus of the *rrl* gene that had been studied in prior research (Dunbar and Jacobson, 2007; Battaglia *et al.*, 2011) was selected for target sequence design. The same locus for all nine bacteria was used for RRL probe design. This locus was ideal because the primer- and probe-binding regions were thought to be

polymorphic enough to enable a suitable level of specificity. An ancillary benefit was that the use of the same locus for all RRL target sequences permitted the use of degenerate primers, which were more cost-effective to synthesize than nine separate primer sets. Within this locus, it would have been preferred to design a unique RRL probe for each species, but again this was precluded in some cases by a high degree of homology in the probe-binding region. A unique RRL probe was successfully designed for the two species in the Clostridium genus (Cdif/RRL and Cbot/RRL) and the two species in the Yersinia genus (Yent/RRL and Ypes/RRL). However, a unique RRL probe could not be designed for the two species in the Bacillus genus or the two species in the Francisella genus. Instead, generic Bacillus (Bac/RRL) and Francisella (Fran/RRL) probes had to be designed due to total homology in the probe-binding sequences of organisms in each of these genera. Confirmatory identification of the four pathogenic species was accomplished through the use of the TOX probes.

B. Probe Design for the TOX Gene

Unlike the RRL probes which were (1) designed to target the same gene locus for all nine bacteria, and (2) designed within a locus that had been previously studied, the TOX probes were each within different genes and designed *de novo* for this study. This was reflected in the extensive background literature search and bioinformatics study undertaken for the TOX genes. After the TOX gene sequences were identified, they were aligned using ClustalW to identify regions in which there was little or no homology with sequences in the genomes of the other pathogenic bacteria used in this study. The TOX gene sequences were uploaded to Primer3 software (Rozen and Skaletsky, 2000). This program designs probes and PCR primers while considering factors such as melting

temperature, length, GC content, secondary structure, and primer-dimer formation. Any homologous regions identified in the ClustalW alignment were excluded from probe design in Primer3. With these excluded regions in mind, Primer3 identified a probe sequence (as well as PCR primers needed to amplify the probe) for each of the four pathogenic bacteria. The TOX probes were all designed with similar melting temperatures (~60ºC) so that a single hybridization temperature could be used in a multiplexed analysis. For the same reason, the PCR primers were also designed with similar melting temperatures.

The candidate probes identified by Primer3 were carefully examined for the specific location of polymorphic bases within the twenty base oligo. Based on recommendations from Luminex technical support and the findings of other researchers (Dunbar and Jacobson, 2005; Battaglia *et al.*, 2011), probes were designed so that there existed a minimum homology to other probes, particularly in the mid region. When the probe and target sequence collide and begin to hybridize, bases in the middle of the sequence bind first, followed by an outward expansion on both sides of the growing duplex. Base complementarity in the middle of a non-specific probe and target sequence may lead to hybridization because the duplex could be irrevocably formed by the time polymorphisms located on the periphery of the binding region are encountered. Polymorphisms close to the center of an oligo are more adept at preventing non-specific binding.

C. Touchdown PCR

A modified method of PCR called touchdown PCR (TD-PCR) was devised in 1991 (Don *et al.*). The technique requires starting the amplification with an annealing

temperature approximately 10°C above the highest primer melting temperature and then incrementally decreasing the annealing temperature (usually 1°C per cycle) until it is just below the lowest melting temperature, at which point the annealing temperature remains constant for the remaining PCR cycles (Korbie and Mattick, 2008). TD-PCR enables the target sequence to be preferentially amplified early on when the higher annealing temperatures are used, thus enhancing amplification of the target and reducing non-specific amplification.

As long as the initial annealing temperature is high enough (while still remaining below the 72°C extension) and provided that it does not drop too far below the lowest primer T_M, it should theoretically pass through the ideal annealing temperature ranges for all primer sets as it descends during the first few cycles. The final annealing temperature should be far enough below the T_Ms so as to permit the primers to anneal, yet not so far below to allow non-specific priming.

TD-PCR is an ideal method for enhancing the specificity of primers whose melting temperature is in the middle or upper part of the range of T_Ms. However, primers whose T_M is at the lower part of the range may be more prone to drop-out and/or non-specific amplification. Amplicons whose primers have higher melting temperatures are preferentially amplified in the first half of the PCR reaction, but amplicons whose primers have lower T_Ms do not experience this benefit. TD-PCR may be a good compromise when primers of various T_Ms are used in a multiplexed PCR reaction, but researchers should also be aware of the nuances of this technique. Imbalances in the proportion of amplicons observed may be partially corrected by modifying other PCR parameters such as primer concentration, template amount, and salt concentration.

D. Monoplex PCR Specificity Study

Proper validation procedures dictated the need for singleplex amplifications to demonstrate that each primer set was capable of amplifying the intended target DNA. Without the singleplex amplifications, there would have been no way to confirm that each primer set successfully amplified the target DNA for which it was designed. It would not have been possible to confirm primer specificity without singleplex amplifications because all target DNA sequences fell within the same general size range (~100-200 bp). Their similar sizes would have made it difficult to differentiate all target DNA sequences using the Agilent 2100 Bioanalyzer.

The touchdown method of PCR as described previously was employed in the singleplex amplifications as a means of further enhancing primer specificity. Ultimately, the goal was to integrate all primers into a single primer mix, which meant that any TD-PCR reaction in which these primers were used would require the use of annealing temperatures consistent with the primer melting temperatures. The melting temperatures of all primers ranged from 64.5 to 56.3°C (see Tables 4.5 and 4.6). Accordingly, a final annealing temperature of 55°C was selected for the monoplex amplifications because this would allow even the primers with the lowest T_Ms to hybridize with the template during the annealing step. The touchdown PCR reaction was programmed to begin with an annealing temperature of 67°C, dropping by 1°C over the first twelve cycles before reaching the final annealing temperature of 55°C for the remaining thirteen cycles.

The expected amplicons were produced when each template was amplified using its designated primer set (Figures 5.4-5.9). The concentrations of the amplicons produced for each template varied somewhat, ranging from 2.43 to 20.69 ng/uL, with a standard

deviation of 6.2 (see Table A3.1 in Appendix III). This was expected because of several variables such as the use of degenerate primers, and the fact that the target sequences were present in higher copy numbers in some bacteria. It was somewhat surprising that the RRL amplicons generally had lower concentrations than the TOX amplicons, because the RRL target sequences were often present in multiple copy numbers whereas the TOX target sequences occurred only once in each respective genome. However, the amplification efficiency of the RRL target sequences could have been compromised by the use of degenerate primers, which may not have amplified the target DNA as efficiently as the TOX primers. The TOX primers were custom designed for each individual pathogenic bacteria and were therefore expected to yield a greater amplification efficiency.

Amplification efficiency could have also been impacted by the various primer melting temperatures. Touchdown PCR begins with an annealing temperature approximately ten degrees above the highest primer melting temperature, descending by about one degree over the course of ten cycles before stabilizing for the remaining cycles. Before the annealing temperature stabilizes, the amplicons whose primers have high melting temperatures are preferentially amplified. Therefore, there could be a correlation between PCR yield and primer melting temperatures. Using the amplicons described in Table 5.4 and Figure 5.9 as an example, yield was often good for *Y. pestis* and *Y. enterocolitica* amplicons, whose primers had T_Ms toward the high end of the range. This supports the statement made above, that amplicons whose primers have high T_Ms are more efficiently amplified using TD-PCR. However, this theory cannot be universally applied to explain all results. For example, yields were often very low for *M. tuberculosis*

amplicons (Table 5.3), whose primers had T_Ms in the middle of the range. Overall these observations might indicate that yield is more dependent on template amount (*M. tuberculosis* has only one *rrl* gene copy) than primer melting temperatures.

Despite a moderate degree of variation in the concentration of amplicons produced, a template concentration study was not undertaken to balance the concentration of all amplicons. The uniformity of all amplicons (in terms of their concentration) was not deemed essential for success of the Luminex assay. The aim of this study was to establish a platform for identifying pathogenic microorganisms. Even if a template concentration study was performed, this would not necessarily be useful as clinical samples would not have balanced ratios. Nevertheless, future researchers may wish to adjust the template input amounts as needed to balance the amplicons, in order to establish a baseline for comparing MFI values for different target sequences.

For quantitations performed on the Agilent 2100 Bioanalyzer, the linear regression line of apparent length (bp) vs. time (sec.) for the ladder should have a coefficient of determination (R^2) near one to ensure optimum sizing accuracy (C. Vanselow, personal communication, 23 August 2011). The reported lengths for the target sequences in the monoplex PCR study (and all subsequent assays) were at or near the accepted sizing accuracy standard for the Agilent Bioanalyzer (Agilent, 2011; data not shown). Occasionally the size of the amplicons reported were slightly outside the accepted sizing accuracy of the Agilent Bioanalyzer (+/- 5% for oligos between 100 and 500 bp; Agilent 2011). This may have been due to the formation of secondary structures within the amplicons, which would affect their migration through the gel. Nevertheless,

any disparities between the reported lengths and the expected lengths were still small enough such that the amplification of the intended targets was able to be confirmed.

A satisfactory level of specificity was observed for the RRL and TOX primers in the monoplex amplification study. Specificity was gauged by the overall absence of superfluous bands when each target DNA was amplified using the designated primers. The specificity of the TOX primers was expected to be especially good because these primers were designed to amplify target sequences found in highly unique toxin-coding genes exclusive to certain pathogenic species. Surprisingly, there was one instance in which a non-specific amplicon was detected – the TOX primers produced a non-specific amplicon when *F. tularensis holarctica* (nonpathogenic species) was used as template (Figure 5.8). This was unexpected because all primers were scrupulously designed to minimize cross-reactivity.

1. Addressing the Non-specific Amplicon Produced in the Amplification of F. tularensis holarctica Template using the Ftul/TOX primer set

Although the initial search of the Ftul/TOX primer and probe sequences produced no significant alignments with species other than *F. tularensis tularensis*, early PCR data indicated non-specific amplification (within the *F. tularensis holarctica* genome) of a PCR product having a similar size as the Ftul/TOX target (Figure 5.8).

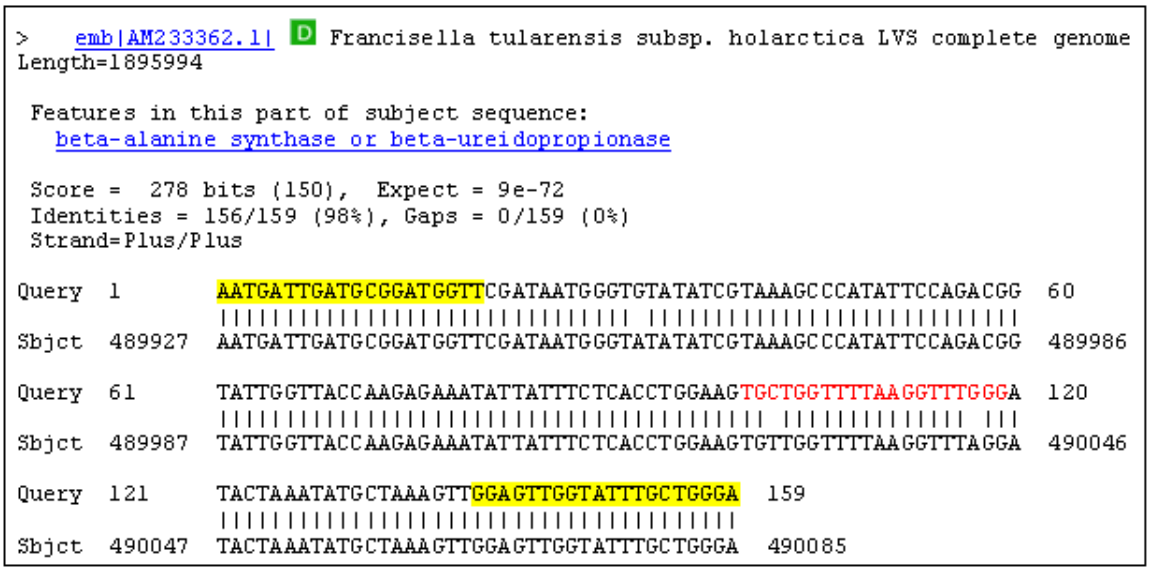
The FTT0435 gene codes for citrulline ureidase, an ammonia-producing enzyme that helps *F. tularensis tularensis* to evade lysosomal degradation in host organisms (Larsson *et al.*, 2005). Given that this enzyme was thought to be unique to *F. tularensis tularensis* (Sjostedt, 2004), it was surprising to find that the TOX primers produced an amplicon of the expected length in the non-toxic species. However, it was considered that

the FTT0435 gene may have been present but undetectable when initial studies with this gene were conducted. For example, techniques used by prior researchers (to infer that FTT0435 is unique to *F. tularensis tularensis*) were based on enzyme activity (Fleming and Foshay, 1955). In other words, because the products (citrulline ureidase and its ammonia byproduct) of the FTT0435 gene were not detected, it might have been erroneously inferred that the gene was absent in *F. tularensis holarctica*.

Another possibility for the FTT0435 gene not being detected in *F. tularensis holarctica* was that the gene may have been present but not expressed. The FTT0435 gene in *F. tularensis holarctica* may be a pseudogene, which is a gene that is still present in the genome but is not transcribed and/or translated, due to a mutation within the gene itself or in the control region. If FTT0435 is a pseudogene in *F. tularensis holarctica*, then the gene would be detectable in a genetic analysis (such as the present study) but the product of this gene would be undetectable (*i.e.*, protein expression for this gene would be negative).

However, having a PCR product of similar size to the Ftul/TOX target does not confirm the presence of the FTT0435 gene in *F. tularensis holarctica*. Furthermore, this non-specific amplification would not necessarily lead to a false positive in the analysis of *F. tularensis holarctica* on the Luminex. If the Ftul/TOX probe-binding region of the target sequence was absent or mutated in the non-specific amplicon from *F. tularensis holarctica*, it would not be detected in a Luminex assay. Nevertheless, it was still puzzling that the Ftul/TOX primers were able to amplify a section of the *F. tularensis holarctica* genome, because all primers underwent an extensive bioinformatics analysis to confirm specificity.

The *F. tularensis holarctica* genome was manually searched for the presence of the Ftul/TOX target sequence. Surprisingly, the Ftul/TOX target sequence possessed 98% similarity to a sequence in *F. tularensis holarctica* beginning at base number 489927 (Figure 6.1). Both the forward and reverse primers for Ftul/TOX were 100% identical to the corresponding sequence found in *F. tularensis holarctica*, thus explaining the non-specific amplicon. Furthermore, the probe sequence was identical to the aligned sequence in the *F. tularensis holarctica* genome at 18 of the 20 bases.



```
>    emb|AM233362.1|  Francisella tularensis subsp. holarctica LVS complete genome
Length=1895994

 Features in this part of subject sequence:
   beta-alanine synthase or beta-ureidopropionase

 Score =  278 bits (150),  Expect = 9e-72
 Identities = 156/159 (98%), Gaps = 0/159 (0%)
 Strand=Plus/Plus

Query  1       AATGATTGATGCGGATGGTTCGATAATGGGTGTATATCGTAAAGCCCATATTCCAGACGG  60
               |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct  489927  AATGATTGATGCGGATGGTTCGATAATGGGTATATATCGTAAAGCCCATATTCCAGACGG  489986

Query  61      TATTGGTTACCAAGAGAAATATTATTTCTCACCTGGAAGTGCTGGTTTTAAGGTTTGGGA  120
               ||||||||||||||||||||||||||||||||||||||| ||||||||||||| |||
Sbjct  489987  TATTGGTTACCAAGAGAAATATTATTTCTCACCTGGAAGTGTTGGTTTTAAGGTTTAGGA  490046

Query  121     TACTAAATATGCTAAAGTTGGAGTTGGTATTTGCTGGGA  159
               |||||||||||||||||||||||||||||||||||||||
Sbjct  490047  TACTAAATATGCTAAAGTTGGAGTTGGTATTTGCTGGGA  490085
```

Figure 6.1. Comparison of the Ftul/TOX target sequence (Query) with a sequence found in the beta-alanine synthase or beta-ureidopropionase gene in *F. tularensis holarctica* (Subject). Yellow highlighted regions indicate PCR primers for the Ftul/TOX target sequence. Bases colored in red font indicate the Ftul/TOX probe. Aligned bases without a dash mark in between indicate a polymorphism between the query and subject sequence.

The sequence producing the significant alignment to the Ftul/TOX target resides in the beta-alanine synthase or beta-ureidopropionase gene (also known as FTL0504) in *F. tularensis holarctica* (Chain *et al.*, 2011). The gene product hydrolyzes amide (carbon-nitrogen) bonds which is similar in function to other hydrolases such as FTT0435 (Mahawar *et al.*, 2009). Although FTT0435 may exhibit reaction catalysis properties that

are similar to other genes such as FTL0504, the literature still supports the belief that FTT0435 is unique to *F. tularensis tularensis*. The role of FTT0435 in conferring toxicity has been corroborated by insertional mutagenesis studies but the results are sometimes complicated to interpret because genomic redundancy evidently compensates for the absence of FTT0435 (Mahawar *et al.*, 2009). It is still possible that FTT0435 is unique to *F. tularensis tularensis* but apparently the particular section of the gene from which the target region originates is not.

It was still puzzling that these significant alignments were not reported when the sequences were originally searched against the nucleotide collection (nr/nt) database maintained by NCBI. The cause may have been related to an algorithm parameter called the Expect (E) value. When the BLAST search engine finds a match to the query sequence, an E value is reported. This value predicts the number of chance matches to the query sequence. Lower E values portend fewer matches by chance, and thus the match becomes more significant.

The default E value threshold for the BLAST software is ten, meaning that any significant matches with an E value higher than ten are not reported. However, the matches for short sequences (such as the primers and probes searched for this study) can sometimes yield E values greater than ten because there is a greater probability of sequences matching by chance when the query sequence is short. This means that significant alignments to the query sequence may not be reported simply because of an E value that is higher than the threshold. In sum, lower E value thresholds enhance the stringency of a search because fewer chance matches are reported, with the caveat that some significant alignments may be unreported. If the E value threshold had been raised

above the default value of ten when the Ftul/TOX primers and probe were searched in the database, the matching sequences in *F. tularensis holarctica* might have been reported as signficant alignments.

Filtering is another possible reason that homologous sequences in *F. tularensis holarctica* might not have produced significant alignments when the Ftul/TOX primers and probes were searched. When a query sequence is uploaded to the BLAST search engine, the software automatically detects regions of low complexity in the sequence. These are sequences with an unusual base composition (usually because of repetition or lack of variability in the bases) that can disrupt the search by producing many insignificant matches. In order to avoid this, the BLAST algorithm uses a default filter called DustMasker (Morgulis *et al.*, 2006) to mask regions of low complexity. The masked regions are excluded from alignments with sequences in the database, thereby potentially preventing a significant match from being made. It is not known for certain whether any part of the Ftul/TOX primers or probe would be considered regions of low complexity, but more significant alignments to these sequences might have been produced if the default filter had been disabled.

The bioinformatics analysis portion of this research began with primer and probe design using Primer3 software, followed by a search of the sequence in the BLAST database. Future researchers may find it more convenient to consolidate these tasks into one program called Primer-BLAST. This is a primer designing tool maintained by NCBI which uses Primer3 for primer design. Additionally, there is an option for the design of an internal hybridization oligo (*i.e.*, probe sequence). The advantage of using Primer-BLAST is that the software automatically reports significant alignments to primer and

probe sequences by searching within the BLAST database. Another advantage of using Primer-BLAST is that the default algorithm parameters are optimized for searches of short sequences. For example, the default E value in Primer-BLAST is 30,000 instead of the default of ten used in the traditional BLAST database search.

The data from the monoplex PCR study strongly indicated that the Ftul/TOX primers lacked specificity and were amplifying a section of the *F. tularensis holarctica* genome. Despite this minor setback, it was still thought that the Luminex might be able to differentiate *F. tularensis tularensis* and *F. tularensis holarctica*. There were two polymorphisms between the Ftul/TOX probe and the aligned section of the *F. tularensis holarctica* genome (Figure 6.1). In lieu of re-designing the Ftul/TOX target, it was decided to test the specificity of the Luminex instrument and determine whether these two polymorphims would be sufficient to enable differentiation of *F. tularensis tularensis* and *F. tularensis holarctica*. Although prior research indicated the ability of the Luminex platform to differentiate sequences that differ by as little as one SNP (Diaz and Fell, 2004; Drago *et al.*, 2009; Deshpande *et al.*, 2010), subsequent research has provided evidence that the location of the SNP(s) within a probe may play an important role in determining specificity (Dunbar and Jacobson, 2005; Battaglia *et al.*, 2011).

E. Multiplex PCR Specificity Study

1. PCR Primer Optimization

The ultimate goal of this study was to create a pathogen detection array, which required the simultaneous amplification of all target sequences that may or may not have been present in the sample. Therefore, PCR primers for all eleven target sequences had to

be combined into a single primer mix and used to amplify each individual template. The primers were combined in very precise ratios based on the abundance of the target sequences for which they were designed. For example, the degenerate RRL primers made up a larger composition of the primer mix because there were seven RRL target sequences. The need for primer concentrations to be well balanced in multiplex reactions has indeed been documented by other researchers (Schoske *et al.*, 2003). Having a primer mix in which the primers are either too scarce or too abundant could have deleterious effects such as amplification drop-out, non-specific amplification, or primer-dimer formation.

When the original primer mix (see Table 4.12) was used to amplify the eleven target sequences (Figure 5.10 and Table A2.1 in Appendix II), there was only one instance of non-specific amplification (the Ftul/TOX target sequence produced when *F. tularensis holarctica* was used as the template). Most primers demonstrated a sufficient degree of specificity when used in a primer mix to amplify their resepective target sequences. Additionally, for the samples in which target sequences were successfully amplified (not including two samples in which all templates were mixed), the concentration of the target sequences was usually adequate (mean amplicon concentration of 13.1 ng/uL; Table A3.2 in Appendix III) and consistent with the results of the monoplex PCR study (mean amplicon concentration of 12.8 ng/uL; Table A3.1). The uniformity among the concentration of target sequences amplified using the original primer mix (standard deviation of 5.1; Table A3.2) was also similar to the concentrations reported in the monoplex PCR study (standard deviation of 6.2; Table A3.1).

Although non-specific amplification was detected infrequently among amplicons from the original primer mix, there were several instances in which amplification drop-out occurred. The expected bands were not always observed, such as in lanes 3, 4, 6, and 8, when *B. cereus*, *B. anthracis*, *C. botulinum*, and *F. tularensis tularensis* were used as template, respectively (Figure 5.10). Most of the instances of amplification drop-out involved the TOX target sequences. In the original PCR primer mix, the primers for the TOX genes were present in 5 uM concentration, whereas the degenerate RRL primers were present in excess (RRL fwd. #1 – 35 uM; RRL fwd. #2 – 10 uM; RRL rev. – 45 uM) to accommodate the five targets which they were designed to amplify. However, these high primer concentrations may have interfered with the ability of the TOX primers to amplify their target (a theory supported by the high volume of primer dimers observed).

The primer-dimer formation observed was unexpected because the initial AutoDimer primer dimer screen did not predict any dimerizations among the primers. First, AutoDimer is a tool for analyzing the tendency for self- and cross-dimerizations, but may not yield predictions that are always observed in practice. Second, the initial AutoDimer primer dimer screen was performed using the default threshold value of seven. Match scores are computed in AutoDimer by comparing two sequences and subtracting the number of mismatching bases from the matching bases. Two sequences whose interaction produces a score below the threshold would not be reported. Although the propensity for cross-reactivity is reduced as the score number drops, this does not imply that interactions with scores below the default threshold value can not occur *in situ*. Such interactions do occur, and future researchers are encouraged to use AutoDimer

threshold values below the default value to obtain a more complete representation of all interactions that could potentially occur, before primers are synthesized and put to use.

Although the occurrence of primer dimers became an inherent and unwanted byproduct, a modified primer mix (see Table 4.13) was devised to minimize their frequency and concentrations, and to ameliorate other problems encountered with the original primer mix (*e.g.*, amplification drop-out). When the modified primer mix was used to amplify the eleven target sequences (Figure 5.11 and Table A2.2 in Appendix II), there was one instance of non-specific amplification, also observed in prior assays (the Ftul/TOX target sequence produced when *F. tularensis holarctica* was used as template). For the single-template samples, the concentration of the target sequences was both adequate (mean amplicon concentration of 12.7 ng/uL; Table A3.2 in Appendix III) and fairly consistent (standard deviation of 6.3; Table A3.2). The mean target sequence concentration produced using the modified primer mix was slightly lower than that observed when the original primer mix was used. Likewise, the standard deviation was also higher when the modified primer mix was used in the PCR reaction. However, these minor differences were overshadowed by the fact that no amplification drop-out was observed in any samples.

Amplification efficiency was probably improved by using an equitably balanced primer mix in which the individual primers were not excessively concentrated. This, in turn, may have reduced the incidences of primer dimer formation. When the original primer mix was used to amplify the individual templates, amplification drop-out was observed in several samples and the average primer-dimer concentration was 12.0 ng/uL (Table A3.2). It came as no surprise that, for samples amplified with the modified primer

mix, the average primer dimer concentration dropped to 6.8 ng/uL and amplification drop-out was totally eliminated (Table A3.2).

A 1:2 diluted version of the modified primer mix (see Table 4.14) was prepared and used to amplify the templates to determine whether primer dimer formation could be further reduced without compromising the specificity of the primers (Figure 5.12 and Table A2.3 in Appendix II). When this diluted primer mix was used to amplify the target sequences, one non-specific amplicon was observed (as in other assays), which most likely corresponded to the Ftul/TOX target sequence amplified when *F. tularensis holarctica* DNA served as template. Technically there were no amplification drop-outs because all target sequences were visible on the gel, although the Myc/RRL target sequence was not concentrated enough to be officially reported. This was not explored further because it was regarded as an aberration, based on successful amplification of the Myc/RRL target in prior amplifications. Also, the concentration of the Myc/RRL target sequence was typically lower than the concentrations of other target sequences in any given assay, most likely because there is only one copy of the *rrl* gene in *M. tuberculosis*.

For single-template samples amplified using the diluted primer mix, the mean target sequence concentration was 10.0 ng/uL (Table A3.2 in Appendix III), not far below the mean target concentrations observed when samples were amplified using the original and modified primer mixes (13.1 ng/uL and 12.7 ng/uL, respectively; Table A3.2). The standard deviation of target sequences amplified with the diluted primer mix was 4.6, compared to 5.1 (original primer mix) and 6.3 (modified primer mix), indicating a superior level of uniformity among the concentrations of target sequences amplified with the diluted primer mix (Table A3.2). The most corroborative evidence collected in

favor of the diluted primer mix was the low frequency of primer dimers. Not only did the diluted primer mix yield the lowest incidence of primer-dimers among all primer mixes tested, but the primer dimers that were observed with the diluted primer mix had the lowest average concentration (3.4 ng/uL) when compared to the average primer dimer concentrations observed with the original and modified primer mixes (12.0 ng/uL and 6.8 ng/uL, respectively; Table A3.2).

2. PCR Template Optimization

When the primers and their templates are abundantly present in the PCR reaction mixture, two things may happen: (1) the primers are so plentiful that non-specific amplification products become more prevalent, or (2) the template DNA is so plentiful that other primers could have an increased tolerance for mismatches and thus bind to non-specific DNA with greater ease. For these reasons, it was important to optimize both primer concentration and template amounts before a standard PCR protocol was established.

The diluted, modified primer mix (see Table 4.14) was used to amplify 10% of the template amounts used in previous studies (Figure 5.13 and Table A2.4 in Appendix II). The purpose was to determine if using 10 ng of template was as effective as using the typical amount (100 ng). There were no non-specific amplicons observed, except for the one Ftul/TOX target sequence that was amplified when *F. tularensis holarctica* was used as template. This phenomenon was observed in other amplifications as well. For single-template samples amplified using the diluted primer mix (and 10 ng of template), the mean target sequence concentration was 9.4 ng/uL, which was very close to the mean

target sequence concentration observed when 100 ng of template was amplifed with the same primer mix (10.0 ng/uL; Table A3.2 in Appendix III). The standard deviation (5.6) of the amplicon concentrations reported when the reduced amount of template was used was in close proximity to the standard deviation (4.6) of the amplicon concentrations reported when the higher amount of template was used (Table A3.2). The number of primer dimers detected in samples amplified with reduced template were the least abundant of any assay performed in the primer concentration study. However, the mean primer dimer concentration observed with reduced template was 6.3 ng/uL, which was slightly higher than the corresponding value observed when 100 ng of template was used (3.4 ng/uL; Table A3.2).

Overall, the data collected with 100 ng and 10 ng template amounts was comparable except that the mean concentration of primer dimers was slightly higher when only 10 ng of template was input into the PCR reaction. These minor artifacts were not expected to be detrimental to the Luminex assay. Also, using less template was preferable given the limited quantity of genomic DNA available from BEI Resources, and more representative of a real-life sample. Even if the genomic DNA sequences were available in unlimited quantities, using 100 ng of template was not necessarily preferable because it often yielded amplicons whose concentration far surpassed the quantities required in the Luminex assay. The standard PCR protocol was ultimately amended to include the use of the diluted, modified primer mix (see Table 4.14) and 10 ng of template per sample.

F. Titration Study

Before probe-coupled beads were used in this study, the attachment of probes to beads was first confirmed. This was accomplished by using reverse probes that were complementary in sequence to their corresponding probe. The verification of probe-bead coupling was paired with a titration study in which the general uniformity in number of probes per bead (across all eleven bead sets) was established by using various amounts of reverse probe, ranging from 0.0 to 250 fmol. Whereas other assays in this study utilized a mixture of all eleven bead sets for multiplexing purposes, the titration study only involved the bead set which had complementarity to the particular probe being studied. This was expected to reduce cross-reactivity that might otherwise prevent an estimate of the number of probes per bead.

The results of the titration study confirmed that all bead sets had the appropriate probe attached. It was also apparent that all eleven bead sets consisted of beads with a generally uniform number of probes attached (approximately 600,000 which was consistent with other estimates; Spiro *et al.*, 2000; Brian Schrader, personal communication, 31 August 2009). This estimate was tabulated by plotting fluorescence observed when various amounts of reverse probe were used. Then, the amount of reverse probe was noted at the approximate location where the curve plateaued. This served as the starting point for a series of calculations used to determine the number of probes attached to each bead. The theory behind starting with the amount of reverse probe at the plateau was that the probes were saturated with reverse probe at this point. An estimate of the number of probes per bead was calculated based on this principle (Figure 5.14).

Although the titration study yielded useful data, the major issue was that fluorescence was sometimes detected for beads that should not have been present. Despite the use of only one bead set per assay, beads from other sets were occasionally detected. Carryover from other assays was not considered as a potential explanation to the problem because, in some instances, beads were being detected from a set that had never been run through the instrument before. Instead, it seemed likely that the instrument was sometimes not able to distinguish one bead set from another.

The titration study for the Yent/RRL bead set, for example, displayed some of the aforementioned aberrant MFI values (Table 10.1 in Appendix X). This assay used only Yent/RRL reverse probe, which meant that all fluorescence observed was expected to correspond to the Yent/RRL bead set. However, fluorescence was also observed for the Bant/TOX bead set. This was puzzling because beads from the Bant/TOX bead set were not used in this assay. These aberrant MFI values were considered a serious problem because if this were a multiplexed assay, the MFI values detected for the Bant/TOX bead set would be false positives.

In order to examine this problem further, the count data from this assay was observed (Table A10.2 in Appendix X). The MFI data represents the median fluorescent intensity for all beads that passed through the channel in a given sample. However, the count data represents the actual number of beads that passed through the channel in a given sample. It was important to refer to the count data because, according to Luminex technical support, MFI values derived from any less than thirty-five beads are unreliable and should not be regarded as significant (C. Norton, personal communication, 05 July 2010). In the Count data table for the Yent/RRL titration study (Table A10.2), note that

reliable MFI values were consistently derived from approximately 500 beads or more. However, the aberrant MFI values often resulted from as few as one or two beads.

Given that the aberrant MFI values originated from less than thirty-five beads, they were regarded as insignificant. But it still did not explain why this phenomenon was being observed. Recall that all beads had two internal dyes (red and infrared) which enabled the instrument to classify them as one particular bead set among one hundred. Each of the one hundred bead sets marketed by Luminex has a very unique spectral signature. If the spectral emission ranges of the two internal dyes for a given bead were plotted against each other, one would expect each bead set to occupy a very precise and exclusive location on a chart of all bead emissions. For example, Figure 6.2 shows the unique spectral location (white ovals) for all 100 Luminex beads when the emission spectra of the internal dyes are plotted against each other.

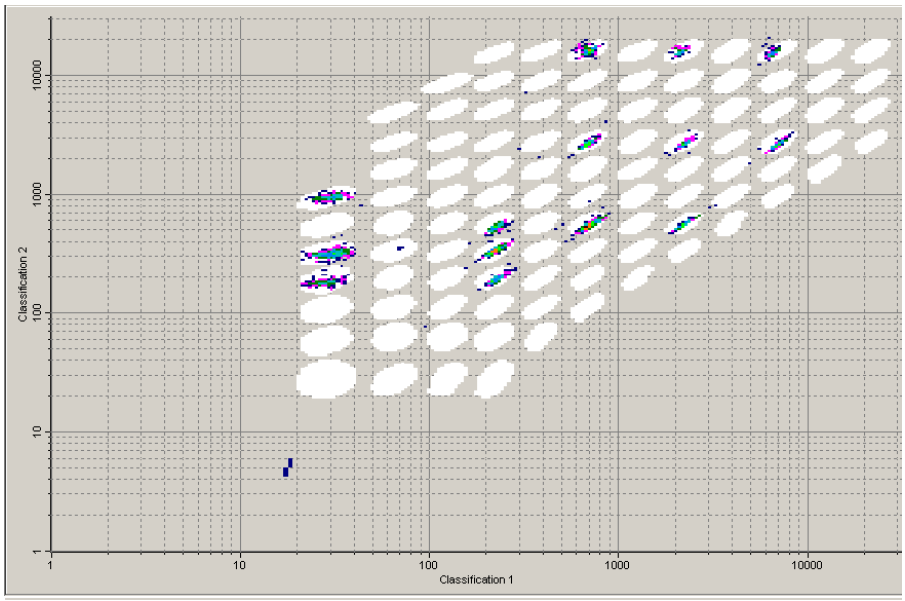


Figure 6.2. Classification plot of two internal dyes for 100 bead sets. The unique spectral location of each bead set is denoted by a white oval. Image adapted from original version found in the Luminex PowerPoint® presentation entitled "Luminex Advanced Applications Training: Nucleic Acid Assay Development". Reprinted with permission from Luminex Corporation. Note: this image was provided by Luminex and is unrelated to any data in this study.

Even though only eleven of the one hundred available bead sets were utilized in this study, each bead set still occupied a unique spectral location. Figure 6.3 shows the classification plot of the eleven bead sets used in this study.

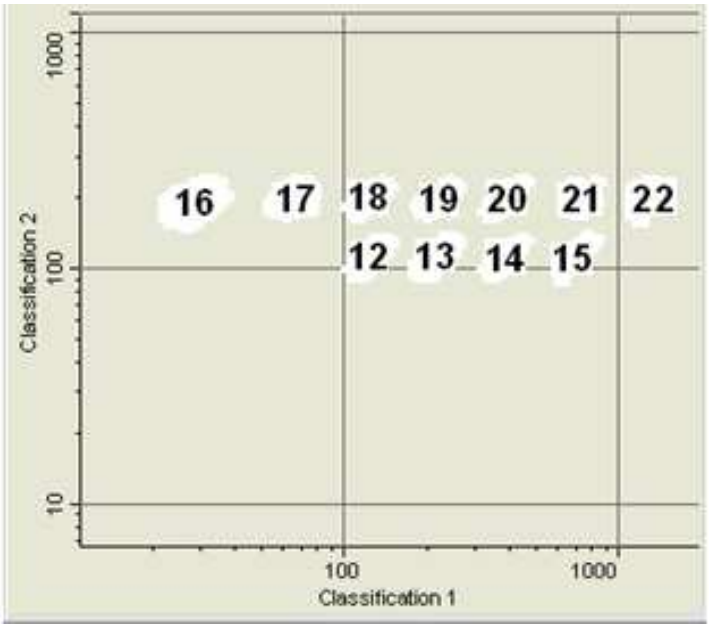


Figure 6.3. Classification plot of two internal dyes for eleven bead sets used in this study.

To obtain more information about the cause of the aberrant MFI values, the classification plot for a selected sample in the Yent/RRL titration study was examined. The selected classification plot is shown in Figure 6.4.

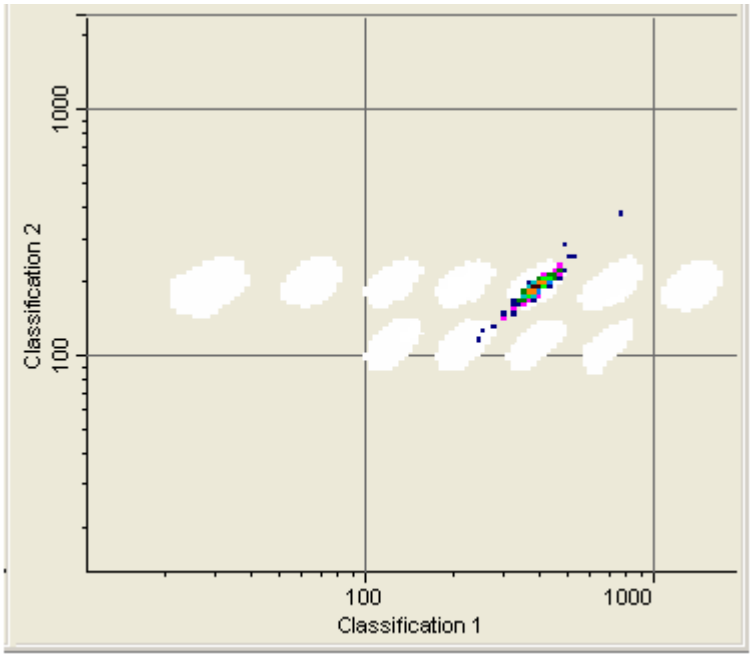


Figure 6.4. Classification plot of two internal dyes for selected sample in Yent/RRL titration study.

In Figure 6.4, notice that most of the beads for the selected sample were classified as belonging to set #20 (Yent/RRL probe). However, some beads were also "bleeding" into a neighboring spectral region, namely that belonging to set #13 (Bant/TOX). Thus, the instrument was apparently misclassifying beads from set #20 as belonging to set #13, possibly because of the close proximity of their spectral regions. This same phenomenon was observed in the titration study for other bead sets as well (data not shown). Misclassifications were more abundantly and frequently observed when the bead set involved in the analysis was diagonally juxtaposed with another bead set, particularly one whose spectral region resided "northeast" or "southwest" of the spectral region for the bead set undergoing analysis . For example, this problem was scarcely observed when bead sets #16, 17, and 18 were analyzed, because these bead sets had no "northeast" or "southwest" neighbors.

Luminex technical support was notified when misclassification problems were first encountered. In order to rule out contaminated buffer as a potential cause for the aberrant MFI values, select samples were re-analyzed with sheath fluid only (no TMAC buffer). However, the same unusual MFI values were observed even when the TMAC buffer was withheld (data not shown). Technical support then produced quality control data to illustrate that the misclassification rate being observed still fell within a range deemed acceptable by Luminex. Given that all misclassifications involved MFI values derived from only one to three beads, technical support advised continuous monitoring of the count data on all samples to ensure that MFI values were considered significant only when derived from more than thirty-five beads. Future researchers are encouraged, if possible, to use bead sets that are spectrally separated as much as possible to minimize the risk of misclassifications.

G. Enzymatic Degradation Study

1. Enzymatic Degradation of B. cereus

The MFI values for the original PCR products (data not shown) and the control group were generally consistent with each other, thus confirming again that the enzyme reaction buffer did not inhibit Luminex detection. For every amount of DNA tested, the MFI values for the experimental group were conspicuously higher than those observed for the control group (Table 5.5). Furthermore, the gap between the net MFIs of a control and experimental sample for a given amount of input DNA steadily widened as the amount of input DNA increased (Figure A5.5 in Appendix V). This was because the net MFI values for the experimental samples had a fast rate of change (linear regression line

slope = 21.9) whereas the net MFI values for the control samples had a slower rate of change (linear regression line slope = 7.8; Figure A5.6).

The MFI values for all samples increased proportionally to the amount of input DNA used, often covering a wide range between the 5 ng and 80 ng samples. For instance, the range of net MFI values for control samples was 620 (see Table 5.5). However, the range of net MFI values was much broader for the experimental group (1651; see Table 5.5), which indicated that the binding of single-stranded target sequences to probes was more efficient in experimental samples. In contrast, the attenuated range observed for control samples was indicative of binding efficiency that was consistent, although not as successful, as binding efficiency observed when the non-complementary PCR strand was absent.

The regression lines for the control and experimental samples displayed R^2 values that were both generally close to 1.0 (0.96 and 0.94, respectively; Figure A5.6). The good fit of the data to the linear trendline in both sample sets indicated that the rate of growth closely resembled a linear model. It was interesting that both control and experimental samples exhibited linear growth instead of logarithmic. This may be evidence that, had larger amounts of input DNA been tested, the curve for the *B. cereus* samples may have eventually plateaued. Nevertheless, the intriguing difference between the control and experimental samples for *B. cereus* was not their pattern of growth, but rather their rate of growth. The rate of growth was substantially higher for the net MFIs of experimental samples. Additionally, the magnitudes of the net MFI values for the experimental samples were higher overall than comparable net MFI values for the control samples. All of these observations provided further support for the belief that the enzymatic

degradation of competing sequences in the reaction mixture enhances the hybridization efficiency of probe and target DNA.

2. B. anthracis

For every amount of DNA tested, the MFI values for the experimental group were noticeably higher than those observed for the control group (Table 5.6). A similar trend was observed for other bacteria in the enzymatic degradation study. Another trend that was observed for the Bac/RRL and Bant/TOX targets was that the gap between the net MFIs for single-stranded DNA (digested) and double-stranded DNA (undigested) widened as larger amounts of DNA were used (Figure A5.7). When the net MFIs for RRL and TOX target sequences were examined, the single-stranded DNA had a faster rate of change than the corresponding double-stranded DNA. For example, the slope of the linear regression line for the digested RRL target was 22.56, but the slope of the line for undigested RRL target was only 5.23 (Figure A5.8). This trend was repeated for the TOX target – the slope of the line for digested target was 27.2, but only 14.8 for undigested target (Figure A5.8).

The MFI values for all samples increased in a manner that was proportional to the amount of input DNA used, often spanning a wide range between the 5 ng and 80 ng samples (Table 5.6). The range of net MFI values of the Bac/RRL and Bant/TOX targets in control samples were 420 and 1430, respectively). However, the range of net MFI values of the same target sequences in the experimental samples was much broader (1806 and 2331, respectively). The breadth of the net MFI ranges in experimental samples was attributed to the enhanced binding efficiency observed when sequences that compete with the probe for access to the target DNA were absent.

A comparison of the coefficients of determination was performed to determine if the data from digested or undigested data was a better fit to the linear model (Figure A5.8). The R^2 value of the net MFIs for digested Bac/RRL target was 0.96, but the R^2 value of the net MFIs for undigested Bac/RRL target was only 0.81. A better linear fit was also observed when the R^2 values of the net MFIs for the Bant/TOX target were examined – digested and undigested DNA had R^2 values of 0.86 and 0.54, respectively. For both target sequences in the control sample, the net MFIs plateaued near 20 ng of input DNA, thus causing them to deviate more from the linear model. In contrast, the rate of change for both target sequences in the experimental sample was more consistent across the entire range of input DNA. This constant rate of change not only yielded a better linear fit, but was also portentous of a better overall performance with larger amounts of input DNA.

3. Y. pestis

With the exception of the 5 ng sample (Ypes/TOX probe) and some samples that yielded zero net MFI for the Ypes/RRL probe, all experimental samples displayed net MFI values that were higher than those observed for the corresponding control samples (Table 5.7). Similar to the trend observed for other bacteria (Figures A5.5 and A5.7 in Appendix V), a large gap was observed between the net MFIs of digested and undigested DNA that further expanded when larger amounts of input DNA were used. However, this gap was not as pronounced and generally uniform for both target sequences in *Y. pestis* (Figure A5.9). The slope of the linear regression line could not be calculated for undigested Ypes/RRL samples (because no net MFI values for these samples exceeded the threshold), but the slope of the line for the net MFI values of digested Ypes/RRL

target was still only 4.62, indicating a very slow rate of change, especially for digested DNA (Figure A5.10). When the net MFIs for the TOX target sequences were examined, the double-stranded DNA had a slope of 27.10, which was only narrowly surpassed by the slope of 28.34 reported for the single-stranded DNA.

The MFI values for all *Y. pestis* samples (except those mentioned above) increased proportionally to the amount of input DNA used. Although the range of MFI values observed for single-stranded DNA was typically greater than the double-stranded DNA, it did not mirror the wide ranges observed with *B. cereus* and *B. anthracis*. The range of net MFI values of the Ypes/RRL and Ypes/TOX targets in control samples were 0 and 2037, respectively; Table 5.7). The range of net MFI values of the same target sequences in the experimental samples was somewhat broader (325 and 2485, respectively; Table 5.7). These ranges were acceptable, but far from being vastly superior to the ranges observed in control samples. The data indicated that the *Y. pestis* probes did bind their target DNA with greater efficiency when competing DNA was absent, but the enzymatic degradation of the competing DNA did not promote hybridization of the probe and target as much as expected, or as much as it had for other bacteria.

The R^2 values of the linear regressions were compared for digested and undigested *Y. pestis* samples to examine the ability of the data to fit the linear model (Figure A5.10). The R^2 value for the undigested Ypes/RRL samples was not computed, but the R^2 value for digested Ypes/RRL samples was 0.97. The R^2 value of the net MFI values for undigested Ypes/TOX DNA (0.99) was better than the R^2 value for digested Ypes/TOX DNA (0.78). The poor R^2 value for the digested Ypes/TOX target was the result of the plateau in the curve (see Figure A5.9). This was surprising because data

from previous assays indicated that the net MFIs for digested DNA maintained a linear trend even when large quantities of input DNA were tested, whereas the net MFIs for undigested DNA tended to exhibit a more logarithmic trend that plateaued at a certain input amount. Although the *Y. pestis* results exhibited some unusual trends, they still corroborated the main conclusion of this study, which was that enzymatically digested DNA binds its probe with greater efficiency when competing sequences are eliminated.

Even though the overall conclusions of the enzymatic degradation study were upheld despite some minor aberrations observed in the data for *Y. pestis*, it was still necessary to investigate the cause of the diminished hybridization efficiency between the Ypes/RRL probe and target DNA. The net MFIs for the Ypes/RRL target were conspicuously lower than those observed for other bacteria in the enzymatic degradation study. It was believed that the diminished binding efficiency for this probe might have been caused by formation of secondary structure in the probe or probe-binding region of the target. Secondary structures might diminish MFI values for a given probe because they prevent optimum hybridization efficiency.

When the primers and probes were initially designed, all sequences were uploaded to an AutoDimer file to check for self-complementarity and cross-reactivity with other sequences. The default threshold value of seven was used which meant that no potential primer dimers or hairpins were reported unless their score exceeded seven (AutoDimer scores are computed by subtracting the number of mismatched bases from the matching ones). However, interactions with scores below the threshold could still occur. In order to further examine the predilection for dimerization or hairpin formation, the Ypes/RRL probe and Ypes/RRL target sequence were uploaded to AutoDimer and

analyzed for hairpins using descending threshold values lower than seven but greater than three. The hairpin screen predicted one hairpin in the Ypes/RRL target sequence (Figure A11.1 in Appendix XI). This hairpin was not detected in the initial AutoDimer screen because the threshold was seven and the aforementioned hairpin returned a score of four.

The propensity for secondary structure formation could have possibly been reduced by raising the hybridization temperature, thus favoring more specific binding. However, the hybridization parameters were optimized to accommodate all sequences in this study, and it was not possible to raise the temperature without going above the melting temperature for other sequences.

Regardless of the cause, it was not felt that the low response for the Ypes/RRL bead set would inhibit an identification of *Y. pestis*. A positive response with the Ypes/TOX bead set was the most reliable indicator of *Y. pestis*, and the Ypes/TOX bead set consistently displayed a strong response to *Y. pestis*.

H. Specificity/Sensitivity Study

1. B. cereus

Enzymatically digested samples of *B. cereus* were analyzed on the Luminex instrument. The LLD for the Bac/RRL probe was 1 ng (Table A6.3 in Appendix VI). There were some false positives but the net MFIs were relatively low and only occurred when large amounts of target DNA were analyzed. For example, there were no false positives for any samples in which 5 ng of target DNA (or less) was loaded into the Luminex well. The most prominent group of false positives (in terms of frequency and magnitude) occurred with the Cdif/RRL bead set.

Overall, the data were consistent with data previously collected for single-stranded target DNA of *B. cereus* in the enzymatic degradation study (Table 5.5). In both this assay and the enzymatic degradation study of *B. cereus*, the Bac/RRL probe displayed net MFI values that reached close to 2000 (for the 80 ng sample). When digested *B. cereus* DNA was analyzed in the enzymatic degradation study, the slope of the linear regression and R^2 value for the net MFIs corresponding to the Bac/RRL probe were 21.9 and 0.94, respectively (Figure A5.6). A similar slope and R^2 value were observed for the same probe in the specificity/sensitivity study of *B. cereus*: 27.0 and 0.85 (Figure A6.3). In addition to verifying the reproducibility of the data, the slopes of the regression lines confirmed the strong sensitivity of the Bac/RRL probe for its complementary target DNA. The R^2 values demonstrated the data's close adherence to the linear model, which was a reflection of the data's consistent rate of change.

2. B. anthracis

Enzymatically digested samples of *B. anthracis* were analyzed on the Luminex instrument. The LLD for the Bac/RRL probe was 1 ng and the LLD for the Bant/TOX probe was less than 0.1 ng (Table A6.4). There were a few false positives but the net MFIs were relatively low and only occurred when very large amounts of target DNA were used. For example, there were no false positives for any samples in which 20 ng of target DNA (or less) was loaded into the Luminex well. Again, the most prominent group of false positives was observed for the Cdif/RRL bead set.

The data were consistent with data previously collected for single-stranded target DNA of *B. anthracis* in the enzymatic degradation study (Figure 5.6). When digested *B. anthracis* DNA was analyzed in the enzymatic degradation study, the slope of the linear

regression and R^2 value for the net MFIs corresponding to the Bac/RRL probe were 22.56 and 0.96, respectively (Figure A5.8). A similar slope and R^2 value were observed for the same probe in the specificity/sensitivity study of *B. anthracis*: 24.0 and 0.82 (Figure A6.4). Likewise, the linear regression line for the Bant/TOX probe in the enzymatic degradation study displayed a slope and R^2 value of 27.2 and 0.86, respectively (Figure A5.8). Although the slope of the linear regression line (41.1; Figure A6.4) indicated better sensitivity to the Bant/TOX probe in the specificity/sensitivity study, the R^2 value (0.56) was demonstrative of an inconsistent rate of change, which was caused by the plateau in net MFIs that was observed around 20 ng of input DNA.

Other evidence supporting the reproducibility of the data was that the Bac/RRL and Bant/TOX probes in both studies of *B. anthracis* displayed net MFI values that reached close to 2000 and 4000, respectively (for the 80 ng sample). One might have expected a stronger fluorescent response for the Bac/RRL probe because the Bac/RRL target sequence is present eleven times in *B. anthracis*, but the Bant/TOX target sequence is present only once. Theoretically this discrepancy would have been at least partially carried over to the digested DNA that was analyzed on the Luminex, because a template concentration study was not performed to ensure the homogeneity of amplicon concentrations. However, the precise amount of Bant/TOX target analyzed is difficult to pinpoint because the amount of DNA analyzed in each Luminex assay was based solely on post-digest quantitation results of the RRL band (not the TOX band).

The fact that the fluorescent response for the Bac/RRL probe was not stronger may have been due to cross-reactivity of the Bac/RRL target sequence with other probes. However, this theory was neither predicted by the AutoDimer primer dimer screen nor

supported by the net MFI values for other probes in this assay, which demonstrated very low levels of cross-reactivity. The attenuated fluorescent response of the Bac/RRL probe might have also been due to tendencies for secondary structure formation within the primer, probe, or target sequences. Even though hairpin formation within these sequences was not predicted in the AutoDimer hairpin screen, other secondary structures (besides hairpins) could have formed within these sequences which might have led to a diminished fluorescent response of the Bac/RRL probe to its target DNA.

3. C. botulinum

Enzymatically digested samples of *C. botulinum* were analyzed on the Luminex instrument. The LLD for the Cbot/TOX bead set was 0.5 ng and the LLD for the Cbot/RRL bead set was 5 ng (Table A6.5). There were some false positives but the net MFIs were relatively low and were generally observed only when large amounts of target DNA were used. For example, there were no false positives for any samples in which 5 ng of target DNA (or less) was loaded into the Luminex well.

Two prominent groups of false positives were observed: those of the Bant/TOX and Cdif/RRL bead sets. False positives with the Cdif/RRL bead set were commonly observed with other samples and were investigated further at the end of this study. False positives with the Bant/TOX bead set may have been due to the Cbot/TOX and/or Cbot/RRL target sequences cross-reacting with the Bant/TOX probe. An AutoDimer primer dimer screen (threshold of 3) of the Bant/TOX probe sequence against the Cbot/TOX and Cbot/RRL target sequences was performed. The results predicted cross-reactivity between the Bant/TOX probe and the Cbot/RRL target sequence, and between the Bant/TOX probe and the Cbot/TOX target sequence (Figure A11.2).

The peak MFI value for the Cbot/TOX probe (4003; Table A6.5) and the slope of the linear regression line (46.4) both support the superior sensitivity of the Cbot/TOX probe (Figure A6.5). The R^2 value (0.67) reflected a moderate deviation of the data from the linear model. This represented a departure from a uniform rate of change, but was likely caused by the plateau in the net MFI values as larger quantities of DNA were used.

The fluorescent response of the Cbot/RRL probe was much weaker, displaying a maximum net MFI value of 298 within the amounts tested (Table A6.5). Although a relatively constant rate of change was confirmed by the R^2 value (0.82), the slope of the linear regression line was just 4.0, providing further evidence for the lack of sensitivity for the Cbot/RRL probe (Figure A6.5).

The unusually low net MFI values for the Cbot/RRL probe were not likely to have been caused by hairpin formation within the probe attached to the bead. The reverse probe was evidently binding without difficulty (see positive control net MFI of 4676; Table A6.5). The low net MFI response for the Cbot/RRL bead set could have been due to formation of hairpins that were undetected by the initial AutoDimer hairpin screen when a threshold of seven was used. The AutoDimer hairpin screen for the Cbot/RRL target DNA was repeated using a threshold of three. The results indicated three possible hairpin-forming structures within the Cbot/RRL target sequence (Figure A11.3).

4. F. tularensis holarctica

Enzymatically digested samples of *F. tularensis holarctica* were analyzed on the Luminex instrument. The LLD for this bacterium was 0.5 ng (Table A6.6). One group of false positives occurred with the Cdif/RRL bead set, a trend that was observed in other assays and discussed further below. A few other scattered false positives were observed,

but these were relatively low and only observed when large amounts of target DNA were used. For example, excluding the Ftul/TOX bead set, there were no false positives for any samples in which 10 ng of target DNA (or less) was loaded into the Luminex well.

The Fran/RRL probe displayed an adequate level of sensitivity, having net MFI values that reached 2614 (for the 80 ng sample) and a linear regression slope of 32.7 (Figure A6.6). The net MFI values pleateaued around 20 ng of input DNA, thus causing the R^2 value (0.62) to reflect an inconsistent rate of change.

The most prominent group of false positives were those observed within the Ftul/TOX bead set, which displayed consistently high MFIs even when as little as 0.5 ng of target DNA was used. Recall from the monoplex and multiplex PCR studies that a non-specific amplicon within the general size range of the Ftul/TOX target sequence was observed when *F. tularensis holarctica* DNA was used as template.

Although initially not detected, further investigations into the fully sequenced genome of *F. tularensis holarctica* revealed that a 98% identical form of the Ftul/TOX target sequence was indeed present in the *F. tularensis holarctica* genome (Figure 6.1). The primer-binding regions for the Ftul/TOX target sequence were identical in *F. tularensis tularensis* and *F. tularensis holarctica*, thus accounting for the non-specific amplicon observed when *F. tularensis holarctica* was amplified with the Ftul/TOX primer set.

Despite the total homology between the two Francisella bacteria in the Ftul/TOX target sequence primer-binding region, there were two polymorphisms between the Ftul/TOX probe and the probe-binding sequence in the non-specific amplicon from *F. tularensis holarctica*. Other researchers (Diaz and Fell, 2004; Deshpande *et al.*, 2010)

have reported their ability to distinguish binding with probes between which there is a single nucleotide polymorphism (SNP). Therefore, it was thought that these two polymorphisms might be sufficient to allow the Luminex platform to prevent non-specific binding.

The non-specific target DNA was able to bind the FtuI/TOX probe with surprisingly great efficiency (despite the two polymorphisms), with the net MFI values ranging from 85 to 3043. However, the net MFI values observed when FtuI/TOX target (generated from *F. tularensis tularensis*) and FtuI/TOX probe bind had a slightly higher range, spanning from 623 to 3827 (Table A6.7). The difference between net MFI values of the actual complement and the non-specific amplicon are noticeable but not nearly sufficient for distinguishing the binding of the true complement from false positives due to *F. tularensis holarctica*.

Ikuta *et al*. (1987) found that the melting temperature for the hybridization of a perfectly matched probe and target sequence may differ by several degrees when compared to the melting temperature for a probe and target sequence with a SNP. However, the difference in melting temperatures between perfectly matched sequences and single base mismatched sequences was difficult to exploit in the current research because a multiplexed assay requires a hybridization temperature that accommodates the melting temperatures of all sequences. Furthermore, the two polymorphisms in the probe-binding region of the *F. tularensis holarctica* non-specific amplicon reside at bases three and eighteen, respectively. Prior research has shown that SNPs are most effective in preventing cross-reactivity when they reside between positions eight and fourteen of a twenty-base probe (Dunbar and Jacobson, 2005).

5. F. tularensis tularensis

Enzymatically digested samples of *F. tularensis tularensis* were analyzed on the Luminex instrument. The LLD for the Fran/RRL bead set was 0.5 ng and the LLD for the Ftul/TOX bead set was also 0.5 ng (Table A6.7). There were some false positives but the net MFIs were relatively low and generally only observed when large amounts of target DNA were used. For example, there were no false positives for any samples in which 5 ng of target DNA (or less) was loaded into the Luminex well. The Cdif/RRL false positive trend was again observed, which was addressed at the end of this section.

Although the R^2 values for the linear regressions of the Fran/RRL and Ftul/TOX net MFIs were mediocre (0.76 and 0.54, respectively), the slopes of the linear regressions (34.5 and 37.4, respectively) indicated a reasonable degree of sensitivity (Figure A6.7). This was further confirmed by the net MFI values for these probes, which reached close to 3500 (for the 80 ng sample). Unfortunately the two-base polymorphism between the non-specific amplicon in *F. tularensis holarctica* and the Ftul/TOX probe was not enough to enable a differentiation of the two Francisella species. However, there was minor difference in the sensitivity of this probe for its real target and the non-specific amplicon, albeit too minute to enable a reliable differentiation. The MFIs for the Ftul/TOX probe when *F. tularensis tularensis* was analyzed had a range of 3204 and reached as high as 3827 (for the 80 ng sample; Table A6.7). In contrast, the net MFIs for the Ftul/TOX probe when *F. tularensis holarctica* was analyzed had a range of 2958 and reached as high as 3043 (for the 80 ng sample; Table A6.6).

6. Y. enterocolitica

Enzymatically digested samples of *Y. enterocolitica* were analyzed on the Luminex instrument. The LLD for this bacterium was 10 ng (Table A6.8). There were some false positives but the net MFIs were relatively low and generally occurred only when very large amounts of target DNA were used. For example, there were no false positives for any samples in which 20 ng of target DNA (or less) was loaded into the Luminex well. The Cdif/RRL false positive trend was again observed in the table below. This trend was further explored at the end of this section.

The uniform rate of change for the net MFIs corresponding to the Yent/RRL probe was confirmed by the R^2 value of 0.99 (Figure A6.8). However, the slope of the linear regression (4.7) and the highest net MFI observed (364 for the 80 ng sample) both attest to the weak sensitivity of this probe. The data did not indicate secondary structure formation within the probe (attached to the bead), because the reverse probe was evidently binding with reasonable efficiency, having a net MFI of 2553 (Table A6.8). These low MFI values may have been due to hairpin formation within the Yent/RRL target sequence. These hairpins may not have initially been predicted when the AutoDimer hairpin screen was performed using the default threshold of seven. An AutoDimer hairpin screen of the Yent/RRL target sequence was repeated using a threshold value of three (Figure A11.4).

The probe sequence is further downstream of the secondary structure in the Yent/RRL target sequence shown in Figure A11.4. Although the probe sequence is not involved in the secondary structure, binding of the probe could still be inhibited

somewhat because of steric hindrance, thus accounting for the low (though not completely absent) MFI response for the Yent/RRL probe.

7. Y. pestis

Enzymatically digested samples of *Y. pestis* were analyzed on the Luminex instrument. The LLD for the Ypes/RRL bead set was 10 ng and the LLD for the Ypes/TOX bead set was 0.5 ng (Table A6.9). There were some false positives but the net MFIs were relatively low and generally observed only when very large amounts of target DNA were used. For example, there were no false positives for any samples in which 40 ng of target DNA (or less) was loaded into the Luminex well. The false positive trend with the Cdif/RRL bead set was again observed, which is discussed at the end of this section.

The data were consistent with data previously collected for single-stranded target DNA of *Y. pestis* in the enzymatic degradation study (Table 5.7). When digested *Y. pestis* DNA was analyzed in the enzymatic degradation study, the slope of the linear regression and R^2 value for the net MFIs corresponding to the Ypes/RRL probe were 4.62 and 0.97, respectively (Figure A5.10). A similar slope and R^2 value were observed for the same probe in the specificity/sensitivity study of *Y. pestis*: 1.28 and 0.80 (Figure A6.9). Likewise, the linear regression line for the Ypes/TOX probe in the enzymatic degradation study displayed a slope and R^2 value of 28.3 and 0.73, respectively (Figure A5.10). The R^2 value of the linear regression line for the Ypes/TOX probe in the specificity/sensitivity study was 0.72, indicating a consistent rate of change in both assays. However, the slope of the linear regression line (41.0) indicated better sensitivity to the Ypes/TOX probe in the specificity/sensitivity study (Figure A6.9).

Other evidence supporting the reproducibility of the data was that the net MFIs for the Ypes/RRL and Ypes/TOX probes in both studies of *Y. pestis* reached close to 200 and 3500 MFI for the 80 ng sample, respectively (Tables 5.7 and A6.9). One might have expected a stronger fluorescent response for the Ypes/RRL probe because the Ypes/RRL target sequence is present six times in *Y. pestis*, but the Ypes/TOX target sequence is present only once. Theoretically this discrepancy would have been at least partially carried over to the digested DNA that was analyzed on the Luminex, because a template concentration study was not performed to make the amplicon concentrations homogeneous. However, because the amount of DNA analyed in each Luminex assay was based solely on post-digest quantitation results of the RRL band (not the TOX band), it was difficult to pinpoint the precise amount of Ypes/TOX target actually analyzed.

One potential cause for the attenuated fluorescent response with the Ypes/RRL probe could have been secondary structure formation within the probe. However, this theory could not be corroborated by the data; the net MFI for the positive control (4000) confirmed that the probe was binding the reverse probe with adequate efficiency (Table A6.9). The low MFI values observed with the Ypes/RRL probe may have been due to hairpin formation within the Ypes/RRL target sequence. These hairpins may not have initially been predicted when the AutoDimer hairpin screen was performed using the default threshold of seven. An AutoDimer hairpin screen of the Ypes/RRL target sequence was repeated using a threshold value of three (Figure A11.1).

As with the hairpin predicted for the Yent/RRL target sequence (Figure A11.4), the probe-binding sequence in the Ypes/RRL target sequence is downstream from the hairpin (Figure A11.1). Although this may not completely prevent binding of the probe to

its complementary sequence, binding of the probe could be inhibited due to steric hindrance imparted by the hairpin.

Note that the secondary structure predicted using the lower threshold in Figure A11.1 has a very low melting temperature (19°C) relative to the perfectly matched probe (56.3°C; see Table 4.3). This was a common observation for most secondary structures depicted in Appendix XI. Duplexes with such low melting temperatures should theoretically be unstable at the hybridization temperature used (48°C). However, the position of complementary bases relative to the center of the binding sequences may be the overriding factor that permits partial formation of the duplex even when the oligo should, in theory, favor the single-stranded state given the high hybridization temperature.

8. M. tuberculosis

Enzymatically digested samples of *M. tuberculosis* were analyzed on the Luminex instrument. The LLD for this bacterium was 0.2 ng, which made the *M. tuberculosis* RRL probe one of the most sensitive in this study (Tables A6.10 and 6.1). There were some false positives in the *M. tuberculosis* specificity/sensitivity study, most notably with the Cdif/RRL bead set. Many of these false positives displayed net MFIs whose magnitudes were greater than those corresponding to the Myc/RRL probe. False positives in the Cdif/RRL bead set were observed in this study for other bacteria. With the small amounts of DNA tested in this sensitivity study, the Cdif/RRL false positives would make it difficult to distinguish between *C. difficile* and *M. tuberculosis*. Perhaps if larger amounts of DNA were available, the ability to distinguish *M. tuberculosis* would be improved.

Analyzing *M. tuberculosis* on the Luminex presented a unique situation because the amount of available digested *M. tuberculosis* target DNA was very low compared to the amounts of target DNA available for other bacteria in this study. This was probably due to the fact that all of the other bacteria in this study had multiple copies of the *rrl* gene (and thus more DNA available for amplification), but *M. tuberculosis* has only one copy. There was so little *M. tuberculosis* target DNA in the enzymatically digested sample that its concentration was not even reported by the Agilent software. However, a band was visible in the gel (Figure A6.2). The LLD for the Agilent assay is 0.02 ng/uL, and because a faint band was observed, it was assumed that approximately 0.05 ng/uL (just slightly above the Agilent LLD) of *M. tuberculosis* target DNA was available for the Luminex assay. This very low concentration did not permit the typical spectrum of concentrations to be tested in the sensitivity study, which usually spanned from 0.1 to 80 ng. Instead, very small increments were tested between 0.05 and 0.85 ng. Theoretically, even the lowest increment (0.05 ng) should yield a positive result in the Luminex assay because the LLD in the titration study for the Myc/RRL probe was 0.25 fmol, and 0.05 ng = 3.85 fmol, which is well above the LLD for this probe.

The superior sensitivity of the Myc/RRL probe was supported not only by the very low LLD but also by the slope of the linear regression line (198.3; Figure A6.10), even though the linear fit was poor ($R^2 = 0.54$). It is questionable whether this extraordinary level of sensitivity would have been maintained beyond the 0.85 ng input DNA amount, because the net MFI for the reverse probe still peaked around the same area as other reverse probes (~5000; Table A6.10).

When this study was initially designed, *M. tuberculosis* was selected as an outlier because it resides within the phylum Actinobacteria whereas other bacteria studied reside in the phlya Firmicutes and Proteobacteria. The distant genetic relatedness of *M. tuberculosis* was not reflected in the specificity/sensitivity study, with some probes cross-reacting with the Myc/RRL target DNA, and with the Myc/RRL probe cross-reacting with some other target sequences. Although these instances of cross-reactivity were generally mild and easily distinguishable from perfect matches (excluding those corresponding to the Cdif/RRL probe), it is still a testament to the high degree of homology in the *rrl* gene. The inability of the Myc/RRL probe to serve as a true outlier prompted its exemption from the mixture studies.

9. C. difficile

Enzymatically digested samples of *C. difficile* were analyzed on the Luminex instrument. The LLD for this bacterium was 1 ng (Table A6.11). There were some false positives but the net MFIs were relatively low and only occurred when large amounts of target DNA were used. For example, there were no false positives for any samples in which 20 ng of target DNA (or less) was loaded into the Luminex well.

The net MFI values for the Cdif/RRL probe reached as high as 1526 (for the 80 ng sample). Although greater sensitivity was observed for some other probes throughout the entire specificity/sensitivity study, the Cdif/RRL probe still displayed a strong response to its target sequence. The slope of the linear regression line for the Cdif/RRL probe was 20.3, indicating a good level of sensitivity for the Cdif/RRL target DNA (Figure A6.11). The data exhibited a consistent rate of change, as supported by the R^2 value of 0.86.

The cross-reactivity of the Cdif/RRL bead set with several other target sequences was an ongoing problem in the specificity/sensitivity study. The source of this cross-reactivity was explored by performing an AutoDimer primer dimer screen (threshold set to 3) in which the Cdif/RRL probe was checked for cross-reactivity with all target sequences used in this study. The results predicted that the Cdif/RRL probe would cross-react with nine target DNA sequences, including Bac/RRL, Cbot/RRL, Cbot/TOX, Fran/RRL, Ftul/TOX, Yent/RRL, Ypes/RRL, and Myc/RRL (Figure A11.5). Many of these predicted secondary structures have melting temperatures that are very low relative to melting temperature for the perfectly matched duplexes (Tables 4.3 and 4.7). Theoretically these secondary structures should have been unstable given the relatively high assay hybridization temperature (48°C) but sequences with partial base complementarity still could have experienced transient or incomplete duplex formation, either of which could negatively impact the specificity of the Cdif/RRL probe.

The cause of the cross-reactivity observed with the Cdif/RRL probe may be related to this probe's GC content, and thus, its melting temperature. All probes in this study had GC contents ranging from 25 to 60% and melting temperatures that ranged from 50.2 to 64.5°C. The Cdif/RRL probe had a GC content of 25% and a melting temperature of 50.2°C, thus placing it at the lower end of the spectrum in both categories. In order to accommodate the broad range of melting temperatures possessed by the various probes, a hybridization temperature just below the lowest melting temperature (48°C) was selected during the optimization phase of this study. The use of a hybridization temperature that was so close to the melting temperature of the Cdif/RRL

probe may have enhanced this probe's tolerance for basepair mismatches, thus having a negative effect on the specificity of the Cdif/RRL probe.

Several options were considered for improving the specificity of the Cdif/RRL probe: (1) extending the length of the probe longer than 20 bases to incorporate more polymorphisms, (2) re-designing the probe in a more polymorphic region of the *rrl* gene, (3) shifting the location of the Cdif/RRL probe to incorporate more Gs and Cs with the intent of raising the GC content enough so that the melting temperature could be closer to the range observed for most other probes in this study, or (4) adjusting hybridization conditions (*e.g.*, reducing salt concentration of the hybridization buffer or lowering the hybridization temperature below 48°C).

The aforementioned option #1 was jettisoned because close inspection of the bases flanking the probe-binding region indicated a very high degree of homology across the eleven microbes. Thus, it was unlikely that extending the length of the probe would improve specificity.

Option #2 was also not considered to be a viable solution because there was a very high degree of homology within the *rrl* gene, owing to its key role in translation and low mutation rate. It would be exceedingly difficult (if not impossible) to find a single region in the *rrl* gene in which eleven bacteria are totally polymorphic. The region of the *rrl* gene in which the probes were designed had been studied extensively and used successfully by other researchers (Dunbar and Jacobson, 2007; Battaglia *et al.*, 2011). The findings of these other researchers as well as the bioinformatics data accumulated in the preliminary phases of this study (not shown) all supported the belief that the region of

the *rrl* gene in which the probe sequences were designed was one of the most polymorphic regions of the *rrl* gene.

Option #3 was a more viable solution to the specificity issues observed with the Cdif/RRL probe. Future researchers should take all possible measures to select highly polymorphic probes that have uniform melting temperatures. If an outlier (in terms of its melting temperature) is observed, shifting the probe-binding region to incorporate more Gs and Cs may be the best alternative. Although option #3 could have potentially resolved the cross-reactivity issues with the Cdif/RRL probe, it was a complicated solution that would have required synthesis of new oligos, re-coupling of the new oligos to the beads, a titration study, etc. Reagent limitations precluded further exploration of this solution but future researchers may wish to bear it in mind.

Option #4 was a fast way of potentially resolving the Cdif/RRL cross-reactivity problems, and it was easy enough to test in a single assay. Lowering the salt concentration should theoretically improve specificity because salts contain positively charged ions which shield the negatively charged phosphate groups on the DNA bases. Thus, a lower salt concentration releases this inhibition on the bases and makes them more available to bind the target DNA. Adjustments to the salt concentration of the hybridization buffer were not attempted because prior tests in which the buffer salt concentration was modified were unsuccessful in reducing cross-reactivity (Battaglia, 2009). However, lowering the hybridization temperature was not only a quick modification that was easy to test, but it also had a strong likelihood of resolving the observed cross-reactivity issues.

The specificity/sensitivity study for *C. difficile* was carried out exactly as before, with the only modification being that the hybridization temperature was lowered to 45°C. This new temperature was well below the Cdif/RRL probe's melting temperature (50.2°C) and was expected to help improve specificity with the Cdif/RRL probe. The only concern was that the specificity of other probes could be compromised. Other probes in this study had melting temperatures as high as 60°C, and specificity is generally best when hybridization temperatures are *right* below the probe's T_M. The data for this assay is shown in Table A6.12 and can be compared to the data already obtained in the first *C. difficile* specificity/sensitivity study (Table A6.11).

The results obtained when the modified hybridization temperature was used, when compared to the data from the first specificity/sensitivity study with *C. difficile*, demonstrated that lowering the hybridization temperature to 45°C improved the specificity of the Cdif/RRL probe. This was supported by the absence of false positives in all samples. However, the net MFI values in the first assay (48°C hybridization temperature) reached as high as 1526 (for the 80 ng sample), whereas the net MFI values in the repeated assay (45°C hybridization temperature) only reached 974 (for the 80 ng sample). The sensitivity of the Cdif/RRL probe was also better for the original assay in which the 48°C hybridization temperature was used. The LLD in the first assay was 1 ng but the LLD in the repeated assay using the 45°C hybridization temperature was 5 ng.

Aside from examining the effect of the new hybridization temperature on the specificity and sensitivity of the Cdif/RRL probe, it was most important to determine how other probes would respond to the lower hybridization temperature. Enhanced specificity was observed with the Cdif/RRL probe, but it was considered that this might not

necessarily translate to similar results with other probes, especially those with melting temperatures much higher than 45°C. When the 45°C hybridization temperature was tested with other probes (data not shown), the results indicated that there was a negative impact on performance and specificity. For instance, false positives more than doubled when the lower hybridization temperature was used. Furthermore, the net MFI values of these false positives were greater in magnitude than originally observed.

Although the 45°C hybridization temperature may have improved the specificity of the Cdif/RRL probe, it was detrimental to the specificity of the other ten probes. The original 48°C hybridization temperature became the standard for all future assays, with the stipulation that the specificity of the Cdif/RRL probe may, at times, show diminished specificity.

10. Comparison of LLDs

Dunbar and Jacobson (2007) reported LLDs ranging from 30 to 80 ng of PCR product. The sensitivities reported by Battaglia *et al.* (2011) were noticeably better, ranging from <0.25 to 2.0 ng of PCR product. The former study separated PCR amplicons on a 4% agarose gel, but concentrations were determined by densitometry of ethidium bromide-stained gels using the GelDoc-IT Imaging System and accompanying software. The latter study also relied on a 4% agarose gel to separate the PCR amplicons, but the concentrations were determined by visual comparison to standards in a DNA ladder. This method of quantitation is simple and cost-effective, but introduces an element of ambiguity that may compromise accuracy. The inherent subjectivity involved in visual interpretation of gels makes it difficult to compare LLDs reported for studies in which more accurate quantitation methods were relied upon.

PCR amplicons in the current study were quantified using the Agilent 2100 Bionanalyzer. This is a microfluidics-based method of quantitation in which samples migrate through microscopic channels filled with a proprietary sieving polymer when subjected to a voltage gradient. A fluorescent dye is used to label the analytes, whose migration times are compared to those of known standards (*i.e.*, ladder). The Agilent DNA 1000 kit has a sensitivity threshold of 0.02 ng/uL and a sizing accuracy of +/- 5% for oligos between 100 and 500 bp (Agilent, 2011).

Table 6.1 provides a summary of the LLDs observed for each probe in the present study. Sensitivities reported by the Agilent method of quantitation in the current study ranged from 0.1 to 10 ng of PCR product. The sensitivity levels were far superior to those reported by Dunbar and Jacobson (2007). Many of the LLDs in the current study were consistent with or better than those reported by Battaglia *et al.* (2011). As mentioned above, it is difficult to compare sensitivity levels for two studies in which different quantitation methods were used. The sensitivity limit in the current study may be better than it appears by comparison to the LLDs in other studies. Sensitivity was conspicuously improved by digesting the non-complementary PCR strand with lambda exonuclease.

	Probe												
	BC	***BA***		***CD***	***CB***		***FH***	***FT***		***YE***	***YP***		***MT***
	rrl	rrl	tox	rrl	rrl	tox	rrl	rrl	tox	rrl	rrl	tox	rrl
LLD (ng)	1	1	0.1	1	5	0.5	0.5	0.5	0.5	10	10	0.5	0.2

Table 6.1. Summary of LLDs for each probe.

I. Mixture Study (1:1)

In another study (Battaglia *et al.*, 2011), target DNA sequences were individually amplified and mixed post-amplification before being analyzed on the Luminex instrument. In order to create a more realistic scenario in the present study, the template DNA sequences were mixed pre-PCR, amplified together, and then analyzed on the Luminex instrument. Although this presented a more realistic situation, it was difficult to control how much DNA was loaded into each Luminex sample well, because all components of a mixed PCR product inevitably had concentrations that did not necessarily reflect the PCR input ratios of the template DNA. Although early optimization studies indicated an approximately uniform ratio of amplicons for the nine bacteria (provided that equal amounts of template were used), this uniformity was not exact nor could it be guaranteed in every assay.

Due to the inability of the Agilent Bioanalyzer to accurately quantify single-stranded DNA, it was determined that the most reliable method for loading enzymatically digested (single-stranded) DNA on to the Luminex plate was to base the concentration of the digested DNA on the concentration of a corresponding control sample, which contained undigested, double-stranded DNA. Another reason that this was thought to be a reliable method was because the amount of target DNA in control and experimental samples was expected to be identical, the only difference between the two being that the non-complementary strand was degraded in the digested sample.

Quantifying the amount of DNA in the control tubes was a reliable method of gauging Luminex input amounts, but the original method as applied in the specificity/sensitivity study could not be exactly replicated in the mixture study. In the

mixture study, the likelihood of having several components, each with a different concentration of amplicons in a single PCR tube made it difficult to know the exact concentration of each component. The concentration of the RRL band(s) in the control tubes was still considered a reliable basis for estimating the concentration of the target DNA in the experimental tubes, provided that slight accommodations were made for each possible scenario: (1) if only one RRL band was detected in a given control tube, the concentration of digested target DNA in the corresponding experimental tube was estimated by dividing the concentration of the sole RRL band by the number of components in the mixture; (2) if one RRL band was detected for each component of the mixture, the concentration of digested target DNA in the experimental tube was estimated by taking the average concentration of all RRL bands observed in the control sample. In either case, the outcome was a value that represented the approximate concentration of *each* component in the mixture (see Tables A7.1, A7.3, and A7.5).

1. Batch #1

The net MFIs (see Table A7.2 and Figures 5.18-5.28) reported by the four TOX probes in Batch #1 (Table 4.17) were particularly reliable. The two probes designated for identifying *B. anthracis* and the two probes designated for identifying *F. tularensis* were some of the most reliable in this study, based on the absence of false negatives and the low occurrence of false positives (the only false positive with these probes was observed for the FT/CB sample which had a net MFI of 14 for the Bant/TOX probe).

The fluorescent response of most RRL probes was weak, but the Cbot/RRL and Ypes/RRL probes, in particular, yielded net MFI values that were substantially lower than the values obtained for other probes. In fact, the only false negatives in this entire

assay were two samples (BA/YP and YP/FT; Figures 5.18 and 5.21) for which the Ypes/RRL probe should have had a positive net MFI response but did not. This may have been due to possible hairpin formation within the Cbot/RRL and Ypes/RRL target sequences, as predicted by AutoDimer hairpin screens (Figures A11.1 and A11.3). Issues with these RRL probes would not be expected to preclude identification of *C. botulinum* or *Y. pestis* because the toxin gene probes for these bacteria consistently yielded reliable MFI data.

The Ypes/TOX probe sometimes reported a positive net MFI when *Y. pestis* DNA was absent but other target DNA sequences were present (*e.g.*, see Ypes/TOX net MFIs for BA/FT, FT/CB, and BA/FT/CB samples in Table A7.2). This anomaly was occasionally observed in the specificity/sensitivity study. An AutoDimer primer dimer screen of the Ypes/TOX probe versus all target sequences was performed using a threshold of three. Three potential interactions between the Ypes/TOX probe and non-specific target sequences were predicted (Figure A11.6).

Although result interpretation could be complicated if the components are not present in a 1:1 mixture, false identifications of *Y. pestis* can generally be avoided by comparing the magnitude of a net MFI value for the Ypes/TOX probe when *Y. pestis* is actually present to a net MFI value when *Y. pestis* has been falsely identified, with the latter MFI being much lower than a truly positive result.

2. Batch #2

In batch #2 (Table 4.17), the only false negatives were the three observed for the Cbot/RRL probe (Table A7.4 and Figures 5.29-5.39). In the initial AutoDimer analysis performed at the beginning of the bioinformatics study (which used a default threshold

value of seven), no probes were identified as dimer-forming. However, the Cbot/RRL probe was re-analyzed with a threshold of three and three potential hairpin-forming sequences were identified (Figure A11.3). The second sequence in Figure A11.3, in particular, would have been amenable to the assay hybridization temperature (48°C) given its melting temperature (46.8°C). The AutoDimer results for the Yent/RRL and Ypes/RRL probes (Figures A11.1 and A11.4) may also explain why the net MFIs for these probes were somewhat lower than the net MFIs observed for the other probes in this assay.

The Ftul/TOX probe consistently yielded false positives when samples containing *F. tularensis holarctica* were analyzed. As discussed in the Specificity/Sensitivity section (also see Figure 6.1), the two polymorphisms between the Ftul/TOX probe and the non-specific amplicon from *F. tularensis holarctica* were not enough to prevent the false positives observed with the Ftul/TOX probe. Aside from this, the two probes with the most abundant false positives were Cdif/RRL and Ypes/TOX. The Cdif/RRL probe was prone to cross-reactivity with several target sequences, as determined by an AutoDimer primer dimer screen of this probe versus all target sequences, with a threshold of three (Figure A11.5). The melting temperatures of the secondary structures predicted in Figure A11.5 are often well below the assay hybridization temperature. However, complementary bases near the center of the binding sequences may have augmented the stability of a partial duplex, despite the low melting temperatures. The Cdif/RRL probe also had the lowest melting temperature, and its specificity may have been compromised by the hybridization temperature that was used in the Luminex assay to accommodate a set of probes with a broad range of melting temperatures, ranging from 50.2 to 64.5°C.

The other prominent group of false positives in Batch #2 were those corresponding to the Ypes/TOX probe. False positives with this probe had been observed sporadically in the specificity/sensitivity study and in Batch #1 of the 1:1 mixture study. An AutoDimer primer dimer screen (threshold of three) in which the Ypes/TOX probe was tested for cross-reactivity with all target sequences predicted cross-reactivity of the Ypes/TOX probe with the Cbot/RRL, Cbot/TOX, and Ypes/RRL target sequences (Figure A11.6). Interestingly, false positives with the Ypes/TOX probe in Batch #2 were most prevalent when target DNA from *F. tularensis* was present. It was considered that this might have been due to a sequence homology between the *Y. pestis* and *F. tularensis* target DNA that was not predicted by the AutoDimer primer dimer screen. The Ypes/TOX probe was visually compared to the other probe sequences to discern any possible homologies. A six-base homology between the Ypes/TOX probe and the Fran/RRL probe was detected (Table 6.2).

Probe Name	Probe Sequence
Fran/RRL	ACACAAGTTGGGTACTCACT
Ypes/TOX	TATGATGAGGGCAAAGGAGG

Table 6.2. Comparison of Fran/RRL and Ypes/TOX probe sequences. Highlighted bases indicate homologies between the two probes.

The homology depicted above would not have been detected by the initial AutoDimer analysis because the threshold value was set to seven. AutoDimer computes scores by subtracting the mismatched bases from the matching bases, but the binding of the Ypes/TOX probe with the Fran/RRL target sequence would not have been predicted using a threshold of seven (or even three) because this interaction would produce a score of -8 in a primer dimer screen in AutoDimer (Figure A11.7).

The complementary bases between the Fran/RRL target sequence and the Ypes/TOX probe were in the middle of the probe sequence, which is most susceptible to non-specific binding because the probe and target DNA hybridize in the middle first and expand outward. Prior research (Diaz and Fell, 2004) indicated that the Luminex platform had a level of specificity sufficient to distinguish single basepair mismatches, but data in the current study have indicated otherwise. A lower tolerance for mismatches was observed throughout the present study.

3. Batch #3

In batch #3 (Table 4.17), several false positives were observed (Table A7.6 and Figures 5.40-5.50), but generally the MFIs were much smaller than those expected for an actual positive result. The three false positives obtained with the Cdif/RRL probe were most likely a consequence of the cross-reactivity of this probe with other target DNA sequences, as predicted by the AutoDimer primer dimer screen (Figure A11.5). The three false positives obtained with the Ypes/TOX probe may have also been the result of cross-reactivity with other target sequences (particularly the Fran/RRL target sequence), as predicted by an AutoDimer primer dimer screen (Figures A11.6 and A11.7).

The probes displaying the weakest performance in terms of net MFI values were the Cbot/RRL, Yent/RRL, and Ypes/RRL. The target sequences to which these probes bind were all previously identified as being prone to secondary structure formation. One consequence of these hairpins was the single false negative result in this assay with the Ypes/RRL probe when the BC/YP mixture was analyzed. This may have been due to dimer-forming sequences or other secondary structures within the Ypes/RRL target

sequence. The propensity of the Ypes/RRL probe-binding sequence to form secondary structures was predicted by the AutoDimer hairpin screen (Figure A11.1).

The BC/YE/CB sample had the most abundant false positives of any sample in Batch #3. False positives were detected for this sample with the Bant/TOX, Cdif/RRL and Ftul/TOX probes. The Cdif/RRL probe was one of the most vulnerable to cross-reactivity because of homologies with other target sequences. Based on data collected for Batches #1 and #2, the Cdif/RRL probe was more prone to false positives when its pathogenic cousin (*C. botulinum*) was present in the mixture, which makes sense because the Cdif/RRL and Cbot/RRL probes are similar in sequence.

Most false positives with the Ftul/TOX probe in Batches #1, #2, and #3 occurred when *F. tularensis holarctica* was a mixture component. This was because amplification of *F. tularensis holarctica* produced a non-specific amplicon differing by only two bases from the probe-binding sequence of the Ftul/TOX probe (Figure 6.1). Although the *F. tularensis holarctica* non-specific amplicon was the main cause of false positives with the Ftul/TOX probe, it was not the exclusive cause. In Batch #2, the Ftul/TOX probe yielded a false positive for the YE/CB and BA/YE/CB mixtures. The net MFI values for these false positives were minute, but they were nevertheless above threshold. Even though mild cross-reactivity with the Ftul/TOX probe and the *Y. enterocolitica* and/or *C. botulinum* target sequences was indicated by the data, there were no AutoDimer results to substantiate this theory. However, there might have still been interactions with this probe and target sequences that would have been undetectable by AutoDimer due to scores below the threshold value.

False positives with the Bant/TOX probe were previously observed with the FT/CB sample in Batch #1 and with the BC/YP/FT and YP/FT/CD samples in Batch #2. Unlike the false positives observed with the Ftul/TOX probe, there was no obvious pattern or potential cause. It was possible that these false positives were caused by the hybridization of the Bant/TOX probe with other target sequences, and that these interactions were unpredicted because of scores below the AutoDimer threshold value. However, there was also the possibility that the random pattern of false positives observed with the Bant/TOX probe were caused instead by bead carryover from previously analyzed samples, despite the use of inter-sample washes. Each of the three batches in the mixture study consisted of samples in which *B. anthracis* was a component, so it was conceivable that Bant/TOX beads might have lingered within the internal parts of the instrument. These beads would have then led to a false positive with the Bant/TOX probe upon their release. The concept of internal bead carryover is explored in greater detail in a later section of the Discussion.

4. Conclusions of the 1:1 Mixture Study

The templates of each mixture component were present in a 1:1 ratio for the amplification step. However, this ratio was rarely reflected in the net MFI values of the mixtures (Tables A7.2, A7.4, and A7.6). The amplification efficiency was certainly not uniform for all templates in the mixture, because some templates had higher copy numbers than others. Amplicon proportions that deviated from the original template input ratios might have been exacerbated by the manner in which Luminex input amounts were determined. First, the precise concentration of the digested DNA could not be determined (due to the inability of the Agilent 1000 kit to accurately quantify single-stranded DNA)

and was therefore approximated using the post-digest quantitation data from the control tubes (undigested). Second, only the RRL band(s) in each control tube were used to estimate the quantity of digested DNA. This was the best way to create a standard method that was applicable to all bacteria, because an RRL target sequence was designed for each of the nine bacteria. However, one flaw with this method was that it ignored the possibility that the concentration of the TOX target sequences did not exactly equal the concentration of the RRL target sequences, on which Luminex input amounts were based.

Although the net MFI values did not always reflect template input amounts, they were seemingly unaffected by the number of components in the mixture. One might have expected MFI values to decline as the number of templates was increased, due to the possibility of overwhelming the PCR primers or *Taq* polymerase. On the contrary, net MFI values for any given probe were often consistent whether the target sequence was one of two, three, or four mixture components. In Batch #1, for example, the net MFI for the Bant/TOX probe in the BA/YE mixture was 1943, and the net MFI for the same probe in the BA/CB/FT/YP mixture was 1980. A similar degree of consistency in the net MFI values of binary, ternary, and quaternary mixtures was observed for many of the probes.

Another trend was that, in most mixtures containing pathogenic bacteria, the average net MFI for the RRL probes was somewhat lower than the average net MFI for the corresponding TOX probe. In Batch #1, for example, the average net MFI for a positive result with the Bant/TOX probe was 2533 but only 875 for the Bac/RRL probe. Similar observations were made for the two probes designated for *C. botulinum*, *F. tularensis*, and *Y. pestis*. The least severe example was often *F. tularensis tularensis*, for

which the fluorescent response from the Ftul/TOX probe was consistently stronger than the Fran/RRL probe, but not to the extent that net MFIs of the other TOX probes greatly surpassed the net MFIs of the RRL probes.

It was unexpected that the TOX probes were more responsive to their target sequences in each individual mixture and overall. However, this trend was consistent with observations in the Enzymatic Degradation study and in the Specificity/Sensitivity Study. The enhanced responsiveness of the TOX probes compared to the RRL probes was somewhat unexpected because the RRL target sequences were present in multiple copy numbers in their respective genomes, whereas the TOX targets were all single-copy sequences. This discrepancy in copy number was expected to promote a higher level of amplification efficiency for the RRL target sequences. Indeed, this was confirmed with the post-amplification quantitation results, which often reported higher concentrations for the RRL amplicons than the TOX amplicons.

The fact that the dominance of the RRL target sequences was not carried over to the Luminex assay indicated that one or more of three possible scenarios were occurring: (1) The RRL target sequences were prevented from binding efficiently to their probe because of secondary structure formation within the probe, (2) the RRL target sequences were cross-reacting with other probes or target sequences, thus leaving fewer RRL target sequences available for binding with the probe with which they had designed complementarity, or (3) secondary structures were forming within the target sequences themselves, thus preventing them from binding efficiently to their complementary probes.

The first possibility was ruled out because net MFIs observed with the positive controls indicated that the reverse probes were binding the probes with adequate

efficiency. The second possibility was conceivable and somewhat corroborated by AutoDimer results collected in the Specificity/Sensitivity Study (*e.g.*, the predicted cross-reactivity of the Cbot/RRL target sequence with the Bant/TOX probe). However, in order to explain the low MFI response of the RRL probes, this theory would have to be substantiated by abundant examples of false positives, which there were not. The most plausible of all three explanations was the third, because the results of the AutoDimer hairpin screens performed in response to data collected in the Specificity/Sensitivity study indicated that several RRL target sequences (specifically Cbot/RRL, Yent/RRL, and Ypes/RRL) were prone to hairpin formation. Other secondary structures having scores below the threshold or consisting of structures besides hairpins could also form within the target sequences without being predicted in the AutoDimer hairpin screen.

Due to the consistently weak performance of the RRL probes, it was more difficult to identify non-pathogenic mixture components, because their identification hinged on the fluorescent response of one RRL probe. In Batch #2, for example, the weak net MFI values returned by the Yent/RRL probe would make it difficult to positively identify *Y. enterocolitica*. Based on the Yent/RRL probe alone, it might also be difficult to distinguish different Yersinia species in the same mixture. Identification of the pathogenic bacteria was also partially (though not totally) compromised by the weak net MFI values of the RRL probes. The Cbot/RRL probe in the FH/CB mixture of Batch #2, for example, had a net MFI of just 18, whereas the Cbot/TOX probe for the same mixture had a net MFI of 2001. The TOX probes were originally intended for confirmatory identification after presumptive identification had already been established using the RRL

probes. However, the deficiencies of the RRL probes often made the TOX probes the sole means of identifying the pathogenic bacteria.

J. Mixture Study 1:1 (Intra-Genus)

The Luminex platform had a sufficient specificity to enable the identification of two organisms in a mixture when both organisms belonged to the same genus (Table A8.2 and Figures 5.51 and 5.52). The net MFI values for the Cbot/RRL, Yent/RRL, and Ypes/RRL probes were somewhat lower than the net MFIs observed for the corresponding TOX probes. These low net MFI values were most likely due to the tendency for hairpin formation (as predicted by the AutoDimer hairpin screens) within the target sequences designed to be complementary to these probes.

The net MFIs for the Cdif/RRL, Cbot/TOX, and Cbot/RRL probes in the 10 ng sample of the Specificity/Sensitivity Study were 480 (Table A6.11), 2392 (Table A6.5), and 125 (Table A6.5), respectively. With the exception of the Cdif/RRL probe (which actually displayed greater sensitivity in the Intra-genus mixture study), the net MFIs for all probes were consistent with those observed in the Specificity/Sensitivity study. The net MFIs for the Ypes/TOX, Yent/RRL, and Ypes/RRL probes in the 10 ng sample of the Specificity/Sensitivity study were 1974 (Table A6.9), 21 (Table A6.8), and 18 (Table A6.9), respectively. The sensitivity for all three probes was better in the Intra-genus mixture study. This enhanced sensitivity in the presence of competing DNA was surprising but may be attributable to imprecision introduced due to assumptions made when interpreting quantitation data for the purpose of determining Luminex input amounts.

There were no false negatives observed in the Intra-genus mixture study. However, several false positives were observed. False positives with the Cdif/RRL bead set were observed in prior assays and may have been the result of cross-reactivity between this bead set and other target sequences. Both false positives observed in this assay generally had a net MFI whose magnitude was easily distinguishable from a positive result.

K. Mixture Study (Ratios)

1. Remarks on Study Design

The first part of the mixture study consisted of binary, ternary, and quaternary mixtures in which components were present in a 1:1 ratio. However, mixture components may not always be present in equal ratios, and it was important to determine if this would compromise the ability of the Luminex platform to make a positive identification. In prior research (Battaglia *et al.*, 2011), a follow-up to the 1:1 mixture study was performed in which the ratio of the mixture components was varied. The purpose of this was to determine if the presence of a major and minor component had any impact on the ability of the Luminex instrument to positively identify all components of the mixture. For simplicity, only ratio mixtures with two contributors (binary mixtures) were analyzed. Here, in Phase II, a similar ratio mixture study was performed.

First, the results were examined for all mixtures analyzed in the 1:1 mixture study. Only binary (not ternary or quaternary) mixtures were examined because only these mixtures were considered for further analysis in the ratio mixture study. A mixture was deemed a strong candidate for further examination in the ratio mixture study if it

displayed few or no false positives and false negatives. It was believed that these samples would prove most informative in terms of gauging the ability of the Luminex platform to identify all components of a ratio mixture, without complicating matters with mixtures possessing components prone to error. There were several strong candidates for further examination, but the BA/FT, BA/YP, YP/FT, and BA/CB mixtures were selected.

There were several options to consider when determining the amount of DNA to be loaded on to the Luminex plate. Ultimately it was decided to base the amount of DNA on the concentration of the RRL band from the minor contributor. Doing so ensured that the amount of DNA from the minor contributor would meet or exceed the LLD for the Luminex (10 ng). Basing the amount of Luminex input DNA on the minor component also ensured that quantities of all components of any given mixture would agree with the sensitivity limitations of the Luminex platform. The only samples for which a slightly modified strategy was used were the 1:1 samples. If a single RRL band was detected for these samples, the concentration of that band was divided in half; if two RRL bands were detected (one for each component), the concentrations of those bands were averaged. After this calculation, the resulting value was used as a basis for determining the amount of digested DNA to load on the Luminex plate.

Although it may seem as if an important piece of data was omitted when determining Luminex input quantities, using only the concentration of the minor component represented the best compromise because: (1) it was better to base the amount of Luminex input DNA on the minor component because it ensured that enough of the minor contributor was being loaded to meet or exceed the LLD of the Luminex; (2) the ratios were, at best, an approximation – it was impossible and impractical to assume that

the precise ratios of template DNA used for the PCR reaction would be reflected in the amplicons, because the PCR reactions were optimized for 10 ng of template DNA, yet some ratio mixtures required the input of as much as 100 ng of template, thus potentially overwhelming PCR reagents such as the polymerase. Furthermore, the target sequence copy numbers varied among the bacteria used in this study, thus introducing the potential for a deviation from the intended ratio of minor/major components. For all the reasons stated above, it was also unreasonable to expect the ratios of template DNA in the mixtures to be precisely reflected in the quantitations, or in the MFI data.

2. BA:FT Ratio Mixtures

There were no false negatives detected in this assay (Table A9.2 and Figure 5.53). However, there were several false positives corresponding to the Cdif/RRL and Ypes/TOX probes, whose AutoDimer analyses demonstrated a proclivity for cross-reacting with target sequences other than those for which these probes were designed (Figures A11.5-11.7). An AutoDimer primer dimer screen of the Cdif/RRL probe predicted that this probe would cross-react with nine target sequences, one of which was the Bac/RRL target sequence (Figure A11.5).

The Ypes/TOX probe may have been vulnerable to cross-reacting with the Fran/RRL target sequence, owing to a six-base homology in the probe-binding region of these two sequences (Table 6.2). This was a plausible explanation because false positives with the Ypes/TOX probe had a higher magnitude when *F. tularensis tularensis* made up a larger proportion of the mixture. Most of the false positives detected in this assay (particularly those pertaining to the Cdif/RRL probe) yielded net MFIs of a much smaller magnitude than a normal positive result, thus making them easily identifiable as false

positives. The weak net MFI values reported for the Ypes/TOX probe, combined with the fact that all samples yielded net MFIs for the Ypes/RRL probe that were below threshold, would make the false identification of *Y. pestis* particularly unlikely.

The results were also consistent with data previously obtained for the BA/FT binary mixture, which was first analyzed in Batch #1 of the 1:1 Mixture Study. The original BA/FT sample from the 1:1 Mixture Study yielded net MFIs for the Bac/RRL, Bant/TOX, Fran/RRL, and Ftul/TOX probes of 671, 3067, 1462, and 2990, respectively (Table A7.2). These were very similar to the net MFI values reported for the same probes in the BA/FT 1:1 sample in the ratio mixture study (853, 2894, 1896, and 3305, respectively; Table A9.2). In both samples the TOX probes displayed a higher level of responsiveness for the complementary target DNA, possibly because of a lower tendency for hairpin formation within the TOX target sequences.

When compared across the various BA:FT mixture ratios studied, the net MFI values for the relevant probes demonstrated a sufficient level of sensitivity despite the various ratios of input DNA used. The minor component was never undetectable, even when overshadowed ten-fold by the major component. In fact, the net MFI values for the Bac/RRL, Bant/TOX, Fran/RRL, and Ftul/TOX probes were all well above threshold, whether the mixture was made up of one part or ten parts of the target sequence. In addition to the minor components consistently having net MFI values above threshold, the net MFIs of the minor components were also ostensibly unaffected by the quantity of the major component. For example, the four net MFIs for the Bac/RRL probe when *B. anthracis* made up one part of the mixture were very uniform, having a coefficient of

variation of 5.5% despite there being two, five, or ten parts *F. tularensis tularensis* (Table A9.9).

The net MFI values generally increased and decreased predictably depending on the amount of target DNA contributed by each mixture component. For example, the net MFIs for the Bac/RRL probe descended from 1373 (ten parts BA) to 1269 (five parts BA) to 919 (two parts BA), before stabilizing around 850 (one part BA). Likewise, the net MFIs for the Fran/RRL probe began around 1600 (one part FT) and ascended to 2184 (two parts FT) to 2745 (five parts FT) to 3017 (ten parts FT). Similar trends were evident for the Bant/TOX and Ftul/TOX probes.

3. BA:YP Ratio Mixtures

One intriguing observation from the quantitation data in this assay was that the concentration of the Bant/TOX target diminished as the amount of *Y. pestis* template DNA was increased (Table A9.3). Furthermore, the concentration of the Ypes/TOX band was consistently high, even when *Y. pestis* template made up only one tenth of the PCR template. These trends were particularly evident in the BA:YP/1:10 sample when the amplification efficiency of the Bant/TOX target sequence was so poor that the concentration was not even reported. In sum, large quantities of *Y. pestis* template DNA were apparently hindering the specificity of the Bant/TOX primers.

It was considered that there might have been a sequence homology between the Bant/TOX and Ypes/TOX primers. Even though suppression of the Bant/TOX primers by the Ypes/TOX primers had not been observed in earlier assays when both templates were present in equal quantities, the abundance of one template could exacerbate any tendencies for cross-reactivity. If present, any sequence homologies between the

Bant/TOX and Ypes/TOX primers might have prevented one primer from amplifying as much of its true target than would otherwise be possible, thus causing the preferential amplification of another target. A visual comparison of the Bant/TOX and Ypes/TOX primers did not reveal any sequence similarities that might have resulted in cross-reactivity. In order to obtain a more thorough estimate of any potential cross-reactivity, an AutoDimer primer dimer screen of all RRL and TOX primers was performed with a threshold of three. The analysis predicted cross-reactivity between the Bant/TOX and Ypes/TOX primers (Figure A11.8).

Evidently, the Bant/TOX fwd primer was the limiting reagent in the PCR reaction. Normally the Bant/TOX primers functioned well enough when optimal quantities of *B. anthracis* template were present, but when overshadowed by excess amounts of other templates (*e.g.*, *Y. pestis*), the ability of this primer pair to yield adequate quantities of amplicons was diminished. Despite the evidence of some mild cross-reactivity between the PCR primers for *B. anthracis* and *Y. pestis*, it was decided to proceed with the Luminex assay for this sample set.

No false negatives occurred in this assay, although several false positives were observed (Table A9.4 and Figure 5.54). Most of the false positives corresponded to the Cdif/RRL probe, whose propensity for cross-reactivity was discovered in the Specificity/Sensitivity study. In particular, the Bac/RRL and Ypes/RRL were among the nine target sequences predicted to cross-react with the Cdif/RRL probe (Figure A11.5). The only other false positive detected in this assay was for the Yent/RRL probe when the BA:YP/1:10 sample was analyzed. The net MFI for this false positive was 2, making it difficult to mistake for a genuinely positive result.

The results were also consistent with data previously obtained for the BA/YP binary mixture, which was first analyzed in Batch #1 of the 1:1 Mixture Study (Table A7.2). The original BA/YP sample from the 1:1 Mixture Study yielded net MFIs for the Bac/RRL, Bant/TOX, Ypes/RRL, and Ypes/TOX probes of 355, 1899, 0, and 2565, respectively. These were similar to the net MFI values reported for the same probes in the BA/YP 1:1 sample in the ratio mixture study (588, 1487, 123, and 2992, respectively). In both samples the TOX probes displayed a higher level of responsiveness for the complementary target DNA, possibly because of a lower tendency for hairpin formation within the TOX target sequences.

The net MFI values of the Bac/RRL and Bant/TOX probes in the BA:YP ratio mixtures were somewhat lower than the net MFIs for the same probes in the BA:FT ratio mixtures. The diminished responsiveness of these two probes in the BA:YP ratio mixtures may have stemmed from the low quantities of target sequences from *B. anthracis*. Optimum amplification efficiency for the *B. anthracis* target sequences may have been prevented by large quantities of *Y. pestis* template.

When compared across the various BA:YP mixture ratios studied, the net MFI values for the relevant probes demonstrated a sufficient level of sensitivity despite the various ratios of input DNA used. The minor component was never undetectable, even when overshadowed ten-fold by the major component. In fact, the net MFI values for the Bac/RRL, Bant/TOX, and Ypes/TOX probes were all well above threshold, whether the mixture was made up of one part or ten parts of the target sequence. Although all values were above threshold, the net MFIs for the Ypes/RRL probe were considerably lower than the net MFI values typically observed for other probes. Low net MFI values for the

Ypes/RRL probe were previously observed in the Enzymatic Degradation Study, the Specificity/Sensitivity Study, and the 1:1 Mixture Study. An AutoDimer hairpin screen predicted hairpin formation within the Ypes/RRL target sequence, which may have been responsible for the consistently low net MFI values observed for the Ypes/RRL probe (Figure A11.1).

In addition to the minor components consistently having net MFI values above threshold, the net MFIs of the minor components were generally consistent despite large quantities of the major component. For example, the four net MFIs for the Bac/RRL probe when *B. anthracis* made up one part of the mixture were reasonably uniform, displaying a coefficient of variation of 33% irrespective of there being two, five, or ten parts *Y. pestis* (Table A9.9). However, the net MFIs of the minor component were not as impervious to variation as previously observed in the BA:FT mixtures. The net MFIs of the Bac/RRL probe in the BA:FT ratio mixtures in which *B. anthracis* made up one part of the mixture had a coefficient of variation of just 5.5% (Table A9.9).

The net MFI values generally increased and decreased predictably depending on the amount of target DNA contributed by each mixture component. For example, the net MFIs for the Bac/RRL probe descended from 1446 (ten parts BA) to 1441 (five parts BA) to 866 (two parts BA), before stabilizing around 700 (one part BA). Likewise, the net MFIs for the Ypes/RRL probe began around 200 (one part YP) and ascended to 269 (two parts YP) to 498 (five parts YP) to 835 (ten parts YP). Similar trends were evident for the Bant/TOX and Ypest/TOX probes. Although there was clearly a relationship between the ratio of input DNA and fluorescent response for the BA:YP ratio mixtures, this relationship was not as well defined as it was for the BA:FT ratio mixtures.

4. YP:FT Ratio Mixtures

No false negatives occurred in this assay, and only one false positive was detected (Table A9.6 and Figure 5.55). The sole false positive corresponded to the Cdif/RRL probe when the YP:FT/1:5 sample was analyzed. An AutoDimer primer dimer screen predicted that the Cdif/RRL probe would cross-react with nine target sequences, including Ypes/RRL, Fran/RRL, and Ftul/TOX (Figure A11.5).

The results were also consistent with data previously obtained for the YP/FT binary mixture, which was first analyzed in Batch #1 of the 1:1 Mixture Study (Table A7.2). The original YP/FT sample from the 1:1 Mixture Study yielded net MFIs for the Ftul/TOX, Fran/RRL, Ypes/TOX, and Ypes/RRL probes of 2713, 1193, 3068, and 0, respectively. These were similar to the net MFI values reported for the same probes in the YP/FT 1:1 sample in the ratio mixture study (2580, 1465, 3543, and 201 respectively). In both samples the TOX probes displayed a higher level of responsiveness for the complementary target DNA, possibly because of a lower tendency for hairpin formation within the TOX target sequences.

When compared across the various YP:FT mixture ratios studied, the net MFI values for the relevant probes demonstrated a sufficient level of sensitivity despite the various ratios of input DNA used. The minor component was never undetectable, even when overshadowed ten-fold by the major component. In fact, the net MFI values for the Ftul/TOX, Fran/RRL, Ypes/TOX, and Ypes/RRL probes were all well above threshold, whether the mixture was made up of one part or ten parts of the target sequence. Although all values were above threshold, the net MFIs for the Ypes/RRL probe were appreciably lower than the net MFI values typically observed for other probes. Low net

MFI values for the Ypes/RRL probe were previously observed in the Enzymatic Degradation Study, the Specificity/Sensitivity Study, the 1:1 Mixture Study, and the BA/YP ratio mixtures. An AutoDimer hairpin screen predicted hairpin formation within the Ypes/RRL target sequence (Figure A11.1), which may have been responsible for the consistently low net MFI values observed for the Ypes/RRL probe.

In addition to the minor components consistently having net MFI values above threshold, the net MFIs of the minor components were generally consistent despite large quantities of the major component. For example, the four net MFIs for the Fran/RRL probe when *F. tularensis tularensis* made up one part of the mixture were reasonably uniform, displaying a coefficient of variation of 15% irrespective of there being two, five, or ten parts *F. tularensis* (Table A9.9). Although moderately consistent, the net MFIs of the minor component in YP:FT ratio mixtures displayed a slightly greater degree of vulnerability to various quantities of the major component than that observed for other mixtures. For example, the net MFIs of the Fran/RRL probe in the BA:FT ratio mixtures in which *F. tularensis tularensis* made up one part of the mixture had a coefficient of variation of just 12.7%. Similarly, the Ypes/RRL probe in YP:FT ratio mixtures where *Y. pestis* was the minor component had a coefficient of variation of 76.8%, whereas the same probe in BA:YP ratio mixtures where *Y. pestis* was the minor component had a coefficient of variation of just 30% (Table A9.9).

The net MFI values for all four probes were only very roughly proportional to the ratios of target DNA contributed by each mixture component. For example, the net MFIs for the Ypes/RRL probe (starting with the YP:FT/1:10 sample and progressing towards the YP:FT/10:1 sample) started around 200 and stabilized around 500. Furthermore, the

relationship between net MFIs and input ratios was not always linear, often displaying moderate swings or minor deviations from the expected trend. If the net MFI values for the Ypes/RRL probe had been like those observed in the BA/YP ratio mixtures, they would have been expected to start around 230 and steadily climb to approximately 800.

Overall, the relationship between the ratio of input DNA and fluorescent response could have been more well defined for the YP:FT ratio mixtures. The poor correlation may have been caused by assumptions made during the interpretation of quantitation data which failed to account for the concentrations of all target sequences. For example, Luminex input amounts for this assay (and all other ratio mixture assays) were based on the concentration reported for the RRL band of the minor contributor. Minor aberrations in the input ratios may have also been caused by the potency of the Ypes/TOX target sequence, which was not only amplified very efficiently, but evidently had a very high affinity for its complementary probe as well. Regardless of the source of the poor correlation, this slight flaw was at least partially overcome by a superior level of specificity, which was supported by the unusually low frequency of false positives in the YP:FT ratio mixtures.

5. BA:CB Ratio Mixtures

No false negatives occurred in this assay, although several false positives were observed (Table A9.8 and Figure 5.56). Most of the false positives corresponded to the Cdif/RRL probe, whose propensity for cross-reactivity was discovered in the Specificity/Sensitivity study. In particular, the Bac/RRL, Cbot/RRL, and Cbot/TOX target sequences were among the nine predicted to cross-react with the Cdif/RRL probe (Figure A11.5).

Several other false positives were observed with the Ypes/TOX probe, which were most likely due to the cross-reactivity of this probe with the Cbot/RRL and Cbot/TOX target sequences, as predicted by an AutoDimer primer dimer screen (Figure A11.6). The polymorphisms between the Ypes/TOX probe and the cross-reacting target sequences were sometimes located at the end(s) of the probe sequence, which enhance non-specific binding because complementary bases in the middle of the probe sequence may be irrevocably hybridized before the polymorphism is encountered. However, sometimes the polymorphisms are in the middle of the probe sequence, which helps to prevent non-specific binding because the bases tend to hybridize first in the middle of the probe sequence followed by outward expansion. The various locations of polymorphisms between the probe and target DNA sequences may explain why the false positives observed with the Ypes/TOX probe in this assay were not ubiquitous, but merely observed in just three of seven samples.

Fortunately, the false positive results observed with the Ypes/TOX probe would not lead to a false identification of *Y. pestis* because the net MFIs for the Ypes/RRL probe were all below threshold whenever *Y. pestis* target DNA was absent (*Y. pestis* is only positively identified when the net MFIs for both the Ypes/RRL and Ypes/TOX are above threshold).

The results of the BA:CB ratio mixture study were consistent with data previously obtained for the BA/CB binary mixture, which was first analyzed in Batch #1 of the 1:1 Mixture Study. The original BA/CB sample from the 1:1 Mixture Study yielded net MFIs for the Bac/RRL, Bant/TOX, Cbot/RRL, and Cbot/TOX probes of 1297, 3752, 100, and 2480, respectively (Table A7.2). These were very similar to the net MFI values reported

for the same probes in the BA/CB 1:1 sample in the ratio mixture study (1449, 3946, 123, and 2301, respectively; Table A9.8). In both samples the TOX probes displayed a higher level of responsiveness for the complementary target DNA, possibly because of a lower tendency for hairpin formation within the TOX target sequences.

The net MFI values of the Bac/RRL and Bant/TOX probes in the BA:CB ratio mixtures were somewhat higher than the net MFIs for the same probes in the BA:FT (Table A9.2) and BA:YP (Table A9.4) ratio mixtures. The enhanced responsiveness of these two probes in the BA:CB ratio mixtures may have been because of a reduced tendency of the *B. anthracis* target sequences to dimerize with the *C. botulinum* target sequences, thus making more of the DNA available for binding the probes.

When compared across the various BA:CB mixture ratios studied, the net MFI values for the relevant probes demonstrated a sufficient level of sensitivity despite the various ratios of input DNA used. The minor component was never undetectable, even when overshadowed by a ten-fold excess of the major component. In fact, the net MFI values for the Bac/RRL, Bant/TOX, and Cbot/TOX probes were all well above threshold, whether the mixture was made up of one part or ten parts of the target sequence. Although all values were above threshold, the net MFIs for the Cbot/RRL probe were lower than the net MFI values typically observed for other probes. Low net MFI values for the Cbot/RRL probe were previously observed in the Specificity/Sensitivity Study and in the 1:1 Mixture Study. An AutoDimer hairpin screen predicted hairpin formation within the Cbot/RRL target sequence, which may have been responsible for the consistently low net MFI values observed for the Cbot/RRL probe (Figure A11.3). The second hairpin shown in Figure A11.3 would be more stable than the other predicted

hairpins, because its melting temperature (46.8°C) was in close proximity to the assay hybridization temperature (48°C).

In addition to the minor components consistently having net MFI values above threshold, the net MFIs of the minor components were generally consistent despite large quantities of the major component. For example, the four net MFIs for the Bac/RRL probe when *B. anthracis* made up one part of the mixture were reasonably uniform, displaying a coefficient of variation of 12.2% irrespective of there being two, five, or ten parts CB (Table A9.9). This value, as well as the coefficient of variation (15.3%) for the Bant/TOX probe in samples where *B. anthracis* was the minor component was comparable to the degree of variation for these probes observed in the BA:FT and BA:YP ratio mixtures. The net MFI values for the Cbot/RRL probe displayed a moderately large coefficient of variation (43.6%) for samples in which *C. botulinum* was the minor component. This larger degree of variation shows that the Cbot/RRL probe was more vulnerable than the other probes to fluctuations in the proportion of the major contributor.

The net MFI values generally increased and decreased predictably depending on the amount of target DNA contributed by each mixture component. For example, the net MFIs for the Bac/RRL probe descended from 2866 (ten parts BA) to 2365 (five parts BA) to 1750 (two parts BA), before stabilizing around 1600 (one part BA). Although similar results were observed with other probes, the Bac/RRL probe was the only one that displayed an uninterrupted upward trend in net MFI values as the proportion of *B. anthracis* in the mixture was increased. For other probes, there was a clear difference in the net MFI values observed at either end of the spectrum of ratios examined, but the data in between was not always linear. For example, the net MFI values for the Bant/TOX

probe in the BA:CB/2:1, BA:CB/5:1, and BA:CB/10:1 samples were 4609, 4522, and 4885, respectively. Although there was an overall increase in net MFIs between the 2:1 and 10:1 samples, the slight decrease in the net MFI for the 5:1 sample was unexpected. These minor aberrations may have been further examples of the inherent inaccuracies of the quantitation process.

6. Summary of Ratio Mixture Results

Twenty-eight ratio mixtures were analyzed in which the components were present in different proportions. Both components were successfully identified in 100% of all mixtures. The results indicated a strong ability of the Luminex platform to identify both components of a binary mixture, even when the minor contributor was overshadowed two-fold, five-fold, or ten-fold by the major contributor.

The mean, standard deviation, and coefficient of variation were calculated for net MFIs corresponding to the minor contributor of each sample (see Table A9.9). The purpose of this was to determine if the net MFIs of the minor contributor fluctuated as the proportion of the major contributor was varied. Although some minor contributors were more vulnerable to high proportions of the major contributor, the generally low coefficients of variation indicated overall that net MFIs are moderately stable despite changes in the ratio of the major contributor.

B. anthracis was a component in three of the four binary ratio mixtures analyzed. This provided a good opportunity to determine (with a specified degree of statistical certainty) if the variation observed in the net MFIs for *B. anthracis* was due to chance alone or perhaps dependent on the major contributor. An ANOVA was performed on net MFIs collected for the Bac/RRL and Bant/TOX probes (Table A9.10). The null

hypothesis was rejected at a 95% level of confidence in both analyses, indicating that the identity of the major contributor may determine the degree to which net MFIs of the minor contributor fluctuate. This finding may also provide support for the belief that cross-reactivity with non-specific target sequences plays a pivotal role in determining the specificity of a probe for its complement.

L. Inter-Sample Washes

As might be expected, it was very important that when each sample was analyzed on the Luminex, only the beads from that sample were analyzed and not beads that might have been remaining from a prior sample. Beads from a prior sample may sometimes remain on the inside of the instrument sample probe (termed "internal carryover") and beads from a prior sample may also adhere to the outside of the sample probe (termed "external carryover"). Luminex reports that a maximum of 3% carryover (~1% internal and ~2% external) are within the instrument specifications. Carryover problems with Luminex immunoassays have been reported before, but are predicted to be somewhat less severe with DNA-based assays because beads complexed with nucleotides may not adhere to the internal steel and plastic parts of the instrument with great affinity (Hanley, 2007).

A series of wash steps (using 1X TE buffer) were incorporated into the standard procedures to minimize sample-to-sample bead carryover. Initially, it was found that three inter-sample washes were sufficient for removing any extraneous beads. However, evidence of carryover became more apparent as the Luminex was used more frequently.

Lingering beads were undesirable because if they had a high fluorescence, the net MFI data for the subsequently analyzed sample could be skewed.

A fourth wash step was implemented for the purpose of removing any beads that might have been lingering after the first three washes. Adding a fourth inter-sample wash had a positive impact on the data. The count data illustrated that, in most cases, extraneous beads were completely removed by the end of the fourth wash (data not shown). As a result, fewer false positives were observed. Of the false positives that did occur, they had a much smaller magnitude than those in comparable assays with only three inter-sample washes. The additional inter-sample wash was universally adopted for all subsequent assays (ratio mixture studies only).

One final note on carryover: Lingering beads did not always consistently decrease with each progressing wash step, because the beads may adhere to the probe (internally or externally) and a series of wash steps showing a descending bead count may suddenly spike as residual beads are dislodged. Researchers should recognize this phenomenon, and accept the fact that it is difficult to confirm the total removal of beads prior to the next sample, even when the final wash displays a zero bead count. It must always be considered that a false positive result may not necessarily stem from cross-reacting oligos, but instead may have been caused by lingering beads that failed to leave the system. Consulting the bead count data may help to distinguish false positives due to cross-reactivity from false positives due to bead carryover.

M. Statistical Considerations

1. The Product Rule

In mathematics, the product rule is applied when one wishes to know the probabilitity of multiple independent events occurring at once. It is particularly useful in DNA typing because it enables analysts to determine the random match probability of multiple loci by multiplying their individual frequencies. The random match probability for a single locus might be very high (*i.e.*, there is a strong likelihood that a randomly chosen individual would have the same DNA sequence as the unknown), but if the individual probabilities of several loci are multiplied together, the chance of a random match becomes infinitely smaller as the number of loci examined is increased.

In order for the product rule to be correctly applied, the individual loci whose probabilities are to be multiplied should be in linkage equilibrium (Lander, 1989) and Hardy-Weinberg equilibrium. That is, the loci should be independently assorted such that the inheritance of one locus is not correlated with the inheritance of another. With respect to bacteria (which have a single circular chromosome), many reproduce by a form of cloning called binary fission. However, genetic recombination can occur via other processes such as conjugation, transduction, or transformation. Furthermore, many bacteria have one or more plasmids with different copy numbers. The chromosome and plasmid patterns of inheritance differ widely among bacteria, and thus the product rule cannot be applied as in human genetics.

Even though the product rule cannot be used to calculate random match probabilities for bacterial DNA sequences, there are other ways of establishing the significance of a match. Prior studies in which pathogens were analyzed with Luminex

technology have used as few as one probe to identify each bacteria. Recently, other researchers have begun to use multiple probes for pathogen detection (Deshpande *et al.*, 2010). This dissertation research sought to improve the certainty of an identification by using two probes for each of the four core pathogenic bacteria. The use of multiple loci increases the chances of successfully identifying bacterial strains that have undergone genetic modification. Although a positive reaction with two probes certainly inspires a greater degree of confidence in the results, it may not be sufficient for a positive identification. This is especially true if one or more probes are prone to error, as was the case with some RRL probes in the current study. Using more than two probes may help to minimize the impact of a problematic probe while enhancing confidence in the conclusion.

2. Reproducibility

One of the greatest flaws of pioneering research in criminalistics was the lack of attention to statistical significance. Studies were often conducted without regard to the need for multiple trials, which are necessary to gauge validity by assessing reproducibility. This deficiency was identified early on by Kirk (1963a), and today, many more researchers in forensic science are inculcated with the need to carry out statistically sound studies. Advocates for studies having a strong statistical foundation have generally declared that the ability to draw meaningful conclusions from the data declines proportionally with fewer sample repetitions (N. Petraco, personal communication, 20 September 2011). This is probably true in many instances, but unfortunately small sample sizes cannot always be avoided. For example, multiple trials in the present study were precluded by fiscal limitations and heavily controlled resources such as the genomic

DNA from pathogenic bacteria. Although a smaller study with greater sample repetition might have been possible, it was decided to design a study having a larger scope so that the Luminex technology could be more thoroughly evaluated using multiple bacteria in a range of possible scenarios. With the expanded footprint of the study came the caveat that future research using multiple trials would be needed to corroborate the findings. Critics may disagree with this decision by arguing that having fewer trials widens the gap between the methodology and the conclusions drawn from it. While this may limit the potency of the conclusions, it certainly does not invalidate them.

3. Limitations

The absence of multiple trials may be justified in certain circumstances, but there comes a time when conclusions must be supported by probability and statistics. This study used the terms "presumptive" and "confirmatory" to classify the two different types of probes used. However, it should be noted that the term "confirmatory" while ostensibly implying a definitive and absolute identification, was only confirmatory within the limitations of the study design. For example, if a negative result was observed when the Bant/TOX probe was introduced to some unknown target DNA, it might be said with a certain degree of confidence that the unknown target DNA did not originate from *B. anthracis* (*i.e.*, the unknown was excluded). If, on the other hand, a positive result was observed when the Bant/TOX probe was introduced to some unknown target DNA, it could not be said with absolute certainty that the unknown was *B. anthracis* (*i.e.*, inclusion). It would be impossible to say this without the support of statistical data. Thus, for the purposes of this study, the term "negative" with respect to identification can be taken literally. However, the term "positive" must be qualified by noting that it was

positive only within the limits of this study, and that a target sequence yielding a positive reaction to a certain probe could not truly be regarded as positive without some way of quantifying the likelihood that another randomly chosen DNA sequence would not yield a similar result.

Kirk (1963b) said that "until [the determination of guilt or innocence] is reduced to a reasonable mathematical exactness, the administration of justice will be correspondingly hampered." There must be a way to affix a level of certainty to positive and negative results if Luminex technology is to continue its proliferation in the field of forensic biology. There is a great need for the establishment of a normalized reference database consisting of net MFI values for known quantities of common pathogenic bacteria. Only then will the conclusions drawn from a Luminex analysis be truly meaningful.

CHAPTER VII. CONCLUSIONS

One of the most labor intensive aspects of this study was the preliminary bioinformatics research necessary before commencing benchwork. It must be verified that all sequences originate from regions of their respective genomes that are highly polymorphic, thermodynamically stable, and nonhairpin-forming (Summerbell *et al.*, 2005). Additionally, sequences should reside in genes that are well-conserved and/or essential for life, thus making them less prone to mutation. Evaluating the utility of candidate sequences can be very tedious, especially in a large study where there are numerous sequences to assess. However, it is emphatically stressed that background bioinformatics research should be thoroughly and meticulously conducted. Latent mistakes or omissions made at this stage will become evident down the line, and therefore it is important to invest the appropriate amount of time in the beginning. Researchers need not become burdened with repetitive bioinformatics work, as the initial background research can serve as the basis for many subsequent assays.

Non-specific binding among sequences may be difficult if not impossible to avoid, but can at least be minimized through the use of software programs such as Primer3 and AutoDimer. The specificity of all sequences must also be confirmed by searching a database such as NCBI BLAST. With a comprehensive collection of over 1700 bacterial genomes, this database enables researchers to avoid the use of non-specific sequences. However, this database is neither all-inclusive nor infallible, and success in identifying highly similar sequences depends on the search parameter settings.

Future researchers may have greater success using Primer-BLAST, which not only combines primer/probe design and the search for similar sequences into one step,

but also uses parameters that are specially modified for short sequences. Even if Primer-BLAST becomes the paradigm in future studies, researchers are admonished to remember that this program checks only for self-reactivity, and thus there is still a need for programs such as AutoDimer that check for cross-reactivity. The default threshold of seven was often inadequate for predicting cross-reactivity that occurs *in situ*, so it is advisable to lower the threshold beyond the default value.

Perhaps the single greatest finding of this research was that sensitivity could be improved by using lambda exonuclease to enzymatically degrade the non-complementary strand of the PCR product. Overall, single-stranded (when compared to double-stranded) target DNA in experimental samples was much more available for binding with the probes – reassociation of the complementary target strands was prevented and the effective concentration of the target DNA was increased. Enzymatic digestion of target sequences should be incorporated into Luminex standard protocols as a means of improving sensitivity.

The efficacy of the enzymatic digestion step was corroborated by the findings of the specificity/sensitivity study, in which LLDs as low as 0.1 ng were reported. The sensitivity of most probes matched or surpassed the sensitivity levels observed in previous work (Dunbar and Jacobson, 2007; Battaglia *et al.*, 2011). A diminished level of sensitivity was observed for some probes, possibly due to secondary structure formation within the corresponding target sequence. Specificity was adequate for most probes, but somewhat compromised for others due to moderate homologies with non-specific target sequences.

Thirty-three binary, ternary, and quaternary mixtures were examined in which all components were present in a 1:1 ratio. There was an 80% success rate for identifying all components, indicating that competition does not generally preclude positive identification. Four binary mixtures were subjected to further study in which the two components were combined in various ratios. Both components were detected in 100% of the twenty-eight mixtures analyzed, even when the minor component was overshadowed by a ten-fold excess of the major component. Some false positives were observed in both mixture studies, but typically their weak fluorescent response made them easily distinguishable from a genuine positive result. Statistical data and an ANOVA indicated that the net MFIs of the minor component depend on the identity of the major component.

Future researchers should first seek to create a standard protocol for the analysis of bacterial DNA using the Luminex platform. The current study began with the *a priori* assumption that the genomic DNA had already been extracted and quantified. However, the application of Luminex technology to actual casework would necessarily have to be preceded by an extraction and quantitation protocol designed for clinical samples which assumes that the cells or spores themselves are the starting material. The Luminex instrument has recently been combined with an automated, pre-detection sample processing format (Altman, 2011). However, creating a standard operating procedure for casework is expected to be a task reserved for a very elite laboratory because a Biosafety Level IV status would be required due to the hazardous nature of the samples. However, much of the validation work on these procedures could be performed on non-pathogenic bacteria in a Biosafety Level II laboratory.

After extraction and quantitation protocols are standardized, a normalized reference database of net MFI results for known samples should be assembled, to which future results can be compared. Once data for a set of reference samples has been accumulated, a blind study should be undertaken in which each individual result is compared to the knowns using hypothesis testing. This would attach a statistical significance (within a specified confidence interval) to each result. It would also enable a more thorough investigation into the frequency of type I errors (false positives) and type II errors (false negatives).

CHAPTER VIII. CONTRIBUTIONS TO FORENSIC SCIENCE

There are many pathogenic microorganisms with a tremendous potential for destruction, based on their ability to cause disease and death in humans (CDC, 2011). Previous instances of bioterror have demonstrated the glaring deficiencies in our current methods of rapidly and accurately identifying these pathogens. Luminex technology offers one of the best solutions for filling in this gap. If a bioterror event should occur in the future, a pathogen detection array would provide forensic scientists with a faster means of (1) verifying that the attack has in fact been caused by a pathogenic microorganism, and (2) identifying the pathogenic agent. Identifying the pathogen that has been released will subsequently enable first responders to modify their approach accordingly. It will also permit health care personnel to take the appropriate cautions while administering the most applicable treatments for those who are infected, as well as preventing further infection by quarantine or vaccine.

The efficacy of the Luminex platform as a pathogen detection array was most recently established by Battaglia *et al*. (2011). The results of this study have corroborated prior analyses and expanded our ability to react to a potential bioterror event. It has created a platform upon which other researchers may build, so that a comprehensive array of numerous biological pathogens may one day be available for use with the Luminex method. Once complete, this array will enable bioterror attacks to be expeditiously mitigated, while at the same time contributing to source attribution efforts.

This research has also helped to fortify the scientific underpinnings of the Luminex technology. This could be relevant as Luminex technology becomes more widely applied towards the resolution of scientific problems pertaining to the law.

Scientific techniques (especially new ones) must undergo rigorous scrutiny in order to satisfy evidence admissibility standards. Some states follow the *Frye* standard, which declared that in order for admissibility to be granted, a technique must be generally accepted by the relevant scientific community (*Frye v. U.S.*, 293 F. 1013). Other states and federal courts follow the precedent established by *Daubert*, in which it was held that the Federal Rules of Evidence (Rule 702) supplanted the *Frye* "general acceptance" test for evidence admissibility (*Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579). The court's ruling in *Daubert* was intended to shift some responsibility away from the scientific community by appointing trial judges as the primary gatekeepers. A trial judge may assess the validity of the proffered scientific technique by considering a non-exhaustive list of factors such as testability, peer review, publication, error rate, and general acceptance. Other factors have also been suggested, such as whether or not the science is litigation-driven or whether there is too much of an analytical gap between the methodology and the conclusions (Advisory Committee, 2000). Regardless of whether the *Frye* or *Daubert* standards are used as the benchmark, novel technologies are critically evaluated before admissibility is granted. If the Luminex technology is to be appraised with these standards in mind, then it is important for early research to lay the groundwork that will one day help to establish its veracity in the courtroom.

This research was an important step in the evolution of criminalistics, which has successfully partnered with fields such as molecular biology in recent years. Most importantly, the outcome of this project is poised to help other researchers realize the ability of criminalistics to contribute to the resolution of global-scale criminal acts such as bioterrorism.

APPENDIX I. INFORMATION ON LYOPHILIZED GENOMIC DNA

Organism	BEI Number	See Footnote
Bacillus anthracis str. Sterne	NR-10293	a
Bacillus cereus	NR-12314	b
Clostridium botulinum	NR-2713	c
Clostridium difficile	HM-88D	d
Francisella tularensis subsp. *tularensis*	NR-3015	e
Francisella tularensis subsp. *holarctica*	NR-3018	f
Yersinia pestis	NR-2717	g
Yersinia enterocolitica	NR-3066	h
Mycobacterium tuberculosis	NR-14865	i

Table A1.1. Lyophilized genomic DNA acquired from BEI Resources.

[a]Reagent contributed by P. Hanna, University of Michigan for distribution by the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: Genomic DNA from *Bacillus anthracis*, Strain Sterne BA851 (Δ*asbA*), NR-10293.

[b]Reagent obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: Genomic DNA from *Bacillus cereus*, Strain E33L, NR-12314.

[c]Reagent obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: Genomic DNA from *Clostridium botulinum*, Strain Walls 8G (VPI 4404), NR-2713.

[d]Reagent obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH as part of the Human Microbiome Project: Genomic DNA from *Clostridium difficile*, Strain NAP7 CDC#2007054, HM-88D.

[e]Reagent obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: Genomic DNA from *Francisella tularensis* subsp. *tularensis*, Strain SCHU S4 (FSC237), NR-3015.

[f]Reagent obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: Genomic DNA from *Francisella tularensis* subsp. *holarctica*, Strain LVS (FSC155), NR-3018.

[g]Reagent obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: Genomic DNA from *Yersinia pestis*, Strain CO92, NR-2717.

[h]Reagent obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: Genomic DNA from *Yersinia enterocolitica* subsp. *enterocolitica*, NR-3066.

[i]Reagent obtained through the NIH Biodefense and Emerging Infections Research Resources Repository, NIAID, NIH: Genomic DNA from *Mycobacterium tuberculosis*, Strain H37Rv, NR-14865.

APPENDIX II. MULTIPLEX PCR SPECIFICITY STUDY: DATA TABLES

Lane	Template	Expected Bands (bp)	Observed Bands (bp)	Concentration (ng/uL)
1	All 9 templates	All RRL + TOX	19	7.77
			29	16.0
			122	28.62
			168	1.44
			184	3.59
2	All 9 templates	All RRL + TOX	19	8.09
			29	13.77
			122	29.16
			168	1.50
			183	3.71
3	B. cereus	111 (RRL)	†	n/a
4	B. anthracis		19	7.79
			29	19.03
		111 (RRL)	*	n/a
		230 (TOX)	*	n/a
5	C. difficile		23	11.04
		110 (RRL)	117	10.04
6	C. botulinum		19	8.04
			29	14.57
		100 (TOX)	*	n/a
		109 (RRL)	126	11.42
7	F. tularensis holarctica		18	8.14
			29	18.19
		102 (RRL)	114	19.30
			170	2.88
8	F. tularensis tularensis		22	9.53
			26	5.93
			46	20.90
		102 (RRL)	*	n/a
		159 (TOX)	*	n/a
9	Y. enterocolitica		19	7.45
			30	14.59
		107 (RRL)	120	20.03
10	Y. pestis		19	6.32
			30	11.76
		107 (RRL)	119	14.56
		169 (TOX)	177	6.17
11	M. tuberculosis		19	7.23
			29	13.76
		110 (RRL)	121	10.43
12	None		19	8.12
			30	16.32

Table A2.1. Summary of quantitation data for samples amplified using original primer mix. An asterisk (*) symbol in the "Observed Bands" column indicates that a band was neither visible nor detected by the instrument. A dagger (†) symbol in the "Observed Bands" column indicates that a faint band was visible but not concentrated enough to be reported. In samples containing only one template, expected band sizes are noted adjacent to the observed band most likely to correspond.

Lane	Template	Expected Bands (bp)	Observed Bands (bp)	Concentration (ng/uL)
1	All 9 templates	All RRL + TOX		
2	All 9 templates	All RRL + TOX	23	7.35
			115	18.20
			163	1.24
			176	3.28
3	B. cereus		22	2.85
			28	4.69
			46	8.42
		111 (RRL)	129	0.90
			140	22.81
4	B. anthracis		23	7.09
		111 (RRL)	123	15.24
		230 (TOX)	230	5.37
5	C. difficile		19	7.01
			31	8.55
		110 (RRL)	124	14.60
6	C. botulinum		19	6.32
			30	7.13
		100 (TOX)	114	3.79
		109 (RRL)	127	12.09
7	F. tularensis holarctica		23	7.15
		102 (RRL)	109	13.53
			164	5.52
8	F. tularensis tularensis		19	6.09
			30	7.40
		102 (RRL)	115	13.87
		159 (TOX)	169	5.86
9	Y. enterocolitica		19	6.89
			30	7.18
		107 (RRL)	121	23.25
10	Y. pestis		19	5.92
			30	7.34
		107 (RRL)	118	17.07
		169 (TOX)	177	8.35
11	M. tuberculosis		19	7.50
			31	8.21
		110 (RRL)	121	8.31
12	None		19	7.38
			31	8.93

Table A2.2. Summary of quantitation data for samples amplified using modified primer mix. In samples containing only one template, expected band sizes are noted adjacent to the observed band most likely to correspond.

Lane	Template	Expected Bands (bp)	Observed Bands (bp)	Concentration (ng/uL)
1	All 9 templates	All RRL + TOX	116	27.13
			164	2.56
			178	7.38
2	All 9 templates	All RRL + TOX		
3	B. cereus		22	3.54
		111 (RRL)	122	11.21
4	B. anthracis		20	1.39
			23	3.03
		111 (RRL)	122	11.24
		230 (TOX)	228	4.34
5	C. difficile		22	3.95
		110 (RRL)	117	11.96
6	C. botulinum		23	3.41
		100 (TOX)	107	2.96
		109 (RRL)	120	7.57
7	F. tularensis holarctica		23	3.49
		102 (RRL)	109	10.27
			164	5.63
8	F. tularensis tularensis		23	3.82
		102 (RRL)	108	10.57
		159 (TOX)	162	5.88
9	Y. enterocolitica		22	2.58
		107 (RRL)	114	18.08
10	Y. pestis		23	3.52
		107 (RRL)	112	17.42
		169 (TOX)	170	8.58
11	M. tuberculosis		23	4.82
		110 (RRL)	†	n/a
12	None		23	4.93

Table A2.3. Summary of quantitation data for samples amplified using modified primer mix diluted 1:2. A dagger (†) symbol in the "Observed Bands" column indicates that a faint band was visible but not concentrated enough to be reported. In samples containing only one template, expected band sizes are noted adjacent to the observed band most likely to correspond.

Lane	Template	Expected Bands (bp)	Observed Bands (bp)	Concentration (ng/uL)
1	All 9 templates	All RRL + TOX	115	16.37
			162	1.47
			176	3.70
2	All 9 templates	All RRL + TOX	23	5.36
			114	3.77
3	B. cereus	111 (RRL)	23	6.65
			109	1.15
			120	11.75
4	B. anthracis	111 (RRL)	23	5.72
			120	9.35
		230 (TOX)	231	6.35
5	C. difficile	110 (RRL)	22	5.19
			116	13.20
6	C. botulinum	100 (TOX)	23	6.49
			107	1.01
		109 (RRL)	119	2.54
7	F. tularensis holarctica	102 (RRL)	23	7.99
			107	11.57
			163	6.27
8	F. tularensis tularensis	102 (RRL)	23	9.39
			97	1.14
			107	13.74
		159 (TOX)	162	8.46
9	Y. enterocolitica	107 (RRL)	23	5.24
			114	19.29
10	Y. pestis	107 (RRL)	22	4.57
			112	14.50
		169 (TOX)	170	6.99
11	M. tuberculosis	110 (RRL)	22	5.80
			114	1.54
12	None		23	7.68

Table A2.4. Summary of quantitation data for samples amplified using 10 ng of Template and modified primer mix diluted 1:2. In samples containing only one template, expected band sizes are noted adjacent to the observed band most likely to correspond.

APPENDIX III. STATISTICAL DATA FOR MONOPLEX AND MULTIPLEX PCR STUDIES

Monoplex PCR Specificity Study				
Template	**Target Sequence**	**Concentration**		
B. cereus	Bac/RRL	12.04		
B. anthracis	Bac/RRL	8.58		
B. anthracis	Bant/TOX	17.16		
C. difficile	Cdif/RRL	10.29		
C. botulinum	Cbot/RRL	7.96		
C. botulinum	Cbot/TOX	13.83		
F. tularensis holarctica	Fran/RRL	7.41		
F. tularensis tularensis	Fran/RRL	6.63		
F. tularensis tularensis	Ftul/TOX	20.64		
Y. enterocolitica	Yent/RRL	18.58		
Y. pestis	Ypes/RRL	20.69		
Y. pestis	Ypes/TOX	20.24		
M. tuberculosis	Myc/RRL	2.43		
		12.80615385		Avg
		6.159811332		Std. Dev.
		2.43		Min.
		20.69		Max.

Table A3.1. Statistical data for amplicons produced by their specific primer set in the monoplex PCR study.

Targets Observed with Original Primers			
Lane	Observed Size (bp)	Concentration (ng/uL)	
5	117	10.04	
6	126	11.42	
7	114	19.3	
9	120	20.03	
10	119	14.56	
10	177	6.17	
11	121	10.43	
		13.13571429	**average**
		5.097832406	**std. dev.**

Targets Observed with Modified Primers			
Lane	Observed Size (bp)	Concentration (ng/uL)	
3	129/140	23.71	
4	123	15.24	
4	230	5.37	
5	124	14.6	
6	114	3.79	
6	127	12.09	
7	109	13.53	
8	115	13.87	
8	169	5.86	
9	121	23.25	
10	118	17.07	
10	177	8.35	
11	121	8.31	
		12.69538462	**average**
		6.339084602	**std. dev.**

Targets Observed with Diluted, Modified Primers			
Lane	Observed Size (bp)	Concentration (ng/uL)	
3	122	11.21	
4	124	11.24	
4	228	4.34	
5	117	11.96	
6	107	2.96	
6	120	7.57	
7	109	10.27	
8	108	10.57	
8	162	5.88	
9	114	18.08	
10	112	17.42	
10	170	8.58	
		10.00666667	**average**
		4.612988251	**std. dev.**

Targets Observed with Diluted, Modified Primers and Reduced Template			
Lane	Observed Size (bp)	Concentration (ng/uL)	
3	109/120	12.9	
4	120	9.35	
4	231	6.35	
5	116	13.2	
6	107	1.01	
6	119	2.54	
7	107	11.57	
8	97/107	14.88	
8	162	8.46	
9	114	19.29	
10	112	14.5	
10	170	6.99	
11	114	1.54	
		9.429230769	**average**
		5.630780676	**std. dev.**

Primer-Dimers Observed with Original Primers			
Lane	Observed Size (bp)	Concentration (ng/uL)	
4	19	17.79	
4	29	19.03	
5	23	11.04	
6	19	8.04	
6	29	14.57	
7	18	8.14	
7	29	18.19	
8	22	9.53	
8	26	5.93	
8	46	20.9	
9	19	7.45	
9	30	14.59	
10	19	6.32	
10	30	11.76	
11	19	7.23	
		12.034	**average**

Primer-Dimers Observed with Modified Primers			
Lane	Observed Size (bp)	Concentration (ng/uL)	
3	22	2.85	
3	28	4.69	
3	46	8.42	
4	23	7.09	
5	19	7.01	
5	31	8.55	
6	19	6.32	
6	30	7.13	
7	23	7.15	
8	19	6.09	
8	30	7.4	
9	19	6.89	
9	30	7.18	
10	19	5.92	
10	30	7.34	
11	19	7.5	
11	31	8.21	
		6.808235294	**average**

Primer-Dimers Observed with Diluted, Modified Primers			
Lane	Observed Size (bp)	Concentration (ng/uL)	
3	22	3.54	
4	20	1.39	
4	23	3.03	
5	22	3.95	
6	23	3.41	
7	23	3.49	
8	23	3.82	
9	22	2.58	
10	23	3.52	
11	23	4.82	
		3.355	**average**

Primer-Dimers Observed with Diluted, Modified Primers and Reduced Template			
Lane	Observed Size (bp)	Concentration (ng/uL)	
3	23	6.65	
4	23	5.72	
5	22	5.19	
6	23	6.49	
7	23	7.99	
8	23	9.39	
9	23	5.24	
10	22	4.57	
11	22	5.8	
		6.337777778	**average**

Table A3.2. Summary of statistical data for multiplex PCR specificity study. Highlighted values were used to compare the effectiveness of various primer concentrations and template amounts.

APPENDIX IV. TITRATION STUDY DATA AND CHARTS

Reverse Probe (fmol)	MFI
0	83
0.25	620
0.5	907
1	1114
10	5392
50	10460.5
100	10227
150	10898.5
200	11881
250	12365

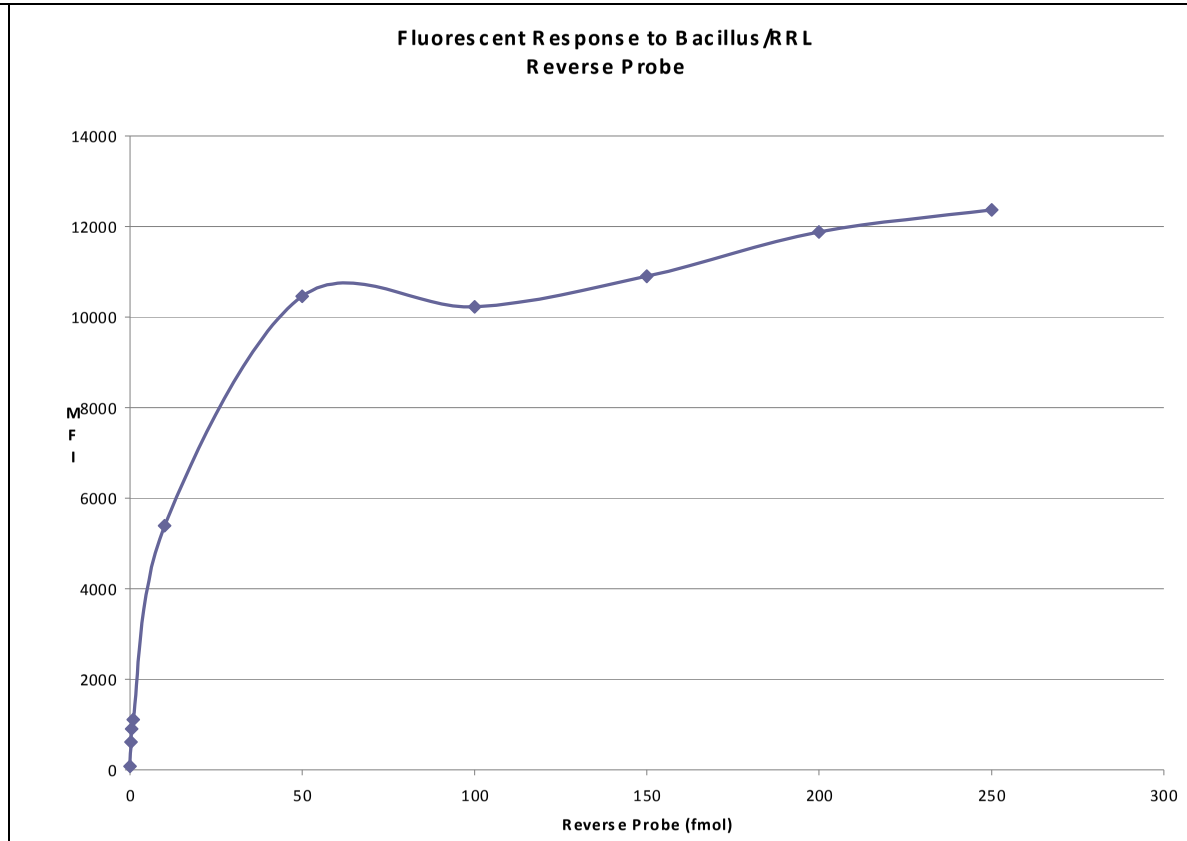


Figure A4.1. Titration study data and chart for Bac/RRL probe.

Reverse Probe (fmol)	MFI
0	50.0
0.25	745.0
0.5	1221.5
1	1339.5
10	9027.0
50	17719.0
100	17830.0
150	17901.0
200	18319.0
250	18312.0

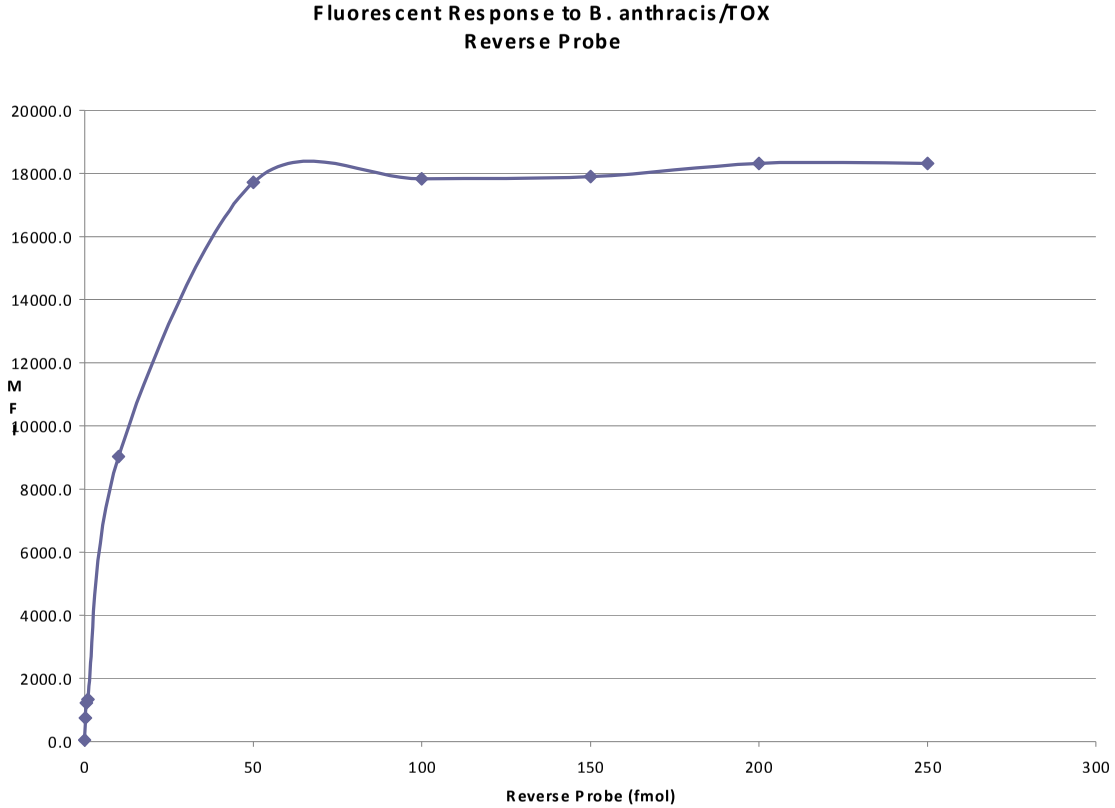


Figure A4.2. Titration study data and chart for Bant/TOX probe.

Reverse Probe (fmol)	MFI
0	62
0.25	919
0.5	663
1	836
10	4741
50	8268.5
100	8061.5
150	8932
200	8623
250	11318

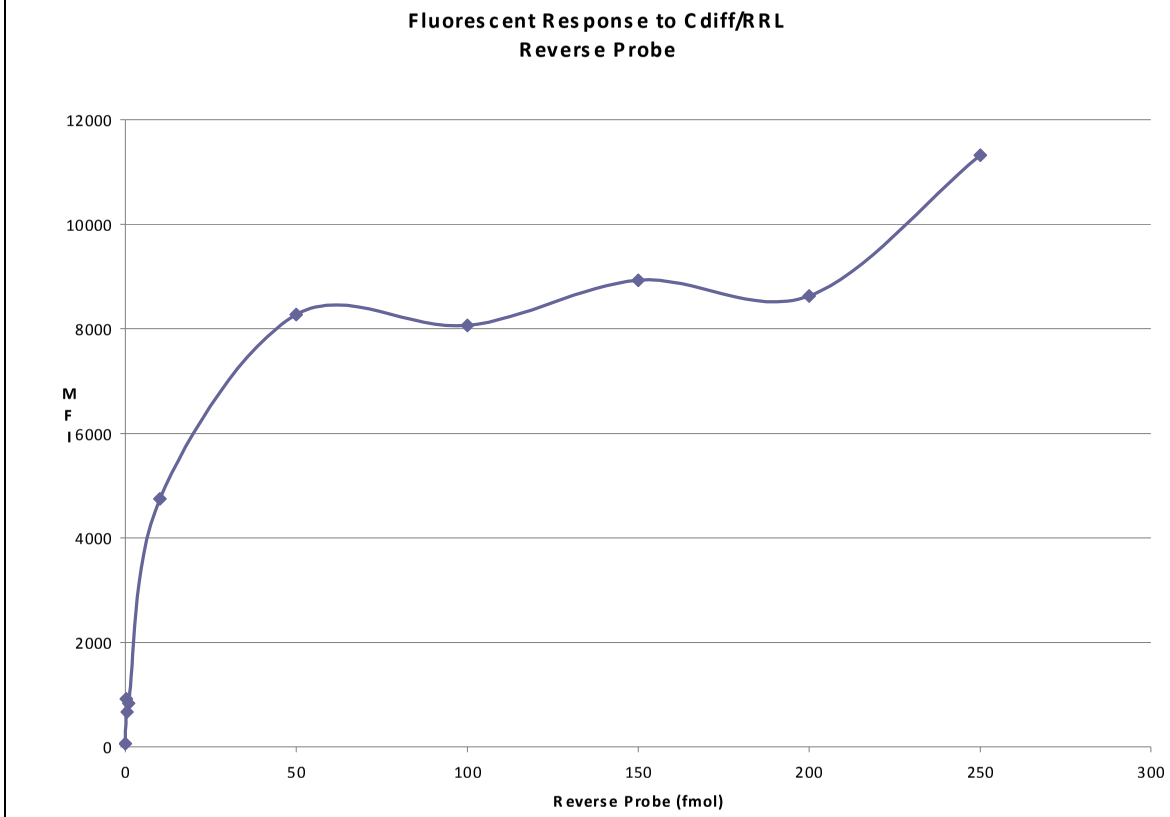


Figure A4.3. Titration study data and chart for Cdif/RRL probe.

Reverse Probe (fmol)	MFI
0	57
0.25	781
0.5	1068
1	1611.5
10	6005.5
50	10509
100	11498
150	11934
200	11645
250	13222

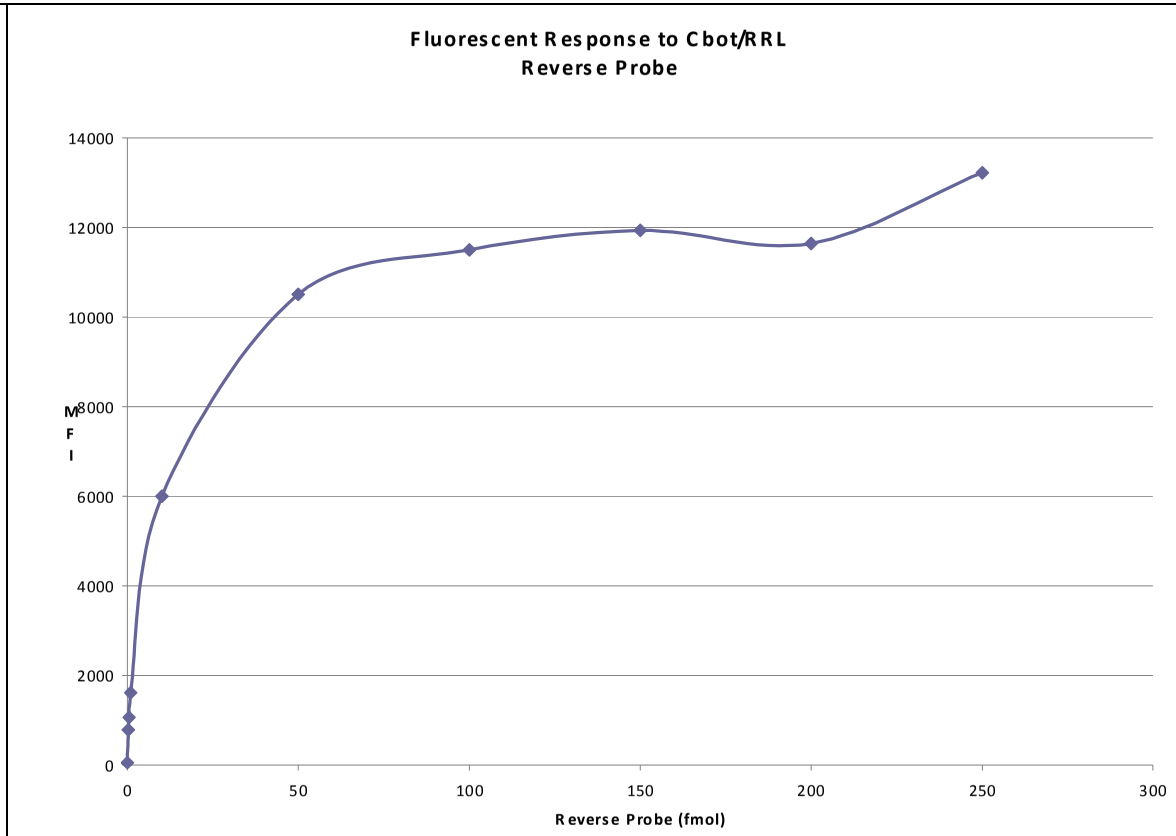


Figure A4.4. Titration study data and chart for Cbot/RRL probe.

Reverse Probe (fmol)	MFI
0	44.0
0.25	1912.0
0.5	1251.0
1	2022.0
10	8881.0
50	12746.0
100	13601.0
150	13571.5
200	13787.5
250	14715.0

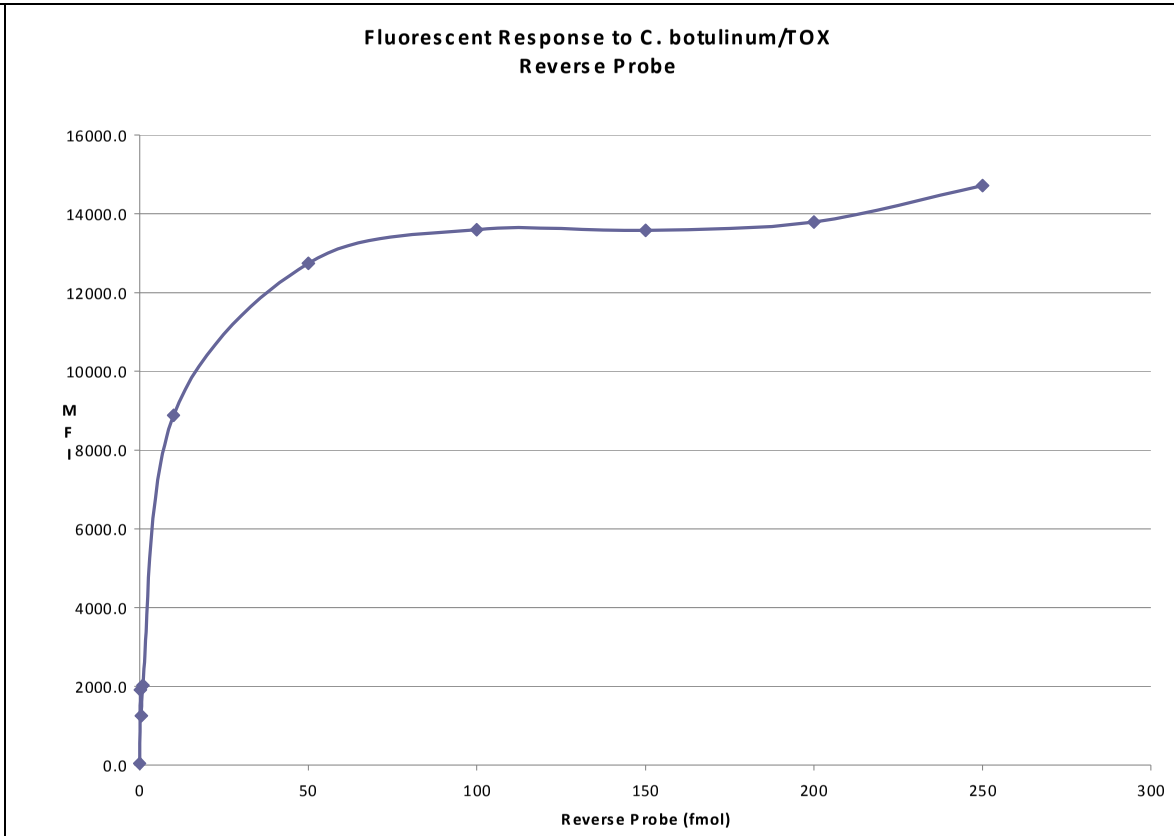


Figure A4.5. Titration study data and chart for Cbot/TOX probe.

Reverse Probe (fmol)	MFI
0	58
0.25	742.5
0.5	1090.5
1	1435
10	7633
50	10609
100	11234
150	11975.5
200	11751.5
250	12488

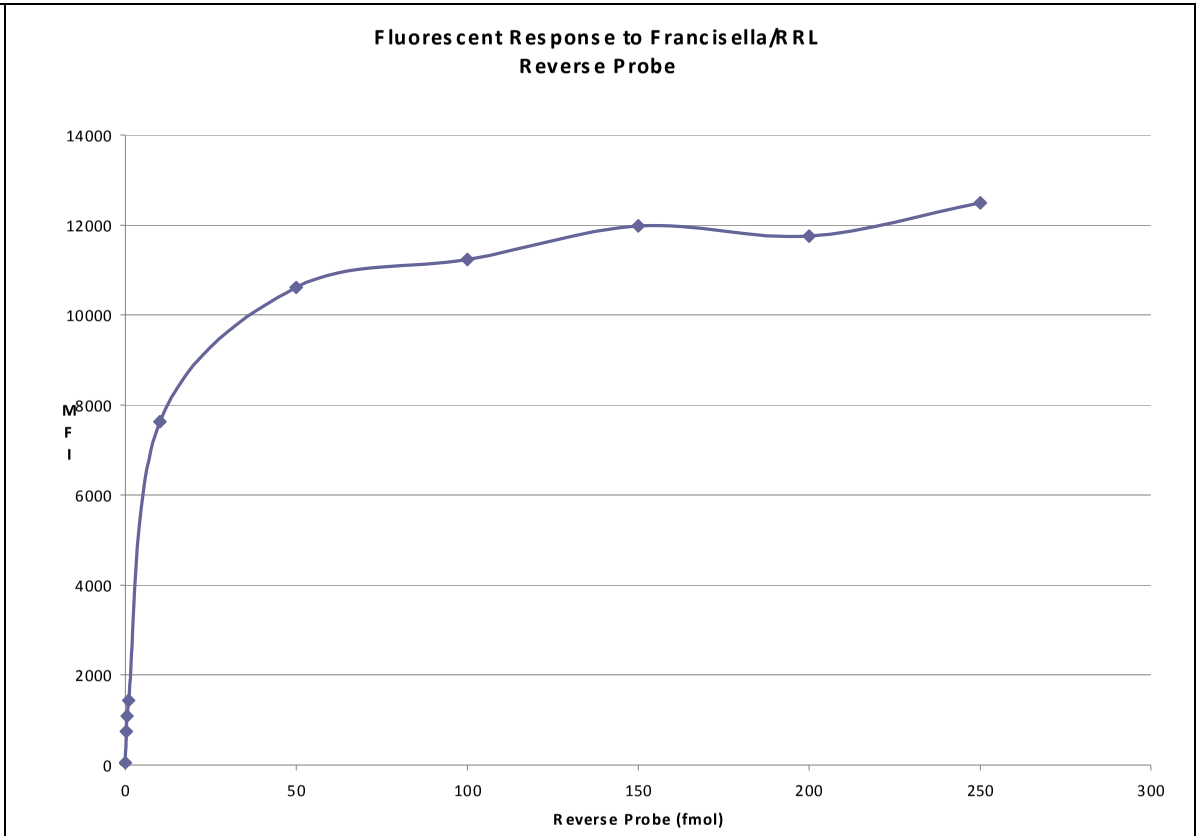


Figure A4.6. Titration study data and chart for Fran/RRL probe.

Reverse Probe (fmol)	MFI
0	64.0
0.25	1106.0
0.5	1998.0
1	2310.0
10	9081.0
50	12272.0
100	12789.0
150	13010.0
200	13255.0
250	13897.5

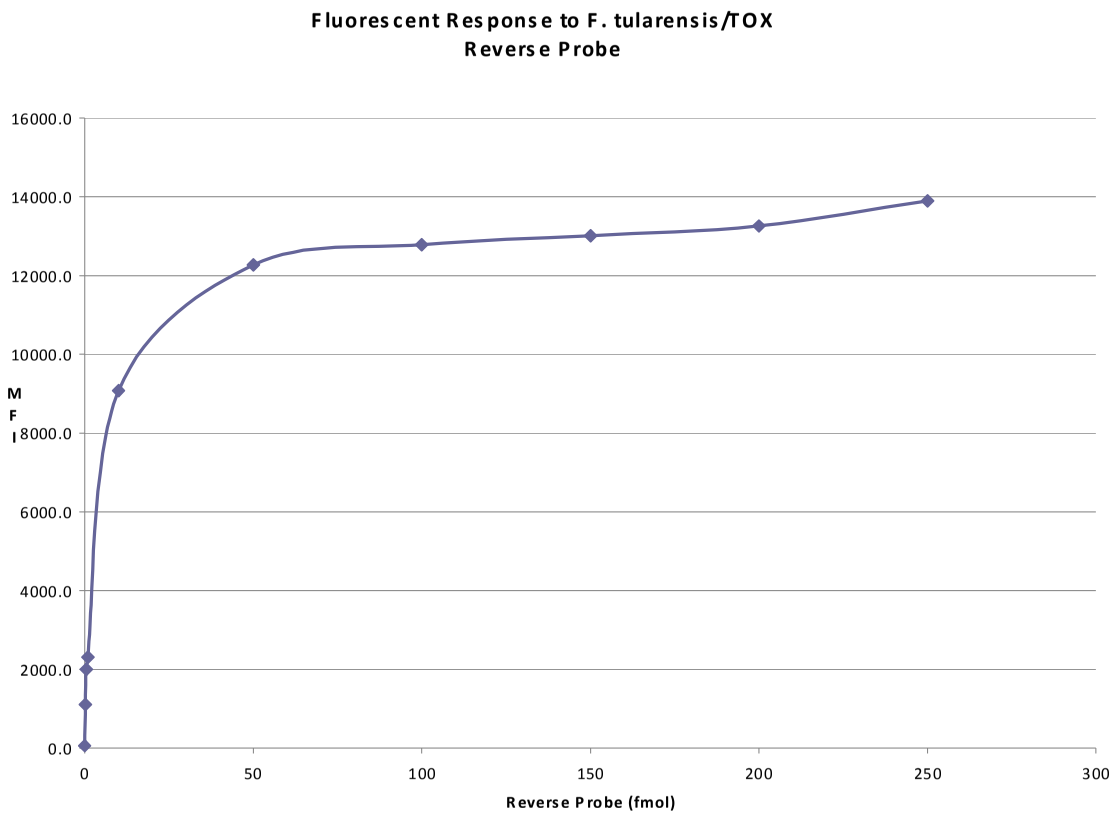


Figure A4.7. Titration study data and chart for Ftul/TOX probe.

Reverse Probe (fmol)	MFI
0	72
0.25	799
0.5	903
1	1025
10	4536
50	8359
100	8929
150	9344
200	9505
250	11499

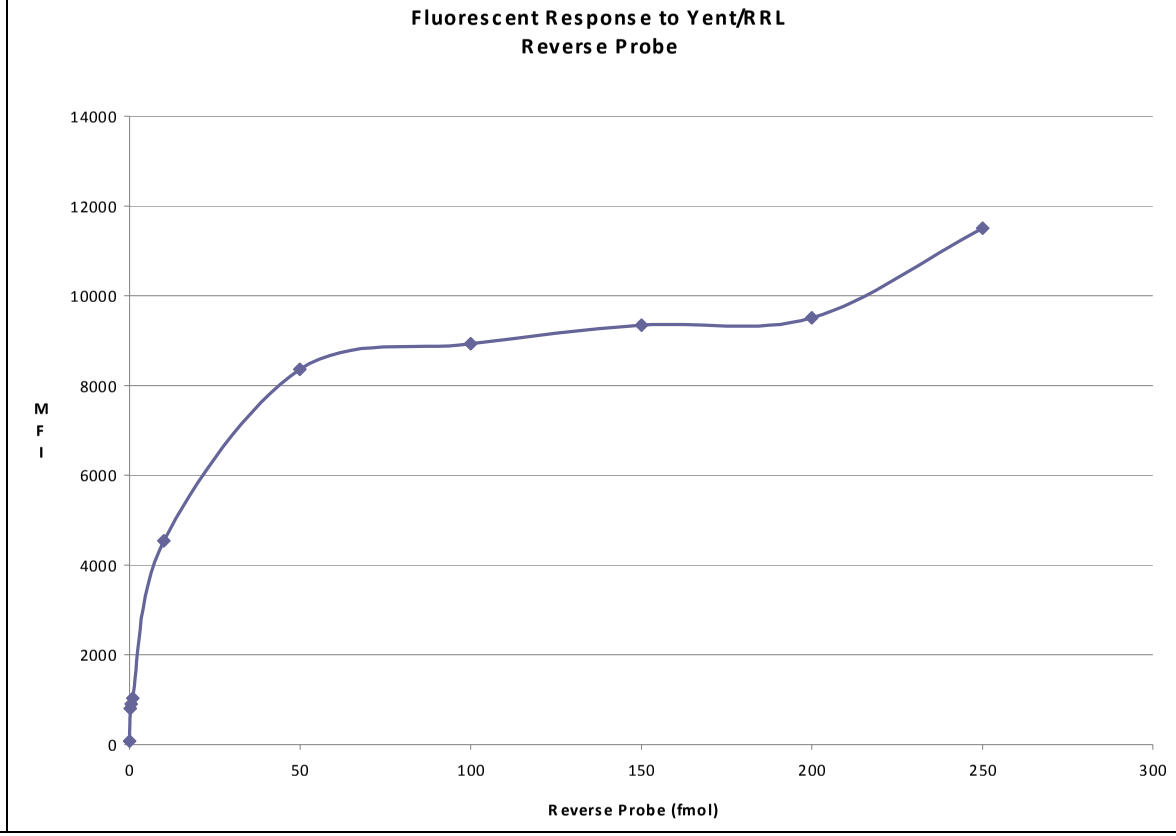


Figure A4.8. Titration study data and chart for Yent/RRL probe.

Reverse Probe (fmol)	MFI
0	97
0.25	755
0.5	685
1	1022
10	4082
50	7813
100	8019
150	7950.5
200	8134
250	10061.5

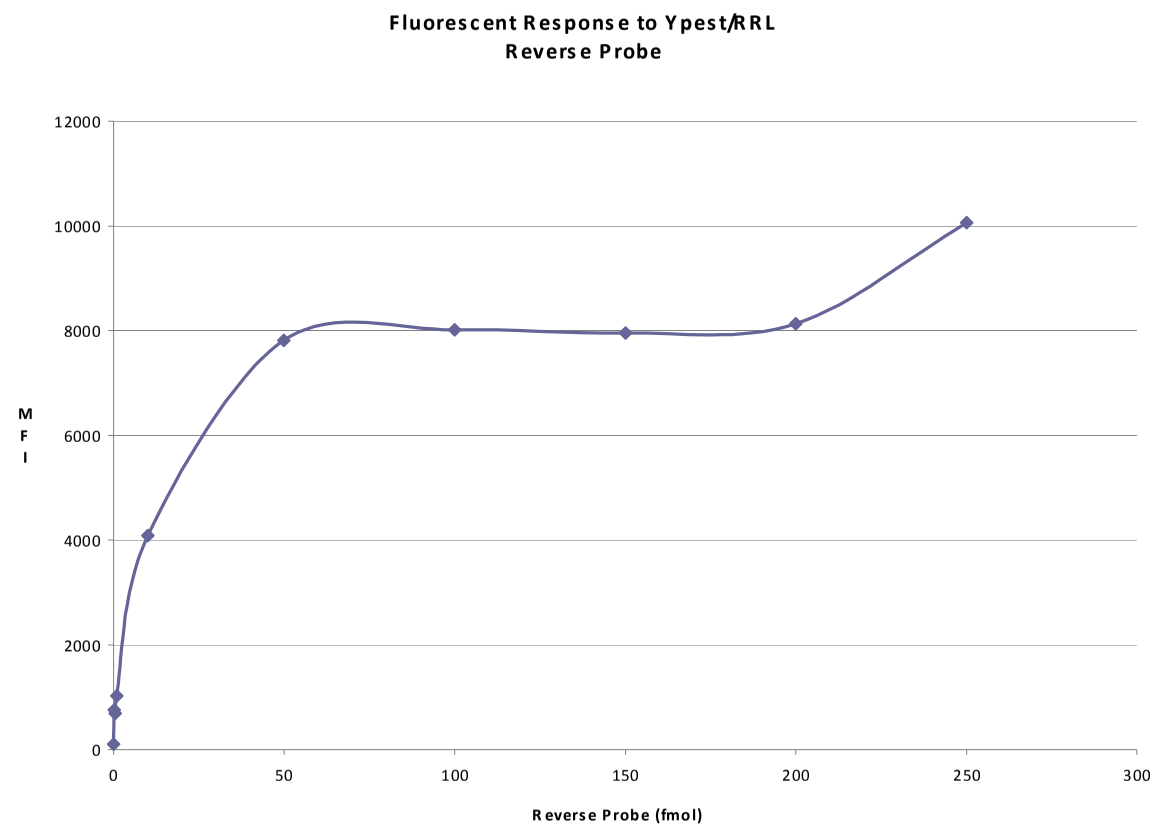


Figure A4.9. Titration study data and chart for Ypes/RRL probe.

Reverse Probe (fmol)	MFI
0	74.0
0.25	1094.0
0.5	1086.0
1	1702.0
10	8289.0
50	13926.0
100	14543.0
150	14707.0
200	14804.0
250	15518.0

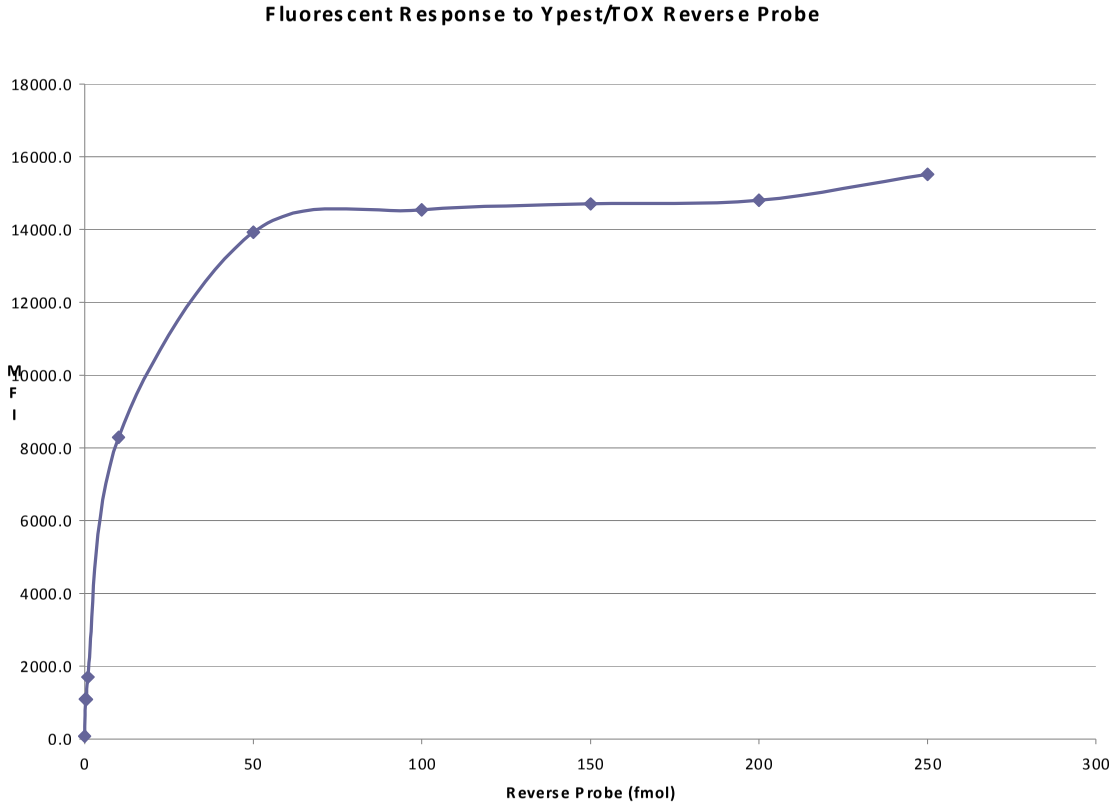


Figure A4.10. Titration study data and chart for Ypes/TOX probe.

Reverse Probe (fmol)	MFI
0	140
0.25	1096
0.5	3006
1	1606.5
10	5513
50	9924
100	9855
150	10315
200	11024
250	11485

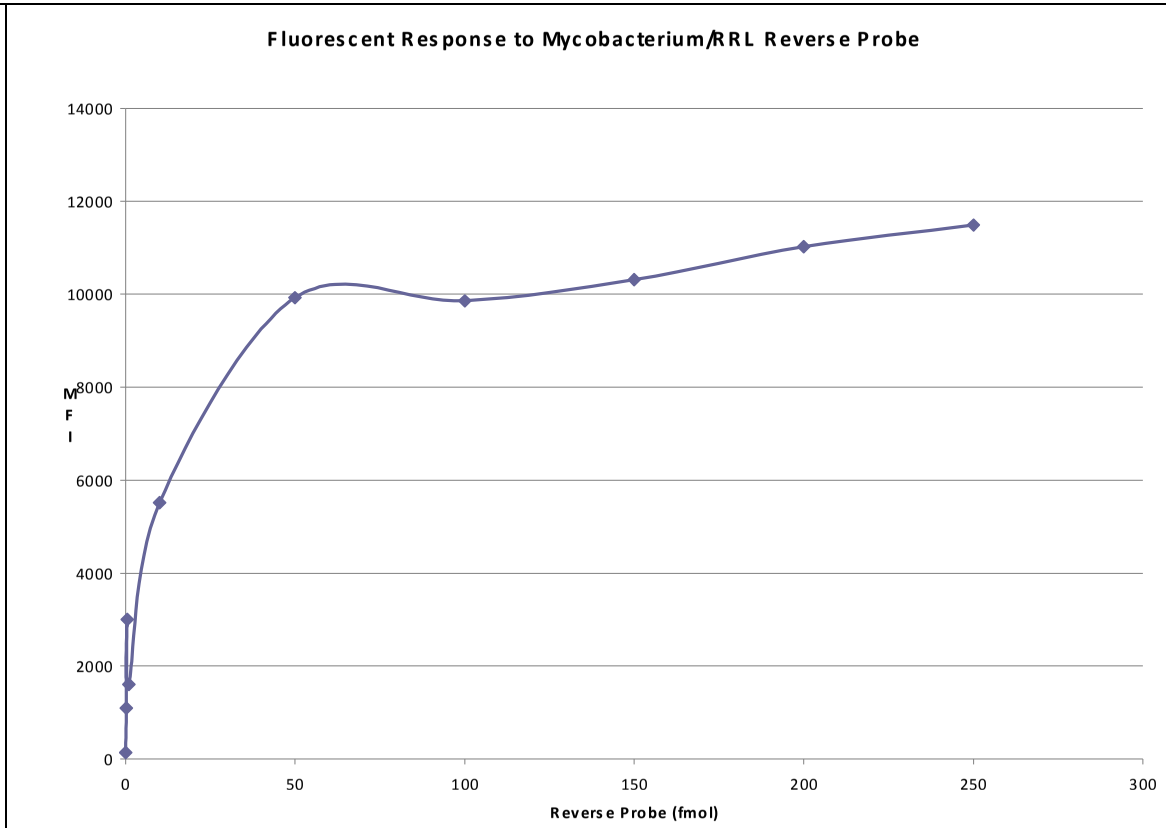


Figure A4.11. Titration study data and chart for Myc/RRL probe.

APPENDIX V. ENZYMATIC DEGRADATION STUDY DATA AND CHARTS

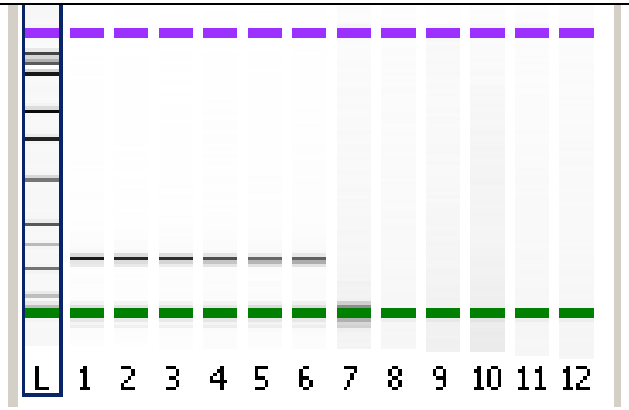


Figure A5.1. Gel image of *B. cereus* amplicons produced before enzymatic digestion. Lane L = ladder; lanes 1-3 = control; lanes 4-6 = experimental; lane 7 = PCR neg. cont.; lanes 8-12 = empty.

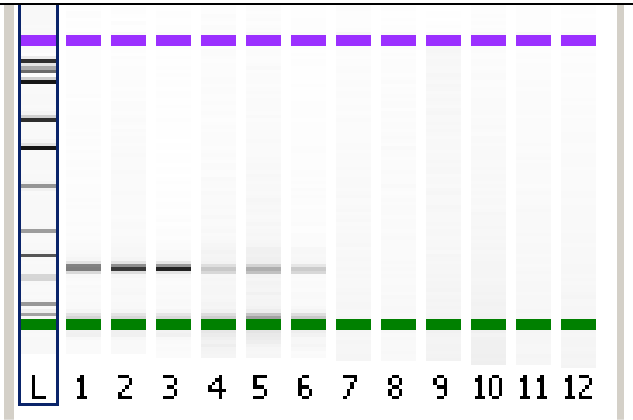


Figure A5.2. Gel image of *B. cereus* amplicons produced after enzymatic digestion. Lane L = ladder; lanes 1-3 = control; lanes 4-6 = experimental; lane 7 = PCR neg. cont.; lanes 8-12 = empty.

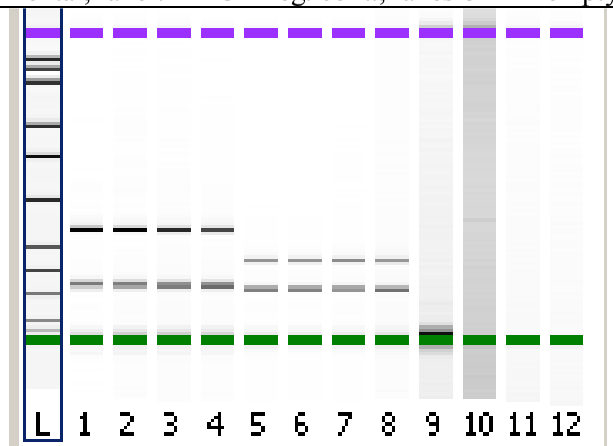


Figure A5.3. Gel image of *B. anthracis* and *Y. pestis* amplicons produced before enzymatic digestion. Lane L = ladder; lanes 1-2 = *B. anthracis* control; lanes 3-4 = *B. anthracis* experimental; lanes 5-6 = *Y. pestis* control; lanes 7-8 = *Y. pestis* experimental; lane 9 = PCR neg. cont.; lanes 10-12 = empty

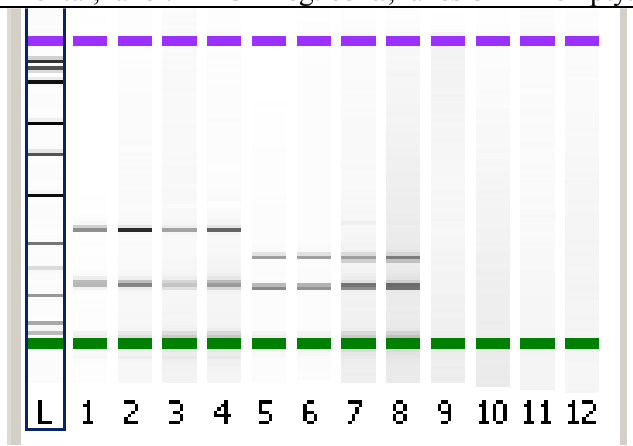


Figure A5.4. Gel image of *B. anthracis* and *Y. pestis* amplicons produced after enzymatic digestion. Lane L = ladder; lanes 1-2 = *B. anthracis* control; lanes 3-4 = *B. anthracis* experimental; lanes 5-6 = *Y. pestis* control; lanes 7-8 = *Y. pestis* experimental; lane 9 = PCR neg. cont.; lanes 10-12 = empty

Lane	Template	Sample Type	Expected Band (bp)	Observed Band (bp)	Concentration (ng/uL)
1	B. cereus	Control	111 (RRL)	122	10.42
2	B. cereus	Control	111 (RRL)	121	11.01
3	B. cereus	Control	111 (RRL)	121	13.38
4	B. cereus	Experimental	111 (RRL)	120	11.17
5	B. cereus	Experimental	111 (RRL)	119	10.43
6	B. cereus	Experimental	111 (RRL)	119	11.01
7	None	PCR Neg. Cont.	None	None	n/a
8-12	Empty				

Table A5.1. Summary of quantitation data for *B. cereus* amplicons produced before enzymatic digestion.

Lane	Template	Sample Type	Expected Band (bp)	Observed Band (bp)	Concentration (ng/uL)
1	B. cereus	Control	111 (RRL)	122	5.76
2	B. cereus	Control	111 (RRL)	121	7.21
3	B. cereus	Control	111 (RRL)	120	6.50
4	B. cereus	Experimental	111 (RRL)	†	n/a
5	B. cereus	Experimental	111 (RRL)	†	n/a
6	B. cereus	Experimental	111 (RRL)	†	n/a
7	None	PCR neg. cont.	None	None	n/a
8-12	Empty				

Table A5.2. Summary of quantitation data for *B. cereus* amplicons produced after enzymatic digestion. Highlighted samples indicate those that were selected for analysis in the Luminex assay. A dagger (†) symbol in the "Observed Bands" column indicates that a faint band was visible but not concentrated enough to be reported. Yellow highlighted cells denote samples whose concentration was used to determine Luminex input amounts.

Lane	Template	Sample Type	Expected Band (bp)	Observed Band (bp)	Concentration (ng/uL)
1	B. anthracis	Control	111 (RRL)	119	14.95
			230 (TOX)	238	12.16
2	B. anthracis	Control	111 (RRL)	119	8.85
			230 (TOX)	237	6.11
3	B. anthracis	Experimental	111 (RRL)	118	9.43
			230 (TOX)	236	6.78
4	B. anthracis	Experimental	111 (RRL)	117	11.92
			230 (TOX)	235	8.33
5	Y. pestis	Control	107 (RRL)	111	17.07
			169 (TOX)	171	9.73
6	Y. pestis	Control	107 (RRL)	111	16.93
			169 (TOX)	171	9.69
7	Y. pestis	Experimental	107 (RRL)	111	16.29
			169 (TOX)	170	9.75
8	Y. pestis	Experimental	107 (RRL)	111	15.09
			169 (TOX)	170	9.65
9	None	PCR Neg. Cont.	None	None	n/a

Table A5.3. Summary of quantitation data for *B. anthracis* and *Y. pestis* amplicons produced before enzymatic digestion.

Lane	Template	Sample Type	Expected Band (bp)	Observed Band (bp)	Concentration (ng/uL)
1	B. anthracis	Control	111 (RRL)	122	7.81
			230 (TOX)	232	7.10
2	B. anthracis	Control	111 (RRL)	121	4.52
			230 (TOX)	231	3.53
3	B. anthracis	Experimental	111 (RRL)	†	n/a
			230 (TOX)	231	0.94
4	B. anthracis	Experimental	111 (RRL)	121	2.95
			230 (TOX)	231	1.66
5	Y. pestis	Control	107 (RRL)	115	10.47
			169 (TOX)	176	5.57
6	Y. pestis	Control	107 (RRL)	115	9.57
			169 (TOX)	176	5.05
7	Y. pestis	Experimental	107 (RRL)	116	4.02
			169 (TOX)	170	1.65
8	Y. pestis	Experimental	107 (RRL)	115	4.31
			169 (TOX)	170	1.18
9	None	PCR Neg. Cont.	None	None	n/a

Table A5.4. Summary of quantitation data for *B. anthracis* and *Y. pestis* amplicons produced after enzymatic digestion. A dagger (†) symbol in the "Observed Bands" column indicates that a faint band was visible but not concentrated enough to be reported. Yellow highlighted cells denote samples whose concentration was used to determine Luminex input amounts.

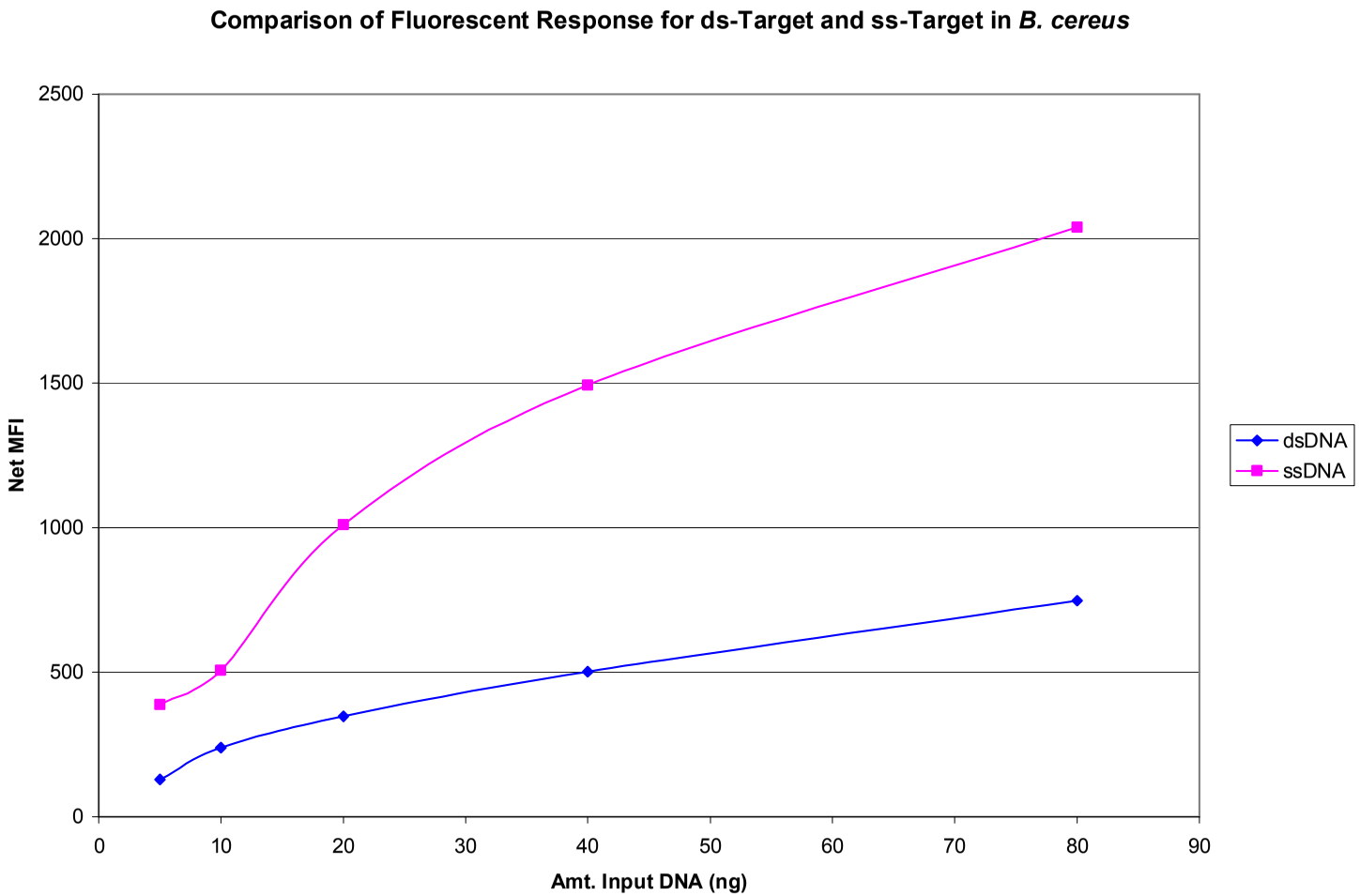


Figure A5.5. Comparison of fluorescent response for ds-target and ss-target in *B. cereus*.

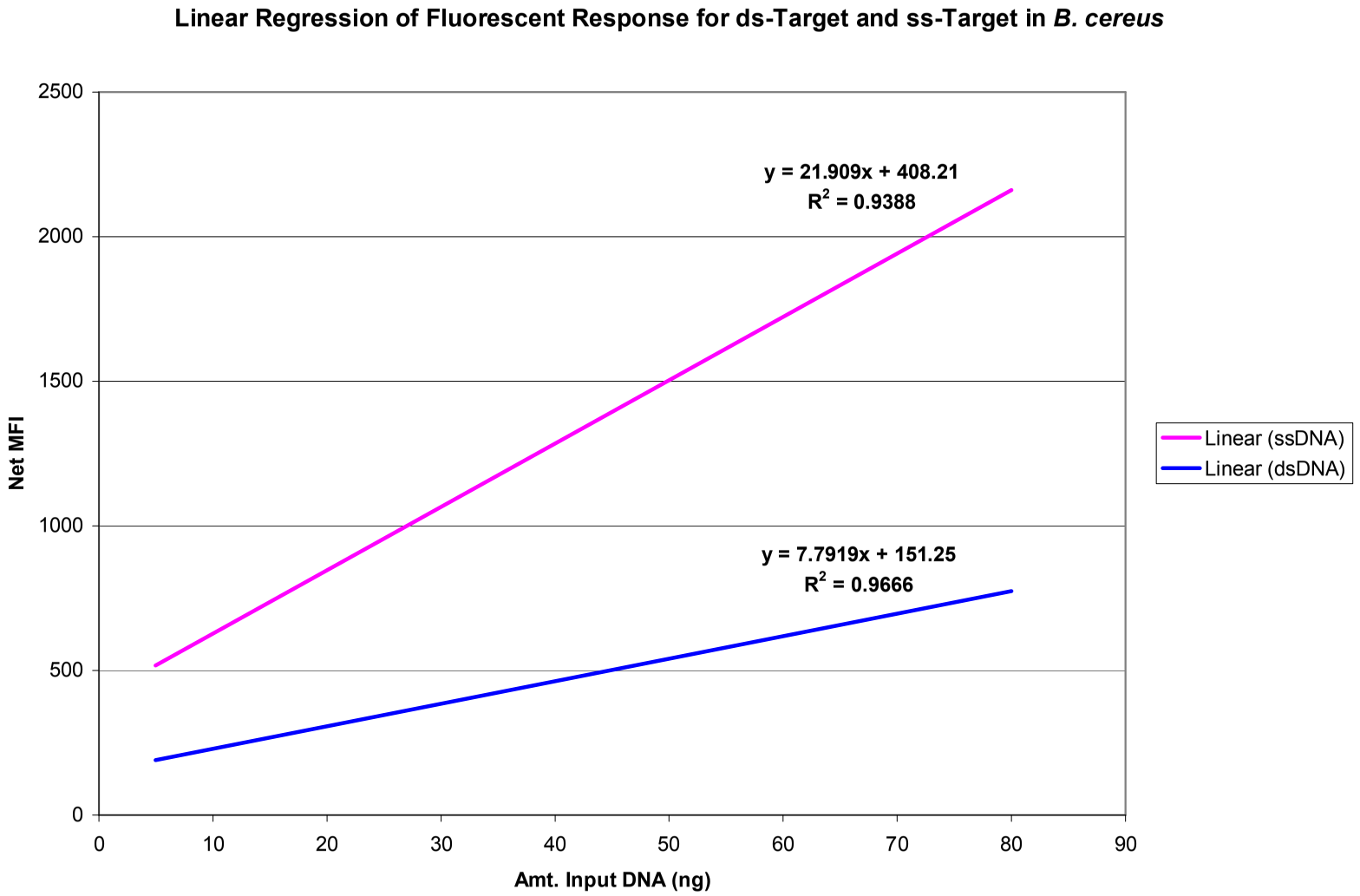


Figure A5.6. Linear regression of fluorescent response for ds-target and ss-target in *B. cereus*.

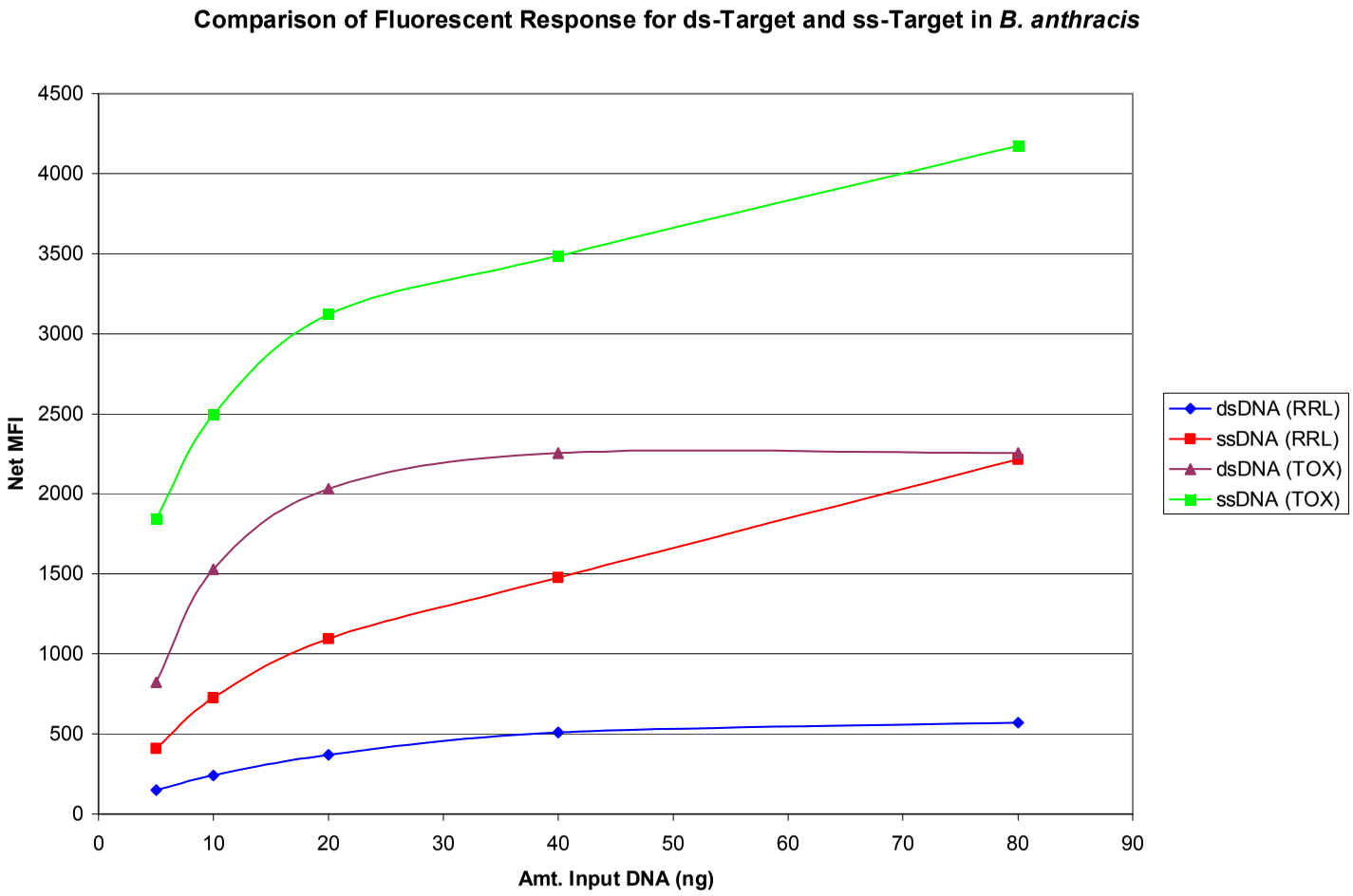


Figure A5.7. Comparison of fluorescent response for ds-target and ss-target in *B. anthracis.*

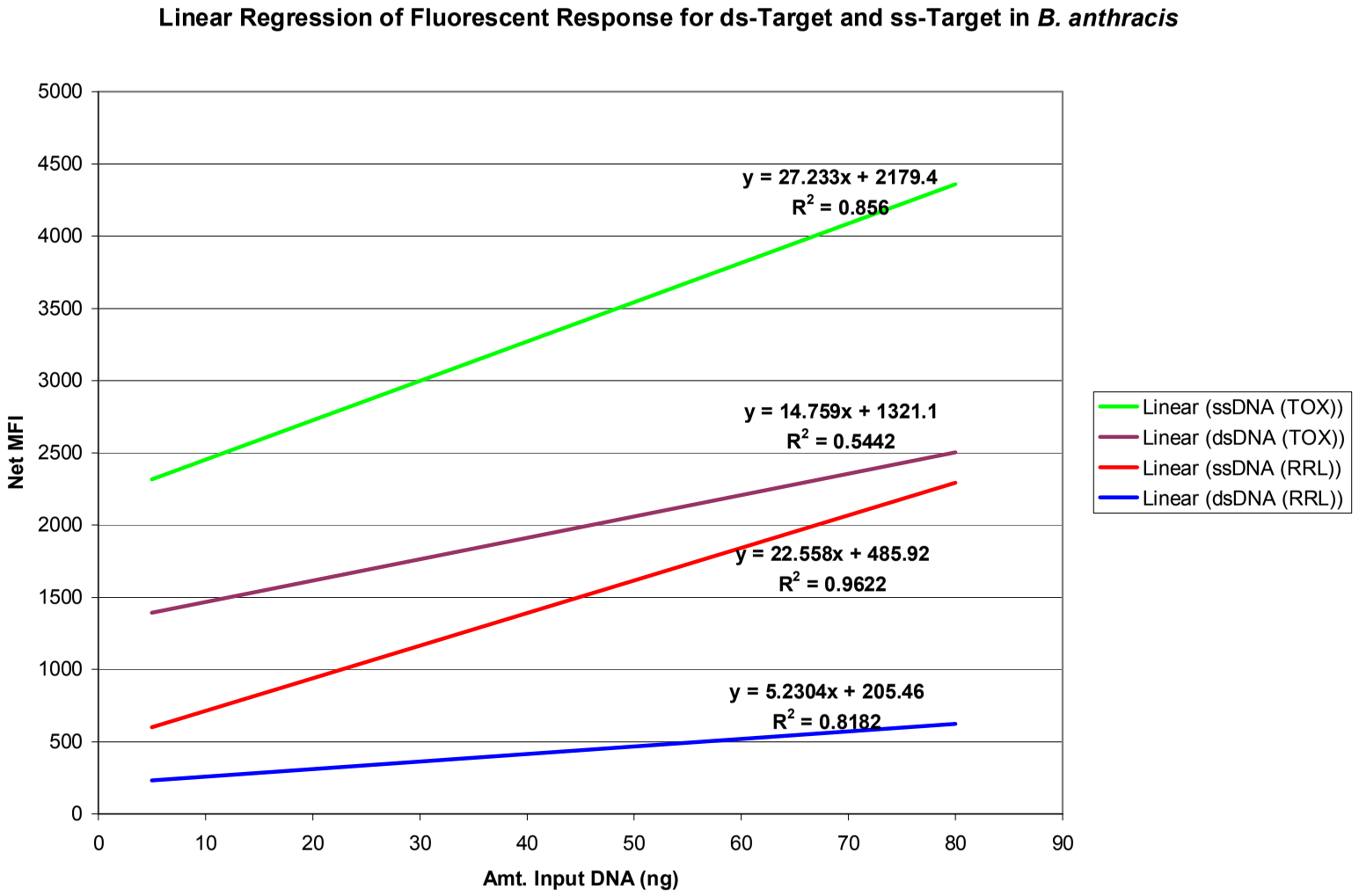


Figure A5.8. Linear regression of fluorescent response for ds-target and ss-target in *B. anthracis*.

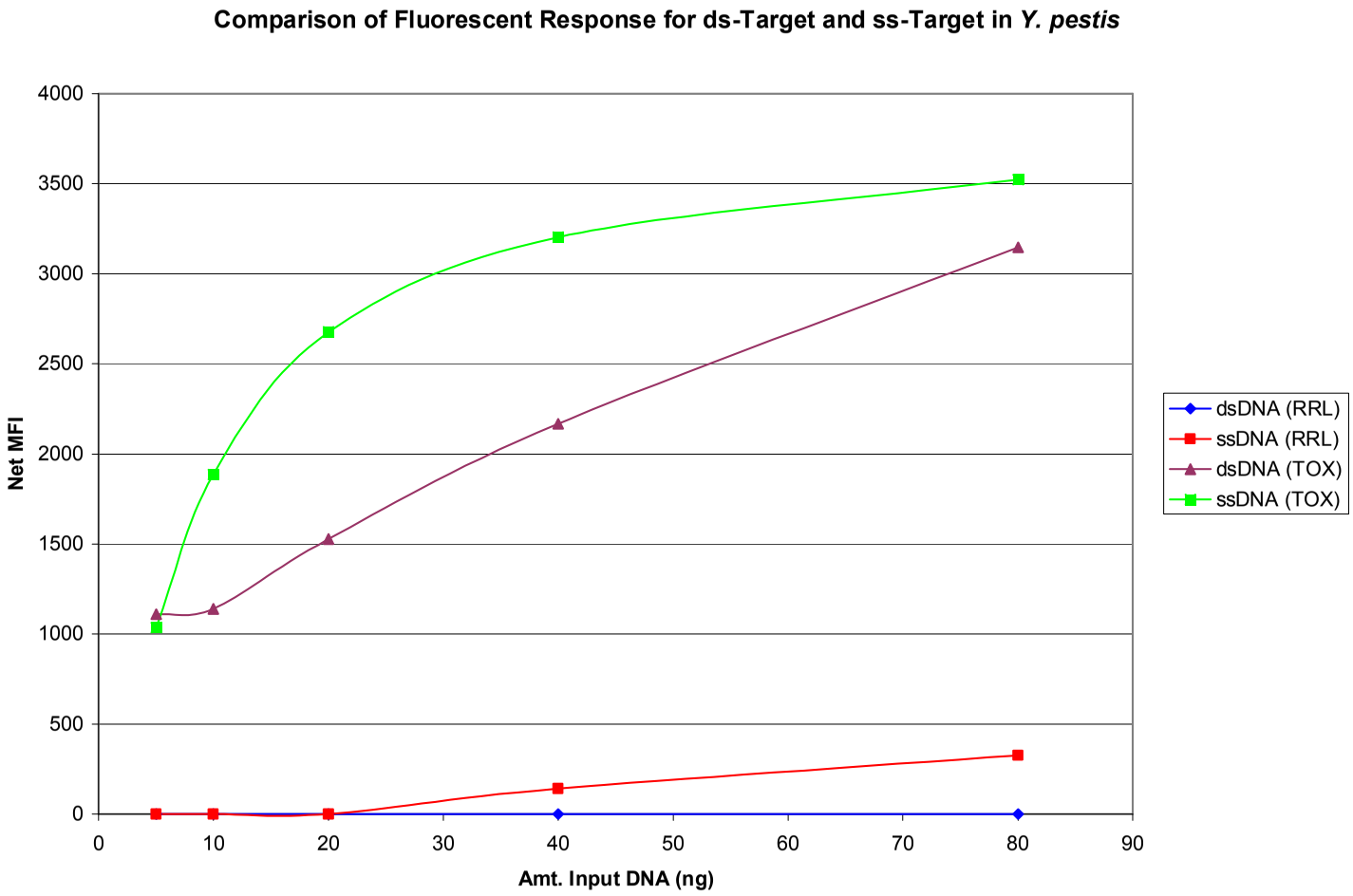


Figure A5.9. Comparison of fluorescent response for ds-target and ss-target in *Y. pestis*.

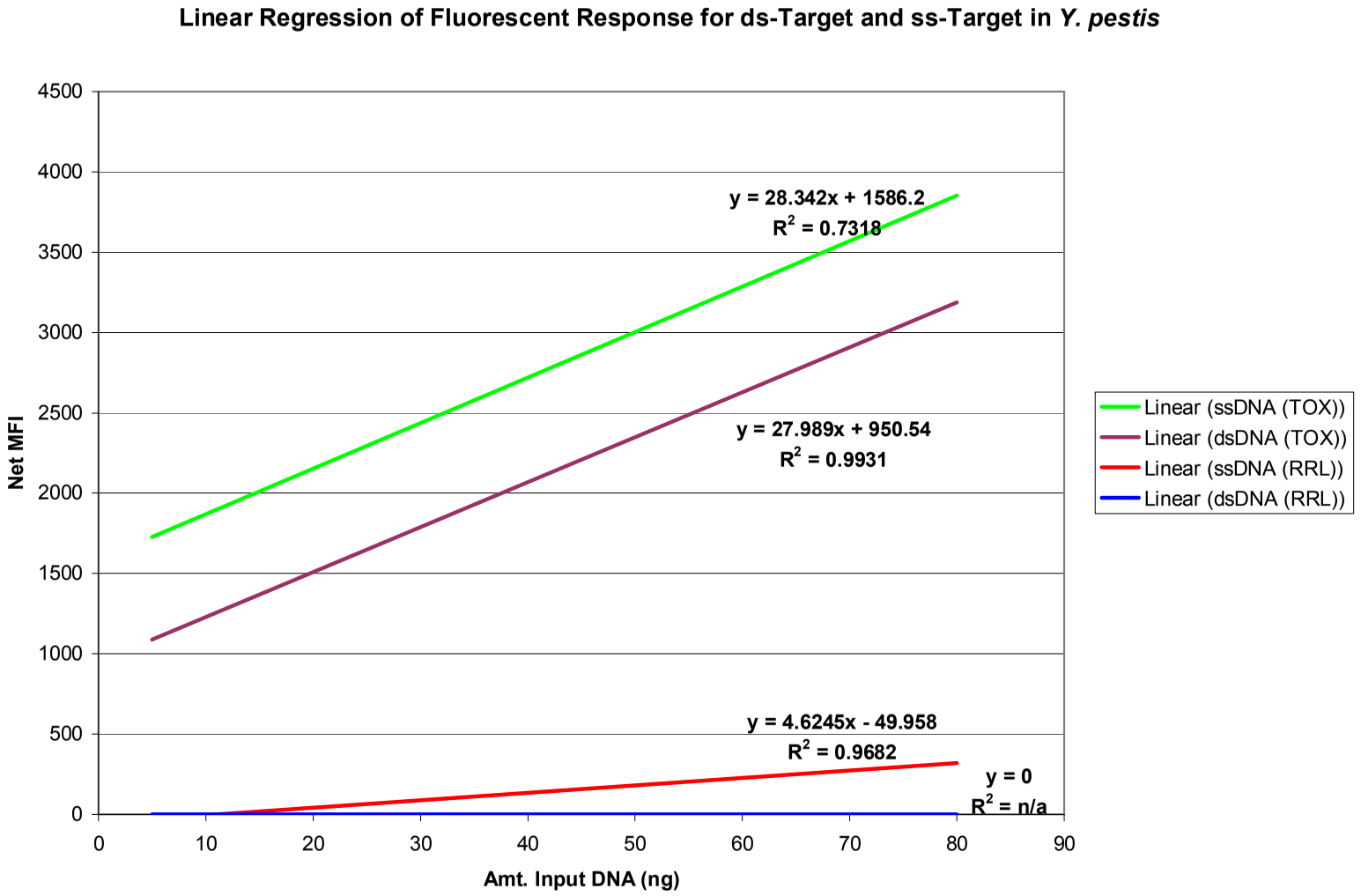


Figure A5.10. Linear regression of fluorescent response for ds-target and ss-target in *Y. pestis*.

APPENDIX VI. SPECIFICITY/SENSITIVITY STUDY DATA AND CHARTS

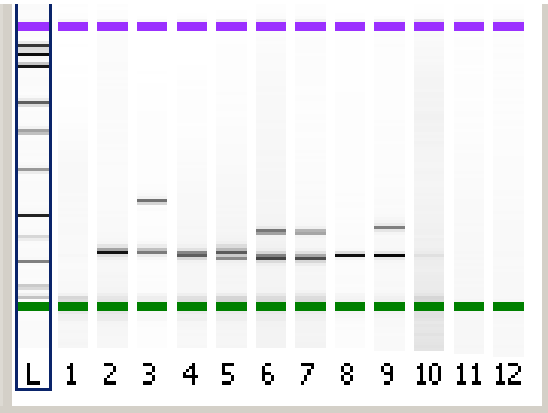


Figure A6.1. Gel image of pre-digest enzymatic degradation control samples for specificity/sensitivity study.

Lane	Template	Sample Type	Expected Band (bp)	Observed Band (bp)	Concentration (ng/uL)
1	None	Control	None	None	n/a
2	B. cereus	Control	111 (RRL)	121	5.63
3	B. anthracis	Control	111 (RRL)	121	7.40
			230 (TOX)	232	5.77
4	C. difficile	Control	110 (RRL)	115	4.24
5	C. botulinum	Control	109 (RRL)	120	3.38
			100 (TOX)	109	1.73
6	F. holarctica	Control	102 (RRL)	109	4.86
				164	2.94
7	F. tularensis	Control	102 (RRL)	108	3.84
			159 (TOX)	163	2.45
8	Y. enterocolitica	Control	107 (RRL)	115	10.47
9	Y. pestis	Control	107 (RRL)	114	6.78
			169 (TOX)	171	4.32
10	M. tuberculosis	Control	110 (RRL)	†	n/a

Table A6.1. Summary of pre-digest quantitation data for enzymatic degradation control samples for specificity/sensitivity study. A dagger (†) symbol in the "Observed Bands" column indicates that a faint band was visible but not concentrated enough to be reported.

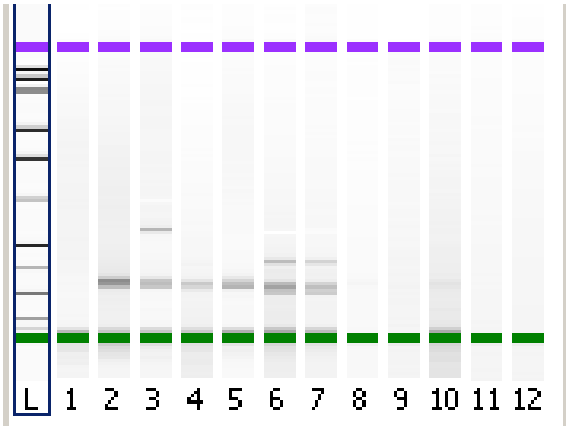


Figure A6.2. Gel image of post-digest enzymatic degradation experimental samples for specificity/sensitivity study.

Lane	Template	Sample Type	Expected Band (bp)	Observed Band (bp)	Concentration (ng/uL)
1	None	Experimental	None	None	n/a
2	B. cereus	Experimental	111 (RRL)	123	3.97
3	B. anthracis	Experimental	111 (RRL)	122	1.75
			230 (TOX)	233	0.94
4	C. difficile	Experimental	110 (RRL)	†	n/a
5	C. botulinum	Experimental	109 (RRL)	114	1.21
			100 (TOX)	†	n/a
6	F. holarctica	Experimental	102 (RRL)	112	2.33
7	F. tularensis	Experimental	102 (RRL)	†	n/a
			159 (TOX)	†	n/a
8	Y. enterocolitica	Experimental	107 (RRL)	†	n/a
9	Y. pestis	Experimental	107 (RRL)	†	n/a
			169 (TOX)	†	n/a
10	M. tuberculosis	Experimental	110 (RRL)	†	n/a

Table A6.2. Summary of post-digest quantitation data for enzymatic degradation experimental samples for specificity/sensitivity study. A dagger (†) symbol in the "Observed Bands" column indicates that a faint band was visible but not concentrated enough to be reported.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	61	28	23	21	16	29	21	23	23	35	21
3X NC	183	84	69	63	48	87	63	69	69	105	63
	Net MFI*										
BC/0.1ng	0	0	0	0	0	0	0	0	0	0	0
BC/0.5ng	0	0	0	0	0	0	0	0	0	0	0
BC/1ng	86	0	0	0	0	0	0	0	0	0	0
BC/5ng	447	0	0	0	0	0	0	0	0	0	0
BC/10ng	984	0	115	0	0	1	10	8	0	0	0
BC/20ng	962	0	0	0	0	0	0	0	0	0	0
BC/40ng	1761	0	142	0	0	8	2	46	0	0	17
BC/80ng	2132	0	335	10	3	66	89	94	16	0	100
BC – PC	4720	41	62	36	25	40	4	51	40	41	20

Table A6.3. Net MFI data for *B. cereus* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

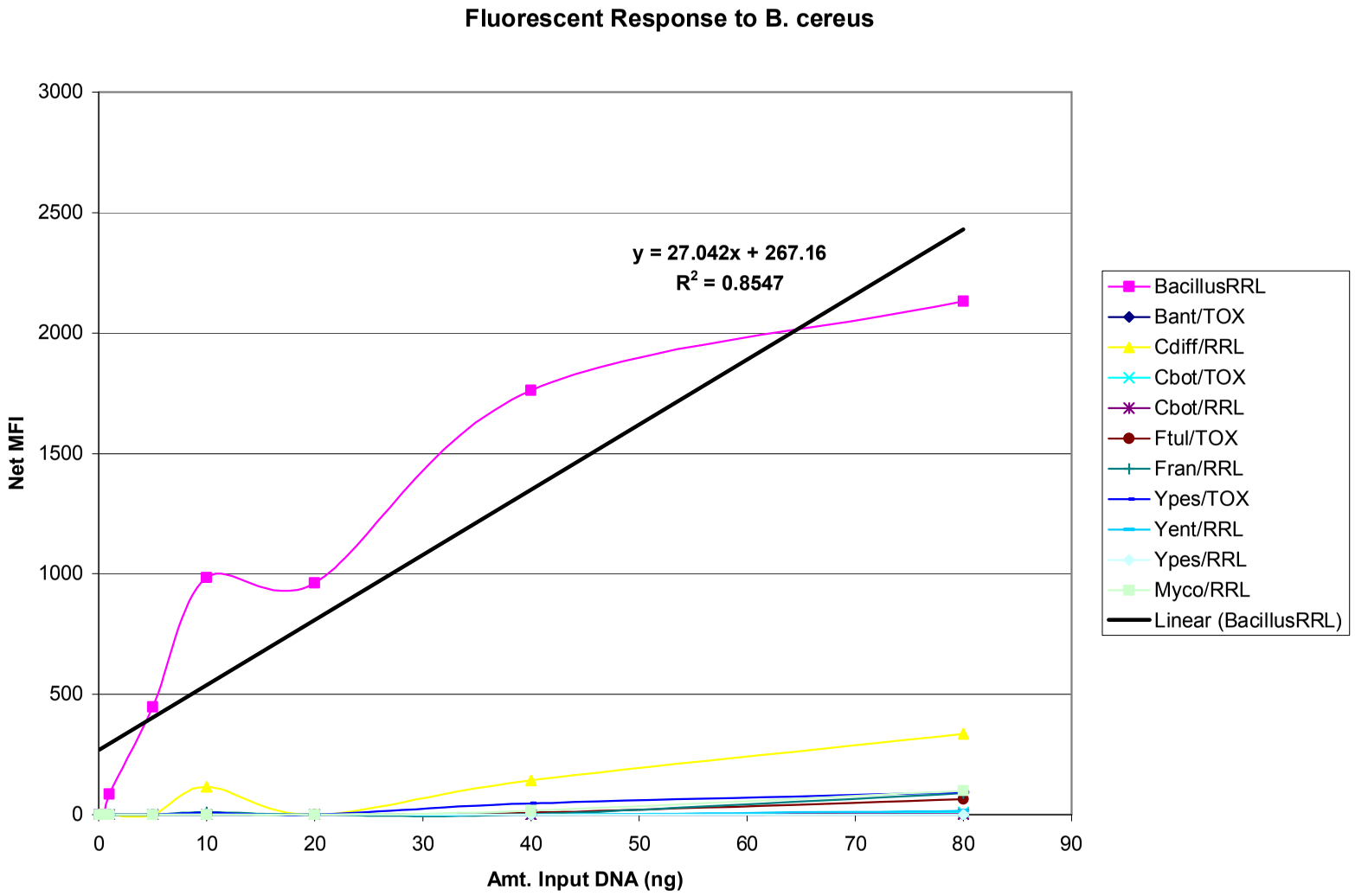


Figure A6.3. Fluorescent response of eleven probes to *B. cereus* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
BA – NC	35	22	38	31	28	41	21	29	28	43	20
3X NC	105	66	114	93	84	123	63	87	84	129	60
	Net MFI*										
BA/0.1ng	0	94	0	0	0	0	0	0	0	0	0
BA/0.5ng	0	409	0	0	0	0	0	0	0	0	0
BA/1ng	82	823	0	0	0	0	0	0	0	0	0
BA/5ng	455	2122	0	0	0	0	0	0	0	0	0
BA/10ng	862	2857	0	0	0	0	0	0	0	0	0
BA/20ng	1141	3379	0	0	0	0	0	0	0	0	0
BA/40ng	1575	3859	116	0	0	0	0	6	0	0	0
BA/80ng	1898	3767	71	0	0	0	9	16	0	0	4
BA – PC	5003	7223	0	0	0	0	3	0	8	0	0

Table A6.4. Net MFI data for *B. anthracis* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

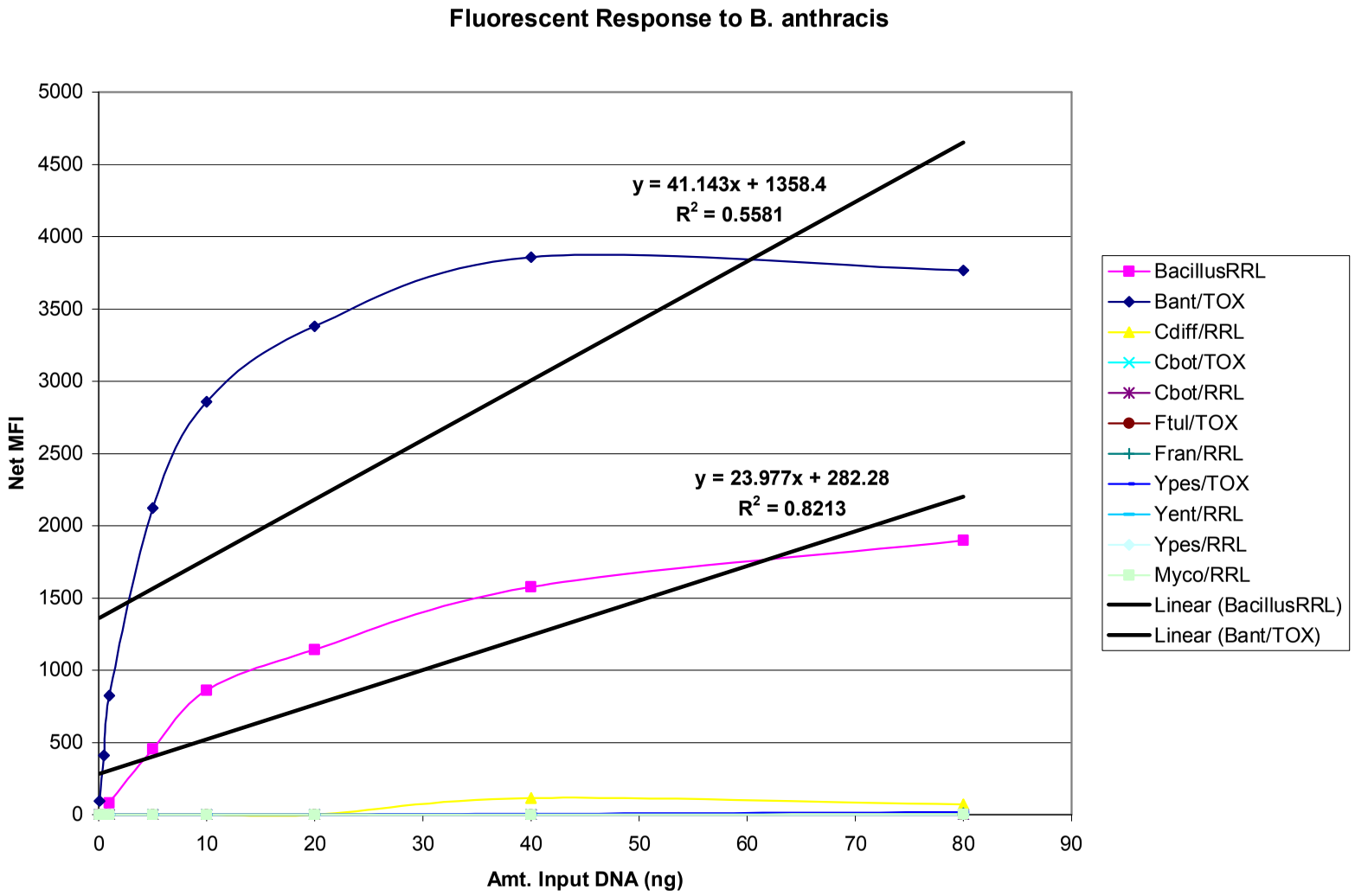


Figure A6.4. Fluorescent response of eleven probes to *B. anthracis* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
CB – NC	46	42	46	245	30	36	36	45	55	47	38
3X NC	138	126	138	735	90	108	108	135	165	141	114
	Net MFI*										
CB/0.1ng	0	0	0	0	0	0	0	0	0	0	0
CB/0.5ng	0	0	0	132	0	0	0	0	0	0	0
CB/1ng	0	0	0	607	0	0	0	0	0	0	0
CB/5ng	0	0	0	1749	48	0	0	0	0	0	0
CB/10ng	0	15	10	2392	125	0	0	0	0	0	0
CB/20ng	0	50	45	3208	159	0	0	0	0	0	0
CB/40ng	0	90	149	3550	279	0	0	0	0	0	0
CB/80ng	0	113	188	4003	298	0	5	0	0	0	0
CB – PC	0	171	52	4718	4676	0	0	0	0	0	0

Table A6.5. Net MFI data for *C. botulinum* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

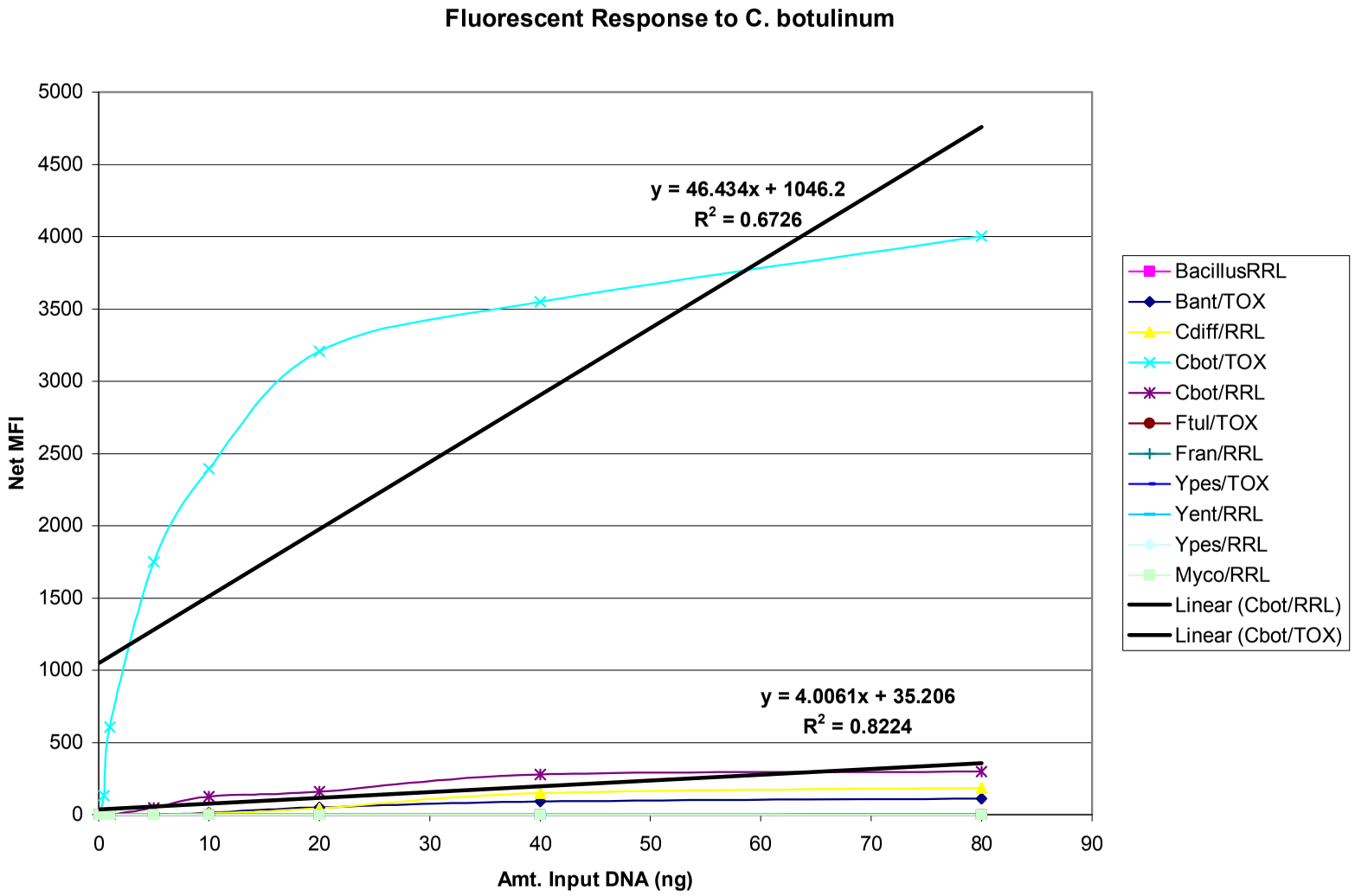


Figure A6.5. Fluorescent response of eleven probes to *C. botulinum* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
FH – NC	56	41	52	36	39	81	40	56	56	54	40
3X NC	168	123	156	108	117	243	120	168	168	162	120
	Net MFI*										
FH/0.1ng	0	0	0	0	0	0	0	0	0	0	0
FH/0.5ng	0	0	0	0	0	85	144	0	0	0	0
FH/1ng	0	0	0	0	0	194	393	0	0	0	0
FH/5ng	0	0	0	0	0	978	1090	0	0	0	0
FH/10ng	0	0	0	0	0	1497	1555	0	0	0	0
FH/20ng	0	0	41	0	0	2090	2337	0	0	0	0
FH/40ng	0	0	188	0	0	2584	2959	32	0	0	1
FH/80ng	0	0	91	0	0	3043	2614	0	0	0	0
FH – PC	0	0	0	0	0	0	3765	0	0	0	0

Table A6.6. Net MFI data for *F. tularensis holarctica* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

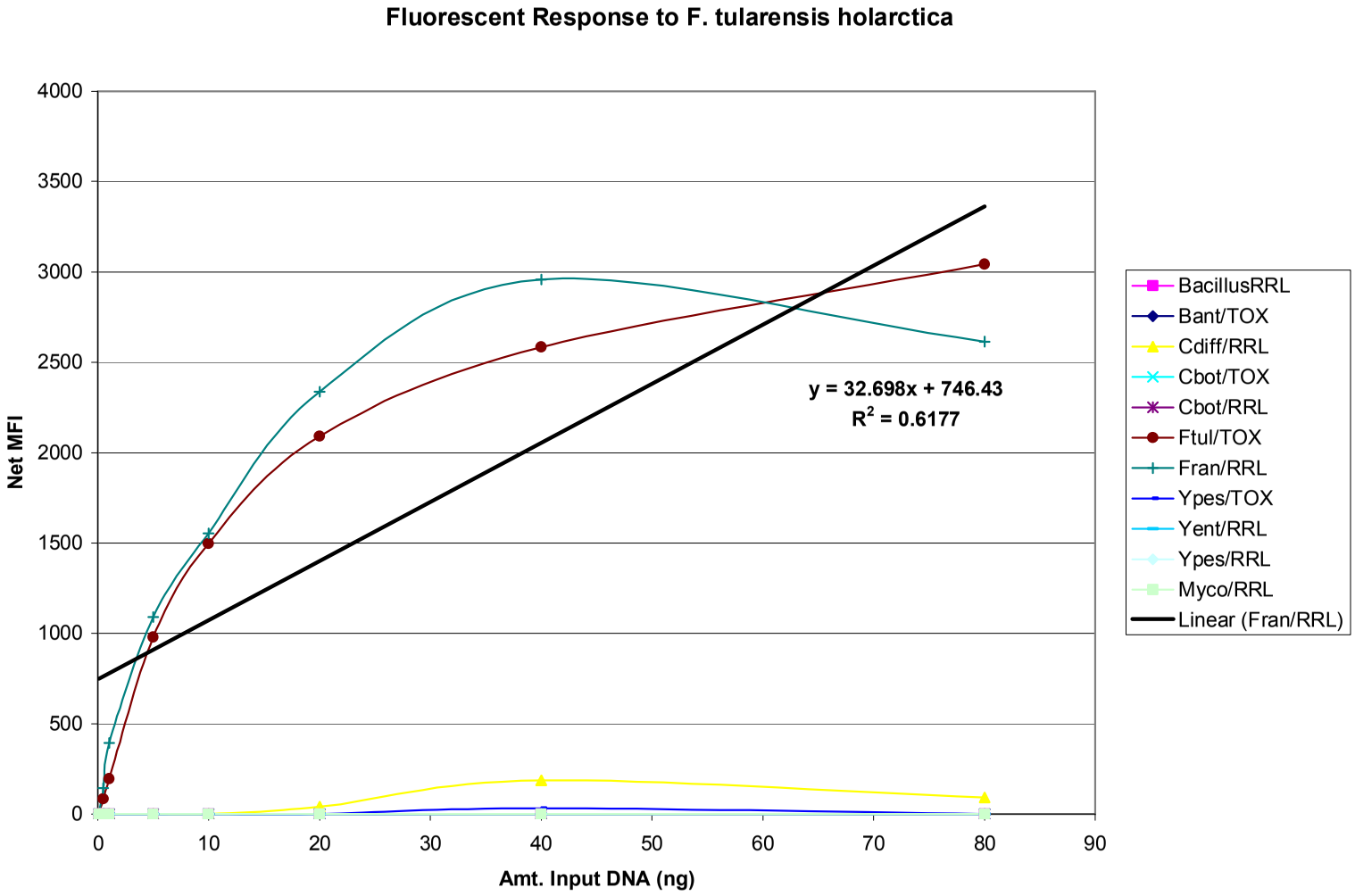


Figure A6.6. Fluorescent response of eleven probes to *F. tularensis holarctica* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
FT – NC	57	41	42	39	36	182	75	63	55	49	29
3X NC	171	123	126	117	108	546	225	189	165	147	87
	Net MFI*										
FT/0.1ng	0	0	0	0	0	0	0	0	0	0	0
FT/0.5ng	0	0	0	0	0	623	37	0	0	0	0
FT/1ng	0	0	0	0	0	1511	400	0	0	0	0
FT/5ng	0	0	0	0	0	2418	1055	0	0	0	0
FT/10ng	0	0	5	0	0	2864	1552	0	0	0	0
FT/20ng	0	0	46	0	0	3259	1996	36	0	0	0
FT/40ng	0	0	93	0	0	3627	2398	31	0	0	0
FT/80ng	0	0	226	17	0	3827	2926	127	0	0	52
FT – PC	0	0	6	10	0	6962	3816	36	0	0	0

Table A6.7. Net MFI data for *F. tularensis tularensis* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

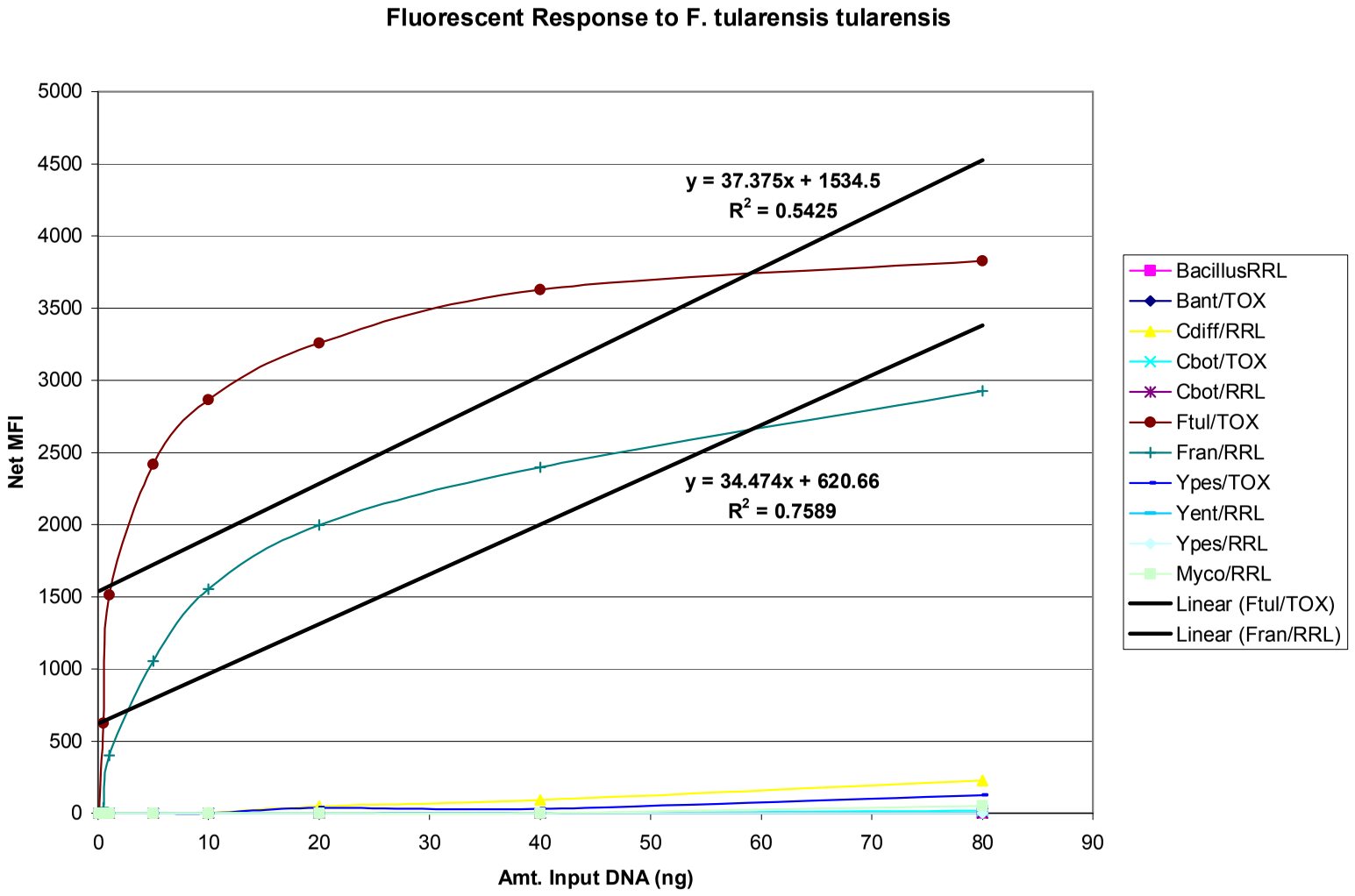


Figure A6.7. Fluorescent response of eleven probes to *F. tularensis tularensis* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
YE – NC	45	44	40	48	36	43	38	35	55	48	30
3X NC	135	132	120	144	108	129	114	105	165	144	90
	Net MFI*										
YE/0.1ng	0	0	0	0	0	0	0	0	0	0	0
YE/0.5ng	0	0	0	0	0	0	0	0	0	0	0
YE/1ng	0	0	0	0	0	0	0	0	0	0	0
YE/5ng	0	0	0	0	0	0	0	0	0	0	0
YE/10ng	0	0	0	0	0	0	0	0	21	0	0
YE/20ng	0	0	0	0	0	0	0	0	92	0	0
YE/40ng	0	0	52	0	0	0	0	0	185	0	0
YE/80ng	9	0	114	0	0	0	17	9	364	0	3
YE – PC	0	0	5	0	0	0	134	6	2553	24	0

Table A6.8. Net MFI data for *Y. enterocolitica* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

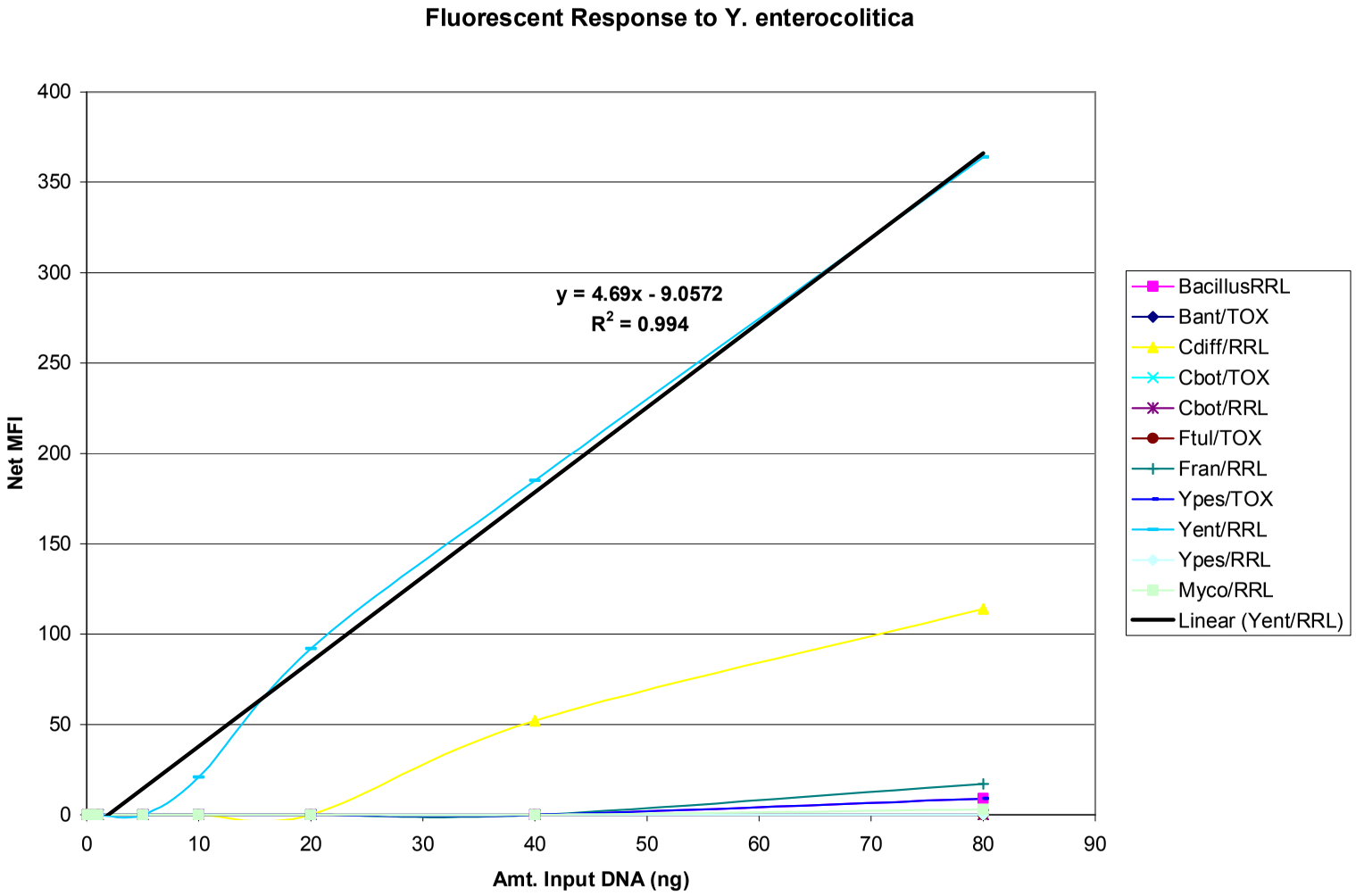


Figure A6.8. Fluorescent response of eleven probes to *Y. enterocolitica* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
YP – NC	42	34	45	41	36	33	32	118	46	33	32
3X NC	126	102	135	123	108	99	96	354	138	99	96
	Net MFI*										
YP/0.1ng	0	0	0	0	0	0	0	0	0	0	0
YP/0.5ng	0	0	0	0	0	0	0	253	0	0	0
YP/1ng	0	0	0	0	0	0	0	601	0	0	0
YP/5ng	0	0	0	0	0	0	0	1601	0	0	0
YP/10ng	0	0	0	0	0	0	0	1974	0	18	0
YP/20ng	0	0	0	0	0	0	0	2661	0	19	0
YP/40ng	0	0	125	0	0	0	0	3110	0	95	0
YP/80ng	0	0	123	0	0	0	7	3599	0	86	0
YP – PC	0	0	0	0	0	0	52	5404	374	4000	0

Table A6.9. Net MFI data for *Y. pestis* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

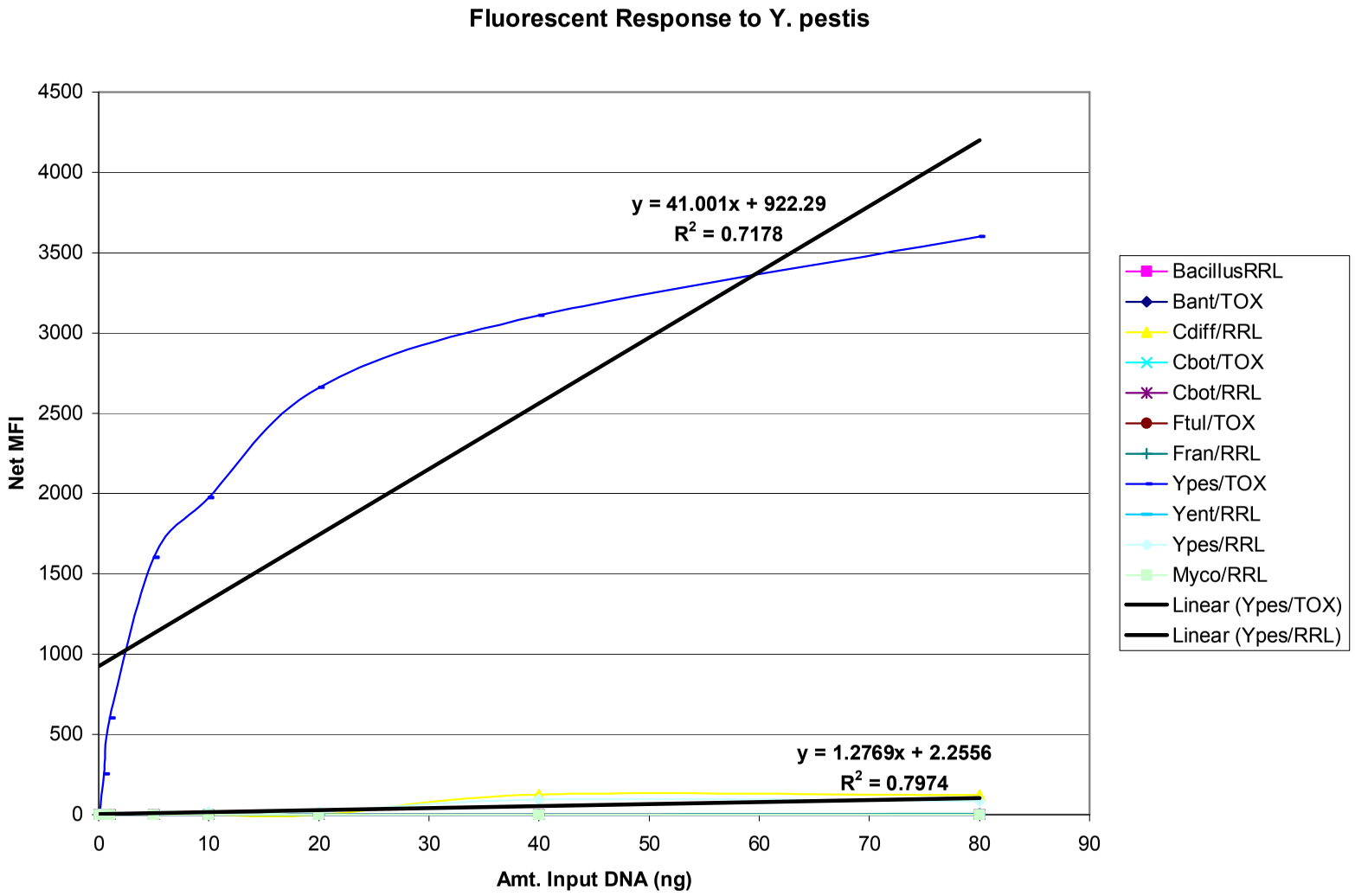


Figure A6.9. Fluorescent response of eleven probes to *Y. pestis* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
MT – NC	58	43	42	48	42	49	49	53	64	36	34
3X NC	174	129	126	144	126	147	147	159	192	108	102
	Net MFI*										
MT/0.05ng	0	0	0	0	0	0	0	0	0	0	0
MT/0.2ng	0	0	136	0	0	0	0	0	0	0	80
MT/0.4ng	28	0	421	0	0	0	4	63	0	0	215
MT/0.6ng	0	0	148	0	0	0	0	0	0	0	126
MT/0.85ng	0	0	201	0	0	0	0	11	0	0	184
MT – PC	0	0	0	0	0	0	0	0	0	12	5929

Table A6.10. Net MFI data for *M. tuberculosis* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

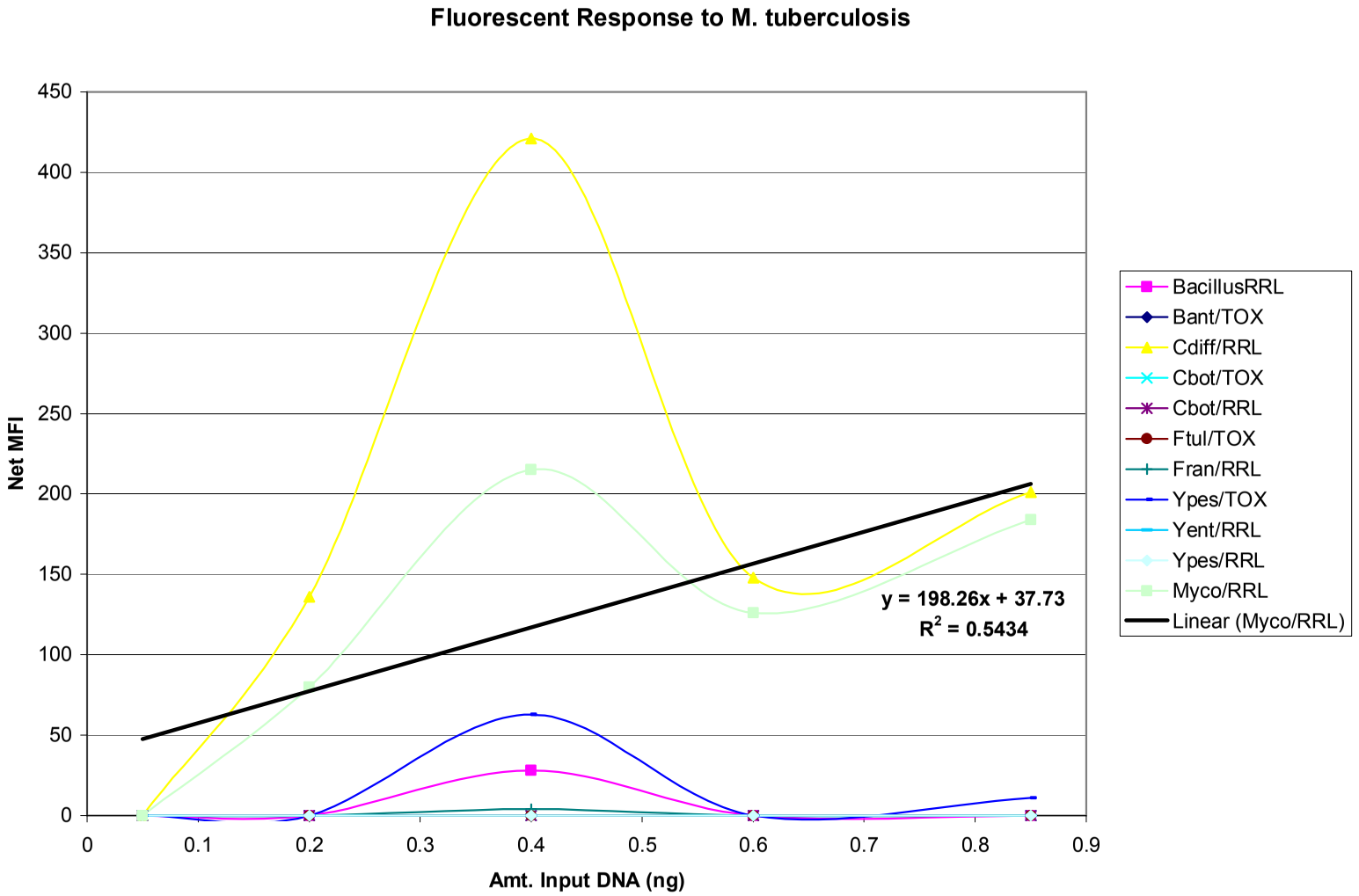


Figure A6.10. Fluorescent response of eleven probes to *M. tuberculosis* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
CD – NC	62	44	45	44	34	43	30	43	50	62	32
3X NC	186	132	135	132	102	129	90	129	150	186	96
	Net MFI*										
CD/0.1ng	0	0	0	0	0	0	0	0	0	0	0
CD/0.5ng	0	0	0	0	0	0	0	0	0	0	0
CD/1ng	0	0	18	0	0	0	0	0	0	0	0
CD/5ng	0	0	251	0	0	0	0	0	0	0	0
CD/10ng	0	0	480	0	0	0	0	0	0	0	0
CD/20ng	0	0	852	0	0	0	0	0	0	0	0
CD/40ng	0	0	1307	0	0	0	2	48	0	0	4
CD/80ng	0	0	1526	0	0	0	38	81	0	0	36
CD – PC	0	36	3622	0	10	79	5	68	62	52	76

Table A6.11. Net MFI data for *C. difficile* specificity/sensitivity study. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

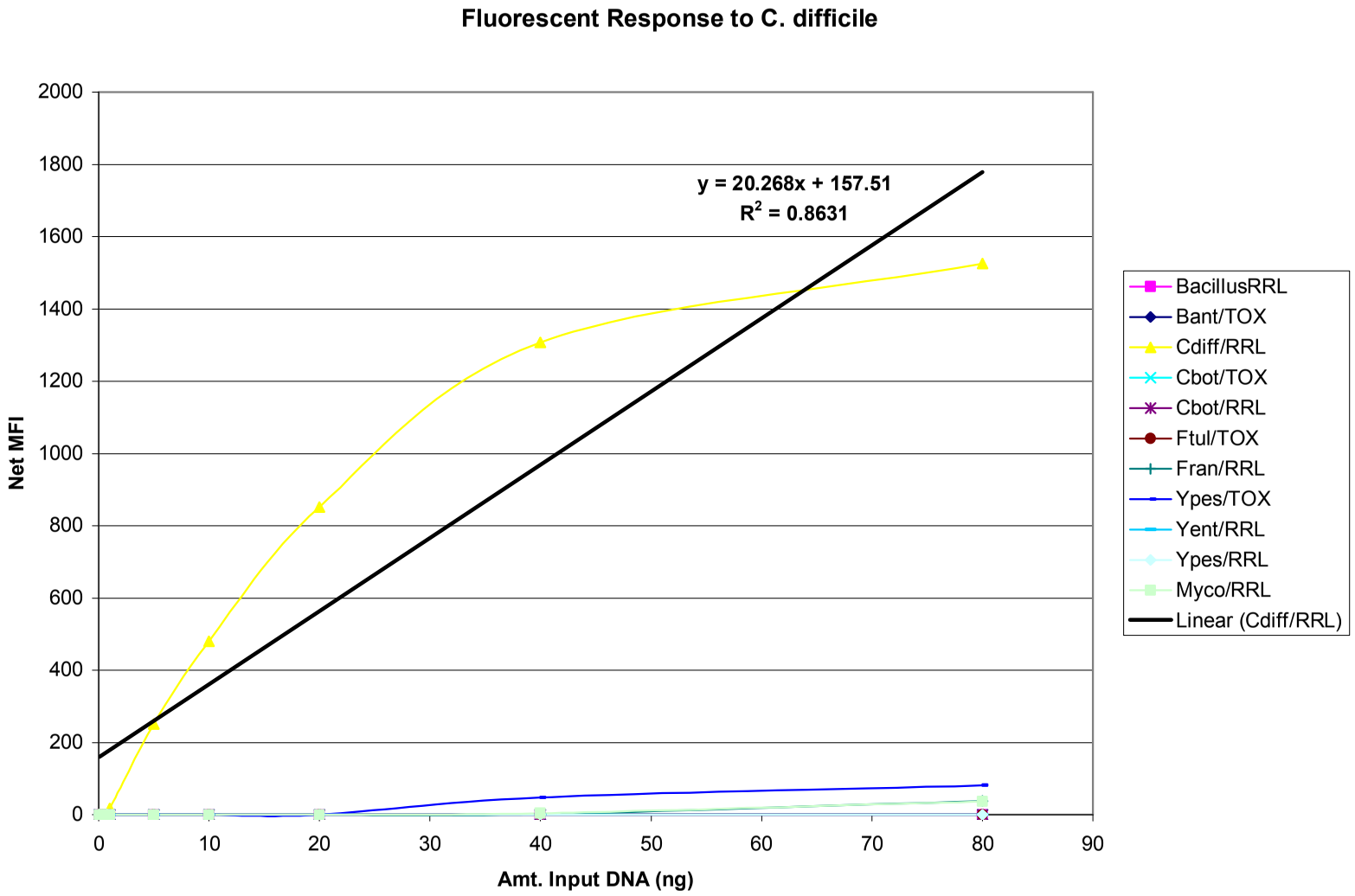


Figure A6.11. Fluorescent response of eleven probes to *C. difficile* target sequence.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
CD – NC	37	39	47	34	41	37	57	53	38	42	39
3X NC	111	117	141	102	123	111	171	159	114	126	117
	Net MFI*										
CD/0.1ng	0	0	0	0	0	0	0	0	0	0	0
CD/0.5ng	0	0	0	0	0	0	0	0	0	0	0
CD/1ng	0	0	0	0	0	0	0	0	0	0	0
CD/5ng	0	0	177	0	0	0	0	0	0	0	0
CD/10ng	0	0	333	0	0	0	0	0	0	0	0
CD/20ng	0	0	485	0	0	0	0	0	0	0	0
CD/40ng	0	0	824	0	0	0	0	0	0	0	0
CD/80ng	0	0	974	0	0	0	0	0	0	0	0
CD – PC	6	0	3178	0	0	0	0	0	23	0	0

Table A6.12. Net MFI data for *C. difficile* specificity/sensitivity study using modified hybridization temperature (45°C). *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

APPENDIX VII. MIXTURE STUDY (1:1) DATA

Lane	Template	Observed band (bp)	Predicted Source of Band	Concentration (ng/uL)
1	BA + YP	116	**BA+YP (RRL)**	**14.86**
		175	YP (TOX)	5.92
		239	BA (TOX)	3.17
2	BA + FT	110	**FT (RRL)**	**4.57**
		122	**BA (RRL)**	**4.21**
		140	FT (TOX)	0.45
		159		1.12
		164		2.94
		239	BA (TOX)	4.61
3	BA + CB	107	CB (TOX)	1.33
		121	**BA+CB (RRL)**	**5.82**
		239	BA (TOX)	3.64
4	YP + FT	114	**YP+FT (RRL)**	**12.53**
		158	FT (TOX)	0.77
		163		2.19
		173	YP (TOX)	5.50
5	YP + CB	108	CB (TOX)	0.97
		114	**YP+CB (RRL)**	**11.48**
		173	YP (TOX)	5.65
6	FT + CB	108	CB (TOX)	4.63
		119	**FT+CB (RRL)**	**2.13**
		158	FT (TOX)	0.61
		163		2.24
7	BA + YP + FT	114	**BA+YP+FT (RRL)**	**9.55**
		163	FT (TOX)	1.21
		173	YP (TOX)	3.51
		238	BA (TOX)	0.99
8	BA + YP + CB	114	**BA+YP+CB (RRL)**	**10.23**
		173	YP (TOX)	4.14
		238	BA (TOX)	0.89
9	BA + FT + CB	108	CB (TOX)	4.38
		120	**BA+FT+CB (RRL)**	**4.25**
		157	FT (TOX)	0.71
		163		2.50
		238	BA (TOX)	2.59
10	YP + FT + CB	113	**YP+FT+CB (RRL)**	**11.80**
		161	FT (TOX)	1.26
		170	YP (TOX)	3.86
11	BA+CB+FT+YP	112	**BA+CB+FT+YP (RRL)**	**12.10**
		160	FT (TOX)	1.07
		169	YP (TOX)	4.46
		235	BA (TOX)	0.75
12	None	n/a	n/a	n/a

Table A7.1. Summary of post-digest quantitation data for enzymatic degradation control samples for mixture study (1:1) batch #1. Yellow highlighted cells denote samples whose concentration was used to determine Luminex input amounts.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	50	45	73	35	34	52	63	55	73	68	38
3X NC	150	135	219	105	102	156	189	165	219	204	114
	Net MFI*										
BA/YP	355	1899	0	0	0	0	0	2565	0	0	0
BA/FT	671	3067	0	0	0	2990	1462	4	0	0	0
BA/CB	1297	3752	0	2480	100	0	0	0	0	0	0
YP/FT	0	0	0	0	0	2713	1193	3068	0	0	0
YP/CB	0	0	0	1631	31	0	0	3281	0	55	0
FT/CB	0	14	9	3075	94	3950	2616	139	0	0	0
BA/YP/FT	839	1936	0	0	0	2684	1733	3902	0	125	0
BA/YP/CB	840	1725	0	1539	23	0	0	4396	0	165	0
BA/FT/CB	1392	3703	62	1983	125	3612	2766	184	0	0	0
YP/FT/CB	0	0	0	1601	45	2635	1660	3790	0	171	0
BA/CB/FT/YP	732	1648	0	1251	22	2074	1617	4075	0	184	0
BA – PC	4967	7389	0	0	0	0	0	0	0	0	0
CB – PC	33	274	65	6757	5470	0	0	30	0	0	54
FT – PC	18	0	0	22	29	7038	5003	115	0	0	0
YP – PC	25	0	0	21	0	3	26	7677	500	4301	19

Table A7.2. Summary of net MFI data for mixture study (1:1) samples, batch #1. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

Lane	Template	Observed band (bp)	Predicted Source of Band	Concentration (ng/uL)
1	BA + YE	117	BA + YE (RRL)	13.75
		233	BA (TOX	3.17
2	BA + FH	110	FH (RRL)	4.49
		121	BA (RRL)	3.54
		161		0.73
		166		2.76
		231	BA (TOX)	3.74
3	BC + FT	110	FT (RRL)	3.75
		121	BC (RRL)	3.34
		160	FT (TOX)	0.72
		165		2.29
4	YE + CB	110	CB (TOX)	0.97
		116	YE + CB (RRL)	12.87
5	FT + CD	109	FT (RRL)	2.74
		117	CD (RRL)	5.15
		164	FT (TOX)	2.96
6	FH + CB	109	CB (TOX), FH (RRL)	4.88
		119	CB (RRL)	1.54
		165		2.77
7	BC + YP + FT	115	BC+YP+FT (RRL)	8.46
		164	FT (TOX)	0.87
		175	YP (TOX)	3.08
8	BA + YE + CB	n/a	CB (TOX)	n/a
		116	BA+YE+CB (RRL)	11.19
		232	BA (TOX)	1.49
9	BA + FH + CB	109	FH (RRL) + CB (TOX)	3.19
		121	BA + CB (RRL)	2.69
		164	FH (TOX)	2.16
		231	BA (TOX)	1.91
10	YP + FT + CD	114	YP+FT+CD (RRL)	8.53
		162	FT (TOX)	0.76
		173	YP (TOX)	2.95
11	BA+CB+FH+YE	108	CB (TOX)	1.64
		114	BA+CB+FH+YE (RRL)	9.47
		161		1.34
		229	BA (TOX)	0.84
12	None	n/a	n/a	n/a

Table A7.3. Summary of post-digest quantitation data for enzymatic degradation of control samples for mixture study (1:1) batch #2. Yellow highlighted cells denote samples whose concentration was used to determine Luminex input amounts.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	43	45	46	44	44	31	45	37	40	45	42
3X NC	129	135	138	132	132	93	135	111	120	135	126
	Net MFI*										
BA/YE	399	1943	0	0	0	0	0	0	89	0	0
BA/FH	596	2911	0	0	0	1768	1469	24	0	0	0
BC/FT	713	0	0	0	0	3179	1496	78	0	0	0
YE/CB	0	0	0	1623	0	9	0	74	115	0	0
FT/CD	0	0	518	0	0	3050	1334	80	0	0	0
FH/CB	0	0	33	2001	18	1756	1902	25	0	0	0
BC/YP/FT	845	161	101	105	0	1805	1458	4566	0	338	0
BA/YE/CB	778	1920	74	1662	0	134	57	31	252	0	0
BA/FH/CB	885	3027	34	2017	0	2071	2144	80	0	0	0
YP/FT/CD	99	36	776	0	0	1868	1521	4402	0	370	0
BA/CB/FH/YE	844	1980	67	1726	6	695	1697	47	400	0	0
BC – PC	4155	0	7	0	0	4	0	0	0	0	0
BA – PC	4582	6503	0	7	0	26	182	145	82	24	0
CD – PC	6	0	3622	0	0	206	101	0	0	0	3
CB – PC	76	224	125	6531	4293	52	13	26	0	1	20
FH – PC	64	21	57	0	25	50	5218	77	34	42	18
FT – PC	108	168	119	91	38	6400	4869	173	51	20	7
YE – PC	330	332	433	301	250	491	447	416	3347	495	329
YP – PC	248	293	166	207	152	407	256	6714	776	3871	265

Table A7.4. Summary of net MFI data for mixture study (1:1) samples, batch #2. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

Lane	Template	Observed band (bp)	Predicted Source of Band	Concentration (ng/uL)
1	BC + YP	117	BC + YP (RRL)	14.89
		178	YP (TOX)	5.86
2	BA + CD	120	BA + CD (RRL)	5.98
		232	BA (TOX)	2.30
3	BC + CB	109	CB (TOX)	2.01
		121	BC + CB (RRL)	6.33
4	YP + FH	115	YP + FH (RRL)	9.96
		165	FH (TOX)	1.71
		175	YP (TOX)	4.24
5	YE + FT	110	FT (RRL)	1.45
		116	YE (RRL)	5.23
		164	FT (TOX)	1.43
6	YP + CD	116	YP + CD (RRL)	10.00
		175	YP (TOX)	3.44
7	BC + YE + CB	n/a	CB (TOX)	n/a
		116	BC + YE + CB (RRL)	12.08
8	BC + YP + CD	116	BC + YP + CD (RRL)	9.03
		175	YP (TOX)	3.33
9	BA + YE + CD	116	BA + YE + CD (RRL)	12.31
		232	BA (TOX)	0.96
10	BC + FT + CD	109	FT (RRL)	2.46
		119	BC + CD (RRL)	5.48
		163	FT (TOX)	2.98
11	BA + YE + FT + CD	114	BA + YE + FT + CD (RRL)	8.62
		161	FT (TOX)	1.14
		229	BA (TOX)	0.70
12	None	n/a	n/a	n/a

Table A7.5. Summary of post-digest quantitation data for enzymatic degradation control samples for mixture study (1:1) batch #3. Yellow highlighted cells denote samples whose concentration was used to determine Luminex input amounts.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	43	31	38	51	34	42	50	35	42	42	31
3X NC	129	93	114	153	102	126	150	105	126	126	93
	Net MFI*										
BC/YP	374	0	0	0	0	0	0	2256	0	0	0
BA/CD	815	2868	593	0	0	0	0	0	0	0	0
BC/CB	891	0	1	1886	14	0	0	0	0	0	0
YP/FH	0	0	0	0	0	948	1354	3420	0	86	0
YE/FT	0	0	8	0	0	2965	1508	39	250	0	0
YP/CD	0	0	733	0	0	0	0	3326	0	131	0
BC/YE/CB	655	131	19	1467	2	167	0	0	204	0	0
BC/YP/CD	865	0	699	0	0	0	0	3966	0	264	0
BA/YE/CD	737	1462	660	0	0	0	0	0	357	0	0
BC/FT/CD	904	0	752	0	0	3451	1882	71	0	0	0
BA/CD/FT/YE	792	1965	722	0	0	2223	1658	21	339	0	0
BC – PC	4993	14	8	118	129	0	0	0	0	0	0
BA – PC	3859	6139	0	0	0	0	0	0	0	0	0
CD – PC	0	0	2537	0	0	0	0	0	0	0	0
CB – PC	49	191	86	5630	4221	0	0	14	148	1	25
FH – PC	13	44	74	0	86	35	6594	56	23	13	41
FT – PC	50	81	98	23	55	6390	3845	221	56	38	57
YE – PC	267	324	357	201	229	378	376	374	3875	598	275
YP – PC	277	356	378	254	219	445	245	7230	964	4427	321

Table A7.6. Summary of net MFI data for mixture study (1:1) samples, batch #3. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

APPENDIX VIII. INTRA-GENUS MIXTURE STUDY (1:1) DATA TABLES

Lane	Template	Observed band (bp)	Predicted Source of Band	Concentration (ng/uL)
8	CD:CB-1:1	109	CB (TOX)	1.39
		118	CD + CB (RRL)	5.79
9	YE:YP-1:1	116	YE + YP (RRL)	6.98
		162	YP (TOX)	0.72
		175		5.04
10	None	n/a	n/a	n/a

Table A8.1. Summary of post-digest quantitation data for enzymatic degradation control samples in intra-genus mixture study (1:1). Yellow highlighted cells denote samples whose concentration was used to determine Luminex input amounts.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	36	22	47	42	28	34	26	44	44	33	33
3X NC	108	66	141	126	84	102	78	132	132	99	99
	Net MFI**										
CD:CB-1:1	0	47	1013	2055	143	0	5	0	0	0	0
YE:YP-1:1*	0	0	54	0	0	0	0	4568	275	214	0
CD – PC	23	23	3005	0	0	82	93	31	10	23	8
CB – PC	5	278	62	6667	5137	32	32	0	0	22	3
YE – PC	18	59	20	0	61	47	367	29	4350	223	29
YP – PC	85	127	91	73	76	168	141	6466	348	3017	79

Table A8.2. Summary of net MFI data for intra-genus mixture study (1:1). *Data from re-analyzed sample in a separate assay. **Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

APPENDIX IX. MIXTURE STUDY (RATIOS) DATA TABLES AND STATISTICAL ANALYSES

Lane	Template	Observed band (bp)	Predicted Source of Band	Concentration (ng/uL)
1	BA:FT-10:1	111	FT (RRL)	2.59
		124	BA (RRL)	6.19
		161	FT (TOX)	0.70
		167		2.46
		234	BA (TOX)	5.19
2	BA:FT-5:1	110	FT (RRL)	2.79
		123	BA (RRL)	5.29
		165	FT (TOX)	3.03
		234	BA (TOX)	4.71
3	BA:FT-2:1	110	FT (RRL)	3.78
		122	BA (RRL)	4.32
		165	FT (TOX)	3.26
		233	BA (TOX)	4.13
4	BA:FT-1:1	109	FT (RRL)	3.49
		121	BA (RRL)	2.67
		165	FT (TOX)	2.64
		234	BA (TOX)	2.53
5	BA:FT-1:2	110	FT (RRL)	4.71
		122	BA (RRL)	2.84
		165	FT (TOX)	3.62
		234	BA (TOX)	2.98
6	BA:FT-1:5	110	FT (RRL)	5.02
		122	BA (RRL)	2.04
		166	FT (TOX)	3.42
		234	BA (TOX)	1.37
7	BA:FT-1:10	111	FT (RRL)	4.86
		122	BA (RRL)	1.68
		166	FT (TOX)	3.37
		234	BA (TOX)	1.59
8	Sample from a separate study (Inter-genus 1:1 mixtures); results presented elsewhere			
9	Sample from a separate study (Inter-genus 1:1 mixtures); results presented elsewhere			
10	None	n/a	n/a	n/a

Table A9.1. Summary of post-digest quantitation data for enzymatic degradation control samples for BA/FT ratio mixture study. Highlighted cells indicate samples whose concentration was used to determine Luminex input amounts.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	36	22	47	42	28	34	26	44	44	33	33
3X NC	108	66	141	126	84	102	78	132	132	99	99
	Net MFI*										
BA:FT-10:1	1373	3688	0	0	0	3183	1406	19	0	0	0
BA:FT-5:1	1269	3829	0	0	0	3358	1566	29	0	0	0
BA:FT-2:1	919	2982	0	0	0	3030	1703	9	0	0	0
BA:FT-1:1	853	2894	0	0	0	3305	1896	33	0	0	0
BA:FT-1:2	913	3095	17	0	0	3670	2184	55	0	0	0
BA:FT-1:5	835	2419	0	0	0	3818	2745	75	0	0	0
BA:FT-1:10	801	2870	10	0	0	4139	3017	107	0	0	0
BA – PC	4808	7811	0	0	0	0	14	0	0	0	0
FT – PC	0	54	346	6	60	6516	4884	428	22	146	3

Table A9.2. Summary of net MFI data for BA/FT ratio mixtures. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

Lane	Template	Observed band (bp)	Predicted Source of Band	Concentration (ng/uL)
1	BA:YP-10:1	115	BA + YP (RRL)	13.17
		177	YP (TOX)	4.51
		235	BA (TOX)	3.93
2	BA:YP-5:1	115	BA + YP (RRL)	8.36
		176	YP (TOX)	3.02
		234	BA (TOX)	1.77
3	BA:YP-2:1	114	BA + YP (RRL)	7.99
		175	YP (TOX)	3.26
		234	BA (TOX)	1.20
4	BA:YP-1:1	114	BA + YP (RRL)	9.62
		175	YP (TOX)	4.12
		234	BA (TOX)	1.44
5	BA:YP-1:2	115	BA + YP (RRL)	10.64
		175	YP (TOX)	4.61
		234	BA (TOX)	0.69
6	BA:YP-1:5	115	BA + YP (RRL)	11.07
		175	YP (TOX)	4.91
		†	BA (TOX)	n/a
7	BA:YP-1:10	115	BA + YP (RRL)	9.40
		175	YP (TOX)	4.03
		†	BA (TOX)	n/a
8	None	n/a	n/a	n/a

Table A9.3. Summary of post-digest quantitation data for enzymatic degradation control samples for BA/YP ratio mixture study. A dagger (†) symbol in the "Observed Bands" column indicates that a faint band was visible but not concentrated enough to be reported. Highlighted cells indicate samples whose concentration was used to determine Luminex input amounts.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	87	112	46	51	68	38	61	188	31	48	42
3X NC	261	336	138	153	204	114	183	564	93	144	126
	Net MFI*										
BA:YP-10:1	1446	2957	43	0	0	0	0	3961	0	201	0
BA:YP-5:1	1441	2862	41	0	0	0	0	4373	0	271	0
BA:YP-2:1	866	2076	0	0	0	0	0	3641	0	217	0
BA:YP-1:1	588	1487	0	0	0	0	0	2992	0	123	0
BA:YP-1:2	425	1208	0	0	0	0	0	3907	0	269	0
BA:YP-1:5	754	1002	42	0	0	0	0	4863	0	498	0
BA:YP-1:10	953	1298	90	0	0	0	0	5832	2	835	0
BA – PC	5584	6506	0	0	0	12	0	0	36	0	0
YP – PC	0	0	19	0	0	61	3	5893	383	3133	0

Table A9.4. Summary of net MFI data for BA:YP ratio mixtures. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

Lane	Sample	Template	Observed band (bp)	Predicted Source of Band	Concentration (ng/uL)
1	459	YP:FT-10:1	110	FT (RRL)	1.80
			117	YP (RRL)	13.02
			*	FT (TOX)	n/a
			180	YP (TOX)	5.67
2	460	YP:FT-5:1	110	FT (RRL)	2.15
			117	YP (RRL)	11.36
			167	FT (TOX)	1.12
			178	YP (TOX)	5.18
3	461	YP:FT-2:1	109	FT (RRL)	2.51
			116	YP (RRL)	11.22
			166	FT (TOX)	1.16
			177	YP (TOX)	5.67
4	462	YP:FT-1:1	110	FT (RRL)	2.69
			116	YP (RRL)	9.69
			166	FT (TOX)	2.21
			178	YP (TOX)	5.11
5	463	YP:FT-1:2	110	FT (RRL)	2.99
			116	YP (RRL)	10.62
			167	FT (TOX)	2.89
			178	YP (TOX)	5.54
6	464	YP:FT-1:5	116	YP + FT (RRL)	12.97
			167	FT (TOX)	3.19
			178	YP (TOX)	5.02
7	465	YP:FT-1:10	111	FT (RRL)	2.90
			116	YP (RRL)	7.94
			166	FT (TOX)	2.81
			177	YP (TOX)	4.00
8	466	None	n/a	n/a	n/a

Table A9.5. Summary of post-digest quantitation data for enzymatic degradation control samples for YP/FT ratio mixture study. An asterisk (*) symbol in the "Observed Bands" column indicates that a band was neither visible nor detected by the instrument. Highlighted cells indicate samples whose concentration was used to determine Luminex input amounts.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	57	51	64	66	48	44	81	63	48	43	50
3X NC	171	153	192	198	144	132	243	189	144	129	150
	Net MFI*										
YP:FT-10:1	0	0	0	0	0	984	1447	5746	0	489	0
YP:FT-5:1	0	0	0	0	0	2660	1976	5586	0	474	0
YP:FT-2:1	0	0	0	0	0	1751	1639	4984	0	454	0
YP:FT-1:1	0	0	0	0	0	2580	1465	3543	0	201	0
YP:FT-1:2	0	0	0	0	0	2337	1212	2926	0	111	0
YP:FT-1:5	0	0	13	0	0	4127	3051	5295	0	519	0
YP:FT-1:10	0	0	0	0	0	3172	1834	3297	0	144	0
YP – PC	0	0	0	0	0	0	0	6383	214	3269	0
FT – PC	22	31	19	0	41	6377	5068	104	0	71	0

Table A9.6. Summary of net MFI data for YP:FT ratio mixtures. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

Lane	Template	Observed band (bp)	Predicted Source of Band	Concentration (ng/uL)
1	BA:CB-10:1	111	CB (TOX)	1.11
		124	BA + CB (RRL)	7.64
		233	BA (TOX)	5.12
2	BA:CB-5:1	110	CB (TOX)	1.22
		122	BA + CB (RRL)	7.44
		231	BA (TOX)	5.25
3	BA:CB-2:1	109	CB (TOX)	1.31
		121	BA + CB (RRL)	6.30
		231	BA (TOX)	4.49
4	BA:CB-1:1	109	CB (TOX)	1.40
		121	BA + CB (RRL)	6.57
		232	BA (TOX)	4.36
5	BA:CB-1:2	110	CB (TOX)	1.47
		122	BA + CB (RRL)	6.60
		232	BA (TOX)	3.84
6	BA:CB-1:5	†	CB (TOX)	n/a
		121	BA + CB (RRL)	1.81
		232	BA (TOX)	0.63
7	BA:CB-1:10	†	CB (TOX)	n/a
		121	BA + CB (RRL)	1.79
		232	BA (TOX)	0.79
8	None	n/a	n/a	n/a

Table A9.7. Summary of post-digest quantitation data for enzymatic degradation control samples for BA/CB ratio mixture study. A dagger (†) symbol in the "Observed Bands" column indicates that a faint band was visible but not concentrated enough to be reported. Highlighted cells indicate samples whose concentration was used to determine Luminex input amounts.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	51	71	52	64	38	51	66	45	43	75	45
3X NC	153	213	156	192	114	153	198	135	129	225	135
	Net MFI*										
BA:CB-10:1	2866	4885	270	2761	143	0	0	5	0	0	0
BA:CB-5:1	2365	4522	90	2391	44	0	0	0	0	0	0
BA:CB-2:1	1750	4609	48	2274	160	0	0	0	0	0	0
BA:CB-1:1	1449	3946	0	2301	123	0	0	0	0	0	0
BA:CB-1:2	1740	4253	41	2975	307	0	0	183	0	0	0
BA:CB-1:5	1567	2905	128	2758	114	0	0	13	0	0	0
BA:CB-1:10	1914	3802	142	6179	372	0	0	0	0	0	0
BA – PC	5988	9596	4	21	10	0	0	112	44	0	0
CB – PC	87	455	60	7969	5272	0	0	402	0	56	0

Table A9.8. Summary of net MFI data for BA:CB ratio mixtures. *Net MFI = Raw MFI – 3X Neg. control for selected bead set; net MFIs ≤ 0 reported as "0". Yellow highlighted cells indicate designed complementarity. Net MFIs in Red font indicate a false negative. Net MFIs in Blue font indicate a false positive.

BA:FT																
	Bac/RRL	Bant/TOX	Cdif/RRL	Cbot/RRL	Cbot/TOX	Yent/RRL	Ypes/RRL	Ypes/TOX	Fran/RRL	Ftul/TOX	Myc/RRL					
NC	36	22	47	28	42	44	33	44	26	34	33		*Variation in the Minor Components*			
3X NC	108	66	141	84	126	132	99	132	78	102	99					
	Net MFI*												Bac/RRL	Bant/TOX	Fran/RRL	Ftul/TOX
BA:FT-10:1	1373	3688	0	0	0	0	0	19	1406	3183	0	Avg.	850.5	2819.5	1642.75	3219
BA:FT-5:1	1269	3829	0	0	0	0	0	29	1566	3358	0	S.D.	46.91482	285.4242	207.9317	145.7555
BA:FT-2:1	919	2982	0	0	0	0	0	9	1703	3030	0	%CV	5.516145	10.12322	12.65754	4.527975
BA:FT-1:1	853	2894	0	0	0	0	0	33	1896	3305	0					
BA:FT-1:2	913	3095	17	0	0	0	0	55	2184	3670	0					
BA:FT-1:5	835	2419	0	0	0	0	0	75	2745	3818	0					
BA:FT-1:10	801	2870	10	0	0	0	0	107	3017	4139	0					
BA:YP																
	Bac/RRL	Bant/TOX	Cdif/RRL	Cbot/TOX	Cbot/RRL	Ftul/TOX	Fran/RRL	Yent/RRL	Ypes/RRL	Ypes/TOX	Myc/RRL					
NC	87	112	46	51	68	38	61	31	48	188	42					
3X NC	261	336	138	153	204	114	183	93	144	564	126					
	Net MFI*															
BA:YP-10:1	1446	2957	43	0	0	0	0	0	201	3961	0		Bac/RRL	Bant/TOX	Ypes/RRL	Ypes/TOX
BA:YP-5:1	1441	2862	41	0	0	0	0	0	271	4373	0	Avg.	680	1248.75	203	3741.75
BA:YP-2:1	866	2076	0	0	0	0	0	0	217	3641	0	S.D.	226.1961	201.4404	61.16644	582.7586
BA:YP-1:1	588	1487	0	0	0	0	0	0	123	2992	0	%CV	33.26413	16.13137	30.13125	15.57449
BA:YP-1:2	425	1208	0	0	0	0	0	0	269	3907	0					
BA:YP-1:5	754	1002	42	0	0	0	0	0	498	4863	0					
BA:YP-1:10	953	1298	90	0	0	0	0	2	835	5832	0					
YP:FT																
	Bac/RRL	Bant/TOX	Ypes/RRL	Ypes/TOX	Yent/RRL	Cdif/RRL	Cbot/TOX	Cbot/RRL	Fran/RRL	Ftul/TOX	Myc/RRL					
NC	57	51	43	63	48	64	66	48	81	44	50					
3X NC	171	153	129	189	144	192	198	144	243	132	150					
	Net MFI*															
YP:FT-10:1	0	0	489	5746	0	0	0	0	1447	984	0		Ypes/RRL	Ypes/TOX	Fran/RRL	Ftul/TOX
YP:FT-5:1	0	0	474	5586	0	0	0	0	1976	2660	0	Avg.	243.75	3765.25	1631.75	1993.75
YP:FT-2:1	0	0	454	4984	0	0	0	0	1639	1751	0	S.D.	187.2278	1050.893	245.2881	788.6911
YP:FT-1:1	0	0	201	3543	0	0	0	0	1465	2580	0	%CV	76.81141	27.91031	15.03221	39.55817
YP:FT-1:2	0	0	111	2926	0	0	0	0	1212	2337	0					
YP:FT-1:5	0	0	519	5295	0	13	0	0	3051	4127	0					
YP:FT-1:10	0	0	144	3297	0	0	0	0	1834	3172	0					
BA:CB																
	Bac/RRL	Bant/TOX	Fran/RRL	Ftul/TOX	Cdif/RRL	Yent/RRL	Ypes/RRL	Ypes/TOX	Cbot/RRL	Cbot/TOX	Myc/RRL					
NC	51	71	66	51	52	43	75	45	38	64	45					
3X NC	153	213	198	153	156	129	225	135	114	192	135					
	Net MFI*															
BA:CB-10:1	2866	4885	0	0	270	0	0	5	143	2761	0		Bac/RRL	Bant/TOX	Cbot/RRL	Cbot/TOX
BA:CB-5:1	2365	4522	0	0	90	0	0	0	44	2391	0	Avg.	1667.5	3726.5	117.5	2431.75
BA:CB-2:1	1750	4609	0	0	48	0	0	0	160	2274	0	S.D.	203.192	579.0639	51.28028	225.1272
BA:CB-1:1	1449	3946	0	0	0	0	0	0	123	2301	0	%CV	12.18543	15.53908	43.64279	9.257826
BA:CB-1:2	1740	4253	0	0	41	0	0	183	307	2975	0					
BA:CB-1:5	1567	2905	0	0	128	0	0	13	114	2758	0					
BA:CB-1:10	1914	3802	0	0	142	0	0	0	372	6179	0					

Table A9.9. Evaluation of consistency in net MFIs of minor components as ratio of major component was varied. Yellow highlighted cells indicate designed complementarity. Values used to derive average, standard deviation, and percent coefficient of variation for each probe in a particular batch are outlined in similar color borders.

H_O= variations in the net MFIs for the minor contributor are due to chance alone and unrelated to the identity of the major contributor

H_A = net MFIs for the minor contributor vary depending on the identity of the major contributor

Bac/RRL MFI Values for Ratio Mixtures in which BA was the Minor Contributor

BA:FT	BA:YP	BA:CB
853	588	1449
913	425	1740
835	754	1567
801	953	1914

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
853	3	2549	849.6667	3297.333
588	3	2132	710.6667	71104.33
1449	3	5221	1740.333	30102.33

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1872822	2	936410.8	26.88158	0.001012	5.143253
Within Groups	209008	6	34834.67			
Total	2081830	8				

Result: P value = 0.001

Conclusion: P value ≤ 0.05, therefore reject the null hypothesis at 95% confidence level

Interpretation: Variation in net MFIs for the Bac/RRL probe is not due to chance alone but may depend on the identity of the major contributor

Bant/TOX MFI Values for Ratio Mixtures in which BA was the Minor Contributor

BA:FT	BA:YP	BA:CB
2894	1487	3946
3095	1208	4253
2419	1002	2905
2870	1298	3802

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
2894	3	8384	2794.667	118500.3
1487	3	3508	1169.333	23025.33
3946	3	10960	3653.333	470852.3

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	9549273	2	4774636	23.39063	0.001469	5.143253
Within Groups	1224756	6	204126			
Total	10774029	8				

Result: P value = 0.001

Conclusion: P value ≤ 0.05, therefore reject the null hypothesis at 95% confidence level

Interpretation: Variation in net MFIs for the Banth/TOX probe is not due to chance alone but may depend on the identity of the major contributor

Table A9.10. ANOVA for select ratio mixtures. Shown is the ANOVA for the Bac/RRL probe net MFI values in ratio mixtures in which *B. anthracis* was the minor contributor (left) and for the Bant/TOX probe net MFI values in ratio mixtures in which *B. anthracis* was the minor contributor (right).

APPENDIX X. SAMPLE MFI AND COUNT DATA FOR TITRATION STUDY

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	0	31	0	0	0	0	0	0	72	0	0
	Raw MFI										
0.25 fmol	0	579	0	0	0	0	0	0	799	0	0
0.5 fmol	0	399	0	0	0	0	0	0	903	0	0
1 fmol	0	0	0	0	0	0	0	0	1025	0	0
10 fmol	0	5654	0	0	0	0	0	0	4536	0	0
50 fmol	0	0	0	0	0	0	0	0	8359	0	0
100 fmol	0	0	0	0	0	0	0	0	8929	0	0
150 fmol	0	7838	0	0	0	0	0	0	9344	0	0
200 fmol	0	7428	0	0	0	0	0	0	9505	0	0
250 fmol	0	7541	0	0	0	0	0	0	11500	0	0

Table 10.1. Raw MFI data for Yent/RRL titration study. Yellow highlighted cells indicate designed complementarity. Net MFIs in Blue font indicate a false positive.

	Probe										
	Bac/ RRL	**Bant/ TOX**	**Cdif/ RRL**	**Cbot/ TOX**	**Cbot/ RRL**	**Ftul/ TOX**	**Fran/ RRL**	**Ypes/ TOX**	**Yent/ RRL**	**Ypes/ RRL**	**Myc/ RRL**
NC	0	2	0	0	0	0	0	0	729	0	0
	Count										
0.25 fmol	0	2	0	0	0	0	0	0	706	0	0
0.5 fmol	0	1	0	0	0	0	0	0	606	0	0
1 fmol	0	0	0	0	0	0	0	0	760	0	0
10 fmol	0	1	0	0	0	0	0	0	597	0	0
50 fmol	0	0	0	0	0	0	0	0	589	0	0
100 fmol	0	0	0	0	0	0	0	0	557	0	0
150 fmol	0	2	0	0	0	0	0	0	582	0	0
200 fmol	0	1	0	0	0	0	0	0	511	0	0
250 fmol	0	2	0	0	0	0	0	0	632	0	0

Table 10.2. Count data for Yent/RRL titration study. Yellow highlighted cells indicate designed complementarity. Bead counts in Blue font indicate a false positive.

APPENDIX XI. AUTODIMER ANALYSES

```
Number of Hits = 1
Number of Sequences = 1
Na+ = 0.085 Ct = 1.0

Y. pestis RRL Target Seq.
Matches = 7
Score = 4
NTNAGTNCCC
est. tm =  19.0 oC
DeltaG @37 degrees =  -4.89 kcal/mole

 TAGATTCATGGG-5'
{  ×|×|||×|||
 GTTTCAGTTCCCCCGGTTCGCCTCGTTAGACTATGTATTCATCTAACGATAGTGCAACGTATTGCACTGGGTTTCCCCATTCGGGTATCGTCGGT-3'
```

Figure A11.1. AutoDimer hairpin screen (threshold of 3) of Ypes/RRL target sequence (complementary strand).

```
Number of Hits = 2
Number of Sequences = 3
Na+ = 0.085 Ct = 1.0
```

```
B. anthracis TOX probe ATGGGAGAACAAATGAAGCG versus C. botulinum RRL Target Seq.
GGGTACTTAGATGTTTCAGTTCCCCGGGTTTACCCTCATACACCTATGTATTCAGTGCATGATACATGGGGTTACCCATGTGAGTTTCCTCATTCGGAAATCCCTGGAT
Matches = 12
Score = 4
NGNTNCATNNGNTNNCCCAT
est. tm =  less than zero
DeltaG 37 degrees =  greater than zero

                    5'-
GGGTACTTAGATGTTTCAGTTCCCCGGGTTTACCCTCATACACCTATGTATTCAGTGCATGATACATGGGGTTACCCATGTGAGTTTCCTCATTCGGAAATCCCTGGAT-3'
                                                                                    x|x|x|||xx|x|xx|||||
                                                                                 3'-GCGAAGTAAACAAGAGGGTA-5'
```

```
B. anthracis TOX probe ATGGGAGAACAAATGAAGCG versus C. botulinum TOX Target Seq.
TTGTATCATCATTAACAGGGTCATTATAATTAAAACTATTTATTACAACTGGCATAAAATCCCCCCTATACTTCTTTCAATGTCCATAAATATGCCTTTT
Matches = 6
Score = 3
TTNTNNCAT
est. tm =  less than 0 oC
DeltaG 37 degrees =  > 0 kcal/mole

                    5'-
TTGTATCATCATTAACAGGGTCATTATAATTAAAACTATTTATTACAACTGGCATAAAATCCCCCCTATACTTCTTTCAATGTCCATAAATATGCCTTTT-3'
                      ||x|xx|||
         3'-GCGAAGTAAACAAGAGGGTA-5'
```

Figure A11.2. AutoDimer primer dimer screen (threshold of 3) of Bant/TOX probe against Cbot/RRL and Cbot/TOX target sequences (complementary strand).

```
Number of Hits = 3
Number of Sequences = 1
Na+ = 0.085 Ct = 1.0
```

```
C. botulinum RRL Target Seq.
Matches = 7
Score = 4
NTNAGTNCCC
est. tm =  19.0 oC
DeltaG @37 degrees =  -4.89 kcal/mole

 TAGATTCATGGG-5'
{  X|X|||X|||
 GTTTCAGTTCCCCGGGTTTACCCTCATACACCTATGTATTCAGTGCATGATACATGGGGTTACCCATGTGAGTTTCCTCATTCGGAAATCCCTGGAT-3'
```

```
C. botulinum RRL Target Seq.
Matches = 13
Score = 4
NCNCATNNGNAANNCCNTGNAT
est. tm =  46.8 oC
DeltaG @37 degrees =  -11.93 kcal/mole

 TGAGTGTACCCATTGGGGTACATAGTACGTGACTTATGTATCCACATACTCCCATTTGGGCCCCTTGACTTTGTAGATTCATGGG-5'
{  X|X|||XX|X||XX||X||X||
 TTCCTCATTCGGAAATCCCTGGAT-3'
```

```
C. botulinum RRL Target Seq.
Matches = 3
Score = 3
GAT
est. tm =  less than zero
DeltaG @37 degrees =  greater than zero

 CCCTAAAGGCTTACTCCTTTGAGTGTACCCATTGGGGTACATAGTACGTGACTTATGTATCCACATACTCCCATTTGGGCCCCTTGACTTTGTAGATTCATGGG-5'
{  |||
 TGGAT-3'
```

Figure A11.3. AutoDimer hairpin screen (threshold of 3) of Cbot/RRL target sequence (complementary strand).

```
Number of Hits = 1
Number of Sequences = 1
Na+ = 0.085 Ct = 1.0

Y. enterocolitica RRL Target Seq.
Matches = 7
Score = 4
NTNAGTNCCC
est. tm =  19.0 oC
DeltaG @37 degrees =  -4.89 kcal/mole

 TAGATTCATGGG-5'
{  ×|×|||×|||
 GTTTCAGTTCCCCCGGTTCGCCTTGCATGGCTATGTATTCACCATGCAATAGTGCAACGAATTGCACTGGGTTTCCCCATTCGGGTATCGTCGGT-3'
```

Figure A11.4. AutoDimer hairpin screen (threshold of 3) of Yent/RRL target sequence (complementary strand).

Figure A11.5. AutoDimer primer dimer screen (threshold of 3) of Cdif/RRL probe against all target sequences (complementary strand).

```
Number of Hits = 3
Number of Sequences = 11
Na+ = 0.085 Ct = 1.0
```

```
Y. pestis TOX probe TATGATGAGGGCAAAGGAGG versus C. botulinum RRL Target Seq.
GGGTACTTAGATGTTTCAGTTCCCCGGGTTTACCCTCATACACCTATGTATTCAGTGCATGATACATGGGGTTACCCATGTGAGTTTCCTCATTCGGAAATCCCTGGAT
Matches = 12
Score = 4
CCNNNTTTNCCCTCATNNNN
est. tm =  9.7 oC
DeltaG 37 degrees =  -2.79 kcal/mole

                     5'-
GGGTACTTAGATGTTTCAGTTCCCCGGGTTTACCCTCATACACCTATGTATTCAGTGCATGATACATGGGGTTACCCATGTGAGTTTCCTCATTCGGAAATCCCTGGAT-3'
                                              ||xxx|||x|||||||xxxx
                                          3'-GGAGGAAACGGGAGTAGTAT-5'
```

```
Y. pestis TOX probe TATGATGAGGGCAAAGGAGG versus C. botulinum TOX Target Seq.
TTGTATCATCATTAACAGGGTCATTATAATTAAAACTATTTATTACAACTGGCATAAAATCCCCCCTATACTTCTTTCAATGTCCATAAATATGCCTTTT
Matches = 8
Score = 3
TNNNNTCATCATN
est. tm =  less than 0 oC
DeltaG 37 degrees =  -0.86 kcal/mole

                     5'-
TTGTATCATCATTAACAGGGTCATTATAATTAAAACTATTTATTACAACTGGCATAAAATCCCCCCTATACTTCTTTCAATGTCCATAAATATGCCTTTT-3'
                       |xxxx|||||||x
             3'-GGAGGAAACGGGAGTAGTAT-5'
```

```
Y. pestis TOX probe TATGATGAGGGCAAAGGAGG versus Y. pestis RRL Target Seq.
GGGTACTTAGATGTTTCAGTTCCCCCGGTTCGCCTCGTTAGACTATGTATTCATCTAACGATAGTGCAACGTATTGCACTGGGTTTCCCCATTCGGGTATCGTCGGT
Matches = 12
Score = 4
CCNCNNTTNNCCTCNTNANA
est. tm =  less than zero
DeltaG 37 degrees =  greater than zero

                     5'-
GGGTACTTAGATGTTTCAGTTCCCCCGGTTCGCCTCGTTAGACTATGTATTCATCTAACGATAGTGCAACGTATTGCACTGGGTTTCCCCATTCGGGTATCGTCGGT-3'
                                              ||x|xx||xx||||x|x|x|
                                          3'-GGAGGAAACGGGAGTAGTAT-5'
```

Figure A11.6. AutoDimer primer dimer screen (threshold of 3) of Ypes/TOX probe against all target sequences (complementary strand).

```
Number of Hits = 1
Number of Sequences = 2
Na+ = 0.085 Ct = 1.0

Y. pestis TOX probe TATGATGAGGGCAAAGGAGG versus Francisella RRL Target Seq. CTATATCGAGTGAGTACCCAACTTGTGTCAGGTGGGT
Matches = 6
Score = -8
NNNNNNNTNCCCNNNTNNTN
est. tm =  less than zero
DeltaG 48 degrees =  greater than zero

                    5'-CTATATCGAGTGAGTACCCAACTTGTGTCAGGTGGGT-3'
                            xxxxxxx|x|||xxx|xx|x
                          3'-GGAGGAAACGGGAGTAGTAT-5'
```

Figure A11.7. AutoDimer primer dimer screen (threshold of -9) of Ypes/TOX probe against Fran/RRL target sequence (complementary strand).


```
Number of Hits = 1
Number of Sequences = 33
Na+ = 0.085 Ct = 1.0
```

```
Bacillus ToxGene fwd. primer TGGACATGCTGTGGATGATT versus Yersinia ToxGene rev. primer AGCATCTCCGCCAATAGAGA
Matches = 8
Score = 6
AGCATNTCCN
est. tm =  8.9 oC
DeltaG 37 degrees =  -4.71 kcal/mole

                     5'-AGCATCTCCGCCAATAGAGA-3'
                        |||||x|||x
               3'-TTAGTAGGTGTCGTACAGGT-5'
```

Figure A11.8. AutoDimer primer dimer screen (threshold of 3) of all primer sequences (complementary strand).

CHAPTER IX. LITERATURE CITED

Advisory Committee's Note on Revised Federal Rule of Evidence 702 (2000).

Agilent (2011). Specifications for 2100 Bioanalyzer DNA Kits. Accessed from the World Wide Web address >http://www.genomics.agilent.com/GenericB.aspx?pagetype=Custom&subpagetype=Custom&pageid=2423< on 08Sep2011.

Altman, A. (2011). Fully automated system for the multiplex detection of biothreat agents and emerging infectious diseases. *Proceedings of the 18th International Conference on Biodetection Technologies*, p. 158, Knowledge Foundation, Brookline, MA. Presented at the Annual Conference on 23 June 2011, Washington DC.

Barabote, R.; Xie, G.; Brettin, T.; Hinrichs, S.; Fey, P.; Jay, J.; Engle, J.; Godbole, S.; Noronha, J.; Scheuermann, R.; Zhou, L.; Lion, C.; and Dempsey, M. (2009). Complete genome sequence of *Francisella tularensis* subspecies *holarctica* FTNF002-00. *Public Library of Science One* **4** (9): 1-12.

Barker, J. and Klose, K. (2007) Molecular and genetic basis of pathogenesis in *Francisella tularensis*. *Annals of the New York Academy of Sciences*. **1105**: 138-159.

Baron, G. and Nano, F. (1998). MglA and MglB are required for intramacrophage growth of *Francisella novicida*. *Molecular Microbiology* **29** (1): 247-259.

Battaglia, A. (2009). Pathogen Detection with Luminex Liquid Array Technology (Unpublished Masters Thesis). John Jay College of Criminal Justice.

Battaglia, A.; Schweighardt, A.; and Wallace, M. (2011). Pathogen Detection Using a Liquid Array Technology. *Journal of Forensic Sciences* **56** (3): 760-765.

Biagini, R.; Murphy, D.; Sammons, D.; Smith, J.; Striley, C.; and MacKenzie, B. (2002). Development of multiplexed fluorescence microbead covalent assays (FCMAs) for pesticide biomonitoring. *Buletin of Environmental Contamination and Toxicology* **68**: 470-477.

Biagini, R.; Smith, J.; Sammons, D.; MacKenzie, B.; Striley, C.; Robertson, S.; and Snawder, J. (2004). Development of a sensitivity enhanced multiplexed fluorescence covalent microbead immunosorbent assay (FCMIA) for the measurement of glyphosate, atrazine, and metolachlor mercapturate in water and urine. *Analytical and Bioanalytical Chemistry* **379**: 368-374.

Black, J. (2005). *Microbiology: Principles and Explorations, 6th edition*. John Wiley & Sons, Inc., Hoboken, NJ.

Bless, C.; Palmeter, H.; and Wallace, M. (2006). Identification of *Acer rubrum* using amplified fragment length polymorphism. *Journal of Forensic Sciences* **51** (1): 31-38.

Briese, T.; Palacios, G.; Kokoris, M.; Jabado, O.; Liu, Z.; Renwick, N.; Kapoor, V.; Casas, I.; Pozo, F.; Limberger, R.; Perez-Brena, P.; Ju, J.; and Lipkin, W. (2005). Diagnostic system for rapid and sensitive differential detection of pathogens. *Emerging Infectious Diseases* **11** (2): 310-313.

Bruin, O.; Ludu, J.; and Nano, F. (2007). The Francisella pathogenicity island protein IglA localizes to the bacterial cytoplasm and is needed for intracellular growth. *BMC Microbiology* **7** (1).

Butler, J. (2005). *Forensic DNA Typing*, 2nd edition. Elsevier Academic Press, Amsterdam.

Candlish, A. (1991). Immunological methods in food microbiology. *Food Microbiology* **8** (1): 1-14.

CDC (2011). Bioterrorism Agents/Diseases, Accessed from the World Wide Web address >http://www.bt.cdc.gov/agent/agentlist-category.asp< on 07Sep2011.

Chain, P.; Larimer, F.; Land, M.; Stilwagen, S.; Larsson, P.; Bearden, S.; Chu, M.; Oyston, P.; Forsman, M.; Andersson, S.; Lindler, L.; Titball, R.; and Garcia, E. (2011). Complete genome sequence of *Francisella tularensis* LVS. Accessed from the World Wide Web address >http://www.ncbi.nlm.nih.gov/protein/89143742< on 10Oct2011.

Chandler, D. and Brockman, F. (2000). Renewable microcolumns for solid-phase nucleic acid separations and analysis from environmental samples. *Trends in Analytical Chemistry* **19**: 314-321.

Chandler, D. and Jarrell, A. (2004). Automated purification and suspension array detection of 16S rRNA from soil and sediment extracts by using tunable surface microparticles. *Applied and Environmental Microbiology* **70**: 2621-2631.

Christopher, G.; Cleslak, T.; Pavlin, J.; and Eitzen, E. (1997). Biological warfare: a historical perspective. *Journal of the American Medical Association* **278** (5): 412-417.

Ciammaruconi, A.; Grassi, S.; DeSantis, R.; Faggioni, G.; Pittiglio, V.; D'Amelio, R.; Carattoli, A.; Cassone, A.; Vergnaud, G.; and Lista, F. (2008). Fieldable genotyping of *Bacillus anthracis* and *Yersinia pestis* based on 25-loci multi locus VNTR analysis. *BMC Microbiology* **8**: 21-46.

ClustalW (2010). ClustalW Multiple Sequence Alignment Tool. Accessed from the World Wide Web address >http://align.genome.jp/< on 17 January 2010.

Cole, S.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Eiglmeier, K.; Gas, S.; Barry, C.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M.; Rajandream, M.; Rogers, J.; Rutter, S.; Seeger, K.; Squares, R.; Squares, S.; Sulston, J.; Taylor, K.; Whitehead, S.; and Barrell, B. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**: 537-544.

Cox, R. (2004). Quantitative relationships for specific growth rates and macromolecular compositions of Mycobacterium tuberculosis, Streptomyces coelicolor A3(2) and Escherichia coli B/r: an integrative theoretical approach. *Microbiology* **150**: 1413-1426.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Deshpande, A.; Gans, J.; Graves, S.; Green, L.; Taylor, L.; Kim, H.; Kunde, Y.; Leonard, P.; Li, P.; Mark, J.; Song, J.; Vuyisich, M.; and White, P. (2010). A rapid multiplex assay for nucleic acid-based diagnostics. *Journal of Microbiological Methods* **80** (2): 155-163.

Diaz, M. and Fell, J. (2004). High-throughput detection of pathogenic yeasts of the genus trichosporon. *Journal of Clinical Microbiology* **42** (8): 3696-3706.

Dimo-Simonin, N. and Brandt-Casadevall, C. (1996). Evaluation and usefulness of reverse dot blot DNA-PolyMarker typing in forensic case work. *Forensic Science International* **81** (1): 61-72.

Don, R.; Cox, P.; Wainwright, B.; Baker, K.; and Mattick, J. (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* **19** (14): 4008.

Drago, F.; Karpasitou, K.; and Poli, F. (2009). Microarray beads for identifying blood group single nucleotide polymorphisms. *Transfusion Medicine and Hemotherapy* **36** (3): 157-160.

Duesbery, N.; Webb, C.; Leppla, S.; Gordon, V.; Klimpel, K.; Copeland, T.; Ahn, N.; Oskarsson, M.; Fukasawa, K.; Paull, K.; and Vande Woude, G. (1998). Proteolytical inactivation of MAP-kinase-kinase by Anthrax lethal factor. *Science* **280**: 734-737.

Dunbar, S. and Jacobson, J. (2007) Quantitative, multiplexed detection of *Salmonella* and other pathogens by Luminex® xMAP™ suspension array. *Methods in Molecular Biology* **394**: 1-19.

Dunbar, S. and Jacobson, J. (2005). Rapid screening for 31 mutations and polymorphisms in the cystic fibrosis transmembrane conductance regulator gene by Luminex xMAP suspension array. *Methods in Molecular Medicine* **114**: 114-171.

East, A.; Bhandari, M.; Hielm, S.; and Collins, M. (1998). Analysis of the botulinum neurotoxin type F gene clusters in proteolytic and nonproteolytic *Clostridium botulinum* and *Clostridium barati*. *Current Microbiology* **37**: 262-268.

Ecker, D.; Sampath, R.; Massire, C.; Blyn, L.; Hall, T.; Eshoo, M.; and Hofstadler, S. (2008). Ibis T5000: a universal biosensor approach for microbiology. *Nature Reviews Microbiology* **6**: 553-558.

Ecker, D.; Sampath, R.; Willett, P.; Wyatt, J.; Samant, V.; Massire, C.; Hall, T.; Hari, K.; McNeil, J.; Buchen-Osmond, C.; and Budowle, B. (2005). The microbial rosetta stone database: a compilation of global and emerging infectious microorganisms and bioterrorist threat agents. *BMC Microbiology* **5**: 19.

Edwards, A.; Civitello, A.; Hammond, H.; and Caskey, C. (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* **49** (4): 746-756.

Fleming, D. and Foshay, L. (1955). Studies on the physiology of virulence of *Pasteurella tularensis*. I: citrulline ureidase and deamidase activity. *Journal of Bacteriology* **70**: 345-349.

Folkertsma, R.; Voort, J.; Groot, K.; Zantvoort, P.; Schots, A.; Gommers, F.; Helder, J.; and Bakker, J. (1996). Gene pool similarities of potato cyst nematode populations assessed by AFLP analysis. *Molecular Plant-Microbe Interactions* **9**: 47-54.

Forman, S.; Wulff, C.; Myers-Morales, T.; Cowan, C.; Perry, R.; and Straley, S. (2008). *yadBC* of *Yersinia pestis*, a new virulence determinant for bubonic plague. *Infection and Immunity* **76** (2): 578-587.

Fry, N.; Savelkoul, P.; and Visca, P. (2009). Amplified fragment length polymorphism analysis. *Methods in Molecular Biology* **551**: 89-104.

Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).

Fulton, R.; McDade, R.; Smith, P.; Kienker, L.; and Kettman, J. (1997). Advanced multiplexed analysis with the FlowMetrix™ system. *Clinical Chemistry* **43**: 1749-1756.

Garrity, G.; Boone, D.; and Castenholz, R. (2001). *Bergey's Manual of Systematic Bacteriology, Volume I: The Archaea and the Deeply Branching Phototrophic Bacteria*, 2nd edition. Springer, New York, NY.

Grate, J.; Bruckner-Lea, C.; Jarrell, A.; and Chandler, D. (2003). Automated sample preparation method for suspension arrays using renewable surface separations with multiplexed flow cytometry fluorescence detection. *Analytica Chimica Acta* **478**: 85-98.

Ghosh, T.; Prelas, M.; Viswanath, D.; and Loyalka, S. (eds.) (2002). *Science and Technology of Terrorism and Counterterrorism*. Marcel Dekker, Inc., New York, NY.

Gil, H.; Benach, J.; and Thanassi, D. (2004). Presence of pili on the surface of *Francisella tularensis*. *Infection and Immunity* **72** (5): 3042-3047.

Golovliov, I.; Baranov, V.; Krocova, Z.; Kovarova, H.; and Sjostedt, A. (2003) An attenuated strain of the facultative intracellular bacterium *Francisella tularensis* can escape the phagosome of monocytic cells. *Infection and Immunity* **71** (10): 5940-5950.

Gordon, A.; Hart, P.; and Young, M. (1980). Ammonia inhibits phagosome-lysosome fusion in macrophages. *Nature* **286**: 79-80.

Hanley, B. (2007). Variance in multiplex suspension array assays: carryover of microspheres between sample wells. *Journal of Negative Results in Biomedicine* **6** (6).

Hanna, P.; Acosta, D.; and Collier, J. (1993). On the role of macrophages in anthrax. *Proceedings of the National Academy of Sciences* **90**: 10198-10201.

Hartwell, L.; Hood, L.; Goldberg, M.; Reynolds, A.; Silver, L.; and Veres, R. (2008). *Genetics: from genes to genomes*, 3rd edition. McGraw-Hill Higher Education, New York, NY.

Hindson, B.; Brown, S.; Marshall, G.; McBride, M.; Makarewicz, A.; Gutierrez, D.; Wolcott, D.; Metz, T.; Madabhushi, R.; Dzenitis, J.; and Colston, B. (2004). Development of an automated sample preparation module for environmental monitoring of biowarfare agents. *Analytical Chemistry* **76**: 3492-3497.

Hirsch, J.; Eslamizar, L.; Filanoski, B.; Malekzadeh, N.; Haugland, R.; Beechem, J.; and Haugland, R. (2002). Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: uses for protein labeling, detection, and isolation. *Analytical Biochemistry* **308**: 343-357.

Horan, P. and Wheeless, L. (1977). Quantitative single cell analysis and sorting. *Science* **198**: 149-157.

Hurst, G.; Doktycz, M.; Vass, A.; and Buchanan, M. (1996). Detection of bacterial DNA polymerase chain reaction products by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **10** (3): 377-382.

Ikuta, S.; Takagi, K.; Wallace, R.; and Ikatura, K. (1987). Dissociation kinetics of 19 base paired oligonucleotides-DNA duplexes containing different single mismatched base pairs. *Nucleic Acids Research* **15**: 797-811.

Ishii, H.; Tanoue, J.; Oshima, M.; Chung, W.; Nishimura, K.; Yamaguchi, J.; Nemoto, F.; So, K.; Iwama, T.; Yoshimatsu, H.; Shimizu, M.; and Kozawa, T. (2008). First application of PCR-Luminex system for molecular diagnosis of fungicide resistance and species identification of fungal pathogens. *Journal of General Plant Pathology* **74** (6): 409-416.

Ivanova, N.; Sorokin, A.; Anderson, I.; Galleron, N.; Candelon, B.; Kapatral, V.; Bhattacharyya, A.; Reznik, G.; Mikhailova, N.; Lapidus, A.; Chu, L.; Mazur, M.; Goltsman, E.; Larsen, N.; D'Souza, M.; Walunas, T.; Grechkin, Y.; Pusch, G.; Haselkorn, R.; Fonstein, M.; Ehrlich, S.; Overbeek, R.; and Kyrpides, N. (2003). Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature* **423**: 87-91.

Janssen, P.; Coopman, R.; Huys, G.; Swings, J.; Bleeker, M.; Vos, P.; Zabeau, M.; and Kersters, K. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology* **142**: 1881-1893.

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

Keppie, J.; Harris-Smith, P.; and Smith, H. (1963). The chemical basis of the virulence of *Bacillus anthracis*. *British Journal of Experimental Pathology* **44**: 446-453.

Kettman, J.; Davies, T.; Chandler, D.; Oliver, K.; and Fulton, R. (1998). Classification and properties of 64 multiplexed microsphere sets. *Cytometry* **33**: 234-243.

Kirk, P. (1963a). The ontogeny of criminalistics. *Journal of Criminal Law, Criminology, and Police Science* **54** (2): 235-238.

Kirk, P. (1963b). Criminalistics. *Science* **140**: 367-370.

Klimpel, K.; Arora, N.; and Leppla, S. (1994). Anthrax toxin lethal factor contains a zinc metalloproteinase consensus sequence which is required for lethal toxin activity. *Molecular Microbiology* **13**: 1093-1100.

Koch, R. (1876). Die aetiologie der milzbrandkrankheit, begründet auf die entwicklungsgeschichte des *Bacillus anthracis* (English translation). *Beiträge zur Biologie der Pflanzen* **2**: 277-310.

Korbie, D. and Mattick, J. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols* **3** (9): 1452-1456.

Lai, X.; Golovliov, I.; and Sjostedt, A. (2001). *Francisella tularensis* induces cytopathogenicity and apoptosis in murine macrophages via a mechanism that requires intracellular bacterial multiplication. *Infection and Immunity* **69** (7): 4691-4694.

Lander, E. (1989). Expert's report in *People v. Castro*. Whitehead Institute for Biomedical Research and Harvard University.

Larsson, P.; Oyston, P.; Chain, P.; Chu, M.; Duffield, M.; Fuzelius, H.; Garcia, E.; Halltorp, G.; Johansson, D.; Isherwood, K.; Karp, P.; Larsson, E.; Liu, Y.; Michell, S.; Prior, J.; Prior, R.; Malfatti, S.; Sjostedt, A.; Svensson, K.; Thompson, N.; Vergez, L.; Wagg, J.; Wren, B.; Lindler, L.; Anderson, S.; Forsman, M.; and Titball, R. (2005). The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. *Nature Genetics* **37** (2): 153-159.

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

Leal, N. and Almeida, A. (1999). Diagnosis of plague and identification of virulence markers in *Yersinia pestis* by multiplex-PCR. *Revista do Instituto de Medicina Tropical de Sao Paulo* **41** (6): 339-342.

Leppla, S. (1995). Anthrax toxins, pp. 543-572. In Moss, J.; Iglewski, M.; Vaughn, M.; and Tu, A. (editors), *Bacterial toxins and virulence factors in disease*. Marcel Dekker, New York, NY.

Leppla, S. (1999) The bifactorial *Bacillus anthracis* lethal and oedema toxins, pp. 243-263. In Alouf, J. and Freer, J. (editors), *The Comprehensive Sourcebook of Bacterial Protein Toxins*, 2nd edition. Academic Press, New York, NY.

Lim, D.; Simpson, J.; Kearns, E.; and Kramer, M. (2005). Current and developing technologies for monitoring agents of bioterrorism and biowarfare. *Clinical Microbiology Reviews* **18**: 583-607.

Lindstrom, M. and Korkeala, H. (1998). Laboratory diagnostics of botulism. *Clinicial Microbiology Reviews* **19** (2): 298-314.

Lodish, H.; Berk, A.; Matsudaira, P.; Kaiser, C.; Krieger, M.; Scott, M.; Zipursky, S.; and Darnell, J. (2004). *Molecular Biology of the Cell*, 5th edition. W.H. Freeman and Company, New York, NY.

Maglott, D.; Ostell, J.; Pruitt, K.; and Tatusova, T. (2005). Entrez gene: gene-centered information at NCBI. *Nucleic Acids Research* **33**: D54-58.

Mahawar, M.; Kirimanjeswara, G.; Metzger, D.; and Bakshi, C. (2009). Contribution of citrulline ureidase to *Francisella tularensis* strain *Schu S4* pathogenesis. *Journal of Bacteriology* **191** (15): 4798-4806.

Martin, J.; Christopher, G.; and Eitzen, E. (2007). History of biological weapons: from poisoned darts to intentional epidemics. In *Medical Aspects of Biological Warfare*, Borden Institute, Walter Reed Army Medical Center, Washington DC.

McBride, M.; Masquelier, D.; Hindson, B.; Makarewicz, A.; Brown, S.; Burris, K.; Metz, T.; Langlois, R.; Tsang, K.; Bryan, R.; Anderson, D.; Venkateswaran, K.; Milanovich, F.; and Colston, B. (2003). Autonomous detection of aerosolized Bacillus anthracis and Yersinia pestis. *Analytical Chemistry* **75**: 5293-5299.

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

Milne, J.; Furlong, D.; Hanna, P.; Wall, J.; and Collier, R. (1994). Anthrax protective antigen forms oligomers during intoxication of mammalian cells. *Journal of Biological Chemistry* **269** (32): 20607-20612.

Mitsis, P. and Kwagh, J. (1999). Characterization of the interaction of lamda exonuclease with the ends of DNA. *Nucleic Acids Research* **27** (15): 3057-3063.

Morag, E.; Bayer, E.; and Wilchek, M. (1996). Reversibility of biotin-binding by selective modification of tyrosine in avidin. *Biochemical Journal* **316**: 193-199.

Morgulis, A.; Gertz, E.; Schaffer, A.; and Agarwala, R. (2006). A fast and symmetric DUST implementation to mask low-complexity DNA sequences. *Journal of Computational Biology* **13** (5): 1028-1040.

NCBI (2010). National Center for Biotechnology Information. Accessed from the World Wide Web address >http://blast.ncbi.nlm.nih.gov/Blast.cgi< on 13 March 2010.

Nomura, M.; Gourse, R.; and Baughman, G. (1984). Regulation of the synthesis of ribosomes and ribosomal components. *Annual Review of Biochemistry* **53**: 75-117.

O'Brien, J.; Friedlander, A.; Dreier, T.; Ezzell, J.; and Leppla, S. (1985). Effects of Anthrax toxin components on human neutrophils. *Infection and Immunity* **47** (1): 306-310.

Okinaka, R.; Cloud, K.; Hampton, O.; Hoffmaster, A.; Hill, K.; Keim, P.; Koehler, T.; Lamke, G.; Kumano, S.; Mahillon, J.; Manter, D.; Martinez, Y.; Ricke, D.; Svensson, R.; and Jackson, P. (1999). Sequence and organization of pXO1, the large *Bacillus anthracis* plasmid harboring the anthrax toxin genes. *Journal of Bacteriology* **181** (20): 6509-6515.

Olejnik, J.; Sonar, S.; Krzymanska-Olejnik, E.; and Rothschild, K. (1995). Photocleavable biotin derivatives: a versatile approach for the isolation of biomolecules. *Proceedings of the National Academy of Sciences* **92**: 7590-7594.

Pabbaraju, K.; Wong, S.; Lee, S.; Tellier, R.; Fonseca, K.; Louie, M.; Drews, S. (2011). Comparison of a singleplex real-time RT-PCR assay and multiplex respiratory viral panel assay for detection of influenza 'A' in respiratory specimens. *Influenza and Other Respiratory Viruses* **5**: 99-103.

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

Perry, R. and Fetherston, J. (1997). *Yersinia pestis* – etiologic agent of plague. *Clinical Microbiology Reviews* **10** (1): 35-66.

Pezard, C.; Berche, P.; and Mock, M. (1991). Contribution of individual toxin components to virulence of *Bacillus anthracis*. *Infection and Immunity* **59** (10): 3472-3477.

Poulain, B.; Stiles, B.; Popoff, M.; and Molgo, J. (2006). Attack of the nervous system by clostridial toxins: physical findings, cellular and molecular actions, pp. 348-389. In Alouf, J and Popoff, M (editors) *The Comprehensive Sourcebook of Bacterial Protein Toxins*, 3rd edition. Academic Press, New York, NY.

Prentice, M.; James, K.; Parkhill, J.; Baker, S.; Stevens, K.; Simmonds, M.; Mungall, K.; Churcher, C.; Oyston, P.; Titball, R.; Wren, B.; Wain, J.; Pickard, D.; Hien, T.; Farrar, J.; and Dougan, G. (2001). *Yersinia pestis* pFra shows biovar-specific differences and recent common ancestry with *Salmonella enterica* serovar Typhi plasmid. *Journal of Bacteriology* **183** (8): 2586-2594.

Radnedge, L.; Agron, P.; Hill, K.; Jackson, P.; Ticknor, L.; Keim, P.; and Andersen, G. (2003). Genome differences that distinguish *Bacillus anthracis* from *Bacillus cereus* and *Bacillus thuringiensis*. *Applied and Environmental Microbiology* **69** (5): 2755-2764.

Read, T.; Peterson, S.; Tourasse, N.; Baillie, L.; Paulsen, I.; Nelson, K.; Tettelin, H.; Fouts, D.; Eisen, J.; Gill, S.; Holtzapple, E.; Okstad, O.; Helgason, E.; Rilstone, J.; Wu, M.; Kolonay, J.; Beanan, M.; Dodson, R.; Brinkac, L.; Gwinn, M.; DeBoy, R.; Madpu, R.; Daugherty, S.; Durkin, A.; Haft, D.; Nelson, W.; Peterson J.; Pop, M.; Khouri, H.; Radune, D.; Benton, J.; Mahamoud, Y.; Jian, L.; Hance, I.; Weidman, J.; Berry, K.; Plaut, R.; Wolf, A.; Watkins, K.; Nierman, W.; Hazen, A.; Cline, R.; Redmond, C.; Thwaite, J.; White, O.; Salzberg, S.; Thomason, B.; Friedlander, A.; Koehler, T.; Hanna, P.; Kolsto, A.; and Fraser, C. (2003). The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* **423**: 81-86.

Rosendahl, S. and Taylor, J. (1997). Development of multiple genetic markers for studies of genetic variation in arbuscular mycorrhizal fungi using AFLP. *Molecular Ecology* **6**: 821-829.

Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In: Krawetz, S and Misener, S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa NJ, pp. 365-386.

Saiki, R.; Walsh, P.; Levenson, C.; and Erlich, H. (1989). Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proceedings of the National Academy of Sciences* **86**: 6230-6234.

Santic, M.; Molmeret, M.; Klose, K.; Jones, S.; and Kwaik, Y. (2005). The *Francisella tularensis* pathogenicity island protein IglC and its regulator MglA are essential for modulating phagosome biogenesis and subsequent bacterial escape into the cytoplasm. *Cellular Microbiology* **7** (7): 969-979.

Schoske, R.; Vallone, P.; Ruitberg, C.; and Butler, J. (2003). Multiplex PCR design strategy used for the simultaneous amplification of 10 Y chromosome short tandem repeat (STR) loci. *Analytical and Bioanalytical Chemistry* **375**: 333-343.

Schwieger, F. and Tebbe, C. (1998). A new approach to utilize PCR-single-strand-conformation polymorphism for 16S rRNA gene-based microbial community analysis. *Applied and Environmental Microbiology* **64** (12): 4870-4876.

Sebaihia, M.; Wren, B.; Mullany, P.; Fairweather, N.; Minton, N.; Stabler, R.; Thomson, N.; Roberts, A.; Cerdeno-Tarraga, A.; Wang, H.; Holden, M.; Wright, A.; Churcher, C.; Quail, M.; Baker, S.; Bason, N.; Brooks, K.; Chillingworth, T.; Cronin, A.; Davis, P.; Dowd, L.; Fraser, A.; Feltwell, T.; Hance, Z.; Holroyd, S.; Jagels, K.; Moule, S.; Mungall, K.; Price, C.; Rabbinowitsch, E.; Sharp, S.; Simmonds, M.; Stevens, K.; Unwin, L.; Whithead, S.; Dupuy, B.; Dougan, G.; Barrell, B.; and Parkhill, J. (2006). The multi-drug resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nature Genetics* **38**: 779-786.

Sigma (2010). DNA Calculator. Accessed from the World Wide Web address >http://www.sigma-genosys.com/calc/DNACalc.asp< on 15 February 2010.

Simpson, L. (2004). Identification of the major steps in botulinum toxin action. *Annual Review of Pharmacology and Toxicology* **44**: 167-193.

Sjostedt, A. (2004). Gram-negative aerobic cocci. Family XVII. Francisellaceae, p. 200-210. In Brenner, D. (ed.), *Bergey's Manual of Systematic Bacteriology* . Springer, New York, NY.

Skillman, L.; Toovey, A.; Williams, A.; and Wright, A. (2006). Development and validation of a real-time PCR method to quantify rumen protozoa and examination of variability between Entodinium populations in sheep offered a hay-based diet. *Applied and Environmental Microbiology* **72**: 200-206.

Smith, T.; Hill, K.; Foley, B.; Detter, J.; Munk, A.; Bruce, D.; Doggett, N.; Smith, L.; Marks, J.; Xie, G.; and Brettin, T. (2007). Analysis of the neurotoxin complex genes in *Clostridium botulinum* A1-A4 and B1 strains: BoNT/A3, /Ba4, and /B1 clusters are located within plasmids. *Public Library of Sciences One* **12**: e1271.

Sodeinde, O. and Goguen, J. (1988). Genetic analysis of the 9.5-kilobase virulence plasmid of *Yersinia pestis*. *Infection and Immunity* **56** (10): 2743-2748.

Sodeinde, O.; Subrahmanyam, Y.; Stark, K.; Quan, T.; Bao, Y.; and Goguen, J. (1992). A surface protease and the invasive character of plague. *Science* **258**: 1004-1007.

Song, Y.; Roumagnac, P.; Weill, F.; Wain, J.; Dolecek, C.; Mazzoni, C.; Holt, K.; and Achtman, M. (2010). A multiplex single nucleotide polymorphism typing assay for detecting mutations that result in decreased fluoroquinolone susceptibility in *Salmonella enterica* serovars Typhi and Paratyphi A. *Journal of Antimicrobial Chemotherapy* **65** (8): 1631-1641.

Southern, E. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**: 503-517.

Spiro, A. and Lowe, M. (2002). Quantitation of DNA sequences in environmental PCR products by a multiplexed, bead-based method. *Applied and Environmental Microbiology* **68**: 1010-1013.

Spiro, A.; Lowe, M.; and Brown, D. (2000). A bead-based method for multiplexed identification and quantitation of DNA sequences using flow cytometry. *Applied and Environmental Microbiology* **66**: 4258-4265.

Srivastava, A. and Schlessinger, D. (1990). Mechanism and regulation of bacterial ribosomal RNA processing. *Annual Review of Microbiology* **44**: 105-129.

Stanley, J. and Smith, H. (1961). Purification of factor I and recognition of a third factor of anthrax toxin. *Journal of General Microbiology* **26**: 49-66.

Summerbell, R.; Levesque, C.; Seifert, K.; Bovers, M.; Fell, J.; Diaz, M.; Boekhout, T.; de Hoog, G.; Stalpers, J.; and Crous, P. (2005). Microcoding: the second step in DNA barcoding. *Philosophical Transactions of the Royal Society B* **360**: 1897-1903.

Taniuchi, M.; Verweij, J.; Noor, Z.; Sobuz, S.; Lieshout, L.; Petri, W.; Haque, R.; and Houpt, E. (2011). High throughput multiplex PCR and probe-based detection with Luminex beads for seven intestinal parasites. *American Journal of Tropical Medicine and Hygiene* **84** (2): 332-337.

Taranenko, N.; Hurt, R.; Zhou, J.; Isola, N.; Huang, H.; Lee, S.; and Chen, C. (2002). Laser desorption mass spectrometry for microbial DNA analysis. *Journal of Microbiological Methods* **48** (2-3): 101-106.

Thomson, N.; Howard, S.; Wren, B.; Holden, M.; Crossman, L.; Challis, G.; Churcher, C.; Mungall, K.; Brooks, K.; Chillingworth, T.; Feltwell, T.; Abdellah, Z.; Hauser, H.; Jagels, K.; Maddison, M.; Moule, S.; Sanders, M.; Whithead, S.; Quail, M.; Dougan, G.; Parkhill, J.; and Prentice, M. (2006). The complete genome sequence and comparative genome analysis of the high pathogenicity *Yersinia enterocolitica* strain 8081. *Public Library of Science Genetics* **2**: 2039-2051.

Tong, X. and Smith, L. (1992). Solid-phase method for the purification of DNA sequencing reactions. *Analytical Chemistry* **64**: 2672-2677.

Vallone, P. and Butler, J. (2004). AutoDimer: a screening tool for primer-dimer and hairpin structures. *BioTechniques* **37**: 226-231.

Valjevac, S.; Hilaire, V.; Lisanti, O.; Ramisse, F.; Hernandez, E.; Cavallo, J.; Pourcel, C.; and Vergnaud, G. (2005). Comparison of minisatellite polymorphisms in the *Bacillus cereus* complex: A simple assay for large-scale screening and identification of strains most closely related to *Bacillus anthracis*. *Applied and Environmental Microbiology* **71**(11): 6613-6623.

Viboud, G. and Bliska, J. (2005). Yersinia outer proteins: role in modulation of host cell signaling responses and pathogenesis. *Annual Review of Microbiology* **59**: 69-89.

Vignali, D. (2000). Multiplexed particle-based flow cytometric assays. *Journal of Immunological Methods* **243**: 243-255.

Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; Lee, T.; Hornes, M.; Frijters, A.; Pot, J.; Peleman, J.; Kuiper, M.; and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23** (21): 4407-4414.

Walsh, P.; Fildes, N.; Louie, A.; and Higuchi, R. (1991). Report of the blind trial of the Cetus Amplitype HLA DQ alpha forensic deoxyribonucleic acid (DNA) amplification and typing kit. *Journal of Forensic Sciences* **36** (5): 1551-1556.

Wang, S.; Wen, J.; Zhou, Y.; Zhang, Z.; Yang, R.; Zhang, J.; Chen, J.; and Zhang, X. (2004) Identification and characterization of *Bacillus anthracis* by multiplex PCR on DNA chip. *Biosensors and Bioelectronics* **20**: 807-813.

Wilson, W.; Erler, A.; Nasarabadi, S.; Skowronski, E.; and Imbro, P. (2005). A multiplexed PCR-coupled liquid bead array for the simultaneous detection of four biothreat agents. *Molecular and Cellular Probes* **19**: 137-144.