AMPLIFIED FRAGMENT LENGTH POLYMORPHISM (AFLP) ANALYSIS AS A DNA TYPING TOOL FOR MICROBIAL FORENSICS

By

KATELYN DEANE WEINBRECHT

Bachelor of Science in Biology

Pacific Lutheran University

Tacoma, Washington

2009

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 2011

AMPLIFIED FRAGMENT LENGTH

POLYMORPHISM (AFLP) ANALYSIS AS A DNA

TYPING TOOL FOR MICROBIAL FORENSICS

Thesis Approved:

Dr. Robert Allen
Thesis Adviser

Dr. Tom Glass

Dr. Frank Champlin

Dr. Gerwald Koehler

Dr. Mark E. Payton
Dean of the Graduate College

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

II. REVIEW OF LITERATURE 6

Agriculture in the United States 7
Agricultural Bioterrorism 8
History of Agroterrorism 8
American Response and Preparedness 10
Pseudomonas syringae 12
Microbial Forensics 13
Amplified Fragment Length Polymorphism Analysis 14
Summary 15

III. METHODOLOGY 17

Overview of AFLP Method 17
Bacterial Strains 18
DNA Isolation 19
DNA Quantification 21
DNA Digestion 21
DNA Ligation 22
Preselective Amplification 23
Selective Amplification 25
Capillary Electrophoresis 27
Analysis 29

Chapter Page

IV. FINDINGS....31

Reproducibility....31
Discriminatory Capability....39
Genetic Relatedness....45
Sørensen Similarity Index....45
Agglomerative Hierarchical Clustering....49
Genetic Relatedness....52

V. CONCLUSIONS....55

REFERENCES....60

APPENDICES....63

LIST OF TABLES

Table Page

1 – Species and strains of *P. syringae* and *P. aeruginosa* used in this study…….20

2 – Thermocycler parameters for preselective amplification …….24

3 – Thermocycler parameters for selective amplification ……..…….26

4 – Electrophoretic peak values from *P. syringae* strain F7 JOE electropherograms....38

5 – Haplotype Codes for *P. syringae* strains F7 and F10A...…….….42

6 – Haplotype Codes for *P. syringae* strains NF5 and NF12…….….43

7 – Haplotype Codes for PAO1-A, PAO1-B, and PAO1-original ……...…..….53

LIST OF FIGURES

Figure Page

1 – Ethidium bromide stained gel of digested *P. syringae* DNA…….....22

2 – Ethidium bromide stained gel of successful *P. syringae* DNA ligation…..…….….25

3A – *P. syringae* strain F10A FAM electropherogram….....28

.

3B – *P. syringae* strain F10A JOE electropherogram….....28

3C – *P. syringae* strain F10A NED electropherogram….....29

4A – *P. syringae* strain F12 FAM electropherogram Run 1 33

4B – *P. syringae* strain F12 FAM electropherogram Run 2.....33

4C – *P. syringae* strain F12 FAM electropherogram Run 3.....33

4D – *P. syringae* strain F12 JOE electropherogram Run 1 34

4E – *P. syringae* strain F12 JOE electropherogram Run 2 34

4F – *P. syringae* strain F12 JOE electropherogram Run 3 ……...…..…….34

4G – *P. syringae* strain F12 NED electropherogram Run 1 35

4H – *P. syringae* strain F12 NED electropherogram Run 2.....35

4I – *P. syringae* strain F12 NED electropherogram Run 3 35

5A – *P. syringae* strain F10A FAM electropherogram ….....40

5B – *P. syringae* strain F7 FAM electropherogram ….....40

5C – *P. syringae* strain F10A JOE electropherogram ….....41

5D – *P. syringae* strain 7 JOE electropherogram ….....41

5E – *P. syringae* strain F10A NED electropherogram ….....42

Figure Page

5F – *P. syringae* strain F7 NED electropherogram….....42

6A – Sorensen similarity index comparing *P. syringae* pv. maculicola strains (FAM)..46

6B – Sorensen similarity index comparing *P. syringae* pv. maculicola strains (JOE)…47

6C – Sorensen similarity index comparing *P. syringae* pv. maculicola strains (NED)...48

7 – Dendrogram displaying genetic relatedness of all analyzed microbial strains……..51

CHAPTER I

INTRODUCTION

In 2001, a single case of foot-and-mouth disease (FMD) was detected in a pig in the United Kingdom. Within the next year, over 6 million animals tested positive for FMD and were slaughtered, leading to economic losses exceeding $16 billion. Today, scientists still do not know whether the FMD virus was naturally occurring or deliberately introduced into the United Kingdom (Breeze, Budowle, and Schutzer, 2005). Either way, the devastation that can result from an agricultural disease outbreak is made evident by the FMD outbreak of 2001.

Reviewing the history of natural microbial outbreaks in the farming sector allows researchers, epidemiologists, and microbiologists to better understand the impact of agroterrorism events. In the United States foodborne diseases are common, with almost 1,100 outbreaks reported to the Centers for Disease Control and Prevention (CDC) in the year 2007 alone (CDC, 2010). Although the majority of outbreaks typically only affect a limited portion of the food supply and a minor percentage of the population, more severe disease outbreaks can lead to widespread human illness and large economic losses. As recently as September, 2010 the Food and Drug Administration (FDA) issued a nationwide recall of eggs due to a multistate outbreak of Salmonella. In total, 380

million eggs were recalled and more than 2,700 people fell ill with Salmonella (CDC, 2010). Although this microbial outbreak is believed to have been naturally occurring, the damage to human health and the economy offers a look at the possible repercussions of a deliberate attack.

An attack on the agricultural industry could potentially cost the economy billions of dollars and ultimately lead to a loss of consumer confidence in the government and the farming industry. Since agroterrorism events have not knowingly occurred in modern history, investigators are forced to take advantage of naturally occurring microbial outbreaks to develop preparedness for agricultural bioterrorism events should they occur in the future (Monke, 2004).

The Federal Bureau of Investigation (FBI) defines terrorism as, "the unlawful use of force or violence against persons or property to intimidate or coerce a Government, the civilian population, or any segment thereof, in furtherance of political or social objectives" (Henneberry, 2001). With the growing possibility of biological weapons being used in a terrorism act, *bioterrorism* has emerged with a separate, more distinct definition. The Department of Defense (DOD) defines bioterrorism specifically as, "the use of biological agents in terrorism. This includes the malevolent use of bacteria, viruses, or toxins against people, animals, or plants" (Henneberry, 2001). With the prevalence and availability of pathogens that could be used as biological weapons, an attack on the United States agricultural or farming sector is a very real possibility. The US is currently unprepared for detecting and responding to an agroterrorism event that utilizes naturally occurring microorganisms as bioweapons (Fletcher, Bender, Budowle, Cobb, Gold, *et al*., 2006). Microbial and molecular techniques must be developed and

adapted for use by the forensic community in preparing for and responding to agricultural terrorism threats (Fletcher, Bender, Budowle, Cobb, Gold, *et a.l*, 2006).

In an effort to prepare for possible bioterrorism events in the future, the field of microbial forensics has been developed. Spurring on the creation of microbial forensics is the recent publication of government reports that have mentioned the need for increased focus on agroterrorism research and preparedness. Published in 2002, *The Bioterrorism Preparedness and Response Act*, explicitly states the need for research in the area of microbial identification (Parker, 2003). With recent government actions mentioning the need for agroterrorism and bioterrorism preparedness, the pace of microbial forensics research has increased. A variety of molecular techniques have been explored as investigators attempt to strengthen and focus efforts in developing effective and rapid techniques for microbial identification and attribution (Cummings and Relman, 2002). Techniques that show promise in the field include, but are not limited to, restriction fragment length polymorphism (RFLP) analysis, multilocus sequence typing (MLST), multiple locus variable number tandem repeat (VNTR) analysis (MLVA), and amplified fragment length polymorphism (AFLP) analysis (Fletcher, Bender, Budowle, Cobb, Gold, et al, 2006). Each of these techniques is easily performed in and adapted to a forensic setting with basic DNA analysis capabilities. However, even with enhanced research in microbial detection techniques, a single cohesive method remains to be developed that forensic laboratories can apply to specifically differentiate, characterize, and identify a wide variety of microbial organisms (Cummings and Relman, 2002).

In addition to the complex variety of microbial identification techniques available to the forensic lab, investigators must also become familiar with the wide array of

microorganisms that may be utilized as pathogens in a bioterrorism event. The research presented here focuses on *Pseudomonas syringae* , an opportunistic, gram-negative bacterium that is a common pathogen for over 50 varieties of cash crops in the United States. *P. syringae* outbreaks occur worldwide and pathogenic strains are becoming increasingly virulent and difficult to control (Rudolph, 1997). In addition to having a wide host range, many *P. syringae* species are increasingly becoming bactericide-resistant (Rudolph, 1997), making *P. syringae* an ideal microbial species to be used as an agroterrorism agent (Monke, 2004). Multiple studies have focused on the genetic typing and differentiation of *P. syringae* strains in an effort to better understand the species and its potential pathogenicity. In previous studies (Georgnaras, 1999; Clerc, Manceau, and Nesme, 1998; Taylor, 2009), amplified fragment length polymorphism (AFLP) analysis was explored and showed promise as a molecular tool capable of bacterial strain differentiation. Past successes with AFLP analysis in analyzing microbial strains, including strains of *P. syringae*, indicate that AFLP analysis may be one tool with potential for attribution of microbial agents (Taylor, 2009; Jackson, Hill, and Laker, *et al*, 1999)

AFLP analysis is a molecular technique that combines restriction fragment length polymorphism (RFLP) analysis and polymerase chain reaction (PCR) techniques (Vos, et al, 1995). In past studies AFLP analysis was used to differentiate a wide variety of bacterial strains and species, including strains of *Serratia marcescens, Pseudomonas aeruginosa, Staphylococcus aureus,* and *Bacillus anthracis*. (Allen, 2006; Beauman, 2007; Taylor, 2009; Jackson, Hill, Laker, Ticknor, and Keim, 1999). Although differentiation of bacterial strains using AFLP analysis has been successful, the technique

has never been utilized to specifically identify an unknown bacterium. Thus, more classical species identification techniques must be used in conjunction with AFLP analysis to identify and attribute an unknown strain.

The purpose of this research was to adapt and refine a published method for DNA profiling of microbial strains (i.e. AFLP analysis), which could be used to attribute the source of a plant pathogen recovered as evidence in a biocrime. In addition, the ability of AFLP analysis to potentially characterize an unknown bacterial strain was investigated. Included among the species subjected to genetic analysis were strains of *P. syringae* and *P. aeruginosa.* Resulting AFLP electropherograms obtained for each strain were subsequently translated into a haplotype code based on the presence and size of DNA fragments in the AFLP profile. The generation of a haplotype code for each analyzed strain allowed for the efficient comparison of strains. Also, the generation of a unique code for each analyzed strain allowed for the creation of a haplotype code database. In addition, haplotype codes for strains of *S. marcescens, P. syringae* pv. tomato,and *S. aureus*, were obtained from previous research (Allen, 2006; Beauman, 2007; Taylor, 2009) and were analyzed to determine if species-specific characteristics exist. Conserved, yet distinct species-specific characteristics within the AFLP haplotype codes would suggest the possibility of AFLP analysis being used for the identification of an unknown microbial organism.

CHAPTER II

REVIEW OF LITERATURE

Recent events in the history of the United States have made American citizens aware of terrorism attacks directly affecting human life, health, and well-being. However, few understand the immense devastation that would accompany an attack on the American agricultural industry. The United States has been fortunate to escape a large scale bioterrorism and/or agroterrorism event as of yet. However, knowledge and preparedness must be in place in order to insure a successful investigation if an attack were to occur. Forensic scientists and professionals working in the agricultural sector must be aware of the potential for biocrimes targeting agriculture to occur and must follow specialized procedures and have effective tools available to investigate an incident, should it occur (Budowle, Murch, and Chakraborty, 2004). One approach to developing tools and procedures involves studying natural microbial outbreaks in hopes of identifying characteristics that would help distinguish a natural event from a biocrime. Similarly, having detailed molecular knowledge concerning the wide variety of pathogenic organisms that may be used in a bioterrorism attack is essential in microbial forensic preparedness.

Agriculture in the United States

The agricultural industry in the United States is a key component to the economic success of the country. Put into numbers, agriculture in the U.S. is one- sixth of the gross domestic product (GDP) of the country, which equates to over $1 trillion annually. The agricultural industry provides more jobs for American citizens than any other industry in the United States, a number that translates into one in every eight U.S. citizens being employed by the farming industry (Parker, 2003). Along with being a substantial source of employment, agricultural exports are also quite significant, totaling over $50 billion annually. These numbers, when combined, make the farming sector the single largest positive contributor to the U.S. economy (Parker, 2003). Due to the economic significance of the farming industry, an attack on the agricultural system could be devastating to the economy and also destructive of the public trust in the government's ability to protect a stable food source.

The vast size of the agricultural system of the United States directly relates to the economic success of the farming sector. However, the massive size also contributes directly to the government's inability to guarantee a protected food source. Crops, forests, and rangelands occupy an extensive part of the United States, covering over one billion acres (Fletcher, Bender, Budowle, Cobb, Gold, *et al*., 2006). Providing constant security surveillance to this large amount of land is an impossible task, leaving the farming lands of the United States completely unsecure the majority of the time. Additionally, a variety of potential targets that are vulnerable to a possible terrorist attack exist within the food production chain. Areas of the food production chain that are largely unsecure and open for an attack include, field crops, farm animals, food items in the

processing and distribution chain, market ready foods, and agricultural facilities such as processing plants (Parker, 2003). Because of the vulnerability of the U.S. farming industry as a whole, the federal government has recognized the need for better preparedness for the agricultural system. Former Health and Human Services Secretary, Tommy Thompson, stated, "I cannot understand why terrorists have not attacked our food supply because it is so easy to do" (Halbrook, 2006). Preparing for possible attacks on the agricultural sector must include learning from past biological outbreaks with development of technologies and processes to recognize and respond to future incidents in a timely manner.

Agricultural Bioterrorism

History of Agroterrorism

Agricultural bioterrorism is by no means a new problem on the world stage. In the 20th century alone, nine countries have had documented agricultural bioweapons programs. These countries include Canada, France, Germany, Iraq, Japan, South Africa, the United Kingdom, the United States, and the former USSR. Aside from these known programs, four other countries, Egypt, North Korea, Rhodesia, and Syria, are believed to have agricultural bioweapons programs (Monke, 2004). However, with the Biological and Toxin Weapons Convention of 1972, many countries, including the United States, stopped all military use of biological weapons and destroyed their stockpiles (Monke, 2004).

Although bioterrorism is not a newly emerging problem, the use of bioweapons to target food sources and agriculture has been relatively rare in modern history. In the 20th

century there were only 222 reported cases of bioterrorism or biocrimes, with only 24 of those being confirmed cases. The vulnerability of agricultural targets is illustrated by the fact that 22 of the 24 confirmed instances of bioterrorism or biocrimes directly targeted food or commercial animals and plants (Parker, 2003). Some of the more prominent attacks on agriculture have directly threatened both animal and human targets. In 1915, German intelligence used Anthrax and Glanders to infect U.S. and draft animals and livestock to cripple forces during World War I. Aside from cash crops and livestock, humans have also been the target of bioterrorism. In 1984, the Rajneeshee cult spread Salmonella in the salad bars of a small town in Oregon in an attempt to sway the outcome of a local election (Monke, 2004). The actions of the Rajneeshee cult are considered to be one of the most well-known bioterrorism events in American history. Although agricultural bioterrorism events have occurred in modern history, the events have been isolated, affecting a select, targeted population. While reviewing modern day terrorist attacks contributes to our understanding of the history of bioterrorism, these events do not reflect the immense scope and consequences that could result from a mass release of pathogenic bacteria targeting our agricultural enterprise.

Naturally occurring microbial outbreaks can provide some sense of the possible devastation that would follow a successful biological attack. The 2001 outbreak of Foot-and-Mouth Disease (FMD) in the United Kingdom is one example of a mass biological outbreak that had devastating effects on the agricultural system of a nation. What began with a single case of FMD in a pig resulted in the slaughter of over 6 million animals and economic losses exceeding $16 billion (Fletcher, Bender, Budowle, Cobb, Gold, *et al.,* 2006). While an attack on the United States livestock population would clearly devastate

the country's economy, investigators have concluded that an attack on American crops would have an even greater impact. Crops grown in the United States account for 54% of the value of American commodities and contribute to more exports then American livestock (Parker, 2003). A successful attack on the croplands would result in devastating consequences for the US economy and American trust in the government. By reviewing outbreaks such as the FMD incident in the United Kingdom, researchers can begin to fully comprehend the need for an efficient response plan and effective preparedness in the case of future attacks.

American Response and Preparedness

Awareness of bioterrorism in America has been intensely heightened since the anthrax attacks of 2001. However, the main public concern for bioterrorism has remained largely focused on human and economic protection, leaving threats to U.S. agriculture unpublicized by comparison. Although somewhat forgotten in post-9/11 reports on the state of preparedness for terrorist attacks, protecting U.S. agriculture became a focus in later government actions (Parker, 2003). Federal reports released in the years after 2001 began to emphasize the importance of research in preparing for and responding to bioterrorism events of the future (Monke, 2004). Reports such as the Congressional Research Service (CRS) Report for Congress, "Agroterrorism: Threats and Preparedness," have focused attention on protecting American agriculture. The CRS report explicitly points out that agriculture is an area of U.S. industry and economy that cannot continue to go unprotected (Monke, 2004).

Although recent progress has been made in microbial forensics, relatively little attention is paid to agroterrorism in comparison to terrorist threats that target livestock or humans. Henry Parker (2003) attributes the lack of attention to three main reasons: American's tendency to take food for granted, the decreasing national visibility of agriculture and food sources, and the limited public awareness of bioterrorism against agriculture. In addition to the lack of public awareness concerning food security, the federal government has also been slow to recognize bioterrorism threats towards agriculture. In the 9/11 Commission, a national report on terrorist attacks on the United States, no mention of agroterrorism or attacks on the U.S. food supply was made (Parker, 2003). However, with analysts at the local and national level recognizing that American food sources may be a viable and successful target for a possible terrorist attack, more attention is now focused on the area of preparedness (Monke, 2004).

Recently the federal government has enacted regulation and legislation concerning possible agroterrorism events in the U.S. (Monke, 2004, Parker, 2003). The first legislation concerning agroterrorism, *The Bioterrorism Preparedness and Response Act*, was enacted in 2002. Multiple provisions concerning agroterrorism were cited, including "Agricultural Bioterrorism Research and Development" (Parker, 2003). In 2003, the Senate Committee on Government Affairs held the first congressional hearing devoted entirely to agroterrorism entitled, "Agroterrorism: The Threat to America's Breadbasket." With enhanced focus on agricultural security, the USDA, FDA, and US Department of Homeland Security came together in 2009 to hold the first *International Symposium on Agroterrorism*. The Symposium brought together international officials to discuss agroterrorism preparedness, with a focus on future technologies and

methodologies related to food defense and an effort to support collaborations between security officials and academia in preparedness efforts (www.fbi-isa.org, 2011). A clear call for greater knowledge concerning biological threats led to the creation of the discipline of microbial forensics, a field dedicated to the isolation, identification, and differentiation of pathogens in the aiding of law enforcement.

Pseudomonas syringae

Pseudomonas syringae is a bacterial pathogen capable of infecting a variety of plants, including many cash crops critical to the U.S. farming industry. *P. syringae* is a gram-negative bacterium that consists of over 50 pathovars, which are defined by their host range (Joardar, 2005). The wide variety of host plants that can be infected by *P. syringae* pathovars is only heightened by the lack of current effective management techniques for controlling and eliminating disease outbreaks (Rudolph, 1997). *P. syringae* strains are found worldwide and are considered to be highly destructive pathogens. In addition, outbreaks are increasing in frequency as mutant strains become more prevalent and as management techniques continue to evolve at a slow pace (Rudolph, 1997). In addition to a wide host range, *P. syringae* also has a rapid outbreak and epidemic pattern with both cash crops and mammalian hosts. These two factors, the wide host range and rapid outbreak pattern, give *P. syringae* the reputation of being a highly pathogenic and destructive microorganism.

According to the CRS report on agroterrorism, a successful bioweapon should be infectious against a wide variety of both plant and animal hosts, should be able to survive in the environment, and should be both contagious and virulent (Monke, 2004). *P.*

syringae species encompass all of these factors. As discussed earlier, *P. syringae* pathogens infect a wide variety of plant species and show a destructive, virulent pattern that is often uncontrollable. Many *P. syringae* strains have been identified as being resistant to bactericides and antibiotics, meaning that prevention and treatment may be unavailable in certain disease outbreaks (Rudolph, 1997).

Microbial Forensics

The discipline of forensic microbiology is relatively new within all of the fields of forensic science and has developed in large part as a response to increased bioterrorism threats and actions against the United States in recent years. Microbial forensics is defined by Craig Cummings and David Relman (2002) as "the detection of reliably measured molecular variations between related microbial strains and their use to infer the origin, relationships, or transmission route of a particular isolate." Although microbial forensics has remained largely focused on studying those pathogens that threaten human health or life, the field covers the entire range of microbes that could be used to do harm. Therefore, agroterrorism threats also fall under the umbrella of microbial forensics.

Being able to quickly and efficiently determine the identity of a microbial agent used in a biocrime is critical to the success of the investigation. According to the National Research Council (NRC) report, *Countering Agricultural Bioterrorism*, "aggressive research in both science and technology is needed to improve our ability to prevent, detect, respond to, and recover from biological attacks on plants and animals" (2002). Although some research has been performed using genetic analysis to identify microbial strains, more research is needed to streamline and further validate the most

applicable techniques. When responding to bioterrorism attacks involving microbial organisms, investigators must be able to efficiently isolate and identify the strain in use (Cummings and Relman, 2002). Identification and attribution of a microbial strain are the goals in the investigation of a microbial outbreak, both intentional and those resulting from natural events (Fletcher, Bender, Budowle, Cobb, Gold, *et al*., 2006). Therefore, training lab personnel in methods used in forensic microbial investigations is a critical aspect of bioterrorism preparedness. Current research in the area of microbial identification indicates that traditional DNA typing strategies such as restriction fragment length polymorphism (RFLP) analysis may not be as strong for pathogen identification and attribution as newer PCR based methods such as amplified fragment length polymorphism (AFLP) analysis. It is a fact that the processing of microbial forensic samples may become the responsibility of crime labs. Therefore, PCR and capillary electrophoresis based technology used for the rapid characterization of pathogenic microbial organisms will potentially allow for the involvement of crime lab personnel in microbial forensic analysis.

Amplified Fragment Length Polymorphism (AFLP) Analysis

Amplified fragment length polymorphism (AFLP) analysis is a molecular technique that combines restriction fragment length polymorphism (RFLP) analysis with polymerase chain reaction (PCR) amplification in technique that creates a “fingerprint” of a genome (Lin, Kuo, and Ma, 1996; Janssen, 1996; Vos, *et al*., 1995). Previous studies utilizing AFLP analysis of *P. syringae* and other bacterial species have demonstrated the discriminatory power of the technique to distinguish even pathovars of the same species

(Clerc, Manceau, and Nesme, 1998; Georgnaras, Kenene, von Holy, and Hastings, 1999; Taylor, 2009). AFLP analysis techniques have also been utilized to effectively distinguish closely related strains of *Serratia marcescens, Pseudomonas aeruginosa, Staphylococcus aureus,* and *Bacillus anthracis*. (Allen, 2006; Beauman, 2007; Taylor, 2009; Jackson, Hill, Laker, Ticknor, and Keim, 1999). The ability of AFLP analysis to distinguish among closely related strains of the same bacterial species underscores the high discriminatory power of the method.

In addition to differentiating between closely related strains of the same bacterial species, research has suggested that AFLP analysis may offer the ability to specifically identify different bacterial species (Taylor, 2009). For example, AFLP analysis of *P. syringae* strains has demonstrated that elements of the AFLP profile are conserved among related pathovars. These common elements of the AFLP profile may thus be useful to specifically identify an unknown bacterium as *P. syringae* (Taylor, 2007).

Summary

A major part of preparing for agroterrorism and biowarfare events is being able to quickly and efficiently identify biological agents. While regulations are in place to track sources of biological agents kept in labs across the country, one central method for identifying and databasing biological agents does not currently exist (Cummings and Relman, 2002). One goal of bioterrorism research is to formalize methodology that can be employed in the event of an agroterrorism attack to identify the biological agent in use. The recent need for the ability to identify and trace biological agents has heightened

the development of the field of microbial forensics and put more focus on molecular identification techniques.

Although AFLP analysis has been used for the genetic fingerprinting of many microbial species, the reliability and suitability of this technique as a validated forensic tool is largely unknown and is a central theme of the study presented here using *P. syringae* as a model pathogen. The main goal of this research was to use AFLP analysis for the characterization of genomic DNA isolated from a variety of strains of *P. syringae.* Specific goals of the study were to assess the discriminatory power of AFLP analysis and different pathovars of *P. syringae* and to identify elements of the AFLP profile that are conserved among pathovars, species and even the genus *Pseudomonas,* if possible.

CHAPTER III

METHODOLOGY

Overview of the AFLP Analysis Method

The importance of being able to rapidly and specifically identify and attribute a microorganism has been underscored in recent natural outbreaks of disease throughout the world's agricultural systems. Although the need for efficient molecular tools in microbial forensics is evident, the development of such tools remains an evolving process. As the field of microbial forensics continues to develop, advances in identification technologies must be made available to labs charged with identifying and characterizing pathogens. Epidemiologists and forensic lab personnel must be equipped with the knowledge and tools to identify and attribute microbial agents in the event of a biocrime or a bioterrorism attack.

AFLP is a molecular tool that could provide rapid differentiation and specific characterization of microbial agents utilized in an agroterrorism event. AFLP analysis is a DNA fingerprinting technique that combines RFLP variability in the genome with PCR to produce amplified restriction fragment patterns of genomic DNA that are highly individualized for a particular sample (Vos, *et al*., 1995).

The AFLP process used here begins by digesting the genomic DNA extracted from a bacterial strain with two restriction enzymes, *EcoR1* and *Mse1*. Once digestion is complete, oligonucleotide linkers are ligated to the ends of the restriction fragments that will ultimately serve as targets for PCR primers used to amplify the DNA. Two rounds of PCR amplification, preselective and selective, are then carried out to reduce the number of restriction fragments constituting the genetic fingerprint of the sample. Genetic analysis produces a set of electropherograms for each strain, which are converted into a haplotype code for use in differentiation and characterization of a specific bacterial species and/or strain. The purpose of this study was to evaluate the ability of AFLP analysis to specifically identify and differentiate bacteria in the *Psedomonas* genus.

Bacterial Strains

Nine cultures of *P. syringae* were graciously provided by Jacqueline Fletcher from the Oklahoma State University Department of Entomology and Plant Pathology in Stillwater, Oklahoma. In addition, four cultures of *P. aeruginosa* were kindly provided by Dr. Frank Champlin at Oklahoma State University Center for Health Sciences in Tulsa, Oklahoma. Bacterial strains used in this study are listed in Table 1. In some cases, AFLP profiles from *P. syringe* pv. tomato (provided by Andrew Taylor), *Serratia marcescens* (provided by Charlene Beauman), and *Staphylococcus aureus* (provided by Dr. Robert Allen) were included in the study for comparison purposes.

Each bacterial strain was plated on a separate Mueller-Hinton agar plate and incubated until visual colonies formed. *P. syringae* strains were incubated at room temperature, whereas the *P. aeruginosa* strains were incubated at 37° C. Once bacterial

colonies were apparent, a 2 mL aliquot of Mueller-Hinton growth medium was inoculated with a single colony from each Petri dish. Inoculants were incubated overnight with shaking, at room temperature for all *P. syringae* strains and at 37° C for all *P. aeruginosa* strains. Each of the bacterial cultures was centrifuged at 10,000 xg for three minutes at room temperature to obtain a cell pellet for DNA extraction.

DNA Isolation

Isolation of bacterial DNA began with re-suspension of the cell pellet in 250 µL TNE (10 mM TRIS-Cl pH8.0 + 0.2 M NaCl and 1 mM EDTA) with 15 mg/mL lysozyme, followed by incubation at 37° C for 30 minutes to weaken the cell walls. Once lysozyme digestion was complete, 250 µL TNE with 2% sodium dodecyl sulfate (SDS), 40 µL proteinase K (20 mg/mL in 10 mM TRIS-Cl ph 8.0 + 0.2 M KCl & 50% v/v glycerol), and 2 mg ribonuclease A was added to each sample. The samples were allowed to incubate at 65° C for one hour. After the second incubation was complete, the samples were extracted using an equal volume of phenol:chloroform/isoamyl alcohol (9:0.96:0.04 v/v). Samples were centrifuged at 10,000 xg for two minutes to induce phase separation, with the bacterial DNA located in the aqueous (top) layer. For each sample, the aqueous layer was then removed and placed in a clean 1.8 mL microfuge tube. The samples were extracted a second time with an equal volume of chloroform:isoamyl alcohol (24:1 v/v) and centrifuged at 10,000 xg for two minutes to obtain phase separation. Once again, the aqueous layer was removed to clean 1.8 mL microfuge tubes. Two volumes of 95% ethanol were added to each sample to precipitate the bacterial DNA. Using a sterile inoculation loop, the clot of DNA was retrieved from

each sample and resuspended in a 200 µL aliquot of TE^{-4}(10 mM TRIS-Cl pH 8.0 +0.1 mM EDTA) in a clean 1.8 mL microfuge tube. The isolated DNA samples were stored at 4°C, if not immediately used.

Table 1: Species and strains of *P. syringae* and *P.aeruginosa* used in this study.

Species	Strain
Pseudomonas syringae	F15
Pseudomonas syringae	F18
Pseudomonas syringae	F7
Pseudomonas syringae	F10A
Pseudomonas syringae	F12
Pseudomonas syringae	NF3
Pseudomonas syringae	NF3A
Pseudomonas syringae	NF5
Pseudomonas syringae	FF5
Pseudomonas syringae	F22
Pseudomonas aeruginosa	PA01
Pseudomonas aeruginosa	1211
Pseudomonas aeruginosa	ATCC 10145
Pseudomonas aeruginosa	ATCC 27853

DNA Quantification

The concentration and purity of DNA isolated from the bacterial strains were measured using a Nanodrop ND-1000 microspectrophotometer (Nanodrop Technologies

Inc., Rockland, DE). A calculated 260:280 ratio of 1.8 indicates reasonably pure DNA and was considered acceptable DNA purity for this study. Based on the calculated concentration of DNA using A260 absorption, samples were diluted using TE^{-4} to a final concentration of 300-500 ng/μL of genomic DNA. Agarose gel electrophoresis of genomic DNA showed it to be largely intact and high molecular weight (i.e. > 23 kb) (not shown).

DNA Digestion

DNA from the different bacterial strains was digested with two restriction enzymes, *EcoR1* and *Mse1*. An aliquot of 500 ng of isolated DNA was placed into a clean 1.8 mL mircrofuge tube. In addition, 1.0 μL (10 units) *Mse1* (New England Biolabs, Inc., MA), 1.0 μL 10x NEB Buffer 4 (New England Biolabs, Inc., MA), and enough dH_2O to make a final volume of 10 μL were added to the DNA in each sample. The samples were incubated at 37^0C for 1 hour, followed by a 5 minute incubation at 65^0C to kill enzyme activity. The samples were then put on ice for 5 minutes to insure that the reaction had fully stopped. To 10 μL of the first *Mse1* digest, 1 μL containing 20 U of *EcoR1* (New England Biolabs, Inc.), 2 μL of 10x *EcoR1* buffer (New England Biolabs, Inc.), and 7 μL of dH_2O were added. The samples were once again incubated at 37^0C for 1 hour, followed by a 5 minute incubation at 65^0C. The samples were placed on ice for 5 minutes to insure that the reactions had completely stopped.

To determine the extent of the digestion, an aliquot of each digest was electrophoresed on a1% agarose gel equilibrated in TAE buffer (10 mM Tris-acetate pH 8.3 with 1 mM EDTA). DNA samples were prepared for electrophoresis by mixing 7 μL

of DNA digest with 3 μL of loading dye (5X TAE with 1% ficoll 400 and 0.05% (w/v) each of xylene cyanol & bromphenol blue). The DNA was loaded into the agarose gel and allowed to electrophorese at 75 V until the tracking dye entered the gel, then at 100 V for approximately 45 minutes. Once electrophoresis was complete, the digest was visualized using ethidium bromide staining and a UV light box. Successful DNA digestion was indicated by a smear of restriction fragments in the gel lane (Figure 1). Once complete digestion was confirmed, 6 μL of TE^{-4} was added to 4 μL of digested DNA. If not used immediately, the DNA samples were stored at 4^0C.

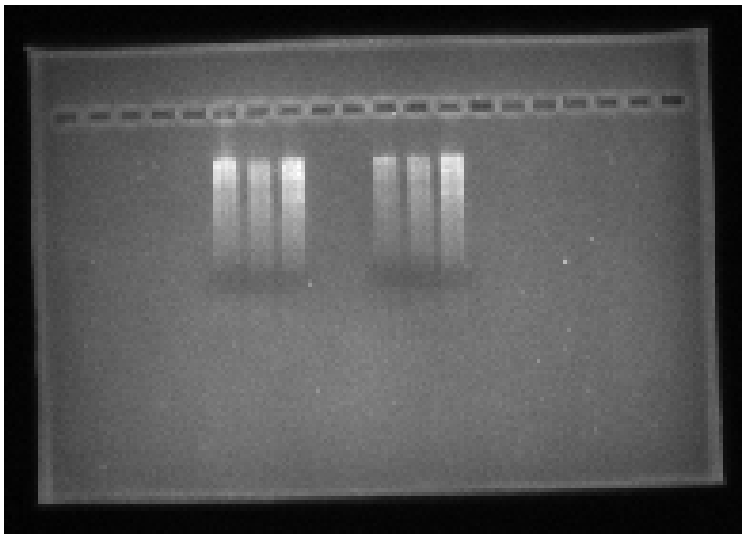

Figure 1: Ethidium bromide stained gel of digested *P. syringae* DNA. At this point in the AFLP process, the microbial DNA samples have been digested with the restriction enzymes, *EcoR1* and *Mse1* as described. From left to right the *P. syringae* strains (in duplicate): F12, F22, FF5, Empty Lane, F12, F22, FF5.

DNA Ligation

Once DNA digestion was confirmed, oligonucleotide adaptor pairs provided with the microbial AFLP genotyping kit (Applied Biosystems Inc., Foster City, CA) were ligated onto the ends of the restriction fragments following instructions provided with the kit. The adaptor pairs represent known DNA sequences that provide target sites

complementary to the PCR primers used in both the preselective and selective PCR steps and also complementary to the sticky ends of the *Mse1* and *EcoR1* fragments produced during restriction digestion. Ligation reactions were created using an enzyme master mix that was prepared by mixing 1 μL T4 DNA Ligase Buffer (Applied Biosystems Inc., Foster City, CA), 1 μL 0.5M NaCl, 1 μL *Mse1*, 2 μL *EcoR1*, 0.5 μL T4 DNA Ligase (Applied Biosystems Inc., Foster City, CA), and 4.5 μL dH_2O. Ligation reactions contained approximately 10-30 ng of DNA (1 μL of diluted digested DNA) with 1 μL T4 DNA Ligase Buffer, 1 μL 0.5M NaCl, 0.5 μL bovine serum albumin (1 mg/mL), 1 μL *Mse1* adapter (supplied with the AFLP kit), 1 μL *EcoR1* adapter (supplied with the AFLP kit), 1.0 μL enzyme master mix, and 4.5 μL dH_2O. The samples were thoroughly mixed and allowed to incubate in a thermocycler at 37^0C for two hours. After incubation was complete, 189 μL TE^{-4} was added to each sample. If not immediately used, the samples were stored at 4^0C.

Preselective Amplification

Preselective amplification is the first of two rounds of PCR in the AFLP technique. Preselective amplification nonspecifically amplifies all DNA fragments that have successfully undergone digestion and ligation to the *EcoR1* and/or *Mse1* oligonucleotide adaptor sequences. Two primers are utilized, an *EcoR1* preselective primer, which targets the DNA fragment ends created by ligation of the *EcoR1* adapter, and an *Mse1* preselective primer, which targets the DNA fragment ends created by ligation of the *Mse1* adapter. It should be noted that because *Mse1* recognizes a 4 basepair sequence and *EcoR1* recognizes a 6 basepair sequence, most restriction

fragments will have *Mse1* sites on each end, many of the restriction fragments produced will have an *EcoR1* site on one end and an *Mse1* site on the other, and few fragments with have *EcoR1* sites on each end. Ultimately, only fragments with *EcoR1* sites on both ends (rare) or an *Mse1* site on one end and an *EcoR1* site on the other (much more common) will be detectable in the AFLP profiles produced. To set up the reactions, 4 μL diluted DNA (containing approximately 100 pg of ligated DNA) was mixed with 0.5 μL *EcoR1* preselective primer (Applied Biosystems Inc., Foster City, CA), 0.5 μL *Mse1* preselective primer (Applied Biosystems Inc., Foster City, CA), and 15 μL AFLP amplification core mix, containing PCR reactants and Taq polymerase (Applied Biosystems Inc., Foster City, CA). The samples were mixed, placed in a thermocycler, and amplified using PCR settings recommended in the instructions for the AFLP genotyping kit (Table 2).

Table 2: Thermocycler parameters for preselective amplification.

HOLD	**CYCLE**			**HOLD**
	Each of 20 Cycles			
72°C 2 min.	94°C 20 sec.	56°C 30 sec.	72°C 2 min.	4°C (forever)

Once amplification was complete, 10 μL of the product was mixed with 190 μL TE^{-4} and stored at 4^0C, if not immediately used. The other 10 μL of PCR product was mixed with 3 μL loading buffer and electrophoresed on a 1% agarose gel a 75 V for one hour to confirm that ligation and preselective PCR were both successful. If the reactions were successful, a smear was visualized on the gel when stained with ethiduim bromide and viewed under UV light (Figure 2).

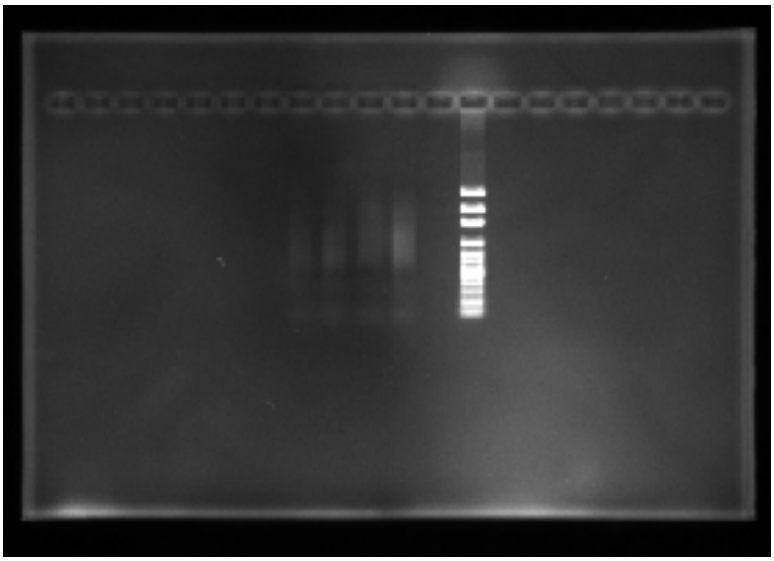

Figure 2: Ethidium bromide stained gel of successful *P. syringae* DNA ligation and preselective amplification. At this point in the AFLP method the microbial DNA samples have been successfully digested with *EcoR1* and *Mse1*, ligated to adaptor pairs, and non-selectively amplified. From left to right the strains are: NF3, NF3A, NF5, and NF12.

Selective Amplification

Selective amplification is the second of two rounds of PCR in the AFLP method (Vos, *et al.*, 1995). Selective amplification reduces the final number of DNA fragments that will be analyzed by utilizing a set of primers that target a subset of non-selectively amplified products that contain a complementary one to two nucleotides within the restriction fragment distal to the *Mse1* and/or *EcoR1* recognition site. The presence of the extra one to two nucleotides in the selective primers reduces the number of initial restriction fragments that will be amplified by a factor of as much as sixteen. In addition, the selective primer targeting the "*EcoR1* end" of the non-selective PCR products is labeled with one of three fluorescent dyes, depending on the selective nucleotide in the primer. The *Mse1* primer that was used in this study has an extra adenine (A) nucleotide at the 3' end of the primer, therefore only amplifying fragments that have the *Mse1* target sequence followed directly by a thymidine (T) nucleotide. Three different *EcoR1* primers

were used for selective amplification, each with a different nucleotide extension: adenosine (A), guanine (G), or cytosine (C). The *EcoR1* primers were coupled with a different fluorescent dye, *EcoR1*-A with FAM (blue), *EcoR1*-G with JOE (green), and *EcoR1*-C with NED (yellow). Three different selective PCR reactions were created for each DNA sample. Every sample was amplified with the *Mse1*-A primer plus one of the *EcoR1* primers, either A, G, or C, using the following cycling conditions (Table 3).

Table 3: Thermocycler parameters for selective amplification.

Hold	Cycle: Selective Amplification			Number of Cycles
94^0C 2 min	94^0C 20 sec	66^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	65^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	64^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	63^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	62^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	61^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	60^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	59^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	58^0C 30 sec	72^0C 2 min	2
	94^0C 20 sec	57^0C 30 sec	72^0C 2 min	2
94^0C 2 min	94^0C 20 sec	56^0C 30 sec	72^0C 2 min	20
60^0C 2 min				1
4^0C Forever				1

Selective PCR reactions contained 1.5 μL of diluted preselective amplification product was mixed with 7.5 μL AFLP core amplification mix, 0.5 μL *Mse1*-A primer, and 0.5 μL *EcoR1*-A, *EcoR1*-G, or *EcoR1*-C. The reactions were mixed, placed into a thermocylcer, and amplified using settings specified in the AFLP kit instructions (Table 3). Once amplification was complete, the samples were stored at 4^0C.

Capillary Electrophoresis

Capillary electrophoresis was performed using an ABI Prism 310 genetic analyzer (Applied Biosystems). The ABI Prism 310 genetic analyzer employs a capillary filled with a sieving polymer connected to high voltage source to conduct electrophoresis, which separates DNA fragments based on size. The DNA, which is labeled with fluorescent dyes, is detected by a laser at the end of the capillary.

To insure successful analysis, each of the three samples (FAM, JOE, and NED) for every DNA sample was prepared in a separate tube for electrophoresis. The Samples were prepared for electrophoresis by mixing 1 μL PCR product with 24.5 μL Hi-Di formamide and 0.5 μL of LIZ labeled size standard (Applied Biosystems Inc., Foster City, CA). Samples were placed in the genetic analyzer and electrophoresed for 24 minutes at 60^0C. Three electropherograms were thus produced for each sample, one for DNA fragments labeled with FAM (blue), one for DNA fragments labeled with JOE (green), and one for DNA fragments labeled with NED (yellow, appearing as black). An example of AFLP results can be seen in Figures 3A, 3B, and 3C, which display the three electropherograms for *P. syringae* strain F10A. The three electropherograms for each

strain constitute the AFLP profile for the sample which was used in analysis of each sample.

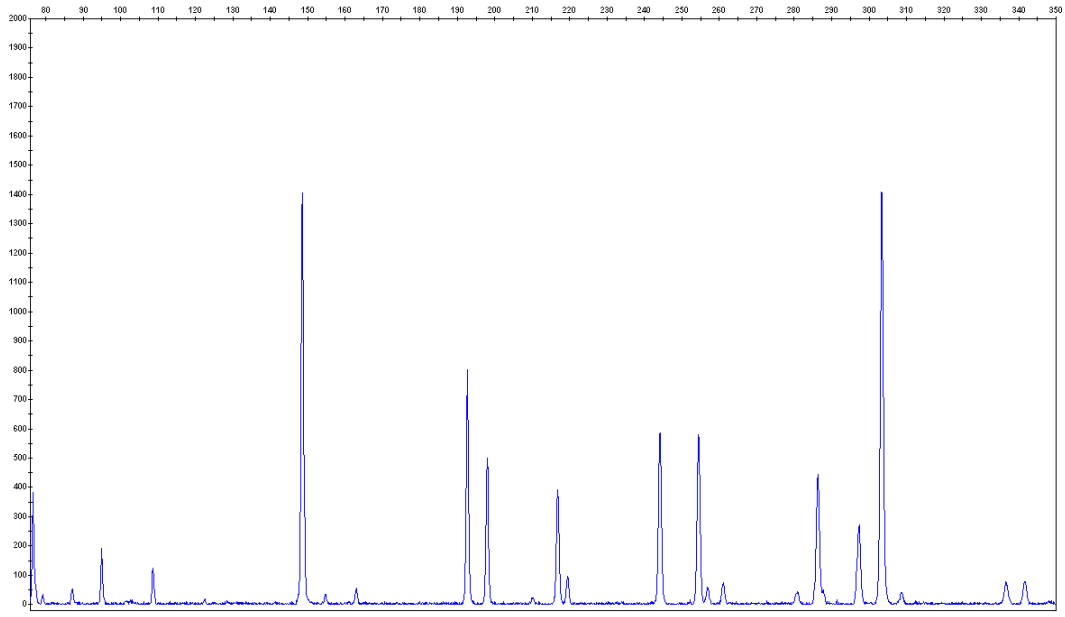

Figure 3A: *P. syrinage* pv. maculicola strain F10A FAM electropherogram. The Y-axis of the histogram represents relative fluorescent units (RFU) and the X-axis represents fragment size in basepairs. Each peak on the FAM histogram represents an amplified DNA restriction fragment.

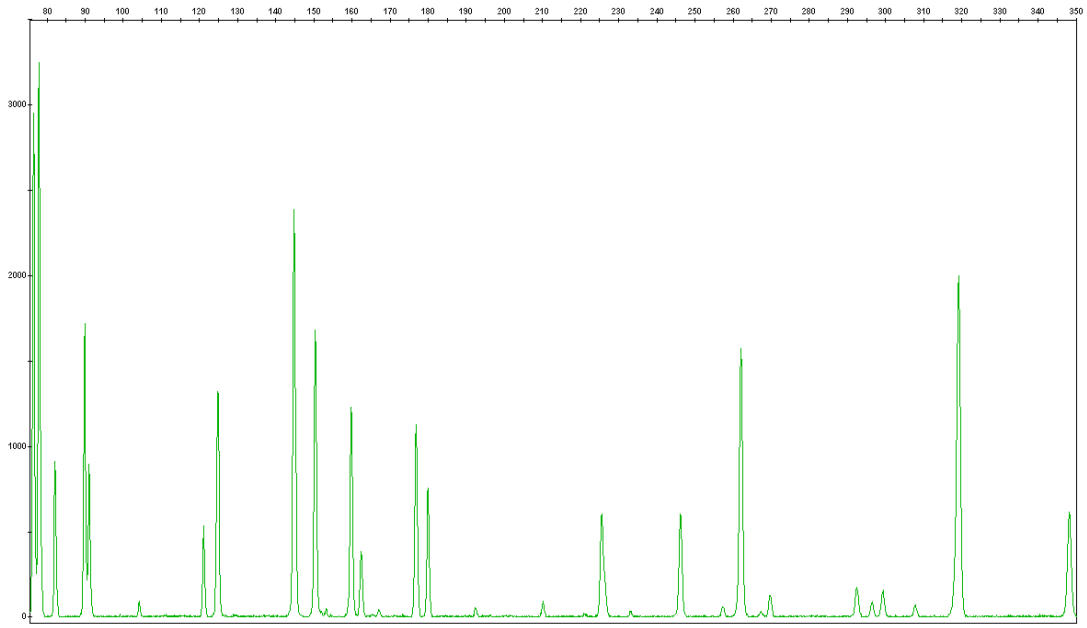

Figure 3B: *P. syrinage* pv. maculicola strain F10A JOE electropherogram. The Y-axis of the histogram represents relative fluorescent units (RFU) and the X-axis represents fragment size in basepairs. Each peak on the histogram represents an amplified DNA restriction fragment.

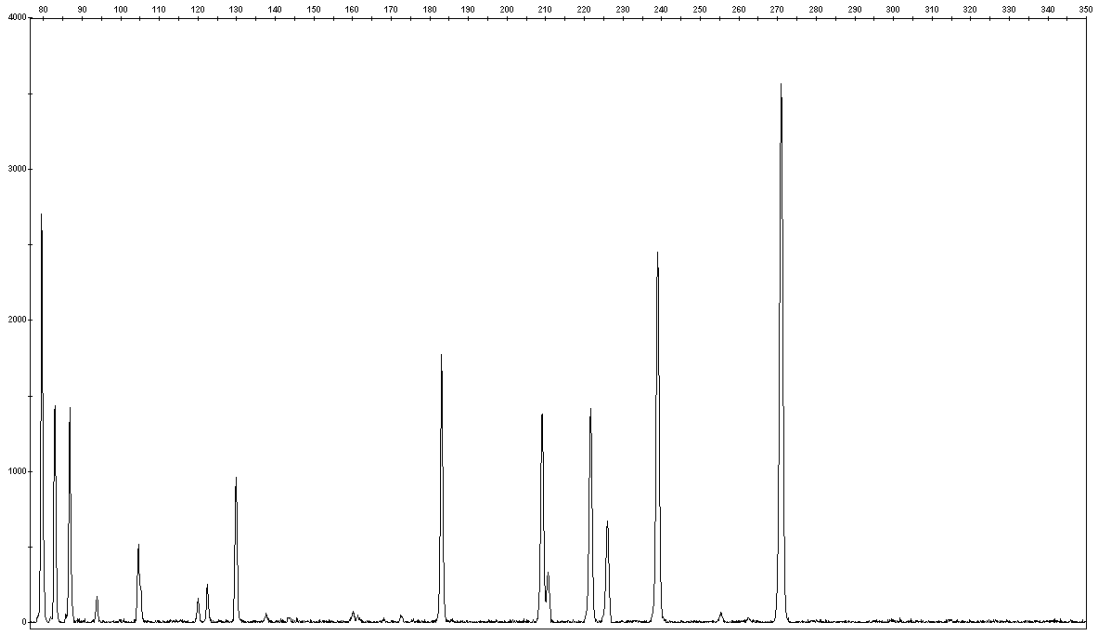

Figure 3C: *P. syrinage* pv. maculicola strain F10A NED electropherogram. The Y-axis of the histogram represents relative fluorescent units (RFU) and the X-axis represents fragment size in basepairs. Each peak on the histogram represents an amplified DNA restriction fragment.

Analysis

AFLP data analysis was performed by evaluating each of the electropherograms for each sample and converting the histogram of fluorescent peaks into a numerical code. Each of the electropherograms consists of a plot of relative fluorescence (RFU, Y-axis) versus fragment size (Basepairs, X-axis). The electropherograms for each sample differ from one another because alterations in the genome of each bacterial strain alter the spatial arrangement of *EcoR1* and *Mse1* restriction sites which determines the characteristics of the AFLP profile produced. Interpreting the unique pattern of fragment sizes corresponding to each bacterial sample allows for the specific characterization and differentiation of each strain.

The numerical haplotype code assigned to each strain was created by encoding the fragments in each electropherogram that fell within the size range of 75 to 350 basepairs. The resulting 275 basepairs of size range was subdivided using a binning system consisting of 28 bins, each spanning 10 nucleotides in size. An analytical threshold of the average relative fluorescent units (RFU) of fluorescent peaks in the analyzed size range divided by two was set to enhance reproducibility and to normalize electrophoretic runs. Only peaks above the set threshold of each electropherogram were encoded in the haplotype code. Final haplotype codes assigned to each bacterial strain consisted of three parts, one code for each of the FAM, JOE, and NED electropherograms produced from each strain.

CHAPTER IV

FINDINGS

A validated forensic microbial DNA typing technique must be reproducible, highly discriminatory, and ideally produce portable results that can be easily shared among laboratories. This study was designed to investigate these areas with regards to AFLP analysis and the results demonstrate AFLP analysis to be proficient in these areas. Additionally, this study addressed the issue of bacterial mutation and the effect that changes in the genome may have on final AFLP results.

Reproducibility

Establishing the reproducibility of an assay is essential for the successful application of any laboratory technique. If differences are observed between results obtained from two samples, an analyst must be confident that the differences are real and due to inherent differences in the samples. To establish AFLP reproducibility, the analysis was performed in triplicate for each of the *P. syringae* pv. maculicola strains and for each of the *P. aeruginosa* strains. The results of the three analyses for each strain allowed the reproducibility of the AFLP assay to be determined. Reproducibility was determined based on the location of peaks (representing DNA restriction fragment size)

observed in the electropherograms for each strain. The reproducibility between each of the three electropherograms for a strain can be easily observed upon visual observation of peak location. As is evident from Figures 4A through 4I, the three FAM, three JOE, and three NED electropherograms for *P. syringae* pv. maculicola strain NF12 appear consistent with one another when peak location is visually compared. Although the relative fluorescence of the peaks does change in each individual run, the peaks that are present for this strain are consistent throughout each of the three electropherograms for each color channel.

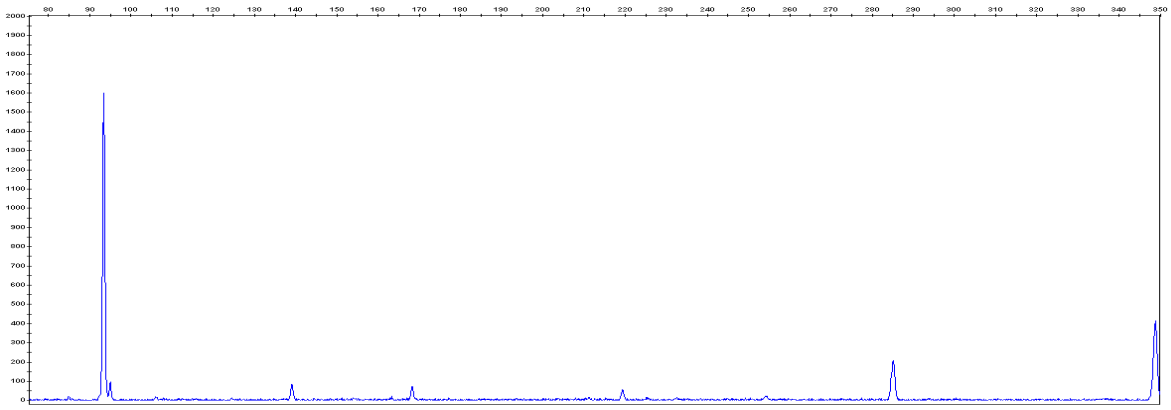

Figure 4A: *P. syringae* strain F12 FAM electropherogram. When compared to other strain F12 FAM electropherograms (Figure 4B and Figure 4C) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

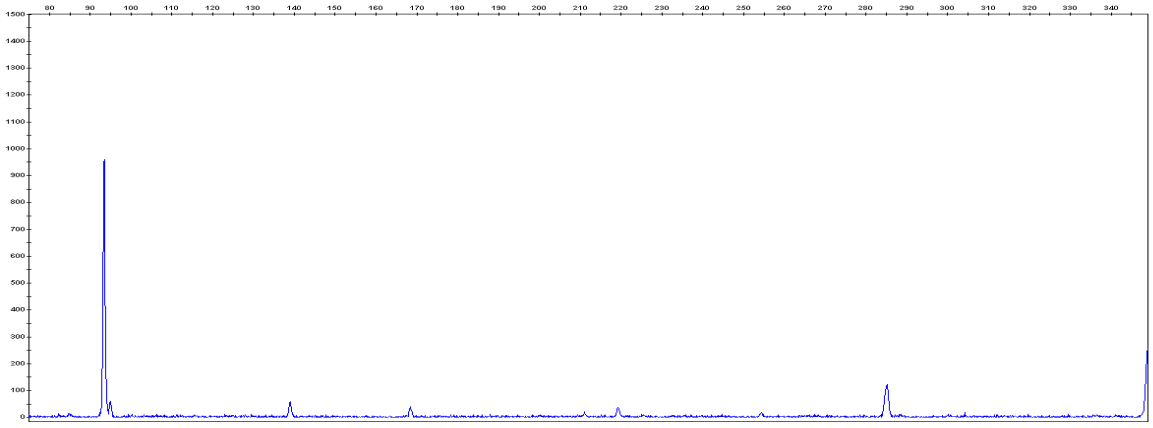

Figure 4B: *P. syringae* strain F12 FAM electropherogram. When compared to other strain F12 FAM electropherograms (Figure 4A and Figure 4C) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

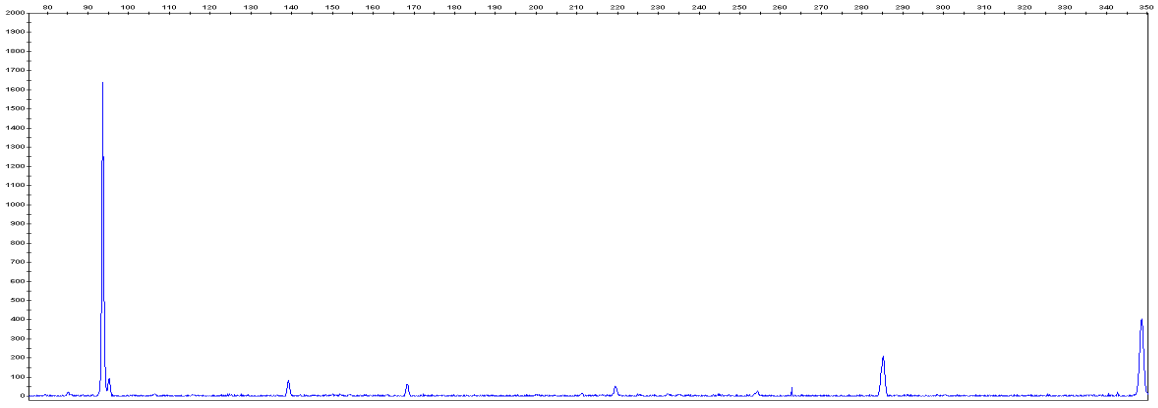

Figure 4C: *P. syringae* strain F12 FAM electropherogram. When compared to other strain F12 FAM electropherograms (Figure 4A and Figure 4B) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

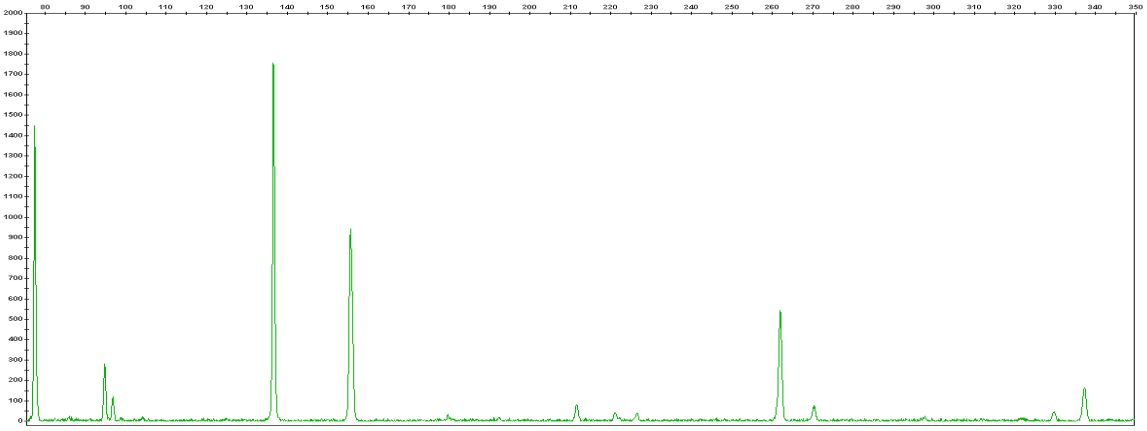

Figure 4D: *P. syringae* strain F12 JOE electropherogram. When compared to other strain F12 JOE electropherograms (Figure 4E and Figure 4F) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

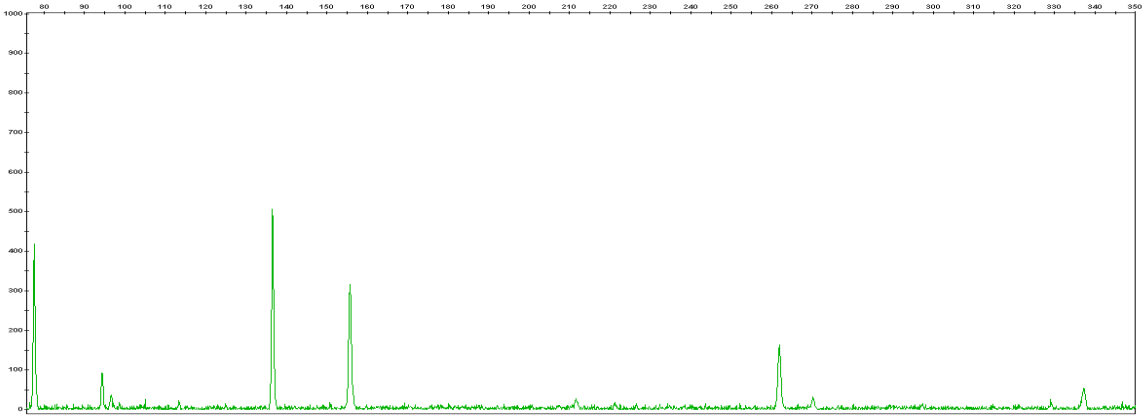

Figure 4E: *P. syringae* strain F12 JOE electropherogram. When compared to other strain F12 JOE electropherograms (Figure 4D and Figure 4F) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

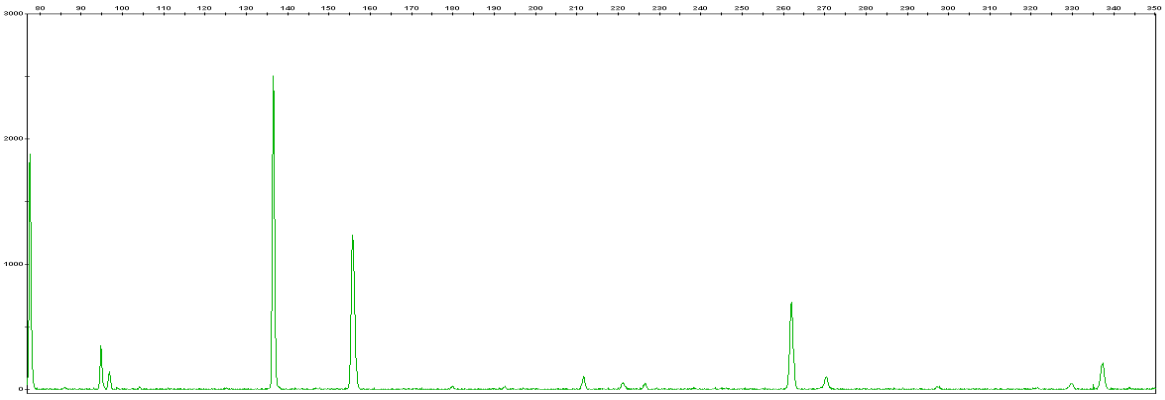

Figure 4F: *P. syringae* strain F12 JOE electropherogram. When compared to other strain F12 JOE electropherograms (Figure 4D and Figure 4E) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

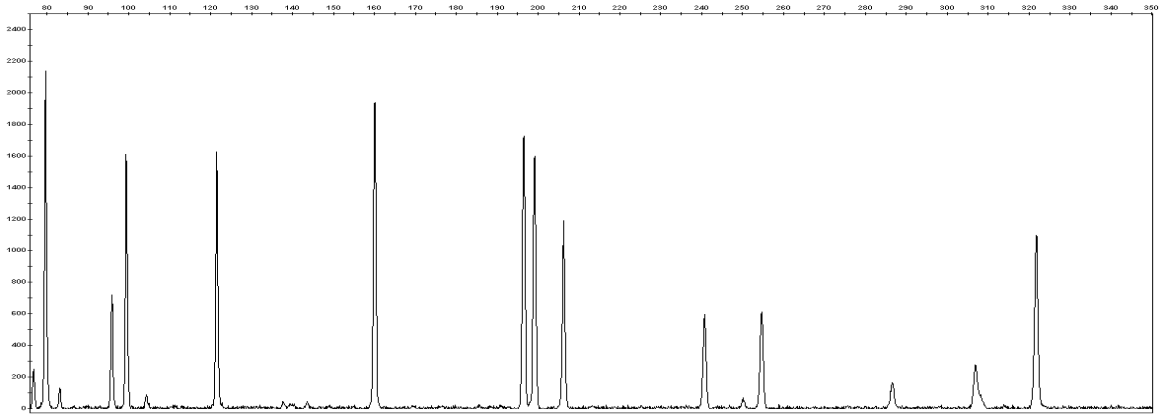

Figure 4G: *P. syringae* strain F12 NED electropherogram. When compared to other strain F12 NED electropherograms (Figure 4H and Figure 4I) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

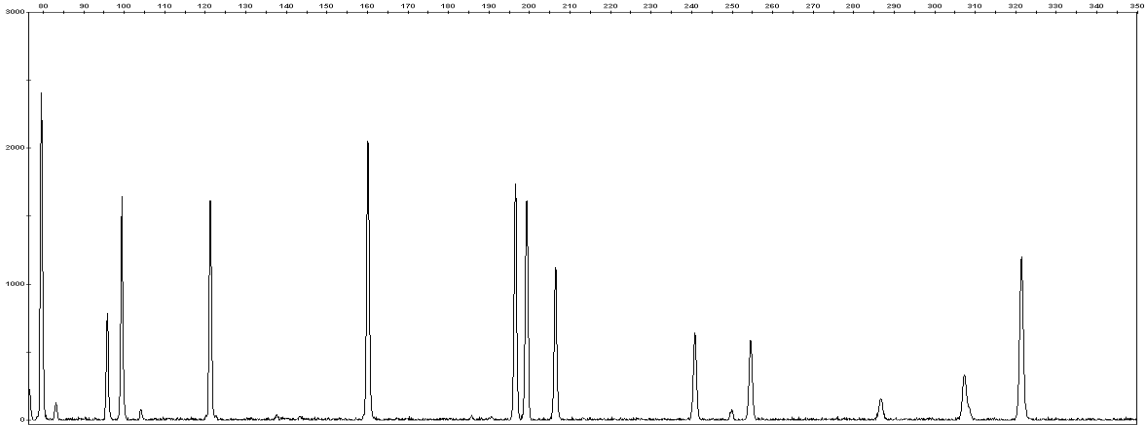

Figure 4H: *P. syringae* strain F12 NED electropherogram. When compared to other strain F12 NED electropherograms (Figure 4G and Figure 4I) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

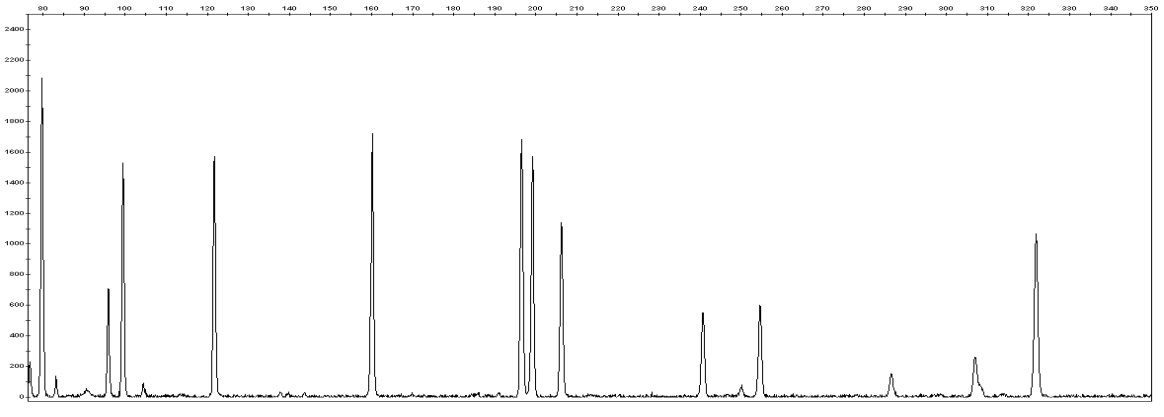

Figure 4I: *P. syringae* strain F12 NED electropherogram. When compared to other strain F12 NED electropherograms (Figure 4G and Figure 4H) the histogram peaks appear to fall in the same locations indicating the reproducibility of the AFLP assay.

One way to capture the molecular characterization of an AFLP profile is to convert peak locations (in terms of size) into a haplotype code (Beauman, 2006; Taylor 2009). A haplotype code was produced for each electropherogram (FAM, JOE, and NED) generated from the analysis of each individual strain. Each individual haplotype code reflects the number of peaks located in the electropherogram within the size range of 75 to 350 basepairs. This basepair size range was subdivided into 28 bins, each spanning 10 basepairs in size. The AFLP analysis method is not concerned with the fluorescent peaks that are weakly fluorescent. These minor components of the profile are eliminated by incorporating an analytical threshold defined as the average fluorescence of all peaks in a profile divided by two. Applying the threshold also greatly enhanced the reproducibility of the assay and normalized results from different runs. Thus, any peak that had a relative fluorescent value greater than the set threshold was included in the haplotype code. This concept is illustrated in Table 4 for results from *P. syringae* pv. maculicola strain F7. The first run is missing a peak at 104 basepairs and at 257 basepairs, both of which are present in the second and third runs. The second and third runs are missing a peak at 290 basepairs that is present in the first run. However, because all of these peaks fall below the set threshholds of the individual runs, the missing peaks do not affect the final JOE haplotype codes and therefore do not affect the final reproducibility for this strain.

Reproducibility was mathematically calculated using the haplotype codes produced from each of the triplicate runs for each strain. When calculating reproducibility, the following formula was used:

$$(Number\ of\ Matching\ Bins \div Total\ Number\ of\ Bins) \times 100 = \%\ of\ Reproducibility$$

The total number of bins was always 28, as there are 28 bins possible in every haplotype code for each individual color channel (FAM, JOE, and NED). The number of matching bins was determined by counting how many bins matched the bins of the code with the largest number of peaks. The reproducibility was first determined for each strain within each color channel. The average of all reproducibility values for all of the strains in each color channel was then calculated to determine an overall reproducibility value for each species.

Table 4: Electrophoretic peak values (Size, Basepairs and Height, RFU) from the *P. syringae* strain F7 JOE electropherogram from each of the three runs.

P. syringae Strain F7					
Run #1		**Run #2**		**Run #3**	
Size (Basepairs)	**Height (RFU)**	**Size (Basepairs)**	**Height (RFU)**	**Size (Basepairs)**	**Height (RFU)**
76.89	1164	77.82	2244	77.91	2137
81.18	341	82.04	695	82.16	652
88.78	676	89.84	1366	89.97	1267
89.98	358	91.03	721	91.09	666
		104.1	70	104.26	68
120.42	191	120.99	381	121.39	367
124.19	526	124.8	1130	125.19	1059
144.23	905	144.83	1745	144.96	1580
149.88	270	150.38	520	150.51	471
159.08	450	159.75	968	159.75	909
162.73	171	162.46	330	162.63	304
179.02	334	179.89	666	179.89	608
225.88	73	226.35	153	226.43	141
246.37	464	246.16	918	246.49	831
		257.28	57	257.64	50
261.83	609	262.02	1173	262.27	1084
268.78	56	269.85	98	269.76	83
290.12	71				
291.43	74	291.11	265	291.29	210
318.93	196	319.18	371	319.16	336
Threshold=	**192.47**	**Threshold=**	**365.03**	**Threshold=**	**337.45**

The lowest reproducibility value calculated for a single strain was 95.6 % for *P. syringae* strain F7. The highest reproducibility value calculated for a single strain was 100.0% for *P. syringae* strain F22. In addition, the NED electropherograms showed the highest overall reproducibility with seven of the ten *P. syringae* strains having a calculated reproducibility value of 100.0%. The overall reproducibility for *P. syringae* pv. maculicola was 97.9% and the overall reproducibility for *P. aeruginosa* was 97.5%. Previous studies utilizing AFLP analysis (Beauman, 2007 and Taylor, 2009) found the assay to be only slightly less reproducible. The results of this study therefore demonstrate that the AFLP assay is reliably reproducible.

Discriminatory Capability

A validated DNA typing tool must demonstrate a high discriminatory capability at the bacterial species and strain level. During an investigation, identifying the exact species and strain of the pathogen at hand can expedite the response efforts and aid in attribution of the microbe to a suspected source. This study demonstrated that AFLP analysis is capable of discriminating between pathogenic species as well as closely related strains of *P. syringa*e and *P. aeruginosa*. All ten strains of *P. syringae* and all four strains of *P. aeruginosa* generated distinct electropherograms and distinct haplotype codes. Visually, the electropherograms for each bacterial strain could be distinguished from one another. As depicted in Figures 5A through 5F, *P. syringae* strains F7 and F10A appear to have numerous peaks at similar locations in the electropherograms of AFLP products in each of the three color channels, yet clear differences in the FAM, JOE, and NED electropherograms for each strain also exist.

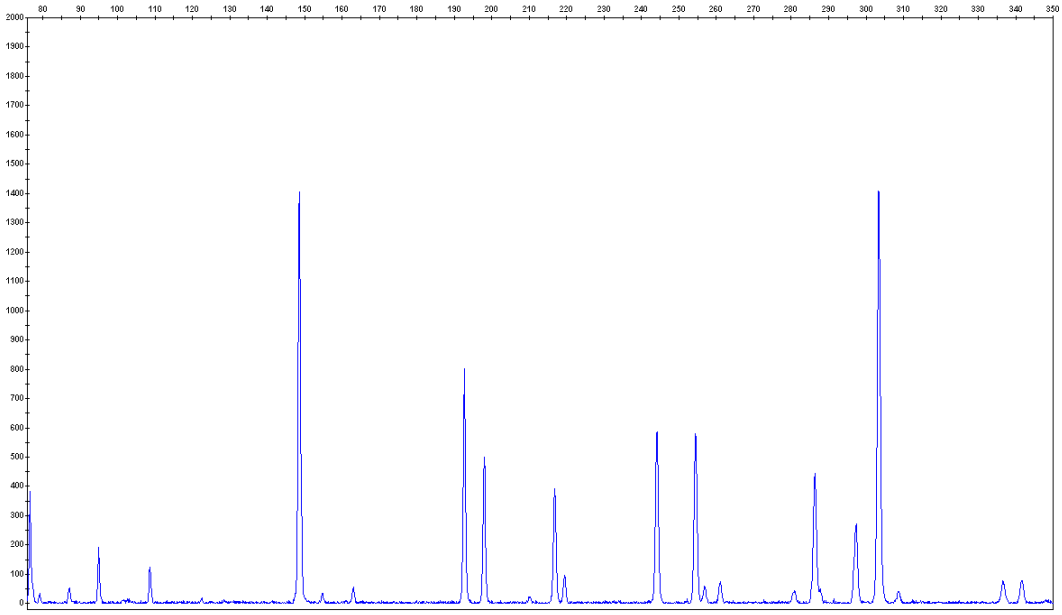

Figure 5A: *P. syrinage* strain F10A FAM electropherogram. When compared to the *P. syrinage* strain F7 FAM electropherogram (Figure 5B), similarities can be observed, however, differences clearly exist that make the AFLP profiles of the two strains unique from one another.

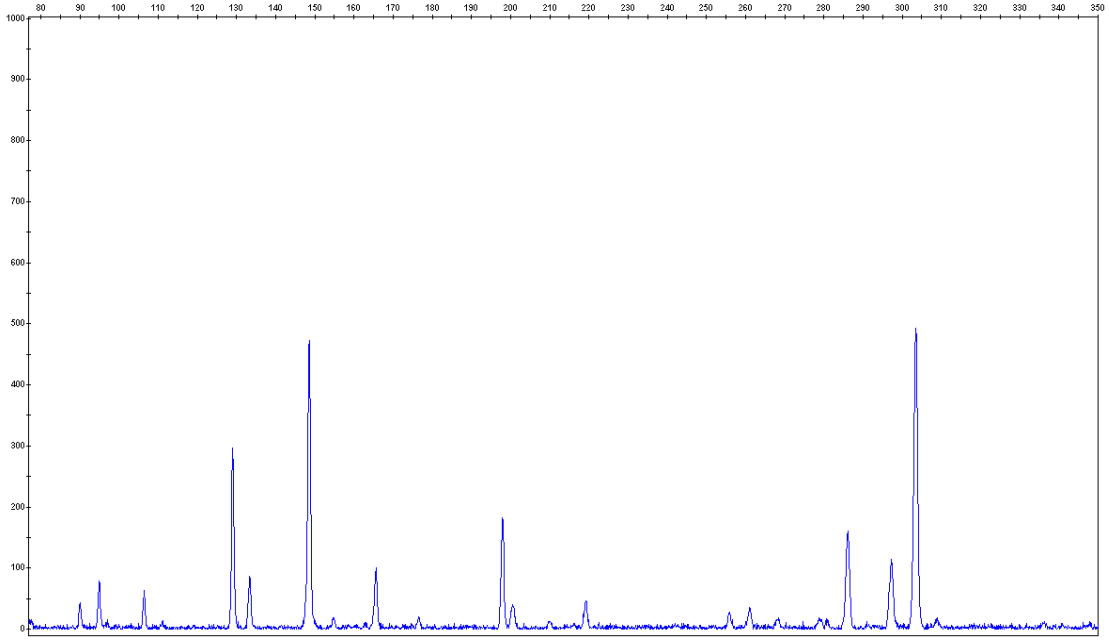

Figure 5B: *P. syrinage* strain F7 FAM electropherogram. When compared to the *P. syrinage* strain F10A FAM electropherogram (Figure 5A), similarities can be observed, however, differences clearly exist that make the AFLP profiles of the two strains unique from one another.

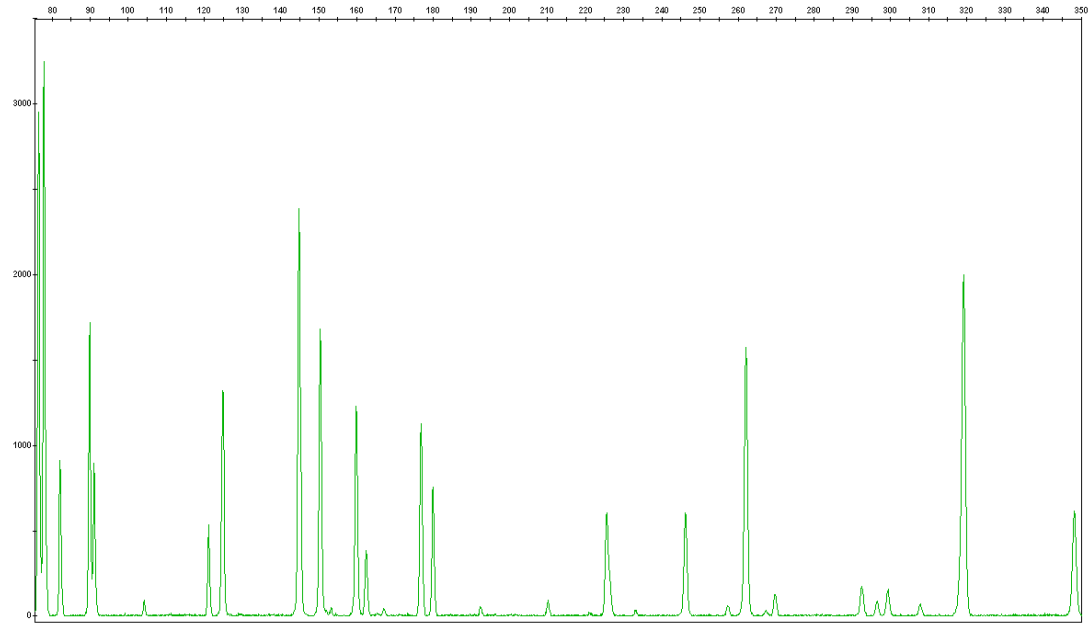

Figure 5C: *P. syringae* strain F10A JOE electropherogram. When compared to the *P. syringae* strain F7 JOE electropherogram (Figure 5D), similarities can be observed, however, differences clearly exist that make the AFLP profiles of the two strains unique from one another.

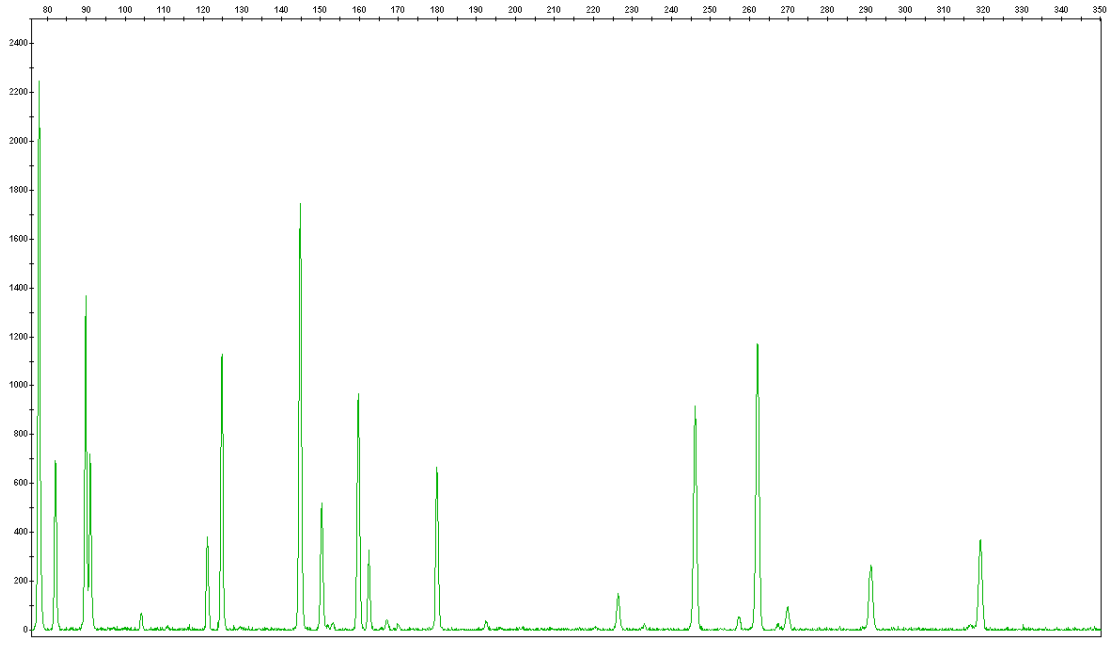

Figure 5D: *P. syringae* strain F7 JOE electropherogram. When compared to the *P. syringae* strain F10A JOE electropherogram (Figure 5C), similarities can be observed, however, differences clearly exist that make the AFLP profiles of the two strains unique from one another.

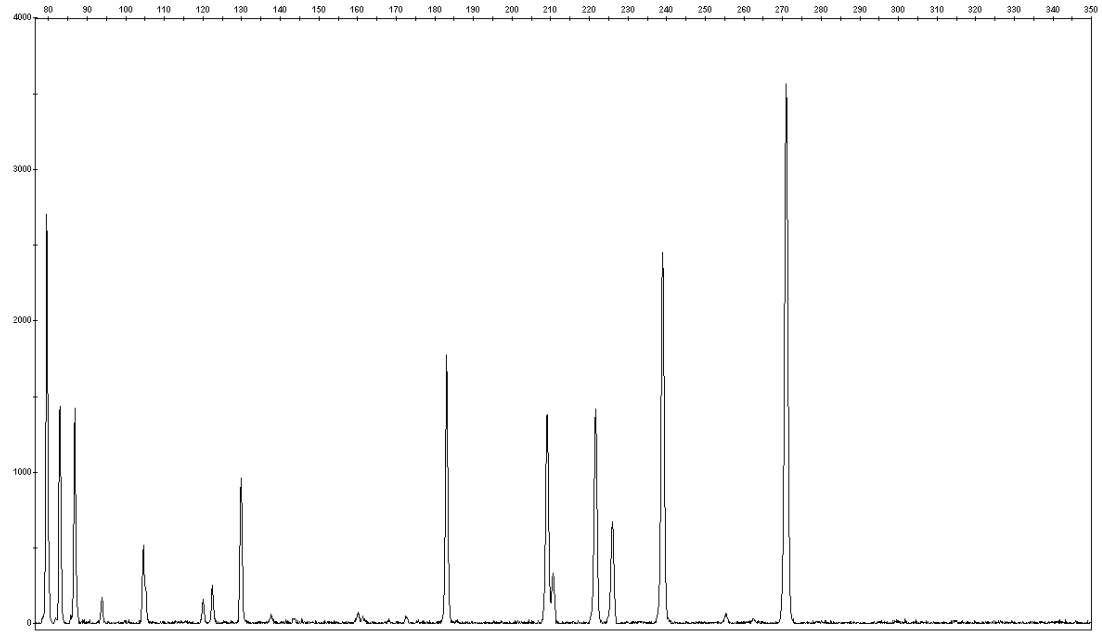

Figure 5E: *P. syrinage* strain F10A NED electropherogram. When compared to the *P. syrinage* strain F7 NED electropherogram (Figure 5F), similarities can be observed, however, differences clearly exist that make the AFLP profiles of the two strains unique from one another.

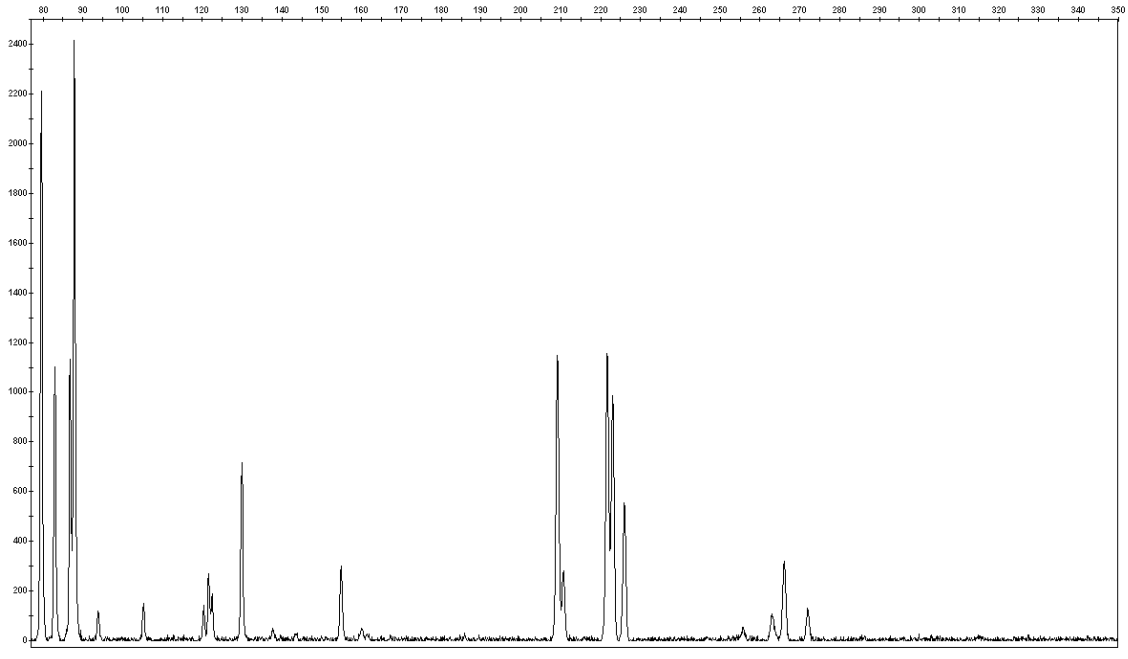

Figure 5F: *P. syrinage* strain F7 NED electropherogram. When compared to the *P. syrinage* strain F10A NED electropherogram (Figure 5E), similarities can be observed, however, differences clearly exist that make the AFLP profiles of the two strains unique from one another.

To facilitate assessing the discriminatory power of the AFLP assay, electropherograms were translated into haplotype codes as described. The creation of a haplotype code for each strain allowed for easy comparison and differentiation of the different bacterial species and even closely related strains. Table 5 displays the haplotype codes for all three color channels of *P. syringae* strains F7 and F10A. Haplotype codes for all other analyzed strains can be found in Appendix 1. The haplotype code based on binning AFLP products makes for easy comparison of the two strains without having to view the electropherograms. Reviewing the haplotype codes for F7 and F10A in Table 5, it is also apparent that the two strains share many of the same profile characteristics, consistent with belonging to the same species and pathovar groups.

Table 5: Haplotype codes for *P. syringae* strain F7 and *P. syringae* F10A.

Bin:	75	85	95	105	115	125	135	145	155	165	175	185	195	205	215	225	235	245	255	265	275	285	295	305	315	325	335	345
F7	0	0	0	0	0	2	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0
F10A	1	1	0	0	0	0	0	1	0	0	0	1	1	0	1	0	1	1	0	0	0	1	2	0	0	0	0	0
F7	2	2	0	0	2	0	1	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0
F10A	3	2	0	0	2	0	1	1	2	0	2	0	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	1
F7	2	2	0	1	1	1	0	0	0	0	0	0	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0
F10A	2	1	1	0	0	1	0	0	0	0	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0

Although every strain can be distinguished from every other strain based upon it's haplotype code, there are instances where two or more strains may have the same code within an individual color channel. Table 6 provides an example of two strains that have the same code within a single color channel. *P. syringae* strain NF5 and *P. syringae* strain NF12 have the same haplotype code in the NED (yellow) color channel. If the NED color channel were taken in isolation, these two strains would be indistinguishable. However,

because the two strains have unique codes in both the FAM (blue) and JOE (green) color channels, the strains can be distinguished from one another.

Table 6: Haplotype codes for *P. syringae* strain NF5 and *P. syringae* NF12

Bin:	75	85	95	105	115	125	135	145	155	165	175	185	195	205	215	225	235	245	255	265	275	285	295	305	315	325	335	345
NF5	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
NF12	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
NF5	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
NF12	1	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
NF5	1	0	2	0	1	0	0	0	1	0	0	0	2	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0
NF12	1	0	2	0	1	0	0	0	1	0	0	0	2	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0

Strains NF5 and NF12 are not the only strains that are indistinguishable from each other in a single color channel. *P. syringae strains* NF3 and NF12 are indistinguishable from each other in the FAM (blue) color channel. *P. syringae* strains NF3A and NF12 and strains F12 and F18 are indistinguishable from one another in the JOE (green) color channel. *P. syringae* strains NF3, NF3A, and NF12 are indistinguishable from one another in the NED (yellow) color channel. Although there are strains that share the same code in a single color channel, when the entire haplotype code (all three color channels included) is taken into account, every strain is unique. The results of this study demonstrate the importance of amplifying restriction fragments and all three selective primers in AFLP analysis in order to maximize the discriminatory power of the assay.

Genetic Relatedness

Sørensen Similarity Index

Genetic relatedness of microbial strains was assessed using the Sørensen similarity index. Botanist, Thorvald Sørensen developed the formula in 1948 in order to analyze the similarity between two data sets (Sørensen, 1948). The Sørensen similarity formula, displayed below, can determine the degree of similarity between any two data sets:

$$QS = 2C \div (A + B)$$

In the formula, A and B are the number of items in the two data sets and C is the number of items shared between the two data sets. In this study, the formula was used to compare the code of each strain within a species to each of the other strains. A and B represent the total number of peaks in each of the analyzed strain haplotype codes and C represents the number of peaks the two strains had in common (peaks located within the same bins in each haplotype code).

Each of the *P. syringae* pv. maculicola strains was compared to every other *P. syringae* pv. maculicola strain individually within each of the three color channels. The Sørensen similarity indices for the *P. syringae* pv. maculicola strains can be found in Figures 6A, 6B, and 6C (other Sørensen similarity indices comparing each of the strains analyzed in this study can be found in Appendix 2). A Sørensen similarity value of "0" indicates that two strains do not share any common fragments. As displayed in Figures 6A, 6B, and 6C, a combined Sørensen similarity value was calculated for each color channel by averaging all of the individual Sørensen similarity values. Sørensen similarity

values are displayed as a decimal that can be converted to a percentage of similarity by multiplying this value by 100.

Similarity Between *P. syringae* pv. maculicola Strains:										
	NF3	**NF3A**	**NF5**	**NF12**	**F7**	**F10A**	**F12**	**F15**	**F18**	**F22**
NF3	1.00	0.67	0.67	1.00	0.18	0.31	0.36	0.18	0.60	0.50
NF3A	0.67	1.00	0.67	0.67	0.00	0.15	0.36	0.00	0.60	0.50
NF5	0.67	0.67	1.00	0.67	0.00	0.15	0.36	0.00	0.60	0.50
NF12	1.00	0.67	0.67	1.00	0.18	0.31	0.36	0.18	0.60	0.50
F7	0.18	0.00	0.00	0.18	1.00	0.56	0.13	0.63	0.13	0.15
F10A	0.31	0.15	0.15	0.31	0.56	1.00	0.33	0.56	0.24	0.13
F12	0.36	0.36	0.36	0.36	0.13	0.33	1.00	0.25	0.40	0.67
F15	0.18	0.00	0.00	0.18	0.63	0.56	0.25	1.00	0.13	0.15
F18	0.60	0.60	0.60	0.60	0.13	0.24	0.40	0.13	1.00	0.67
F22	0.50	0.50	0.50	0.50	0.15	0.13	0.67	0.15	0.67	1.00
							FAM Similarity:			**0.44**

Figure 6A: Sørensen similarity table comparing each of the *P. syringae* pv. maculicola strain FAM haplotype codes to each other. A combined Sørensen similarity value of 0.44 was calculated, this value translates to mean that the *P. syringae* pv. maculicola strains have a 44% similarity between their haplotype codes within the FAM color channel.

\

Similarity Between *P. syringae* pv. maculicola Strains:										
	NF3	**NF3A**	**NF5**	**NF12**	**F7**	**F10A**	**F12**	**F15**	**F18**	**F22**
NF3	1.00	0.89	0.73	0.89	0.47	0.36	0.73	0.47	0.73	0.73
NF3A	0.89	1.00	0.83	1.00	0.56	0.43	0.83	0.56	0.83	0.67
NF5	0.73	0.83	1.00	0.83	0.50	0.48	0.71	0.50	0.71	0.57
NF12	0.89	1.00	0.83	1.00	0.56	0.43	0.83	0.56	0.83	0.67
F7	0.47	0.56	0.50	0.56	1.00	0.84	0.60	1.00	0.60	0.50
F10A	0.36	0.43	0.48	0.43	0.84	1.00	0.48	0.84	0.40	0.48
F12	0.73	0.83	0.71	0.83	0.60	0.48	1.00	0.50	1.00	0.86
F15	0.47	0.56	0.50	0.56	1.00	0.84	0.50	1.00	0.60	0.50
F18	0.73	0.83	0.71	0.83	0.60	0.40	1.00	0.60	1.00	0.86
F22	0.73	0.67	0.57	0.67	0.50	0.48	0.86	0.50	0.86	1.00
						JOE Similarity:	**0.70**			

Figure 6B: Sørensen similarity table comparing each of the *P. syringae* pv. maculicola strain JOE haplotype codes to each other. A combined Sørensen similarity value of 0.70 was calculated, this value translates to mean that the *P. syringae* pv. maculicola strains have a 70% similarity between their haplotype codes within the JOE color channel.

Similarity Between *P. syringae* pv. maculicola Strains:

	NF3	NF3A	NF5	NF12	F7	F10A	F12	F15	F18	F22
NF3	1.00	1.00	1.00	1.00	0.27	0.29	0.90	0.30	0.78	0.70
NF3A	1.00	1.00	1.00	1.00	0.27	0.29	0.90	0.30	0.78	0.70
NF5	1.00	1.00	1.00	1.00	0.27	0.29	0.90	0.30	0.78	0.70
NF12	1.00	1.00	1.00	1.00	0.27	0.29	0.90	0.30	0.78	0.70
F7	0.27	0.27	0.27	0.27	1.00	0.67	0.30	0.60	0.26	0.40
F10A	0.29	0.29	0.29	0.29	0.67	1.00	0.32	0.74	0.36	0.42
F12	0.90	0.90	0.90	0.90	0.30	0.32	1.00	0.22	0.86	0.89
F15	0.30	0.30	0.30	0.30	0.60	0.74	0.22	1.00	0.22	0.33
F18	0.78	0.78	0.78	0.78	0.26	0.36	0.86	0.22	1.00	0.76
F22	0.70	0.70	0.70	0.70	0.40	0.42	0.89	0.33	0.76	1.00

NED Similarity: 0.63

Figure 6C: Sørensen similarity table comparing each of the *P. syringae* pv. maculicola strain NED haplotype codes to each other. A combined Sørensen similarity value of 0.63 was calculated, this value translates to mean that the *P. syringae* pv. maculicola strains have a 63% similarity between their haplotype codes within the NED color channel.

Similarity values for each of the three color channels were averaged to obtain an overall similarity value for the *P. syringae* pv. maculicola strains. The overall similarity was found to be 59%. This value indicates that *P. syringae* pv. maculicola haplotype codes are overall 59% similar to one another.

Sørensen similarity tables were also created to compare the *P. aeruginosa* strains analyzed in this study (Appendix 2). *P. aeruginosa* strains were calculated to have an overall similarity of 54%. Likewise, Sørensen similarity tables were created to compare the genetic similarity of all species and strains analyzed in this study with bacterial strains that have been previously analyzed using AFLP (Appendix 2). Species included for comparison were *P. syringae* pv. tomato (Taylor, 2009), *S. marsescens* (Beauman, 2006), and *S. aureus* (Allen, unpublished observations). *P. syringae* pv. maculicola strains were found to be 40% similar to *P. syringae* pv. tomato strains. A 22% genetic similarity was calculated between *P. syringae* pv. maculicola strains and *P. aeruginosa* strains. *P. syringae* pv. maculicola strains were found to be share 23% genetic similarity with *S. aureus* strains, but only 17% similarity to *S. marsescens*. These results demonstrate that *P. syringae* pv. maculicola strains were most similar to strains of the same pathovar and to the same species and were less similar to strains of different species.

Agglomerative Hierarchical Clustering

Agglomerative Hierarchical Clustering (AHC) was used to further establish the relatedness of strains analyzed using AFLP. Strains included in the AHC analysis were *P. syringae* pv. maculicola and *P. aeruginosa* strains from this study, as well as *P. syringae* pv. tomato strains (Taylor, 2009), *S. marsescens* strains (Beauman, 2006), and *S. aureus* strains (Allen, Unpublished observations) analyzed in previous studies. AHC analysis provides progressive grouping of data based on dissimilarities between the groups being analyzed. The process clusters objects (in this case strain haplotype codes) together based on their similarity to one another and establishes their grouping and

distance from other objects based on their dissimilarity. Results are displayed in a binary clustering tree, referred to as a dendrogram. Objects are placed in a certain order within the dendrogram based on the dissimilarity between them. Objects that have a low dissimilarity value (and thus a high similarity) are placed close to each other in the dendrogram. Individual objects within and between their groupings will be placed closer or farther away from each other based on how dissimilar they are (Microsoft ® XLSTAT).

The dendrogram produced in this study, shown in Figure 8, displays the similarity between each of the analyzed strains. Viewing the dendrogram from top to bottom, the strains are ordered with *P. syringae* pv. tomato strains first, followed by *P. syringae* pv. maculicola strains. All of the *P. syringae* strains are placed in a group together, distinguished by brown lines. Continuing towards the bottom of the dendrogram, the *P. aeruginosa* strains follow the *P. syringae* strains and are placed into a second group distinguished by pink lines. Finally, continuing down the dendrogram, the *S. marcescens* and *S. aureus* strains are placed into a third group distinguished by green lines. The order that the bacterial strains are arranged in within the dendrogram indicates their haplotype code similarity (and thus their genetic similarity) to one another. Due to both a close proximity on the graph to one another and the fact that they share a common group, it is evident that the *P. syringae* pv. maculicola strains show the closest genetic relationship to the *P. syringae* pv. tomato strains. The *P. syringae* pv. maculicola strains also have a close relationship to the *P. aeruginosa* strains which fall in a separate group but do follow the *P. syringae* strains in order within the dendrogram. The *S. marcescens* and *S. aureus* strains appear to be the least genetically similar to the *P. syringae* pv. maculicola strains

as they appear in a separate group, the greatest distance away from the *P. syringae* strains. The results displayed within the dendrogram make biological sense, with the same species of bacteria having a greater genetic similarity to each other than to different species. The results of this study demonstrate that AFLP analysis is proficient at distinguishing bacterial genomes while at the same time revealing regions that are homologous or at least related such that different strains can be grouped together.

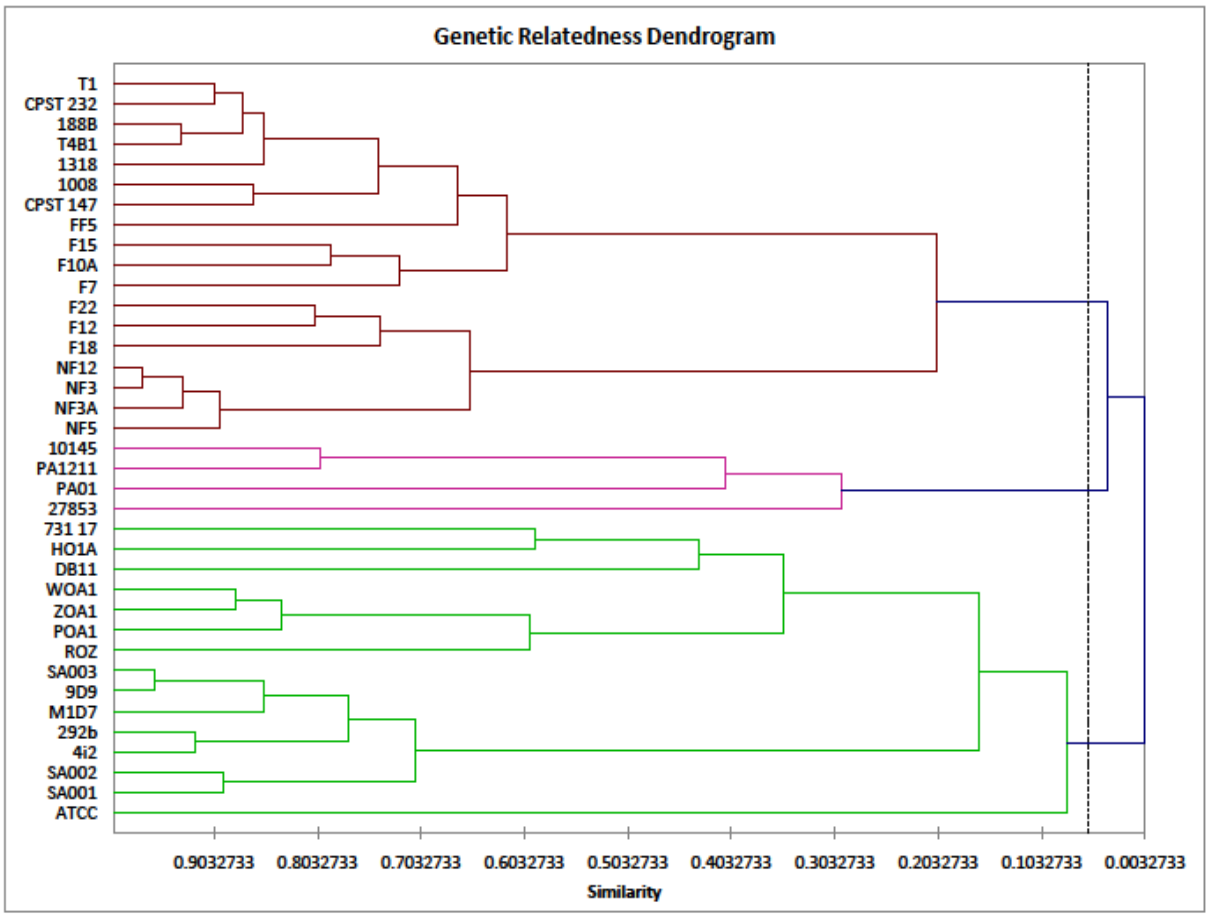

Figure 8: A dendrogram displaying the results of Agglomerative Hierarchical Clustering (AHC). The dendrogram portrays the genetic similarity between (in order from top to bottom) 8 strains of *P. syringae* pv. tomato, 10 strains of *P. syringae* pv. maculicola, 4 strains of *P. aeruginosa*, 7 strains of *S. marcescens*, and 7 strains of *S. aureus*, and 1 strain of *S. marcescens*. The AHC results are separated into three groups (distinguished by line color) based on haplotype code similarity.

Effect of Acquired Mutations on AFLP Results

With many types of mutations occurring in bacterial genomes, including single nucleotide substitutions (SNSs), single nucleotide insertions and deletions, sequence duplications, and sequence inversions, it is important to recognize that the bacterial culture obtained as evidence may be different than the bacterial culture recovered from the suspected source (Velsko, 2005). We therefore investigated the effect of acquired mutations in the bacterial genome on the "stability" of the AFLP profile. In order to establish the effect of acquired mutations on AFLP analysis results and the haplotype code established for a sample, three different cultures of *P. aeruginosa* strain PAO1were analyzed, traceable to an original ATCC reference strain many years ago, using AFLP. The two additional cultures of PAO1, labeled PAO1A and PAO1B, were obtained from two different laboratories at Oklahoma State University. It is important to note that PAO1B was originally cultured from PAO1A several years prior to this study and allowed to grow under separate conditions. Mutations accumulate during cell division as well as during stationary phase, therefore, in the two separate cultures of PA01, allowed to grow and be stored in separate conditions for several years, acquired mutations in the genome of each would be expected. The goal of this study was to determine if mutations acquired by PAO1A and PAO1B would be detected by AFLP analysis and therefore affect the final haplotype code produced for each culture.

The PAO1 comparison study was performed with the same methodology as the previous AFLP analyses. Three replicates were analyzed for each of the PAO1 strains with an overall reproducibility of 99.6%. The results of this study demonstrated that cultures of the same strain, allowed to grow and be stored under separate conditions, do

acquire mutations that are detected by AFLP analysis. As shown in Table 7, haplotype codes obtained for PAO1A and PAO1B are distinct from the haplotype code obtained for the original PAO1 culture used in this study.

Table 7: Haplotype codes for three different cultures of P. *aeruginosa* strain PAO1.

Bin:	75	85	95	105	115	125	135	145	155	165	175	185	195	205	215	225	235	245	255	265	275	285	295	305	315	325	335	345
FAM																												
PA01 A	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
PA01 B	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
PA01 (original)	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
JOE																												
PA01 A	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
PA01 B	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
PA01 (original)	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
NED																												
PA01 A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
PA01 B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
PA01 (original)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0

The haplotype codes for PAO1A and PAO1B are distinct from each other and from the original PAO1 haplotype code. No differences occur in the FAM color channel, however, PAO1B shows a single difference in the JOE color channel and PAO1A and PAO1B both show a difference from the original PAO1 strain in the NED color channel (Table 7).

Both of the observed differences in haplotype codes constitute the loss of a peak (thus the loss of a DNA restriction fragment of a specific size), in bin 255 for the JOE color channel and in bin 275 for the NED color channel. The most likely mutation event to lead to the loss of a restriction fragment in the size ranges of 255 basepairs and 275

basepairs, is the deletion of a restriction site. The deletion of a restriction site could occur through a single point mutation, deleting or inserting a nucleotide at an *EcoR1* or *Mse1* restriction site. The deletion of a restriction site would lead to the creation of a larger restriction fragment, falling in a size range above 350 basepairs, which would in turn lead to a change in the haplotype code produced for that culture. The results of this study demonstrate the possibility that naturally occurring mutation events can alter the haplotype code of a strain. It is therefore possible that the haplotype code for one bacterial culture might not match precisely the haplotype code for a parent strain. Recall however that the three PAO1 strains compared here were separated by years of time and perhaps thousands of generations for mutation events to occur. When using AFLP analysis to type a bacterial strain, analysts must be aware of possible mutation events that may affect the outcome of the analysis and therefore affect any final conclusions that are drawn.

CHAPTER V

CONCLUSIONS

As a DNA typing technique, AFLP has demonstrated consistent reproducibility and a high level of discriminatory power. Similar results have been achieved with the analysis of both prokaryotic and eukaryotic organisms, as well as with both plant, animal, and human pathogens (Allen, 2005; Beauman, 2006; and Taylor, 2009). The goals of this study were to confirm the reliability and discriminatory capability of the AFLP assay to differentiate between closely related strains of a given species and between different species of bacteria. In addition, the results obtained in this study suggest that the AFLP technique may be suitable as both a forensic identification and attribution tool.

Reliability and consistent reproducibility of an assay are both critical elements of a useful DNA typing tool. In that regard, an overall reproducibility of 97.9% for *P. syringae* pv. maculicola strains and an overall reproducibility of 97.5% for *P. aeruginosa* strains was possible for AFLP analysis. These reproducibility values are slightly greater than the 95% reproducibility value generated in a previous study of *P. syringae* pv. tomato (Taylor, 2009). Reproducibility values obtained in this study also exceeded a reproducibility value of 91% reported in earlier work using AFLP analysis to assess the

genetic relatedness of the *Bordetella* genus (Gzyl, *et al*., 2004).

A successful molecular typing tool must also have a high discriminatory power in order to discriminate between closely related strains and species. In this study, the AFLP assay demonstrated a high degree of discriminatory power with all ten of the *P. syringae* pv. maculicola strains and all four of the *P. aeruginosa* strains generating unique haplotype codes. Utilizing the Sørensen similarity index, we calculated a similarity value of 59% for *P. syringae* pv. maculicola strains and an overall similarity value of 54% for *P. aeruginosa* strains was calculated. These similarity values indicate that *P. syringae* pv. maculicola strains share an average of 59% of their haplotype codes and *P. aeruginosa* strains share an average of 54% of their haplotype codes. The similarity values obtained in this study are lower than the similarity value of 81% obtained in a previous study of *P. syringae* pv. tomato (Taylor, 2009). This decrease in similarity could be due to the fact that the *P. syringae* pv. tomato strains had consistent similarity between peak location in each strain, whereas the *P. syringae* pv. maculicola strains did have certain strains that shared no peaks in common with each other, resulting in a Sørensen similarity value of "0" in some instances, lowering the overall similarity value. This decrease in similarity could also be due to the nature of the strains used in the study. The *P. syringae* pv. tomato strains were known to have been collected from the same region and thus would be expected to have a higher genetic similarity to one another than other *P. syringae* strains (Taylor, 2009).

The *P. syringae* pv. maculicola strains were also compared to the four strains of *P. aeruginosa,* eight strains of *S. marsescens* (Beauman, 2006), and seven strains of *S. aureus* (Allen, 2005) to assess the similarity between bacterial species. A similarity of

22% was found between *P. syringae* pv. maculicola strains and *P. aeruginosa* strains. A 17% similarity was calculated between *P. syringae* pv. maculicola strains and *S. marsescens* strains and a 23% similarity was found between *P. syringae* pv. maculicola strains and *S. aureus* strains. Therefore, as might be expected, the similarity indices are higher among related strains of the same species than with strains of unrelated species. Results obtained in this study correlate with the results of an earlier AFLP study, which found that strains within the same species had a 51-100% similarity and strains of different species had less than 25% similarity (Clerc, *et al*., 1998). This conclusion is also supported by the AHC analysis which grouped strains of the same species and even the same genus into clusters more closely spaced in the dendrogram (Figure 8).

Although a high discriminatory power is important for an effective DNA typing assay, the similarities in haplotype codes among strains of the same species can also aid in the identification of a microbial sample and possible attribution to a source. This study established that in addition to each strain being uniquely identifiable based on the strain's haplotype code, strains of the same species also share consistencies within their haplotype codes. It could therefore be possible to exploit these similarities as a bacterial species identification tool.

The reliability, consistent reproducibility, and high discriminatory power of AFLP analysis suggest that this technique holds promise as a microbial forensic DNA analysis technique. The assay is relatively cheap, reasonably fast, and can be performed with basic DNA analysis equipment, typically available at a forensic DNA laboratory. Additionally, the haplotype codes generated in this and previous studies (Allen, 2005; Beauman, 2006; Taylor, 2009) were entered into a database of haplotype codes that could be searched

with the code of an unknown bacterium for the purpose of identification. The creation of a database of AFLP haplotype codes further increases the utility of AFLP analysis in a forensic context.

The creation and use of the database is only possible because AFLP profiles were translated into a haplotype code compatible with Microsoft® Access software. The haplotype codes prepared from several unknown strains were correctly matched against entries in the database created, underscoring the value of a haplotype code database for forensic investigation. The database contains haplotype codes for the ten strains of *P. syringae* pv. maculicola and the four strains of *P. aeruginosa* analyzed in this study, along with eight strains of *P. syringae* pv. tomato (Taylor, 2009), eight strains of *S. marcescens* (Beauman, 2007), and seven strains of *S. aureus* (Allen, 2006) from previous studies. In total, the haplotype code database contains codes for 45 strains of four different species of bacteria. The database was created with the specific ability to enter the haplotype code of an unidentified strain in a query to search the database for any strain with a matching code. If the unidentified strain code matches a microbial strain in the database, a report is generated displaying the species' name, strain, pathovar, host, any available treatments for infections, and the haplotype code of the identified strain. An example of a report generated from a search match in the database can be seen in Appendix 3. Ultimately, the creation and expansion of the haplotype code database allows AFLP analysis to be applied in a broader forensic context, aiding investigators in using this technique for not only the attribution of a specific microbial strain and strain, but also as a tool to identify the species of an unknown bacterium.

The AFLP assay has proven to be consistently reliable, reproducible, discriminatory, and applicable to a forensic setting through the use of a database. However, the mutation study performed using the different subcultures of an initial single strain of *P. aeruginosa* PAO1 revealed that mutations can affect the AFLP profile and subsequent haplotype codes. This fact must therefore be considered when comparing the haplotype code of an evidentiary strain with a suspected source pathogen in the investigation of a biocrimes. However, comparison of the haplotype codes of the three *P. aeruginosa* PAO1 strains differed in only a small proportion of the overall code and so developing a matching algorithm to score an unknown as a possible match would still be of value both for attribution and identification of an unknown. Because of mutation, AFLP analysis should not be considered the sole mechanism for attributing a pathogen used as a weapon. Rather, the technique should be more appropriately considered a screening tool that is effective at narrowing the field of possible candidates as possible sources of a strain involved in a criminal act.

REFERENCES

Allen, R. (2006). [Amplified Fragment Length Polymorphism of *Staphylococcus aureus* Strains]. Unpublished Raw Data.

Beauman, C. (2007). Evaluating the suitability of AFLP technology for genotyping strains of *Serratia marcescens*. Masters Thesis, Oklahoma State University, Tulsa, Oklahoma, December 2007.

Breeze, R.G., Budowle, B., & Schutzer, S.E. (2005). *Microbial Forensics*. London, England: Elsevier, Inc.

Budowle, B., Murch, R., & Charaborty, R. (2005). Microbial forensics: the next forensic Challenge. *International Journal of Legal Medicin,* 119, 317-330.

Centers for Disease Control and Prevention. *Morbidity and Mortality Weekly Report* (2010, August 13). Surveillance for foodborne disease outbreaks. Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5931a1.htm.

Clerc C., Manceau C., & Nesme, X. (1998). Comparison of randomly amplified polymorphic DNA with amplified fragment length polymorphism to assess the genetic diversity and genetic relatedness with genospecies III of *Pseudomonas syringae*. *Applied and Environmental Microbiology*, 64(4), 1180-1187.

Cobiella, Kelly. (2008, July 2). Salmonella scare rocks tomato country. *CBS Evening News*. Retrieved from http://www.cbsnews.com/stories/2008/07/02/eveningnews /main 4229216.shtml.

Cummings, C.A., & Relman, D.A. (2002). Genomics and microbiology: microbial forensic – "cross-examining pathogens". *Science.* (New York, N.Y.), 296(5575), 1976-1979.

Fletcher, J., Bender, C., Budowle, B., Cobb, W.T., Gold, S.E., Ishamaru, C.A., et al. (2006). Plant pathogen forensics: Capabilities, needs and recommendations. *Microbiology and Molecular Biology Reviews,* 70(2), 450-471.

Georgnaras, I., Kenene, N.F., von Holy, A., & Hastings, J.W. (1999). Amplified fragment length polymorphism fingerprinting of *Pseudomonas* strains from a poultry processing plant. *Applied and Environmental Microbiology*, 65(9), 3828-3833.

Gzyl, A., Augustynowicz, E., Mosiej, E., Zawadka, M., Grzegorz, G., Nowaczek, A., & Slusarczyk, J. Amplified fragment length polymorphism (AFLP) versus randomly amplified polymorphic DNA (RAPD) as new tools for inter- and intra-species differentiation within *Bordetella. Journal of Medical Microbiology*, 54(4), 333-346.

Jackson, P.J., Hill, K.K., Laker, M.T., Ticknor, L.O., & Keim, P. (1999). Genetic comparison of*Bacillus anthracis* and its close relatives using amplified fragment length polymorphism and polymerase chain reaction analysis. *Journal of Applied Microbiology*, 87, 263-269.

Janssen, P., Coopman, R., Huys, G., Swings, G., Bleeker, M., Vos, P., et al. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology*, 142, 1881-1893.

Joardar, V., Lindeberg, M., Jackson, R., Selengut, J., Dodson, R., Brinkac, L., et al. (2005). Whole-genome sequence analysis of *Pseudomonas syringae* pv. phaseolicola 1448A reveals divergence among pathovars in genes involved in virulence and transposition. *Journal of Bacteriology,* 187(18), 6488-6498.

Lin, J., Kuo,J., Ma, J. (1996). A PCR-based DNA fingerprinting technique: AFLP for molecular typing of bacteria. *Nucleic Acids Research*, 24(18), 3649-3650.

Monke, J. (2004). Agroterrorism: threats and preparedness (CRS Report for Congress No.RL3251)Library of Congress. Retrieved from www.fas.org/irp/crs/RL32521.pdf.

NRC (National Research Council of the National Academies). Countering Agricultural Bioterrorism. Washington DC, 2002.

Parker, H.S. (2003). Agricultural bioterrorism: A federal strategy to meet the threat. McNair Paper 65. Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?AD= ADA409307& Location=U2&doc= GetTRDoc.pdf.

Rudolph, K., Burr, T.J., Mansfield, J.W., Stead, D., Vivian, A, & Von Kietzell J. (1997). *Pseudomonas syringae* pathovars and related pathogens. *Developments in Plant Pathology*, 9, 575-582.

Taylor, A. (2009). Evaluating the suitability of AFLP genotyping technology for discriminating among strains of *Pseudomonas syringae*. Masters Thesis, Oklahoma State University, Tulsa, Oklahoma, December 2009.

Velsko, S.P. (2005). Resolution in forensic microbial genotyping. *Lawrence Livermore National Laboratory,* UCRL-TR-215305.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M et al. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 1995;23(21):4407-4414.

U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. (2010). Investigation Update: Multistate Outbreak of Human *Salmonella* Enteritidis Infections Associated with Shell Eggs. Retrieved from http://www.cdc.gov/salmonella/ enteritidis/.

APPPENDICES

Appendix 1

Haplotype codes for *P. syringae* pv. maculicola strains (FAM – blue)

	75	85	95	105	115	125	135	145	155	165	175	185	195	205	215	225	235	245	255	265	275	285	295	305	315	325	335	345
NF3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
NF3A	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
NF5	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
NF12	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
F7	0	0	0	0	0	2	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0
F10A	1	1	0	0	0	0	0	1	0	0	0	1	1	0	1	0	1	1	0	0	0	1	2	0	0	0	0	0
F12	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	1	1	1
F15	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	2	0	0	0	1	1	0	0	0	0	0
F18	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1
F22	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1

Appendix 1

Haplotype codes for *P. syringae* pv. maculicola strains (JOE - green)

	75	85	95	105	115	125	135	145	155	165	175	185	195	205	215	225	235	245	255	265	275	285	295	305	315	325	335	345
NF3	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
NF3A	1	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
NF5	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
NF12	1	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
F7	2	2	0	0	2	0	1	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0
F10A	3	2	0	0	2	0	1	1	2	0	2	0	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	1
F12	1	1	0	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
F15	2	2	0	0	2	0	1	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0
F18	1	1	0	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
F22	1	0	0	0	0	0	1	1	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0

Appendix 1

Haplotype codes for *P. syringae* pv. maculicola strains (NED - yellow)

	75	85	95	105	115	125	135	145	155	165	175	185	195	205	215	225	235	245	255	265	275	285	295	305	315	325	335	345
NF3	1	0	2	0	1	0	0	0	1	0	0	0	2	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0
NF3A	1	0	2	0	1	0	0	0	1	0	0	0	2	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0
NF5	1	0	2	0	1	0	0	0	1	0	0	0	2	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0
NF12	1	0	2	0	1	0	0	0	1	0	0	0	2	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0
F7	2	2	0	1	1	1	0	0	0	0	0	0	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0
F10A	2	1	1	0	0	1	0	0	0	0	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0
F12	1	0	2	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
F15	2	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0
F18	1	0	2	0	1	0	0	0	1	0	0	2	1	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0
F22	1	1	1	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0

Appendix 2

Sørensen similarity indices for *P. aeruginosa* strains (FAM – blue)

	PA01	**PA1211**	**27853**	**10145**
PA01	1.00	0.33	0.33	0.40
PA1211	0.33	1.00	0.33	0.40
27853	0.33	0.33	1.00	0.00
10145	0.40	0.40	0.00	1.00

FAM Similarity: 0.48

Appendix 2

Sørensen similarity indices for *P. aeruginosa* strains (JOE - green)

	PA01	**PA1211**	**27853**	**10145**
PA01	1.00	0.50	0.46	0.57
PA1211	0.50	1.00	0.46	0.86
27853	0.46	0.46	1.00	0.33
10145	0.57	0.86	0.33	1.00

JOE Similarity: 0.65

Appendix 2

Sørensen similarity indices for *P. aeruginosa* strains (NED - yellow)

	PA01	**PA1211**	**27853**	**10145**
PA01	1.00	0.40	0.19	0.50
PA1211	0.40	1.00	0.10	0.67
27853	0.19	0.10	1.00	0.11
10145	0.50	0.67	0.11	1.00

NED Similarity: 0.50

Appendix 2

Sørensen similarity indices for *P. aeruginosa* and *P. syringae* pv. maculicola strains (FAM- blue)

	NF3	**NF3A**	**NF5**	**NF12**	**F7**	**F10A**	**F12**	**F15**	**F18**	**F22**
PA01	0.00	0.33	0.00	0.00	0.00	0.15	0.18	0.00	0.20	0.00
PA1211	0.00	0.67	0.33	0.33	0.00	0.15	0.18	0.00	0.40	0.25
27853	0.33	0.33	0.33	0.33	0.18	0.15	0.36	0.18	0.20	0.50
10145	0.00	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.00

FAM Similarity: 0.17

Appendix 2

Sørensen similarity indices for *P. aeruginosa* and *P. syringae* pv. maculicola strains (JOE - green)

	NF3	**NF3A**	**NF5**	**NF12**	**F7**	**F10A**	**F12**	**F15**	**F18**	**F22**
PA01	0.25	0.22	0.36	0.22	0.12	0.18	0.18	0.12	0.18	0.18
PA1211	0.25	0.44	0.36	0.44	0.24	0.18	0.36	0.24	0.36	0.18
27853	0.15	0.29	0.50	0.29	0.27	0.30	0.38	0.27	0.38	0.25
10145	0.29	0.50	0.40	0.50	0.25	0.19	0.40	0.25	0.40	0.20

JOE Similarity: 0.29

Appendix 2

Sørensen similarity indices for *P. aeruginosa* and *P. syringae* pv. maculicola strains (NED - yellow)

	NF3	**NF3A**	**NF5**	**NF12**	**F7**	**F10A**	**F12**	**F15**	**F18**	**F22**
PA01	0.14	0.14	0.14	0.14	0.14	0.15	0.17	0.33	0.13	0.17
PA1211	0.15	0.15	0.15	0.15	0.15	0.33	0.00	0.36	0.14	0.00
27853	0.28	0.28	0.28	0.28	0.28	0.48	0.50	0.40	0.52	0.70
10145	0.00	0.00	0.00	0.00	0.17	0.18	0.00	0.20	0.00	0.00

NED Similarity: 0.20

Appendix 2

Sørensen similarity indices for *P. syringae* pv. tomato and *P. syringae* pv. maculicola strains (FAM- blue)

	CPST 147	CPST 232	RG4	880	T1	TF1	T4B1	1318	30555	B125	3357	PT17	1008	188B
NF3	0.31	0.27	0.31	0.33	0.29	0.27	0.13	0.13	0.36	0.29	0.13	0.31	0.17	0.27
NF3A	0.15	0.13	0.15	0.17	0.14	0.13	0.13	0.13	0.18	0.14	0.13	0.15	0.17	0.13
NF5	0.15	0.13	0.15	0.17	0.14	0.13	0.13	0.13	0.18	0.14	0.13	0.15	0.17	0.13
NF12	0.31	0.27	0.31	0.33	0.29	0.27	0.13	0.13	0.36	0.29	0.13	0.31	0.17	0.27
F7	0.44	0.40	0.67	0.35	0.42	0.50	0.30	0.40	0.50	0.53	0.40	0.56	0.35	0.30
F10A	0.70	0.64	0.70	0.63	0.76	0.64	0.64	0.64	0.89	0.89	0.55	0.70	0.74	0.64
F12	0.33	0.30	0.44	0.35	0.32	0.30	0.30	0.40	0.38	0.32	0.50	0.33	0.35	0.30
F15	0.56	0.60	0.67	0.47	0.63	0.60	0.50	0.50	0.75	0.63	0.50	0.67	0.47	0.50
F18	0.24	0.32	0.24	0.25	0.33	0.32	0.21	0.21	0.27	0.22	0.21	0.24	0.57	0.21
F22	0.13	0.12	0.13	0.14	0.13	0.12	0.12	0.12	0.15	0.13	0.12	0.13	0.14	0.12

FAM Similarity: 0.33

Appendix 2

Sørensen similarity indices for *P. syringae* pv. tomato and *P. syringae* pv. maculicola strains (JOE – green)

	CPST 147	**CPST 232**	**RG4**	**880**	**T1**	**TF1**	**T4B1**	**1318**	**30555**	**B125**	**3357**	**PT17**	**1008**	**188B**
NF3	0.35	0.35	0.38	0.38	0.36	0.36	0.33	0.32	0.35	0.40	0.35	0.36	0.36	0.35
NF3A	0.42	0.42	0.45	0.45	0.43	0.43	0.40	0.38	0.42	0.48	0.42	0.43	0.43	0.42
NF5	0.46	0.38	0.42	0.50	0.48	0.48	0.44	0.43	0.46	0.43	0.46	0.48	0.48	0.46
NF12	0.42	0.42	0.45	0.45	0.43	0.43	0.40	0.38	0.42	0.48	0.42	0.43	0.43	0.42
F7	0.75	0.81	0.73	0.80	0.84	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75
F10A	0.83	0.78	0.76	0.82	0.80	0.72	0.78	0.83	0.83	0.67	0.78	0.72	0.72	0.78
F12	0.54	0.54	0.58	0.50	0.56	0.54	0.54	0.54	0.54	0.46	0.54	0.54	0.54	0.54
F15	0.75	0.81	0.73	0.80	0.77	0.75	0.75	0.81	0.81	0.69	0.81	0.75	0.75	0.81
F18	0.54	0.54	0.58	0.50	0.56	0.54	0.54	0.54	0.54	0.46	0.54	0.54	0.54	0.54
F22	0.54	0.54	0.58	0.50	0.56	0.54	0.54	0.54	0.54	0.46	0.54	0.54	0.54	0.54

JOE Similarity: 0.56

Appendix 2

Sørensen similarity indices for *P. syringae* pv. tomato and *P. syringae* pv. maculicola strains (NED - yellow)

	CPST 147	CPST 232	RG4	880	T1	TF1	T4B1	1318	30555	B125	3357	PT17	1008	188B
NF3	0.29	0.26	0.29	0.40	0.36	0.43	0.43	0.35	0.53	0.36	0.35	0.38	0.40	0.43
NF3A	0.29	0.26	0.29	0.40	0.36	0.43	0.43	0.35	0.53	0.36	0.35	0.38	0.40	0.43
NF5	0.29	0.26	0.29	0.40	0.36	0.43	0.43	0.35	0.53	0.36	0.35	0.38	0.40	0.43
NF12	0.29	0.26	0.29	0.40	0.36	0.43	0.43	0.35	0.53	0.36	0.35	0.38	0.40	0.43
F7	0.10	0.17	0.19	0.10	0.18	0.17	0.17	0.17	0.21	0.18	0.26	0.19	0.10	0.17
F10A	0.19	0.26	0.29	0.20	0.27	0.26	0.26	0.35	0.32	0.27	0.35	0.29	0.20	0.26
F12	0.21	0.29	0.11	0.22	0.20	0.19	0.29	0.19	0.35	0.20	0.19	0.21	0.22	0.29
F15	0.21	0.29	0.21	0.33	0.30	0.29	0.38	0.29	0.35	0.30	0.38	0.32	0.33	0.38
F18	0.36	0.42	0.27	0.38	0.35	0.33	0.42	0.33	0.50	0.35	0.33	0.36	0.38	0.42
F22	0.32	0.38	0.21	0.33	0.30	0.29	0.38	0.29	0.47	0.30	0.29	0.32	0.33	0.38

NED Similarity: 0.32

Appendix 2

Sørensen similarity indices for *S. marsescens* and *P. syringae* pv. maculicola strains (FAM- blue)

	ATCC	HO1A	DB11	73117	POA1	ROZ	ZOA1	WOA1
NF3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NF3A	0.18	0.14	0.00	0.00	0.00	0.00	0.00	0.00
NF5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NF12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F7	0.13	0.11	0.14	0.00	0.00	0.00	0.00	0.00
F10A	0.22	0.19	0.25	0.11	0.29	0.31	0.27	0.29
F12	0.25	0.21	0.14	0.13	0.17	0.18	0.15	0.17
F15	0.13	0.00	0.14	0.00	0.00	0.00	0.00	0.00
F18	0.27	0.33	0.15	0.27	0.18	0.00	0.17	0.18
F22	0.31	0.13	0.00	0.15	0.00	0.00	0.00	0.00

FAM Similarity: 0.08

Appendix 2

Sørensen similarity indices for *S. marsescens* and *P. syringae* pv. maculicola strains (JOE – green)

	ATCC	HO1A	DB11	73117	POA1	ROZ	ZOA1	WOA1
NF3	0.31	0.00	0.13	0.00	0.00	0.22	0.18	0.20
NF3A	0.29	0.14	0.24	0.00	0.00	0.20	0.17	0.18
NF5	0.38	0.00	0.32	0.00	0.36	0.50	0.43	0.46
NF12	0.29	0.14	0.24	0.00	0.00	0.20	0.17	0.18
F7	0.27	0.36	0.32	0.20	0.00	0.11	0.10	0.11
F10A	0.23	0.38	0.41	0.14	0.10	0.18	0.25	0.26
F12	0.25	0.13	0.32	0.14	0.00	0.17	0.14	0.15
F15	0.27	0.36	0.32	0.20	0.12	0.11	0.10	0.11
F18	0.25	0.13	0.32	0.14	0.00	0.17	0.14	0.15
F22	0.13	0.13	0.32	0.29	0.00	0.17	0.14	0.15

JOE Similarity: 0.19

Appendix 2

Sørensen similarity indices for *S. marsescens* and *P. syringae* pv. maculicola strains (NED - yellow)

	ATCC	HO1A	DB11	73117	POA1	ROZ	ZOA1	WOA1
NF3	0.12	0.29	0.32	0.18	0.00	0.46	0.09	0.18
NF3A	0.12	0.29	0.32	0.18	0.00	0.46	0.09	0.18
NF5	0.12	0.29	0.32	0.18	0.00	0.46	0.09	0.18
NF12	0.12	0.29	0.32	0.18	0.00	0.46	0.09	0.18
F7	0.30	0.48	0.42	0.36	0.32	0.46	0.43	0.55
F10A	0.30	0.57	0.42	0.18	0.11	0.31	0.26	0.27
F12	0.13	0.21	0.35	0.20	0.12	0.42	0.19	0.30
F15	0.24	0.27	0.30	0.17	0.10	0.30	0.25	0.26
F18	0.19	0.32	0.35	0.20	0.12	0.58	0.19	0.30
F22	0.19	0.21	0.35	0.20	0.12	0.42	0.19	0.30

NED Similarity: 0.25

Appendix 2

Sørensen similarity indices for *S. aureus* and *P. syringae* pv. maculicola strains (FAM- blue)

	9D9	**SA001**	**SA002**	**SA003**	**M1D7**	**4i2**	**292b**
NF3	0.17	0.17	0.15	0.17	0.20	0.14	0.13
NF3A	0.33	0.33	0.31	0.33	0.40	0.29	0.27
NF5	0.17	0.17	0.15	0.17	0.20	0.14	0.13
NF12	0.17	0.17	0.15	0.17	0.20	0.14	0.13
F7	0.00	0.00	0.11	0.00	0.00	0.11	0.10
F10A	0.21	0.21	0.20	0.21	0.24	0.19	0.18
F12	0.12	0.12	0.22	0.12	0.13	0.21	0.20
F15	0.12	0.12	0.22	0.12	0.13	0.21	0.20
F18	0.38	0.38	3.00	0.38	0.43	0.33	0.32
F22	0.14	0.14	0.27	0.14	0.17	0.25	0.24

FAM Similarity: 0.23

Appendix 2

Sørensen similarity indices for *S. aureus* and *P. syringae* pv. maculicola strains (JOE – green)

	9D9	SA001	SA002	SA003	M1D7	4i2	292b
NF3	0.29	0.33	0.29	0.25	0.29	0.18	0.20
NF3A	0.25	0.29	0.25	0.22	0.25	0.17	0.18
NF5	0.60	0.44	0.60	0.55	0.40	0.43	0.46
NF12	0.25	0.29	0.25	0.22	0.25	0.17	0.18
F7	0.13	0.13	0.13	0.24	0.25	0.40	0.42
F10A	0.20	0.21	0.20	0.29	0.30	0.33	0.35
F12	0.20	0.22	0.20	0.18	0.20	0.29	0.15
F15	0.13	0.13	0.13	0.24	0.25	0.40	0.42
F18	0.20	0.22	0.20	0.18	0.20	0.29	0.15
F22	0.20	0.22	0.20	0.18	0.20	0.43	0.31

JOE Similarity: 0.26

Appendix 2

Sørensen similarity indices for *S. aureus* and *P. syringae* pv. maculicola strains (NED - yellow)

	9D9	SA001	SA002	SA003	M1D7	4i2	292b
NF3	0.11	0.22	0.12	0.21	0.12	0.17	0.10
NF3A	0.11	0.22	0.12	0.21	0.12	0.17	0.10
NF5	0.11	0.22	0.12	0.21	0.12	0.17	0.10
NF12	0.11	0.22	0.12	0.21	0.12	0.17	0.10
F7	0.22	0.33	0.24	0.32	0.24	0.25	0.19
F10A	0.11	0.22	0.12	0.11	0.12	0.17	0.10
F12	0.25	0.38	0.27	0.35	0.27	0.27	0.21
F15	0.11	0.21	0.33	0.10	0.11	0.16	0.09
F18	0.25	0.38	0.27	0.35	0.27	0.27	0.21
F22	0.25	0.50	0.40	0.35	0.27	0.36	0.32

NED Similarity: 0.21

Appendix 3

AFLP haplotype code database query report

Oklahoma State University

Monday, June 13, 2011
10:57:41 AM

AFLP Haplotype Search Results

Search Performed By: Kate Weinbrecht

Species:	Pseudomonas syringae	Strain:	F7
Pathovar:	Maculicola	Host:	Variety of Plant Species
Treatment:	NA		

Haplotype Code:

FAM 75:	0	JOE 75:	2	NED 75:	2
FAM 85:	0	JOE 85:	2	NED 85:	2
FAM 95:	0	JOE 95:	0	NED 95:	0
FAM 105:	0	JOE 105:	0	NED 105:	1
FAM 115:	0	JOE 115:	2	NED 115:	1
FAM 125:	2	JOE 125:	0	NED 125:	1
FAM 135:	0	JOE 135:	1	NED 135:	0
FAM 145:	1	JOE 145:	1	NED 145:	0
FAM 155:	0	JOE 155:	1	NED 155:	0
FAM 165:	1	JOE 165:	0	NED 165:	0
FAM 175:	0	JOE 175:	1	NED 175:	0
FAM 185:	0	JOE 185:	0	NED 185:	0
FAM 195:	1	JOE 195:	0	NED 195:	0
FAM 205:	0	JOE 205:	0	NED 205:	1
FAM 215:	0	JOE 215:	0	NED 215:	2
FAM 225:	0	JOE 225:	0	NED 225:	1
FAM 235:	0	JOE 235:	0	NED 235:	0
FAM 245:	0	JOE 245:	1	NED 245:	0
FAM 255:	0	JOE 255:	1	NED 255:	0
FAM 265:	0	JOE 265:	0	NED 265:	0
FAM 275:	0	JOE 275:	0	NED 275:	0
FAM 285:	1	JOE 285:	0	NED 285:	0
FAM 295:	2	JOE 295:	0	NED 295:	0
FAM 305:	0	JOE 305:	0	NED 305:	0
FAM 315:	0	JOE 315:	1	NED 315:	0
FAM 325:	0	JOE 325:	0	NED 325:	0
FAM 335:	0	JOE 335:	0	NED 335:	0
FAM 345:	0	JOE 345:	0	NED 345:	0

Page 1 of 1

VITA

Katelyn Dean Weinbrecht

Candidate for the Degree of

Master of Science

Thesis: AMPLIFIED FRAGMENT LENGTH POLYMORPHISM (AFLP) ANALYSIS AS A DNA TYPING TOOL FOR MICROBIAL FORENSICS

Major Field: MS in Forensic Science, Emphasis in Forensic Biology/DNA

Biographical:

Education:

Completed the requirements for the Master of Science in Forensic Sciences Oklahoma State University, Tulsa, Oklahoma in July, 2011.

Completed the requirements for the Bachelor of Science in Biology Pacific Lutheran University, Tacoma, Washington in May, 2009.

Completed the requirements for the Bachelor of Arts in Religion Pacific Lutheran University, Tacoma, Washington in May, 2009.

Professional Memberships:

American Academy of Forensic Sciences, Student Affiliate

Name: Katelyn Dean Weinbrecht Date of Degree: July, 2011

Institution: Oklahoma State University Location: Tulsa, Oklahoma

Title of Study: AMPLIFIED FRAGMENT LENGTH POLYMORPHISM (AFLP) ANALYSIS AS A DNA TYPING TOOL FOR MICROBIAL FORENSICS

Pages in Study: 80 Candidate for the Degree of Master of Science

Major Field: Forensic Sciences

Scope and Method of Study:

The objective of this research was to adapt and refine a published method for DNA profiling of microbial strains such that it can be used to attribute a plant pathogen recovered as evidence in a biocrime to a source. Amplified Fragment Length Polymorphism (AFLP) profiles were prepared from a variety of *Pseudomonas* strains pathogenic for plants, humans or both. Included among the species examined were 10 strains of *Pseudomonas syringae* pv. maculicola, and 4 stains of *Pseudomanas aeruginosa.* AFLP profiles were produced and analyzed using capillary electrophoresis with an ABI310 Genetic Analyzer. AFLP profiles were reduced to a numeric haplotype code that conveyed the size characteristics of the AFLP profiles of the different isolates from the different strains and species. Haplotype codes could then be compared and utilized to distinguish the different strains.

Findings and Conclusions:

The AFLP assay in this study exhibited over 97% reproducibility through three replicate analyses performed on each strain. Visual comparison of the electropherograms and use of a numerical haplotype code identifying each strain showed every analyzed strain of *P. syringae* pv. maculicola and *P. aeruginosa* to be unique. Discriminatory power was also assessed across pathovars with comparison to 8 strains of *P. syringae* pv. tomato, and across species with comparison to 8 strains of *Serratia marcescens* and 7 strains of *Staphylococcus aureus.* The discriminatory power of the AFLP assay was further established through pairwise similarity analysis and agglomerative hierarchical clustering analysis of all strains. AFLP analysis performed on three cultures of *P. aeruginosa* PAO1 obtained from three separate laboratories showed that AFLP analysis can detect genetic mutations that accumulate in a single bacterial strain during culturing over many years. Haplotype code discrepancies observed among replicate runs of a strain or among multiple cultures of a strain were minor, with less than 5% of the entire haplotype code being affected in each incident. Haplotype codes were organized into an AFLP database created with the ability to enter the haplotype code of an unknown strain into a query that will search the database for a match. The creation and expansion of an AFLP database allows AFLP analysis to be applied in a broader forensic context. Results of this study determine that AFLP technology is sufficiently reproducible, powerful, and reliable for use as a molecular screening tool in microbial forensics.

ADVISER'S APPROVAL: Dr. Robert Allen