EVALUATING THE SUITABILITY OF AFLP GENOTYPING TECHNOLOGY FOR DISCRIMINATING AMONG STRAINS OF *PSEUDOMONAS SYRINGAE*

By

ANDREW KENNETH TAYLOR

Bachelor of Science in Zoology

University of Oklahoma

Norman, Oklahoma

2007

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
December, 2009

EVALUATING THE SUITABILITY OF

AFLP GENOTYPING TECHNOLOGY FOR

DISCRIMINATING AMONG STRAINS OF

PSEUDOMONAS SYRINGAE

Thesis Approved:

Dr. Robert Allen

Thesis Adviser

Dr. Franklin R. Champlin

Dr. Jacqueline Fletcher

Ms. Jane Pritchard

Dr. A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGMENTS

I would first like to express my sincere gratitude to my advisor, Dr. Robert Allen for all of your guidance. You contributed as much to this project as I did, and I could not have reached this goal without your help.

I would also like to thank the rest of my advisory committee, Dr. Franklin R. Champlin, Dr. Jacqueline Fletcher, and Ms. Jane Pritchard for all of your time and effort in assisting me throughout this process.

To the entire Forensic Science department, most especially Dr. Wagner, Jun, Jay, Penelope, Phylis and Cathy, thank you. All of you have helped me in more ways than you know, and none of this would have been possible without you.

I would like to thank my family and friends, who were there for me before I began this project, and I know will be there for me after.

Finally, and most importantly, I wish to thank my wife, Sarah. Without you, I would not have had the necessary support to have made it this far in life, let alone in my academic career. Words cannot express the amount appreciation I feel, and because of that, I dedicate my thesis to you.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Few realize the importance of America's agricultural industry to our health and our nation's economy. Few also realize the fragility of that industry. A single release of an unknown plant pathogen or a widespread zoological infection could leave our agricultural sector crippled, thereby leading to an economic and health crisis. Stemming from the terrorist attacks on the United States in 2001 and from recent plant pathogen outbreaks, new security challenges are being realized. New defense efforts must focus on the possibility of attacks on, among other things, the agricultural industry.

To date, the majority of defense efforts in the United States have been focused on human safety and economic security. Many security programs are not designed to detect, nor respond to, intentional acts of bioterrorism directed toward contamination of the food supply directly or indirectly through processing and distribution systems (Oberst, 2004). It is because of this lack of security that many think that agriculture is vulnerable to a terrorist attack (Monke, 2004). Possible agroterrorism pathogens include viruses, fungi, and bacteria (Monke, 2004). The consequences of an intentional attack on the agricultural industry could include human casualties due to sickness. More likely however, a successful attack would result in a profound loss of confidence in America’s agricultural security that would lead to economic crisis. Another major potential cause of economic

crisis would be quarantines and/or embargoes on the sale of US crops to major foreign markets. A common problem complicating the detection of an agricultural attack is the high incidence of naturally occurring outbreaks. The intentional release of a pathogen at a single location could easily be mistaken for a natural outbreak and not be recognized as a criminal act (Wheelis, 2000). Additionally, a pathogen could be introduced into animal feed or fertilizer production facilities, thereby creating the potential for a large area to be affected from a single contamination site (Wheelis, 2000).

We have recently seen the problems that human pathogen contaminated food supplies can cause. In September and October of 2006, over 200 people became infected with *Escherichia coli* O157:H7 in 26 states through contact with contaminated spinach (Centers for Disease Control and Prevention, 2006). More recently, in late 2008 and early 2009, nearly 800 people were exposed to *Salmonella typhimurium* through contact with infected products containing peanut butter (Centers for Disease Control and Prevention, 2009). These relatively small-scale events demonstrate the far reaching effects problems with a nation's food supply can have. It is estimated that these two events alone caused at least nine deaths and accrued nearly $1 billion in recalls and other costs associated with controlling the problems (Neale, 2009).

Recent trends in scientific research in support of biodefense suggest that genotyping technologies could be used for the rapid detection and identification of a pathogen used in an agroterrorist attack, and attribution of that crime (Wheelis, 2000). This idea falls within the breadth of microbial forensics, a field encompassing "the detection of reliably measured molecular variations between related microbial strains and their use to infer origin, relationships, or the transmission route of a particular isolate"

(Cummings and Relman, 2002). Currently, there are several programs in existence for the detection of agricultural pests. However, ideal preparedness for an agroterrorism event would include techniques and protocols currently used in forensic laboratories that allow for genotyping of plant pathogenic organisms suspected of being a part of a terrorist incident (Fletcher *et al*., 2006). Some research has been done in this area; however, more is needed given the possible severity of the consequences of an attack on America's agricultural infrastructure.

Amplified fragment length polymorphism (AFLP) mapping technology is a genotyping technique that utilizes restriction enzymes to produce DNA fingerprints unique to specific organisms (Vos *et al*., 1995). This technique has been shown to be reliable and does not require extensive amounts of time or effort. Another advantageous characteristic of this technique that makes it suitable for microbial forensics is that it does not require prior knowledge of the genome of the organism being studied. This feature of the AFLP technique could be very important for rapidly producing a DNA "fingerprint" of a bioweapon so that the attribution process could be accomplished quickly. For these reasons AFLP is commonly used worldwide for genotyping microorganisms (Wittendorp, 2007). AFLP has also proven to be highly reproducible and cost effective, making it a likely choice for future genomic research (Mueller and Wolfenbarger, 1999).

Pseudomonas species are opportunistic bacterial pathogens, which can be highly transmissible, virulent, and robust, all necessary traits for an effective pathological agent (Monke, 2004). *Pseudomonas* species have also been reported to be intrinsically resistant to multiple antibiotics, a characteristic that could prove disastrous if this microorganism were utilized in a destructive manner (Holmes, 1998). *Pseudomonas syringae* was

employed in this study because it is a common, economically important plant pathogen found in the same taxonomic class as several, far more virulent, plant pathogens.

This study features an evaluation of amplified fragment length polymorphism (AFLP) technology for its suitability in genotyping strains of *P. syringae*, an opportunistic pathogen of a wide range of plant species. If AFLP is effective in discriminating *P. syringae* strains in the same way it has proven effective for discriminating strains of *Serratia marcescens* (Beauman, 2007), this study will extend and deepen our understanding of the strengths and weaknesses of AFLP analysis as a forensic tool. We hypothesize that AFLP mapping technology can be used as an aid in forensic microbiological investigations, if it is suitable for correctly and accurately discriminating among *P. syringae* strains. We also propose that any weaknesses revealed in the technology can be overcome to further enhance the suitability of the assay for microbial forensics

CHAPTER II

REVIEW OF LITERATURE

The Importance of Agriculture

In 2004, the Homeland Security Presidential Directive (HSPD) – 9 noted the vulnerability of the United States agriculture and food systems to disease, pest, or poisonous agents delivered by acts of terrorism. In that directive, a national policy to defend the agricultural system against terrorist attack was mandated. Included in the initiative was the establishment of university-based centers of agriculture and food safety research (Bush, 2004). Other sources similarly contend that America's agricultural system is extremely vulnerable and its security is substantially lacking (Fletcher *et al*., 2006). Agriculture makes up an enormous part of the nation's economy. The United States Department of Agriculture (USDA) reported that the U.S. exported $115,450,000,000 worth of agricultural products in 2008 and they expect that number to continue to rise (USDA, 2009). In recent years, the agriculture industry has generated over $1 trillion worth of business for the United States (Monke, 2004). If certain, key food commodities, such as rice, wheat or corn, were to be significantly impacted by a terrorist attack, the results could be catastrophic. The impact of an attack on the food supply would be felt by more than just the farmer. Businesses, such as farm suppliers,

food transportation companies, grocery stores, restaurants, equipment distributors, and finally consumers would all be victims of the attack. Effects of a biological attack on the agricultural sector would not be limited to economic losses, but are likely to be much more far reaching. In the event of an attack, consumer confidence in government officials and their ability to keep America secure could erode, and negative effects on nutrition and the environment could also occur (Fletcher *et al*., 2006). Given America's position as a world power, a successful attack committed against the United States could have global consequences, possibly including "disruption of markets, difficulties sustaining an adequate food supply, and the potential spread of disease and infestations throughout the nation and the world" (NRC, 2002).

Recent Incidents Involving Examples of Agricultural Associated Pathogens

The emergence of agricultural pathogens has been seen recently in several cases. In 2006 and 2007 the human pathogen *E. coli* O157:H7 contaminated American spinach crops and resulted in the hospitalization of over 100 individuals, three of whom died (Centers for Disease Control and Prevention, 2006). The outbreak also was reported to have cost the spinach industry over $300 million in recall and containment costs (Park, 2006). Another event, involving the human pathogen *Salmonella*, occurred in late 2008 and early 2009. Nearly 800 individuals, in 46 states, were infected by *Salmonella*, resulting in at least 3 deaths (Centers for Disease Control and Prevention, 2009). A recent example involving livestock was seen in the United Kingdom when a foot and mouth disease (FMD) outbreak occurred in 2001. With over 2,000 confirmed cases, the English

government ordered the destruction of nearly 4 million animals which resulted in direct and indirect economic losses estimated in the tens of billions of dollars (NRC, 2002).

One problem associated with plant diseases and particularly important when discerning a bioterrorism event is the sheer abundance of naturally occurring diseases every year. In the United State alone, there are over 50,000 known plant pathogens, which account for very significant crop losses (Fletcher *et al*., 2006). Another problem that agricultural biosecurity programs will have to deal with is the amount of cropland that is present in the United States. Land used for agricultural purposes makes up about 442 million acres, or 20% of all land in the United States (USDA, 2002). Finding and recognizing an outbreak could be likened to finding a needle in a haystack.

Agricultural Bioterrorism Throughout History

Biological warfare targeting agricultural infrastructure has existed for centuries. Many world powers, including the former Soviet Union and the United States, have developed extensive biological weapons programs in the past. The United States has researched weaponized anthrax, foot and mouth disease and rice blast (Fletcher *et al*., 2006). There have also been reports of research on the fungus *Fusarium* for the possible destruction of coca plants in Colombia (NRC, 2002). Under Saddam Hussein, in the 1980s and 1990s, Iraqi bioweapons development focused part of their efforts on anticrop weapons, including wheat smut (NRC, 2002). While no documented cases of deliberate use of pathogens to attack American crops or livestock have been seen, preparedness is necessary in light of past, and possibly ongoing, research in this area by potential enemies.

Pseudomonas syringae

A bioweapon that could be used effectively against plants or animals, or transmitted to humans through food must be highly contagious, virulent, and able to survive in the environment (Monke, 2004). Certain *Pseudomonas* species have exhibited all of these traits. In the past 2 decades, *Pseudomonas aeruginosa* has emerged as a human pathogen causing disease outbreaks, particularly among cystic fibrosis (CF) patients (Holmes, 1998). The fact that many *Pseudomonas* species are inherently resistant to multiple antibiotics makes this microbe even more threatening (Holmes, 1998). *P. syringae* is a bacterial species that is known to cause disease in all major groups of higher plants (Hirano, 1990). The species is divided into over 50 pathovars (Clerc *et al.*, 1998). The pathovar concept was conceived by Young *et al.*, in an effort to divide species at the sub species level based upon the plant host range and the symptoms caused (1978). *P. syringae* pathovar (pv.) *tomato*, the model organism used in this study, is the causal pathogen for bacterial speck of tomato, a serious disease that can result in reduced yields and fruit quality.

Microbial Forensics

Defense against *P. syringae*, or any other biological agent, relies on the ability to rapidly identify the organism involved and determine its origin (Monke, 2004). Whether an agent is introduced naturally or disseminated deliberately, identification and determination of source are two critical elements necessary for attribution and containment (Fletcher *et al*., 2006). Laboratory and field resources are strained to deal with naturally occurring outbreaks (NRC, 2002), and due to understaffed and

underfunded laboratories, many believe that preparedness is not adequate. The National Research Council's (NRC) 2002 publication, "Countering Agricultural Bioterrorism" highlights the need for "aggressive research in both science and technology to improve our ability to prevent, detect, respond to, and recover from biological attacks on plants and animals." The knowledge and techniques used for protecting plants and animals against naturally occurring pathogens is a starting point for these efforts. However, they only begin to fill the need. Rapid detection and diagnosis are critical in any investigation, especially ones where dangerous organisms are present (NRC, 2002). Microbial forensics is a field that incorporates the skills necessary to investigate an agroterrorist attack. Genomic information has great potential to aid in investigating a bioterrorist attack because it can suggest geographic origins, microbial sources, and unique complexities of pathogens, all of which can assist with the characterization and attribution process (NRC, 2002).

Detection of plant pathogens may soon become a responsibility of forensic microbiologists at the federal, state, and even local levels who are trained in multiple techniques for microbial genotyping. Techniques that may be used for genotyping include analysis of restriction length fragment polymorphism (RFLP), simple sequence repeats (SSR), and amplified fragment length polymorphism (AFLP) (Rudin and Inman, 2001).

Restriction Fragment Length Polymorphism

RFLP is based on the size differences among DNA fragments in multi-fragment banding patterns that stem from variations in the spatial arrangement of restriction enzyme recognition sites. Electrophoresis, Southern blotting and hybridization to labeled

probes reveal the varying banding patterns in RFLP digests. The presence or absence of restriction sites will define the length of a fragment (Rudin and Inman, 2001).

Simple Sequence Repeats

Simple sequence repeats, also known as microsatellites or SSRs, are DNA sequences that are tandemly repeated a variable number of times and are found at different locations in an organism's genome (Tamaki and Jeffreys, 2005). Assessments of SSRs, unlike RFLP analysis, can be performed with small amounts of DNA. However, the SSR technique requires prior knowledge of the genome sequence (Rudin and Inman, 2002). This characteristic does not make SSR analysis suitable for use in the proposed in this study because a potential agricultural outbreak may be caused by an organism in which the genomic sequence is unknown.

Amplified Fragment Length Polymorphism

In the early 1990s, Keygene N.V. developed AFLP technology (Wittendorp, 2007), a technique that produces DNA fingerprints by combining RFLP analysis of genomic DNA with polymerase chain reaction (PCR) amplification (Vos *et al.*, 1995). AFLP and RFLP analyses are similar, but the former is less labor intensive (Groenewald *et al*., 2005) and reduces the complexity of analysis by utilizing selective amplifications, which reduce the number of amplicons produced. In the AFLP technique, genomic DNA is first cut with two restriction enzymes that generate restriction fragments for amplification. Oligonucleotide "linkers" having target sites for PCR primers that direct their amplification are then attached to the "sticky ends" of the restriction fragments (Vos

et al., 1995). The primers used in AFLP are usually 17 to 21 nucleotides in length and target sequences in the linkers ligated to the restriction fragments (Vos *et al*., 1995). Minor variations in the amplification constraints, such as thermal cycler programming, template concentration, and PCR parameters are reported to not affect AFLP results, thereby enhancing its reliability (Wittendorp, 2007).

AFLP technology has been utilized worldwide and is fast becoming a very popular genotyping technology, especially in the area of plant pathogen genomics. AFLP has been used effectively to reveal genetic differentiation among strains of *Xanthomonas albilineans*, a plant pathogen closely related to *Pseudomonas* (Shaik *et al*., 2008). In South Africa, AFLP technology was used to genotype isolates of *Fusarium oxysporum* f. sp *cubense*, which causes Fusarium wilt of bananas (Groenewald *et al*., 2005). In another study, fragments produced by AFLP on methicillin-resistant and –susceptible strains of *Staphylococcus aureus* were consistent with the known genomic sequences found in GenBank (Savelkoul *et al*., 2007). This example indicates that the AFLP assay can produce very reliable genotyping information. In a study aimed at the genetic characterization of multiple pathovars within the *P. syringae* species, AFLP showed the capability to distinguish the individual pathovars from one another (Sisto *et al.*, 2007), and in a study designed to assess the genetic diversity within *P. syringae*, AFLP proved capable of distinguishing among multiple *P. syringae* pv. *tomato* strains (Clerc *et al.*, 1998). An advantage of AFLP over other techniques is that it can rapidly generate hundreds of highly reproducible genetic markers from any organism (Mueller and Wolfenbarger, 1999). When compared to RFLPs, SSRs and other genotyping methods,

AFLP is time- and cost-efficient, reproducible, highly discriminatory and requires minimal effort.

CHAPTER III

METHODOLOGY

Overview of the AFLP Method

Amplified fragment length polymorphism mapping technology has demonstrated the capability for DNA based discrimination within numerous eukaryotic and prokaryotic species (Mueller & Wolfenbarger, 1999). However, the technology has only recently been explored as a forensic tool for attribution of crimes using bioweapons against agricultural targets (Beauman, 2007). Effective and standardized procedures for investigating potential agricultural crimes, for collecting evidence, and for genetic analysis of suspected agents are lacking (Fletcher *et al*., 2006). Studies such as ours contribute to establishing preparedness in forensic microbiology in general and in agricultural biosecurity in particular. The AFLP process utilizes restriction enzymes, *Mse1* and *EcoR1*, to cut the genomic DNA, adaptors which interact with those restriction sites, and labeled primers that anneal to the adaptor sequences in order to complete the selective amplification.

P. syringae Strains

14 bacterial cultures of *P. syringae* pv. *tomato* (shown below in Table 1) were kindly provided by Dr. Carol Bender, Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, Oklahoma. The specific strains used in this study

were chosen based upon their diverse geographic origins as well as genetic similarities in multiple-locus VNTR (variable number tandem repeat) analysis (MLVA) sequences shown by Dr. Christy Baker among several of the strains (30555, T1, B125, TF1, PT17, T4B1, 188B, and CPST232) (Baker, 2008). These strains should then challenge the ability of AFLP to discriminate between strains that are similar as well as others known to demonstrate genetic differences.

Strain	**Geographical origin**	**Year isolated**
1008	USA	1942
1318	Switzerland	1971
188B	Ontario, Canada	?
3357	New Zealand	1972
30555	Tasmania	1978
880	Yugoslavia	1953
B125	Canada	1981
CPST 147	Czech Republic	1993
CPST 232	Slovakia	1993
PT17	USA	1983
RG4	Venezuela	1985
T1	Canada	1983
T4B1	Canada	1981
TF1	USA	1997

Table 1 *P. syringae* pv. *tomato* strains (including year and location of origin) used in this study.

Upon receipt of the bacterial strains, each was plated on individual blood agar plates until visible colonies formed. An individual colony from each strain was then inoculated into 2 mL of Mueller-Hinton growth medium and shaken at 28 °C overnight to obtain a dense culture. A 1 mL aliquot of each sample was centrifuged at 10,000 xg for three minutes at room temperature to obtain a cellular pellet for DNA extraction.

DNA Extraction

Extraction of DNA from the bacterial cellular pellet began with the re-suspension of the pellet in 500 µL of extraction buffer [0.02g of lysozyme dissolved in 0.5mL of TNE (10mM Tris-Cl pH 8.3, 0.2M NaCl, 1mM EDTA)]. The mixture was incubated at room temperature for ten minutes to weaken the cell wall of the bacteria. 25µL of 20% sodium dodecyl sulfate (SDS), 25 µL of 20 mg/mL of protease K (in 10mM Tris-Cl, 20mM $CaCL_2$, and 50% glycerol), and approximately 5 mg of ribonuclease A (Sigma Chemical Company, St. Louis, MO) were then added to each extract and incubation was continued for 1 hour at 65°C. The samples were extracted with an equal volume of phenol:chloroform/isoamyl alcohol (9:0.96:0.4). Each sample was centrifuged at 10,000 xg for 3 minutes to obtain phase separation. The aqueous (top) layer, which contains the genetic material, was removed and subjected to extraction with an equal volume of chloroform:isoamyl alcohol (24:1), centrifuged at 10,000 xg for 3 minutes and the aqueous layer was again removed and placed in a clean 1.8 mL microfuge tube. Two volumes of 95% ethanol were added to the samples to precipitate the bacterial DNA.

After mixing, a sterile inoculating loop was used to “capture” the fibrous clot of DNA, which was re-suspended in 40 μL of TE^{-4} (10mM Tris-Cl, pH 8.0, 0.1mM EDTA).

DNA Quantitation

The DNA recovered from each bacterial strain was quantified spectrophotometrically at 260 nm and 280 nm using a Nanodrop ND-1000 microspectrophotometer (Thermo Scientific, Wilmington, DE). Recovered DNA was also assessed for purity using the 260/280 ratio, which was generally about 1.80 (data not shown).

DNA Digestion

Aliquots containing 300-500 ng of each DNA sample were digested sequentially with two restriction endonucleases. First, DNA was digested with 10 U of *MseI* (New England BioLabs, Inc., Ipswich, MA) in a total volume of 10 μL following instructions from the supplier. The samples were incubated at 37°C for 1 hour and then in a 65°C heat block for five minutes to halt further enzyme activity. Samples were then chilled on ice for five minutes prior to the second digestion. The second digestion mixture contained 10 μL from the first digest, 1μL containing 20 U of *EcoRI* (New England BioLabs, Inc.), 2 μL of 10x *EcoRI* buffer (New England BioLabs, Inc.), and 8 μL of H_2O for a final volume of 20 μL. Samples were again sequentially incubated for 1 hour at 37°C and 65°C in a heat block for five minutes to inactivate *EcoRI*, after which they were placed on ice for five minutes. The extent of digestion was assessed by electrophoresis in a 1.0% agarose gel, using 1X TAE buffer (10 mM Tris-acetate pH 8.3 with 1 mM EDTA). A

volume of 7 μL of each digest was mixed with 3 μL of a loading buffer-dye mixture (5x TAE and 5% ficoll 400). Electrophoresis was initiated at 50V until the tracking dye entered the gel and then at 70V for 45 minutes. Digested DNA patterns were visualized by ethidium bromide staining (0.2% in dH_2O) and ultraviolet illumination. The stained gel was placed on a UV transilluminator to view and photograph the digest. A smear of ethidium bromide staining over the lower half of the gel track and the absence of a distinct, intensely stained band near the sample well were considered adequate proof of complete sample digestion (Figure 1).

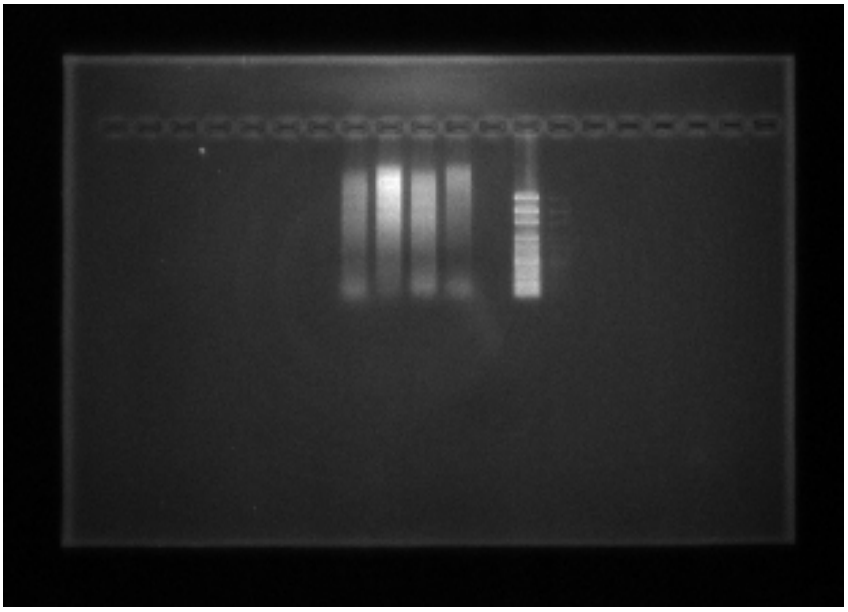

Figure 1 Ethidium bromide stained restriction digests (1st lane – strain 30555, 2nd lane – strain 3357, 3rd lane – strain B125, 4th lane – strain 880, and the last lane – size standard).

DNA Ligation

The *EcoRI* and *MseI* restriction fragments were ligated to oligonucleotide adaptors supplied in the AFLP Microbial Typing Kit (Applied Biosystems). The adaptors

serve as priming sites for PCR amplification of restriction fragments in the first of two amplification reactions. For the ligation, 1 μL of 10x T4 ligase buffer, 1 μL of 0.5M NaCl, 1 μL of 10 U/ μL *MseI*, 2.5 μL of 20 U/μL *EcoRI*, and 2.5 units of T4 DNA ligase at 400 U/μL (all from New England BioLabs, Inc.), and 3.5 μL of dH_2O were combined to create an enzyme master mix. In a separate tube, 1 μL of doubly digested DNA, 1 μL of 10X T4 ligase buffer, 1 μL of 0.5M NaCl, 0.5 μL of bovine serum albumin (New England BioLabs, Inc.) at 10mg/ml, 1 μL of *MseI* oligonucleotide adaptor, 1 μL of *EcoRI* oligonucleotide adaptor, 1 μL of enzyme master mix and 3.5 μL of dH_20 were combined. Each sample was incubated at 37°C in a water bath for 2 hours and then at room temperature overnight. 189 μL of TE^{-4} buffer [10 mM Tris, 1 mM EDTA (pH 8.0)] was then added to each sample. Diluted ligation reaction mixtures were stored at 2-6°C if not immediately used.

AFLP-1 – Pre-selective Amplification

The first of two polymerase chain reaction (PCR) amplification steps is non-selective, meaning that all restriction fragments containing adaptors on each end can be amplified. In a 200 μL PCR tube, 4.0 μL of the diluted restriction-ligation reaction from above, 0.5μL each of the AFLP *EcoRI* and *Mse1* core primer solutions (supplied with the AFLP kit and used as primers for amplification, targeting the oligonucleotide adaptor ligated to each restriction fragment in the previous step), and 15 μL of AFLP amplification core mix (all from Applied Biosystems) were combined for a total volume of 20 μL. The samples were then placed in a 9700 thermal cycler from Applied Biosystems, and were subjected to the PCR cycling parameters shown in Table 2.

Hold	Cycles Pre-selective amplification To be repeated 20 times			Hold
72°C 2 min.	94°C 20 sec.	56°C 30 sec.	72°C 2 min.	4°C

Table 2 Thermal cycler parameters for AFLP pre-selective amplification. (Applied Biosystems, 2007)

Following completion of the cycling program, 10 μL of the preselective product was combined with 190 μL of TE^{-4} and mixed with vortex agitations. The product was stored at 2-6°C. The other 10 μL of preselective product was electrophoresed in a 1% agarose gel in TAE buffer using the same protocol as above to assess the effectiveness of the preselective PCR reaction (Figure 2). Visualization of ethidium staining in a gel track containing pre-selective amplicons provided the first opportunity to confirm ligation and amplification because only ~10 ng of restriction fragments are ligated with adaptors, which are then diluted to ~200 μL final volume (making the fragment concentration ~50 pg/μL). 500 pg of restriction fragments would not likely be detectable in an agarose gel unless amplification was successful, a process dependent on successful restriction digestion and ligation.

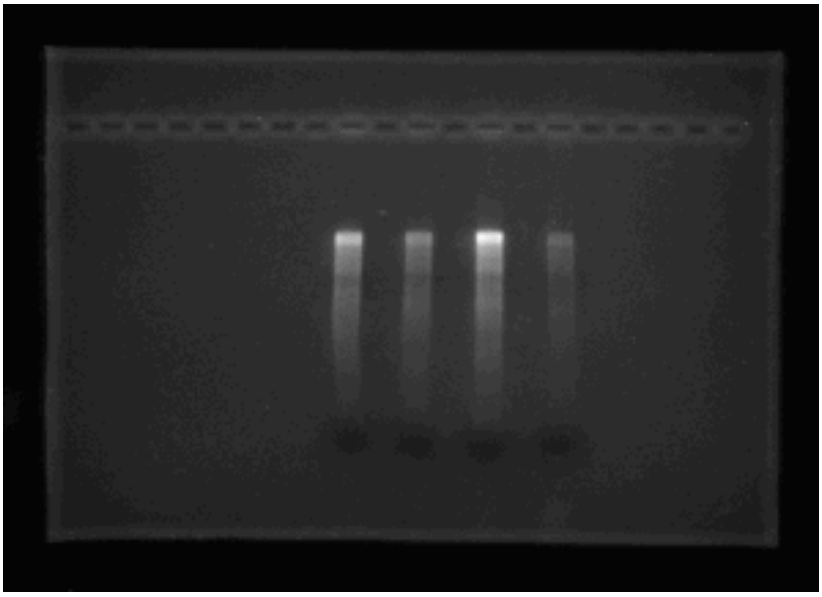

Figure 2 Ethidium bromide stained pre-selective amplification products (1st lane – strain 30555, 2nd lane – strain 3357, 3rd lane – strain B125 and 4th lane - strain 880).

AFLP-2 – Selective Amplification

Bacterial genomes are large, complex, and produce thousands of restriction fragments when digested with *Mse1* and *EcoR1*. It is therefore necessary to reduce the complexity of the AFLP profile to simplify the genetic "fingerprint" produced using AFLP, and thus also simplifying the analysis. A second, more selective PCR reaction, is therefore included as part of the AFLP typing process for microbial genomes. Selectivity in the second reaction is achieved using primers identical to those used in the first non-selective reaction, but extending 1-2 nucleotides beyond the 5' terminus of the adaptor sequence into the flanking nucleotide sequence of the restriction fragment. Successful amplification will occur only with those restriction fragments bearing the adaptor AND having a complimentary nucleotide sequence 1-2 bases upstream from the adaptor. A 4 to 64-fold reduction in the complexity of PCR products compared to the first non-selective

PCR reaction can be expected depending upon the number of selective nucleotides added to the core primer. Selective PCR reactions in this work consisted of 1.5 μL of the preselective amplification product, 0.5 μL *MseI* –A primer (the *Mse1* primer plus an additional adenosine residue at the 3' end of the primer), 0.5 μL selective *EcoRI* primer labeled with one of three fluorescent dyes [FAM (blue), JOE (green), or NED (yellow)], and 7.5 μL of AFLP Core Amplification Mix (all reagents from Applied Biosystems). The primer labeled with NED (yellow) contains a single nucleotide extension of cytosine, that labeled with FAM (blue) has a single nucleotide extension of adenosine and that labeled with JOE (green) has a single nucleotide extension of guanosine. Thus, three distinct PCR reactions are prepared from each pre-selective PCR sample. The components are then combined and subjected to the cycling program shown in Table 3.

Hold	Cycle Selective Amplification			Number of Cycles
94°C 2 min.	94°C 20 sec.	66°C 30 sec.	72°C	2
-	94°C 20 sec.	65°C 30 sec.	72°C	2
-	94°C 20 sec.	64°C 30 sec.	72°C	2
-	94°C 20 sc.	63°C 30 sec.	72°C	2
-	94°C 20 sec.	62°C 30 sec.	72°C	2
-	94°C 20 sec.	61°C 30 sec.	72°C	2
-	94°C 20 sec.	60°C 30 sec.	72°C	2
-	94°C 20 sec.	59°C 30 sec.	72°C	2
-	94°C 20 sec.	58°C 30 sec.	72°C	2
-	94°C 20 sec.	57°C 30 sec.	72°C	2
94°C 2min	94°C 20 sec.	56°C 30 sec.	72°C	20
60°C 2 min		-		1
4°C forever		-		1

Table 3 Thermal cycler parameters for AFLP selective amplification.

Capillary Electrophoresis

Once selective amplification is complete, visualization of amplicons constituting the AFLP profile was accomplished using capillary electrophoresis (CE). CE is used instead of agarose (or other separation) gels, due to multiple advantages including higher resolution and faster analysis (Butler, 2005). A CE system such as the 310 Genetic Analyzer from Applied Biosystems, employs a capillary connected to a high voltage

power supply to conduct electrophoresis. Fluorescent PCR products in individual sample tubes are electro-injected into the capillary, which moves to a buffer solution, which is the cathode used for electrophoretic separation of the PCR product.

To prepare samples for the genetic analyzer, 1.0 μL of PCR product was added to 24.5 μL of Hi-Di formamide containing 0.5 μL of LIZ labeled size standards (all supplied by Applied Biosystems). The selective PCR products amplified by each fluorescent primer (blue, green, and yellow) were prepared individually in separate tubes to eliminate possible fluorescent "pull-up", in which signal from one color "bleeds" into a different color channel. Each tube was placed into a 48-well sample tube rack, which was fitted onto the robotic of the ABI Prism 310 Genetic Analyzer (Applied Biosystems). Each sample was subjected to capillary electrophoresis for 24 minutes at 60°C.

Data Interpretation

Data produced by the 310 Genetic Analyzer were captured as electropherograms (shown in Figures 3A-C for *P. syringae* pv. *tomato* strain T4B1). Electropherograms show the migration of the PCR products throughout the run. The products are detected as they pass a window in the capillary continuously irradiated by the laser (Frazier *et al*., 2000). The electropherograms show the fragment sizes on the abscissa (in base pairs) and the intensity of the fluorescence on the ordinate [expressed as relative fluorescence units (RFU)] (Jain *et al*., 2005). The electropherograms produced are contained within three different color channels (Blue – FAM dye, Green – JOE dye, Yellow – NED dye) based upon the dyes covalently linked to primers provided with the AFLP kit. For genetically distant bacterial strains, one would expect that the relative sizes and intensities of

amplicons in the profile should differ due to the different locations of the restriction sites within each bacterial genome, thereby allowing for discrimination among the different strains. Electropherograms for *P. syringae* pv. *tomato* strain T4B1 are shown in Figures 3A-C.

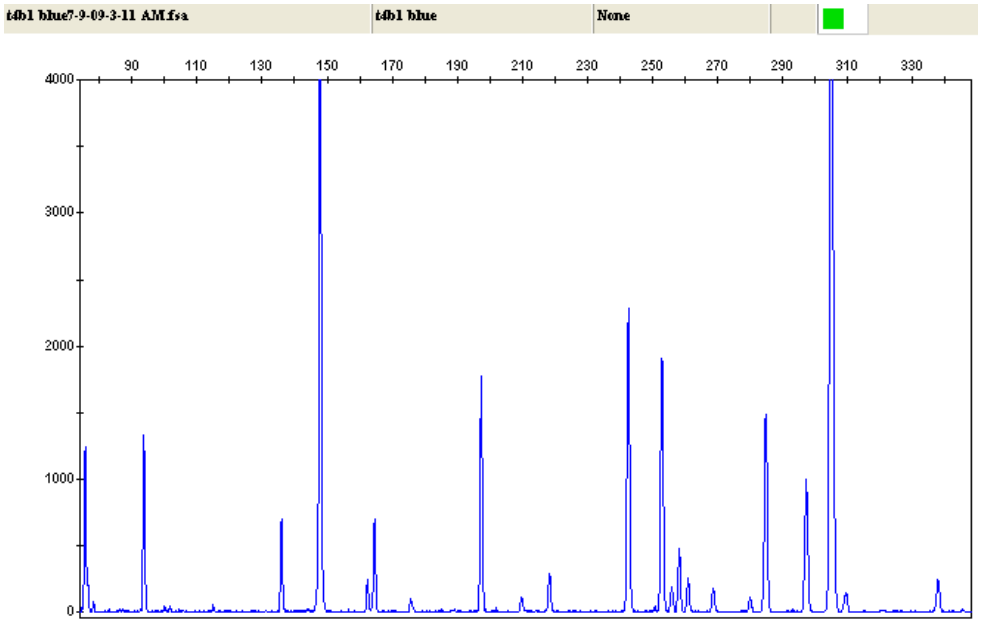

Figure 3A Electropherogram produced by analyzing *P. syringae* pv. *tomato* strain T4B1 with the blue (FAM) labeled primer. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

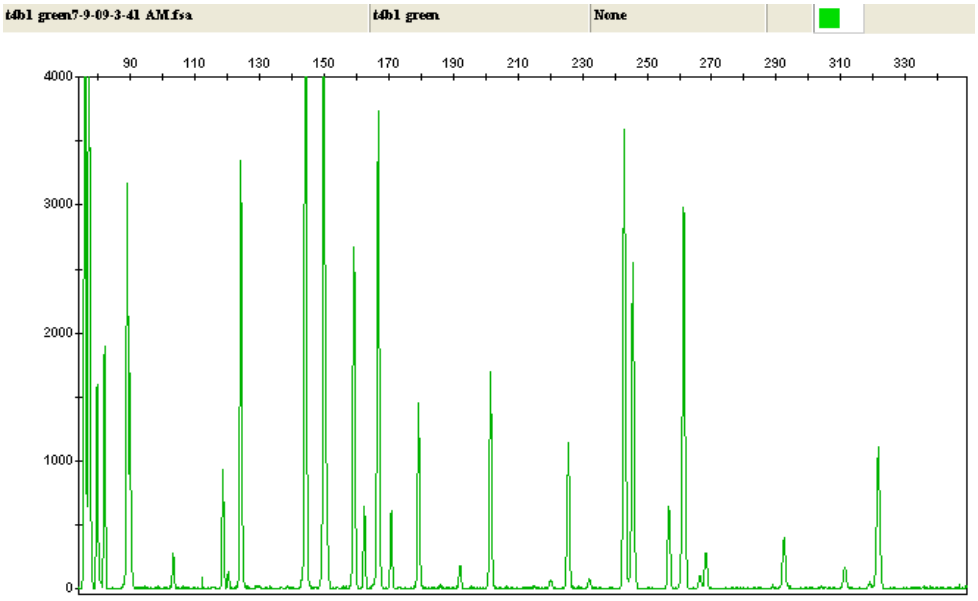

Figure 3B Electropherogram produced by analyzing *P. syringae* pv. *tomato* strain T4B1 with the green (JOE) labeled primer. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

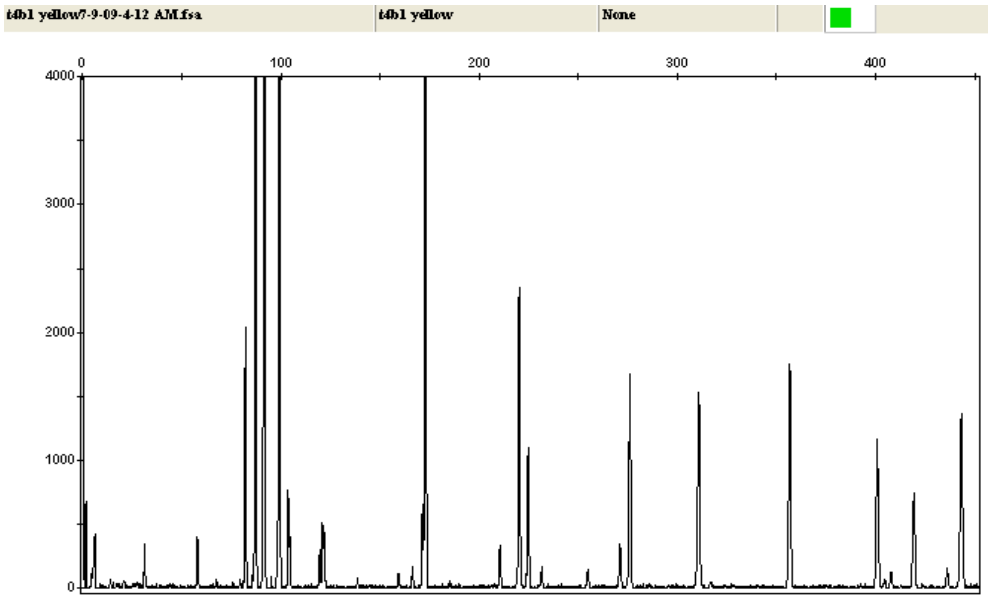

Figure 3C Electropherogram produced by analyzing *P. syringae* pv. *tomato* strain T4B1 with the yellow (NED) labeled primer. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

CHAPTER IV

FINDINGS

Although the goal of this study was to assess the ability of the AFLP technique to be a reliable and powerful tool for a forensic microbiologist charged with the investigation of an outbreak involving an agricultural pathogen, we also evaluated several aspects of the AFLP methodology in an attempt to determine potential causes of poor results.

Key Elements of the AFLP Method

For best results, the AFLP technique requires the use of a known quantity of DNA. This requirement was underscored by the fact that, in this work, altering the amounts of several of the reagents (including both restriction enzymes and T4 DNA ligase) did not correct problems that were seen due to incorrect starting quantities of DNA. The accuracy of the DNA quantitation was then assessed. The Nanodrop spectrophotometer used in this study accurately measures nucleic acids and the results served as the basis for calculating the amount of template DNA processed in the AFLP reaction. However, the relative proportions of RNA and DNA were unknown in this study. RNase A treatment of the nucleic acid extracts greatly reduced the nucleic acid concentration of the extracts and greatly enhanced the quality of the AFLP results

(data not shown), suggesting that a significant portion of the nucleic acid recovered from extracts was RNA and not DNA.

AFLP Assay Optimization

To achieve consistent results, several process parameters were experimentally varied. Varying AFLP reagent concentrations produced no improvement in test performance, so the manufacturer's recommended protocol was largely followed. Likewise, a report from another laboratory working with the AFLP procedure was used as a guide for troubleshooting (Clarke & Meudt, 2008). One report described different storage methods for the AFLP digestion and ligation products than those recommended by the manufacturer of the AFLP typing kit, storing the digestions products at -20°C instead of at 4°C and the ligations products at -70°C instead of 4°C. Although changing this step in the protocol did not lead to the improvement of results, these storage conditions were maintained, as the author believes they help to minimize any unwanted restriction enzyme activity following the restriction and ligation steps.

New shipments of every reagent in the AFLP process were ordered and tested. When the assay was re-run using the newest replacement reagents with the slightly altered protocol mentioned above, consistent and positive results were achieved.

The struggles experienced in this study highlight the importance of verifying products by agarose gel electrophoresis at the end of restriction digestion and after the first non-selective amplification step. When expected products from the first PCR reaction were not visible on a gel, it was clear that subsequent steps in the process would fail as well, so proceeding would be unnecessary.

Reproducibility

Without sufficient reproducibility, the AFLP assay would not be an effective tool for microbial forensics. In the event of a pathogen outbreak at multiple locations, samples from all sites should be compared to one another. If AFLP profile differences were found, one would need to be confident that they were due to genotypic differences between the organisms being compared and not to a lack of reproducibility within the assay. To test the reproducibility of the AFLP process, three replicates were completed for each of the 14 *P. syringae* pv. t*omato* strains tested. Reproducibility was assessed based upon the relative locations of fluorescent restriction fragments visualized with the naked eye in the AFLP electropherograms produced from the different bacterial strains. Although largely qualitative, this approach possesses a quantitative element in that only fragments meeting a specific fluorescent threshold were scored and included in the AFLP profiles.

For human identification testing, the threshold for determining which peaks on an electropherogram to analyze is normally set at 150 relative fluorescent units (RFU). Unfortunately a static threshold of 150 RFU could eliminate useful information in the AFLP profile that falls below the threshold in some electrophoretic runs. This potential problem is illustrated in Figures 4A -D. Figures 4A and 4B depict two separate runs of strain TF1, shown at the same magnification. The second image (Figure 4B) reveals lower peak heights than the first (Figure 4A). Many of these peaks would be eliminated from analysis with a threshold of 150 RFU. However, as shown in Figure 4D, when the electropherogram of the same run is magnified many of the peaks in the AFLP profile are visible, and are consistent with the stronger profile (Figure 4C).

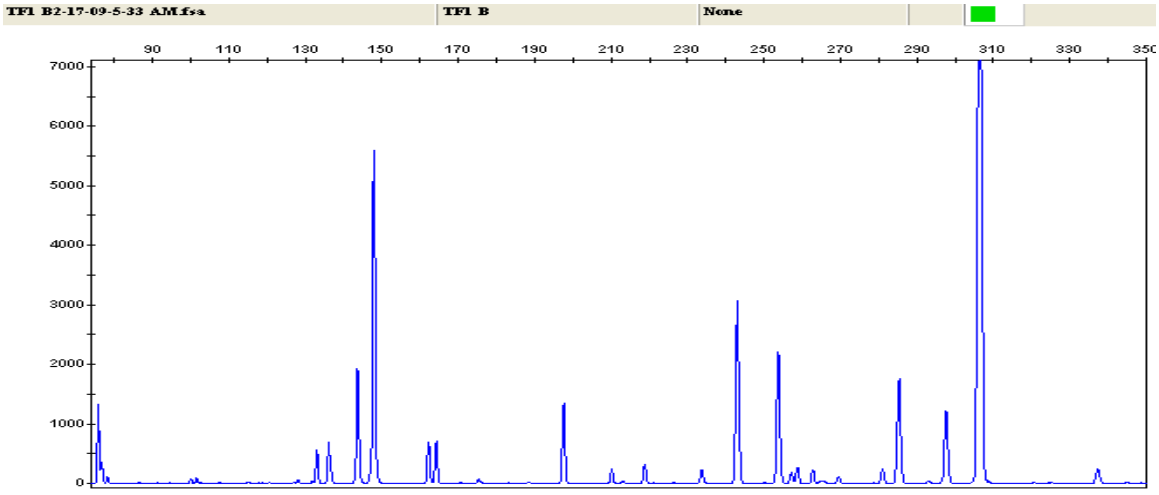

Figure 4A A run of *P. syringae* pv. *tomato* strain TF1 with no magnification. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

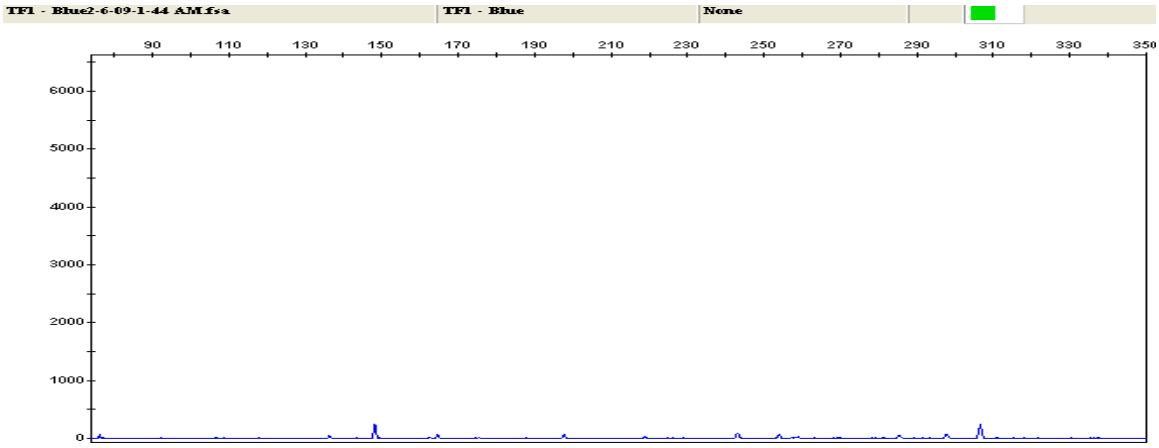

Figure 4B A second run of *P. syringae* pv. *tomato* strain TF1 with no magnification. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

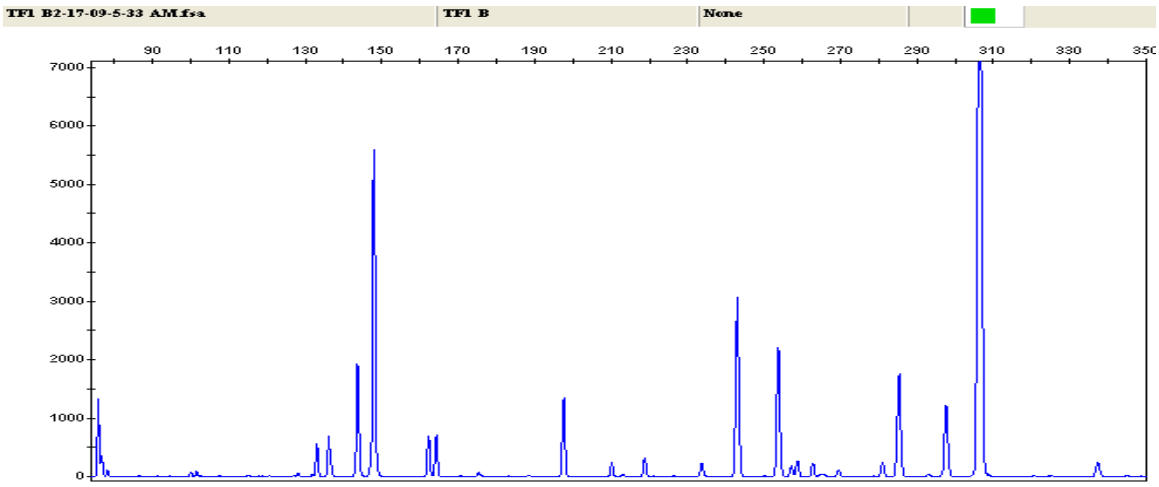

Figure 4C A run of *P. syringae* pv. *tomato* strain TF1 with no magnification The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

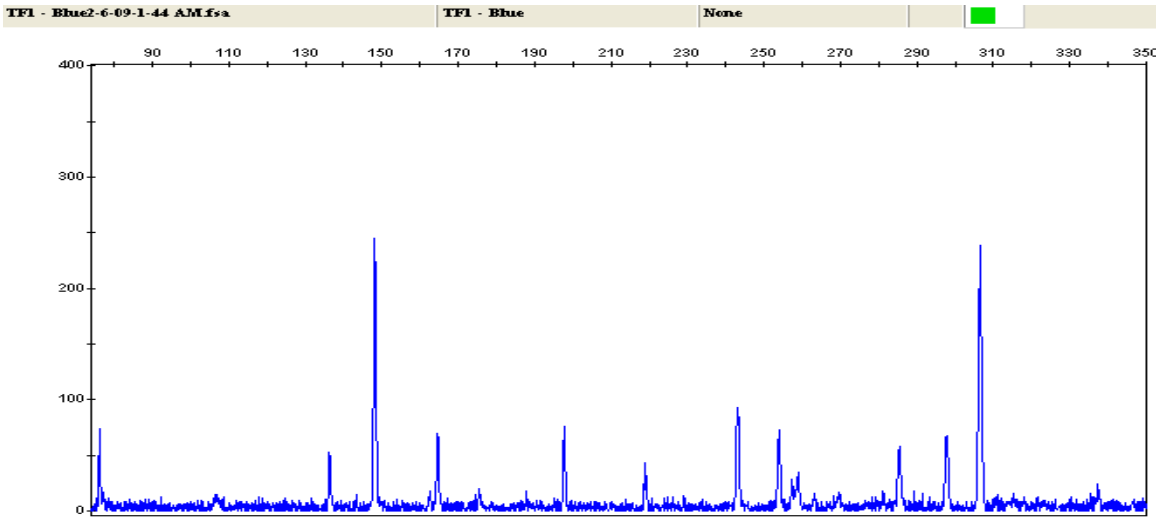

Figure 4D A run of *P. syringae* pv. *tomato* strain TF1 magnified to show the amplicons. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

This example shows that a "set" threshold is less than optimal for data analysis when the level fluorescence varies from run to run. The weak AFLP profile shown in Figures 4B and 4D illustrates some of the inconsistency seen during this study. Some results, of even lower quality than seen in Figures 4B and 4D, were unusable. However,

once suitable reagents were used, the assay proved to be much more robust and reliable. These findings were discussed to show that even if expected peak heights are not obtained, a profile may still be present. In the profile above, 94% of the peaks found in the blue channel were reproducible in all three runs done on strain TF1, despite the fact that one of the runs was clearly of much lower quality than the other two.

To accommodate for the potential variation between runs, a normalized threshold was utilized. One half of the average peak height of all scored peaks (between 75 and 350 basepairs in length) within a run was set as the threshold. A cutoff of 75 and 350 basepairs were used because the few peaks seen outside of this range, proved variable and unreliable in their size estimates. The sliding, normalized threshold allowed for runs yielding different peak height ranges to be compared with one another. In each profile, the same major peaks were included in the analysis regardless of their intensity relative to the other runs. Shown below in Table 4 are the peak sizes and heights of the three replicates of strain TF1. Despite the differences between the first two runs and the third, very similar AFLP profiles were still produced.

	Run 1		Run 2		Run 3	
	Size	Height	Size	Height	Size	Height
	75.95	806	76.03	809.99	76.19	74
	94.03	939	94.15	840.16		
	136.2	444	136.37	454.54	136.33	53
	147.87	2563	148.05	2166.16	148.24	245
	165.02	498	165.22	448.69	165.61	70
	197.64	1037	197.88	1238.28	197.7	76
	243.27	1148	243.57	949.41	243.08	93
	253.41	1191	253.72	1072.98	253.82	73
	258.95	340	259.01	400.19		
	285.72	917	286.07	938.15	285.38	58
	297.36	674	297.97	544.58	297.79	68
	303.54	3061	303.91	3664.75	303.51	239
Threshold (Avg/2)=		335		337		52

Table 4 Three replicates of *P. syringae* pv. *tomato* strain TF1.

The disparity between replicates highlighted in Table 4 was found in the minority of AFLP profiles produced; the majority of the replicates produced very similar peak heights and sizes. Using the normalized threshold simplified greatly the determination of which peaks to analyze and which to disregard. The AFLP amplicons scored for analysis and shown in tables similar to Table 4 shown above were developed for all 14 *P. syringae* strains tested. The best example of run reproducibility occurred with strain 188B where 100% of the AFLP fragments (in all 3 color channels) were found consistently in all three replicates. The data are shown in Table 5 below.

Run 1		Run 2		Run 3	
Size	Height	Size	Height	Size	Height
75.64	711	75.64	565	75.65	1294
93.64	693	93.72	537	93.76	1176
135.88	364	135.99	299	135.97	633
147.7	2274	147.82	1657	147.69	4121
164.5	377	164.51	336	164.56	667
197.3	793	197.41	607	197.38	1458
242.63	936	242.64	760	242.64	2151
252.89	911	252.93	740	253	1591
258.2	256	258.25	213	258.24	432
284.88	671	284.86	498	284.77	1197
297.4	553	297.39	428	297.25	997
304.91	2356	305.04	1705	305.23	4373

Size	Height	Size	Height	Size	Height
75.87	6089	75.85	5602	75.86	4819
77.16	5366	77.24	4921	77.25	4464
79.65	2220	79.69	1985	79.72	1549
81.93	2250	81.94	2142	81.98	1690
88.99	3514	88.9	2972	88.98	2598
89.86	2025	89.76	1714	89.84	1432
118.71	1418	118.65	1243	118.73	958
124.36	4176	124.26	3616	124.34	3042
144.13	5326	144.2	5002	144.28	4356
149.87	5161	149.87	3817	150	3609
159.23	3604	159.24	3172	159.13	2621
166.7	4333	166.75	3817	166.62	3267
179.19	1626	179.24	1520	179.12	1230
201.39	1923	201.49	1708	201.5	1443
225.48	1223	225.57	1051	225.48	923
242.81	3674	242.84	3336	242.89	2934
245.5	2863	245.52	2426	245.57	2008
261.26	3156	261.24	2784	261.32	2300
321.52	1160	321.57	947	321.51	793

Size	Height	Size	Height	Size	Height
81.94	2279	81.94	3072	82.04	1810
86.98	5534	86.97	7315	87.03	4363
91.6	6993	91.59	7146	91.49	7253
98.81	7078	98.7	7331	98.91	7208
103.6	1033	103.61	1345	103.65	798
172.94	5947	172.92	7249	172.84	4918
220.28	2174	220.25	3120	220.27	2002
224.89	1063	224.87	1605	224.81	943
276.25	1484	276.09	2062	276.17	1382
310.84	1573	310.79	2284	310.81	1342

Table 5 Peak sizes and heights for *P. syringae* pv. *tomato* strain 188B. The background color of each table corresponds to the fluorescent dye used to produce the profiles.

Also evident from Table 5 is the fact that there was excellent reproducibility in the size estimation of amplicons constituting each AFLP profile. Reproducible size

estimation is characteristic of capillary electrophoresis based genetic analyzers like the model ABI 310 used, and stems from the presence of a size ladder in each sample.

The reproducibility of replicate runs was determined by dividing the number of unmatching peaks within a color channel, where peaks that differed by 0.5 basepairs were considered unmatching, by the number of total peaks in that color channel. The calculations for each color channel were averaged to give an overall reproducibility figure for each strain.

$$\text{Reproducibility} = \frac{\text{Unmatching peaks between all three runs}}{\text{Peaks in run 1} + \text{Peaks in run 2} + \text{Peaks in run 3}}$$

This calculation produced a number between 0 and 1 that, when multiplied by 100, produces a percentage of difference among the triplicate runs. The triplicate AFLP analyses on all strains produced an overall reproducibility of 95%. Strain 30555 gave the lowest level of reproducibility at 93%, while strain 188B gave the highest level at 100% reproducibility. Reproducibility in this range is consistent with that found in other AFLP literature. The highest variability among replicate AFLP profiles was seen in early analyses that we now know were, in part, generated with less than optimal reagents. Even so, overall reproducibility was sufficiently high to support the use of AFLP analysis for forensic investigations involving the pathogen *P. syringae* pv. *tomato*. Other pathogens, such as *S. marcescens* and *S. aureus* (Allen, unpublished) (data not shown) yielded similar results, thereby suggesting AFLP may be useful for the genetic discrimination of bacterial pathogens in general.

Discriminatory Ability

Several approaches were taken to determine the capabilities of the AFLP assay to discriminate among the 14 *P. syringae* pv. *tomato* strains tested. First, visual discrimination was possible using the AFLP electropherograms. The electropherograms for strains 1008, 188B and T1 are shown in Figures 5A-I. A second discrimination method consisted of a numerical coding system that captures all scored peaks in a profile and places them into fixed "bins", each 10 basepairs in size and spanning a size range for amplicons from 75-350 basepairs. This approach produced an "amplicon haplotype" for each strain, within each color channel.

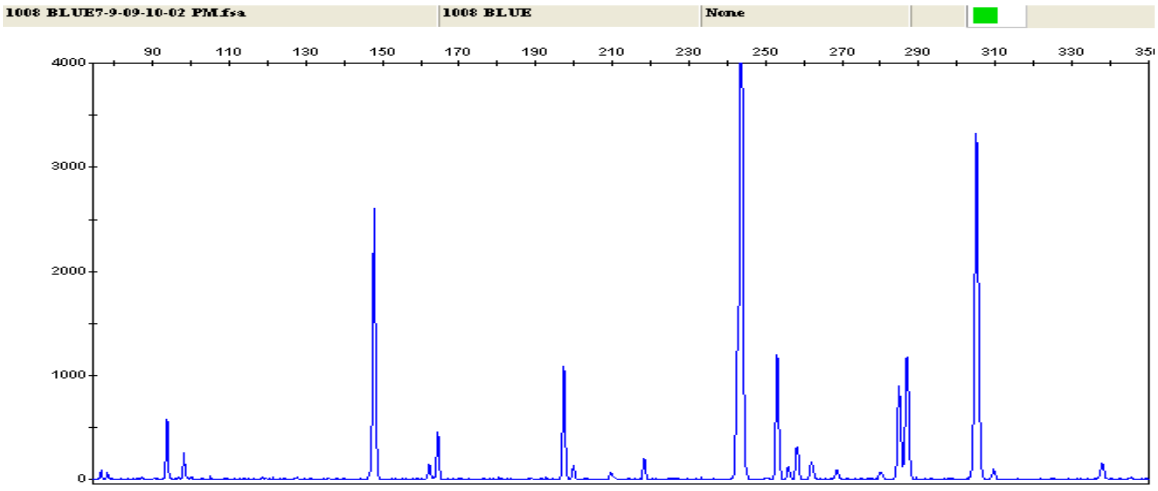

Figure 5A Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain 1008 using the blue-FAM labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

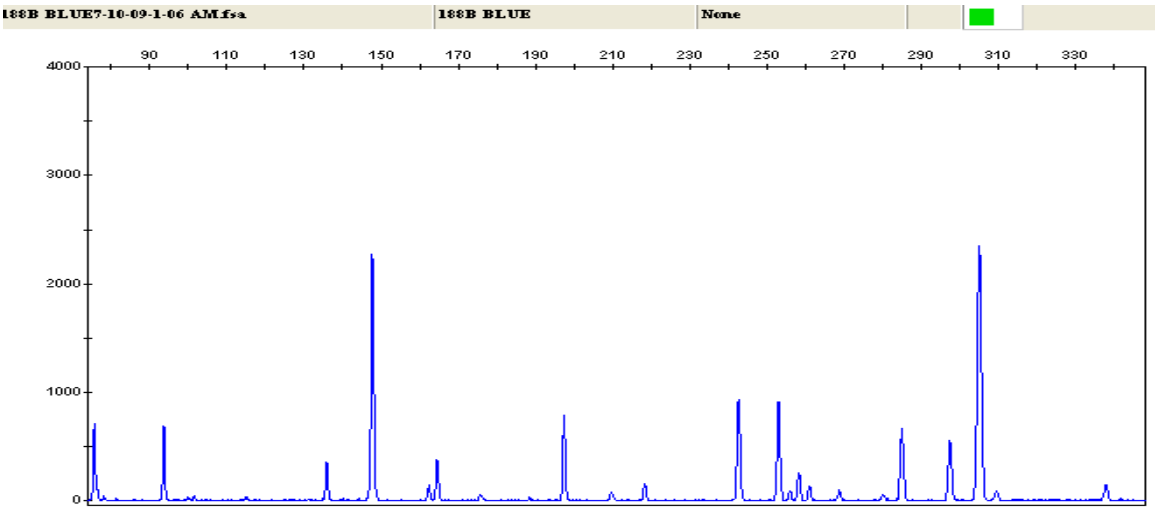

Figure 5B Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain 188B using the blue-FAM labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

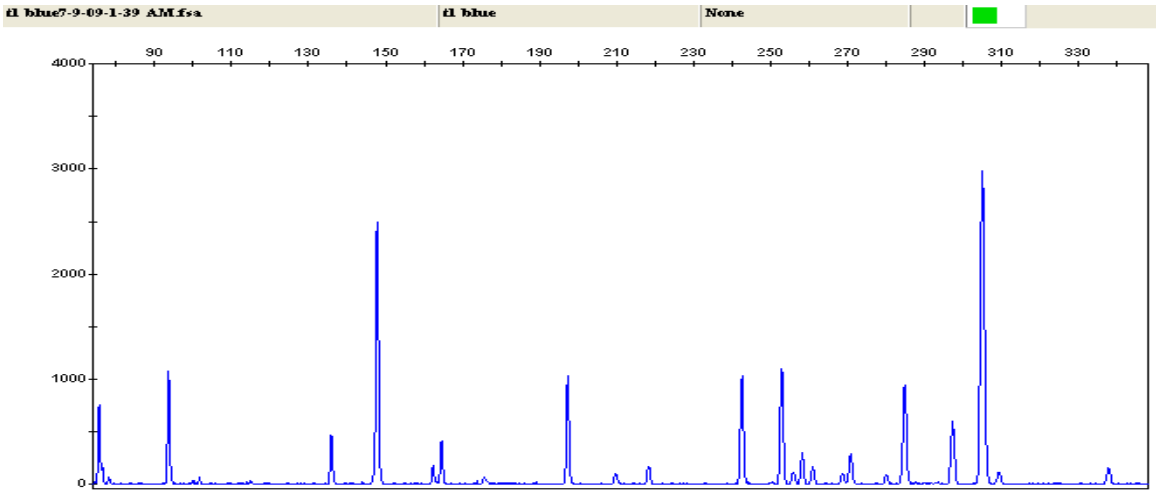

Figure 5C Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain T1 using the blue-FAM labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

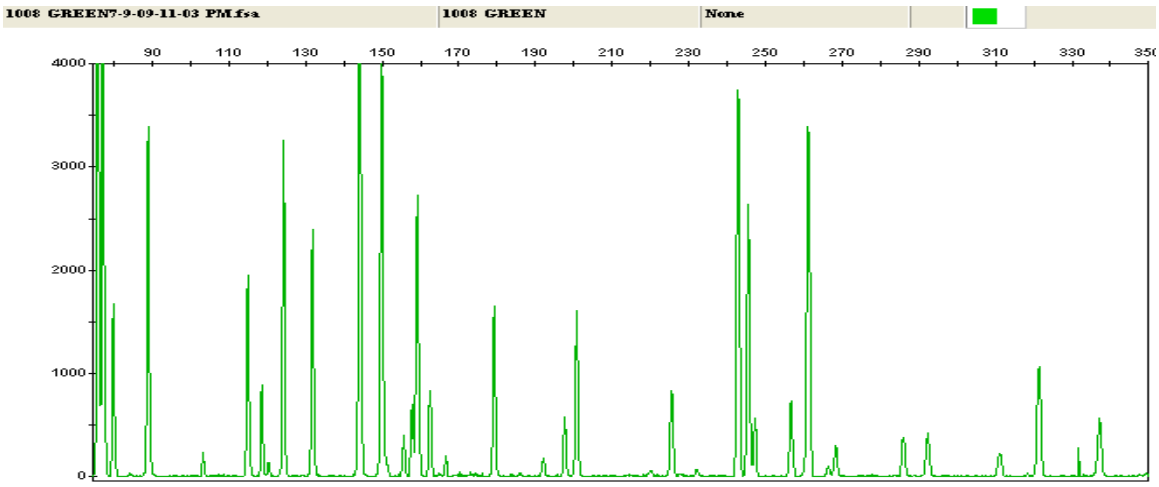

Figure 5D Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain 1008 using the green-JOE labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

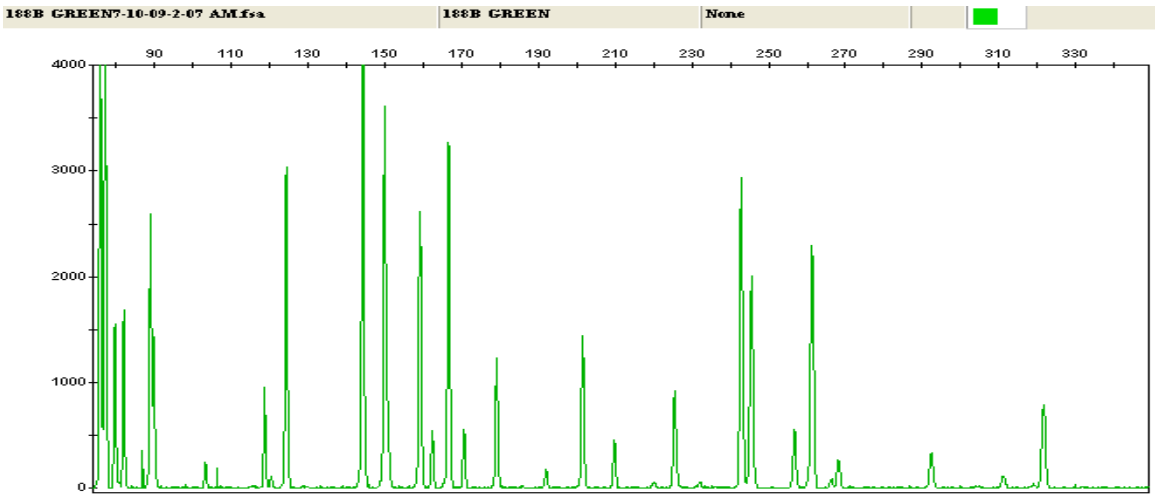

Figure 5E Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain 188B using the green-JOE labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

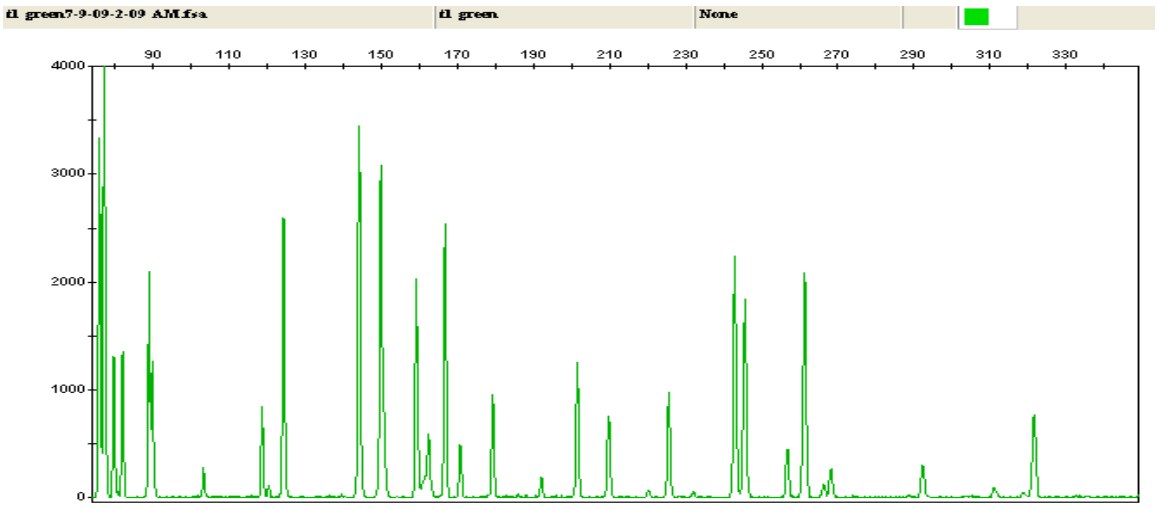

Figure 5F Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain T1 using the green-JOE labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

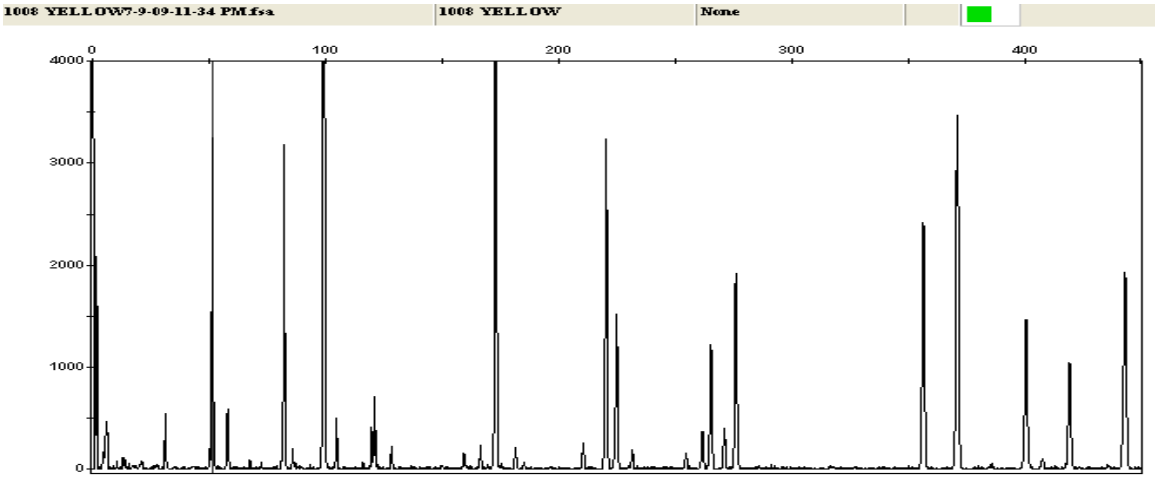

Figure 5G Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain 1008 using the yellow-NED labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

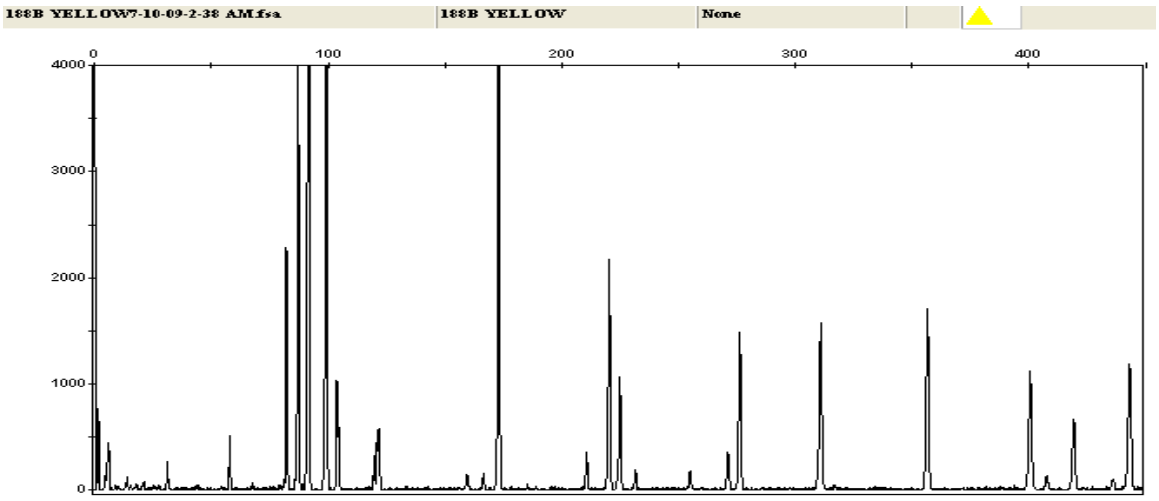

Figure 5H Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain 188B using the yellow-NED labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

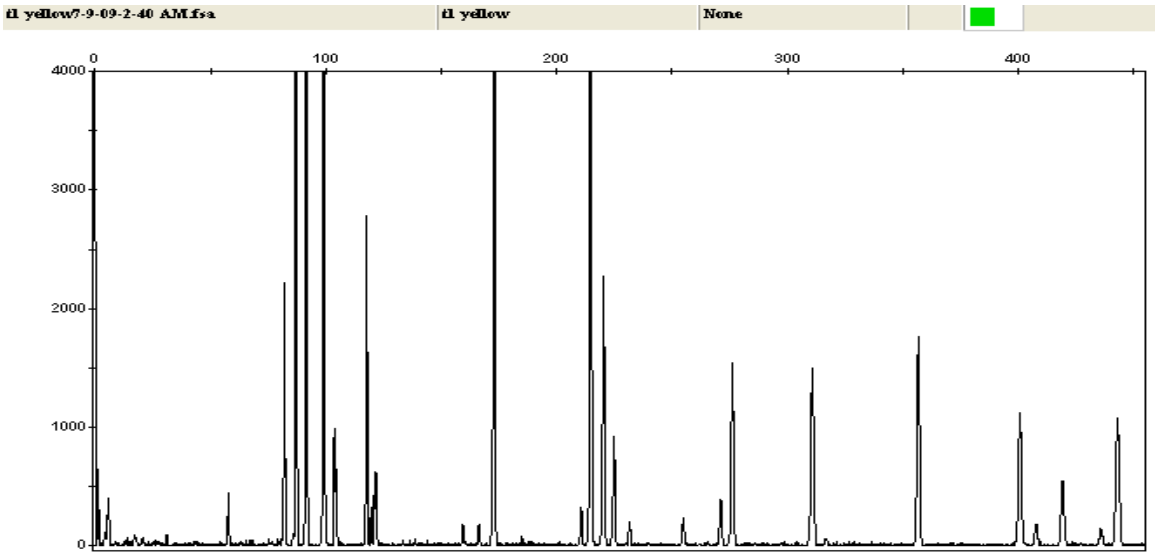

Figure 5I Electropherogram produced by the AFLP process for *P. syringae* pv. *tomato* strain T1 using the yellow-NED labeled primers. The X-axis represents the size (in base pairs) and the Y-axis represents the relative fluorescence units (RFU).

In previous work, a binary coding system was developed that placed a '1' in a bin where one or more amplicons were present, and a '0' in any bin that lacked an amplicon in the AFLP profile. What was produced using this system was a binary code that described the relative location of the peaks found within a particular strain's AFLP

profile. However, using this coding method (Beauman, 2007), a bin containing multiple amplicons of similar size would not be distinguished from a bin with only a single amplicon. In this study, to correct for such limitations, the binary system was replaced by a numerical system that considered number of amplicons within a given bin. This modification allows for a more detailed and accurate description of the AFLP profile, while maintaining much of the code's simplicity. Below, in Table 6, haplotype codes are shown for two bacterial strains, T4B1 and 1318. The table shows the codes produced in each of the three color channels, and highlights the code's ability to convey the different characteristics of the two strains. The complete AFLP haplotype codes for all of the different *P. syringae* pv. *tomato* strains subjected to AFLP analysis are shown in Appendix 1.

T4B1	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	1	1	1	0	1	0	2	0	0	0	0	0
1318	0	1	0	0	0	1	1	1	1	0	1	0	1	0	0	0	1	1	0	0	1	0	2	0	0	0	0	0

T4B1	4	1	0	0	2	0	1	1	2	2	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
1318	4	2	0	1	2	0	1	1	2	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0

T4B1	1	2	2	0	0	0	0	0	0	2	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0
1318	1	2	1	0	1	1	0	0	0	1	0	0	0	0	2	0	1	0	0	0	1	0	0	1	0	0	0	0

Table 6 Haplotype codes produced for *P. syringae* pv. *tomato* strains T4B1 and 1318 using AFLP.

Even when using the numerical system however, identical codes were produced occasionally for strain pairs within each color channel. For example, within the blue color channel, strains 188B and T4B1 were indistinguishable; as were strains B125 and T1 and strains 880 and 1008. In the green channel strains 30555, 3357 and 188B all produced identical codes. And finally in the yellow channel, strains 30555 and B125 could not be distinguished. These data show that all three primer sets are needed for optimum levels of discrimination using the AFLP process. For each instance in which two (or more) strains were identical within a single color channel, they were distinguishable in one or more of the other two channels. Thus, based upon the complete haplotype codes, all *P. syringae* pv. *tomato* strains were distinguishable from one another.

Genetic Relatedness Among *P. syringae* pv. *tomato* Strains

To further characterize the discriminatory capability of the AFLP assay, the genetic relatedness between each of the strains was calculated in a pairwise fashion using the Sorensen similarity index (Sorensen, 1948), a statistical method described by botanist Thorvald Sorensen in 1948 that is used to compare phenotypic (or genotypic) characteristics of two samples. The formula is

$$QS = \frac{2 \cdot C}{A + B}$$

, where ‘A’ and ‘B’ are the numbers of characteristics compared between the two groups. In this case they are the numbers of peaks found within the bins in a single color channel of the AFLP profile. ‘C’ is the number of similar characteristics shared by the two groups, in this case the number of indistinguishable AFLP amplicons scored in the two samples being compared (Sorensen, 1948). This equation ultimately gives a number

between 0 and 1, representing the percentage of genetic relatedness between the two samples being compared. Shown in Table 7 are the Sorensen similarity indices for strain 188B compared to all other *P. syringae* pv. *tomato* strains in the blue color channel. As is evident, strain 188B shares all amplicons with itself, resulting in an index of 1. Also, 188B and T4B1 produced identical codes within the blue channel (column T4B1 has an index of 1 as well). Strains that have indices of 1 within one of the color channels, have lower indices in the other color channels. Sorensen similarity indices for all 14 strains compared pair wise, in all three color channels, and a combined index are shown in Appendix 2.

Strain	CPST 147	CPST 232	RG4	880	T1	TF1	T4B1	1318	30555	B125	3357	PT17	1008	188B
188B	0.73	0.92	0.64	0.67	0.78	0.83	1	0.83	0.73	0.78	0.83	0.64	0.67	1

Table 7 Sorensen similarity indices comparing strain 188B with all other strains of *P. syringae* pv. *tomato.*

To produce a population-wide similarity index, the indices produced by looking at all pair wise combinations of strains were averaged, yielding a value of 0.81 (all 14 strains tested shared 81% of their individual AFLP profiles). This number seems to be a valid view of relatedness within the population, as it indicates that the 14 strains share the majority of their genetic profiles as might be expected when comparing members of the same species and pathovar. Another way to describe the discriminatory power of the AFLP assay is to calculate the probability that the assay will exclude genetically distinct strains of *P. syringae* pv. *tomato*. Based upon the results obtained with the 14 strains studies here, one might conclude that probability to be 100% since no two strains produced identical haplotype codes. However, a different, perhaps more accurate,

approach when extended to an infinite number of strains is to divide the number of Sorensen index values equal to 1.0 by the total number of pairwise comparisons. This value represents the overall probability of falsely grouping two genetically distinct strains. Subtracting this value from 1.0 produces the probability of excluding two distinct strains as being genetically identical. Using this approach, the probability of exclusion for the AFLP assay applied to *P. syringae* pv. *tomato* is 97%, This approach can also be extended to include bacterial strains from other species, including data from *S. marcescens* produced by Beauman (2007) and also *S. aureus* (Allen, unpublished). Inclusion of these species in the comparison results in a "bacterial" probability of exclusion of 99%. Yet another way to quantitate the probability of exclusion would be to only take into account the combined Sorensen similarity indices, since they are based upon the complete haplotype codes, and divide the number of indices equal to 1 by the total number of observations. Since we were able to distinguish among all strains of *P. syringae, S. marcescens* and *S. aureus* our numerator for the calculation would be 0. Thus, in practice we found a probability of exclusion of 100% among *P. syringae* strains, as well as among all three species tested, with the lower level of a 95% confidence interval equaling 98% and 99%, respectively (Hanley & Lippman-Hand, 1983).

Supplemental Findings

Findings outside the scope of the initial research questions led to interesting suggestions about the AFLP assay. For example, Figure 6 depicts the average number of peaks produced per color channel. The fact that the green channel has, on average, nearly double the number of scorable amplicons as the other two colors suggests that *P.*

syringae pv. *tomato* has more sites within its genome that are complementary to the primers having an additional guanosine extension, than sites complementary to the nucleotide extensions on the other two primers. If so, then the areas surrounding the *EcoR1* and *Mse1* restriction sites are richer in cytosine, than in either guanosine or thymine (the bases complementary to the other two primers).

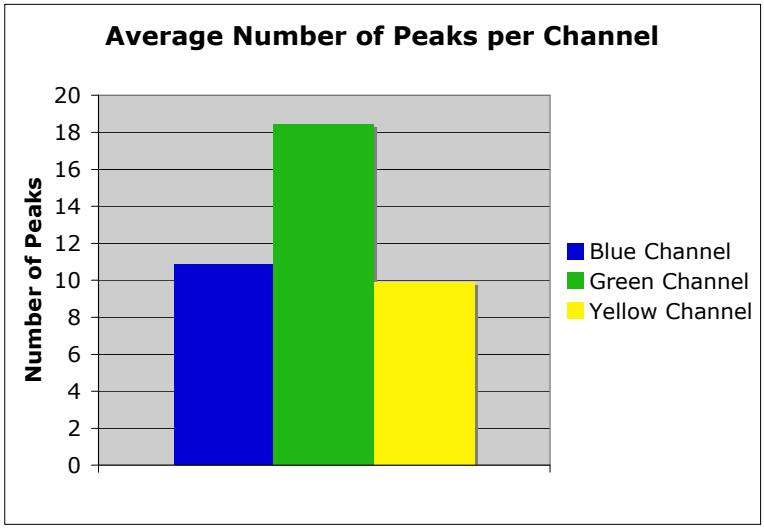

Figure 6 Graph depicting the numbers of peaks observed in each color channel.

Consistent with this conclusion the green dye-labeled primers (with the nucleotide extension G) are the most numerous in the *P. syringae* pv. *tomato* genome, is the fact that the average amplicon size in the green channel is smaller than that in the other two color channels (Figure 7). The significance of this observation is unclear, but the majority of green peaks are in the lower size range, whereas amplicons in the blue channel exhibit consistent amplicon size throughout, in that the average blue channel profile contains equal numbers of large and small peaks, indicating that thymidine (the base

complimentary to the nucleotide extension in the blue primer) is perhaps not “clustered” in proximity to *EcoR1* and *Mse1* restriction sites in the genome.

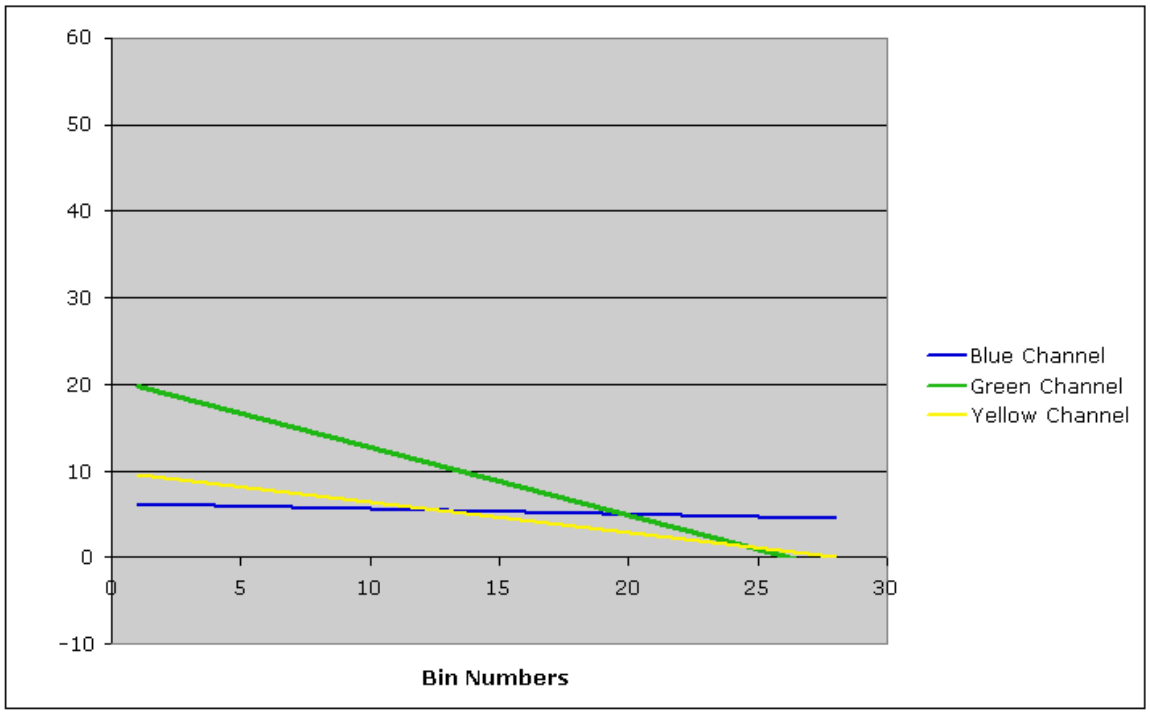

Figure 7 Graph depicting the numbers and locations of the observed peaks. (X-axis represents the in numbers and the Y-axis represents the numbers of peaks found within the bins).

CHAPTER V

CONCLUSIONS

As a molecular genotyping method, AFLP analysis has proven effective in several research applications as a tool to discriminate among members of both eukaryotic and prokaryotic species (Beauman, 2007; Clarke & Meudt, 2008; Groenewald *et al*., 2005; Shaik *et al*., 2008). Whether or not AFLP analysis is sufficiently reliable as a genotyping method to be useful as a forensic tool in the investigation of biocrimes, particularly those targeting crops of agricultural importance, is not known. We sought to quantify both the reliability and the power of the assay to discriminate among microbial populations within a species and among different species.

In previous work from this laboratory, Beauman (2007) demonstrated an 85-90% reproducibility in replicate AFLP profiles produced from strains of *S. marcescens* isolated from different sources. Reproducibility has been improved significantly in this study, perhaps because of the nature of the bacterial species we were working with, the standardization of the starting concentration of DNA (i.e., through the use of RNase to remove contaminating RNA from the genomic DNA preparations), or the lower temperatures at which the products of intermediate steps were stored. The reproducibility of 95% seen in our study also exceeded that reported in other work done with the AFLP

technique in the forensic science field; when applied to producing DNA profiles of red maple trees (*Acer rubrum*), the AFLP reproducibility percentage was 93.8 (Bless *et al*., 2006).

The discriminatory power of the assay reflects its ability to distinguish among members of the population. In the area of human identity testing, current STR genotyping methods are extremely powerful, with a coincidental match probability for DNA profiles from random, unrelated individuals of 2×10^{-18}. In this study, the discriminatory power was measured through the pairwise comparison of AFLP profiles produced from *P. syringae* pv. *tomato* strains using three different selective primers (Sorensen, 1948) and, the assay produced unique profiles for every strain. In addition, a combined similarity index of 0.81 was calculated for all *P. syringae* pv. *tomato* strains tested, signifying that all 14 shared 81% of their AFLP profiles on average.

As mentioned earlier, we chose our fourteen strains of *P. syringae* pv. *tomato* based upon the conclusions of a study completed by Dr. Christy Baker as well as geographic origins of the individual strains. In calculating an average combined Sorensen similarity index for the eight strains that Dr. Baker found to belong to the same MLVA group (30555, T1, B125, TF1, PT17, T4B1, 188B, and CPST232), we found an index of 0.87, which is higher than the average of all *P. syringae* pv. *tomato* strains. This indicates that our AFLP results support Dr. Baker's conclusion that those eight strains are genetically more similar to one another than they are to other strains. Furthermore, in looking at the geographic origins of the strains, one would expect that higher similarity indices would be produced by strains collected from the same region vs. those found in different areas. That was partially shown in our findings, in that our highest pairwise

similarity index of 0.93 was found between strains 188B and T4B1, both of which originated in Canada. This seems to support the claim that strains from similar areas would produce similar AFLP profiles. However, this appears to be a coincidence based upon the other similarity indices. For example, our second highest index (0.91) was seen between strains PT17 and 30555, which were collected from the USA and Tasmania, respectively. This result coupled with many others like it, indicate that geographic area of collection does not coincide with AFLP profiles produced for the 14 *P. syringae* pv. *tomato* strains tested.

AFLP profiles for eight strains of *S. marcescens* and three strains of *S. aureus* were obtained from Charlene Beauman and Dr. Robert Allen respectively (Oklahoma State University – Center for Health Sciences). Those eleven strains were compared to the 14 *P. syringae* pv. *tomato* strains in the same manner as the *P. syringae* pv. *tomato* strains were compared with one another. Amplicon haplotype codes produced for each of the eleven strains of *S. aureus* and *S. marcescens*, shown in Appendix 3, were compared to the haplotype codes produced for the 14 *P. syringae* pv. *tomato* strains. As mentioned above, haplotype codes found among some *P. syringae* pv. *tomato* strains were identical within a single color channel. However, when *P. syringae* pv. *tomato* codes were compared with those of the other two species all codes were different, even within a single color channel. This result is not unexpected, as organisms belonging to a different species would be more genetically distinct than organisms belonging to the same species. The overall probability of exclusion for the AFLP assay, using *P. syringae* pv. *tomato* as the model organism, was, in practice, 100%.

Complete tables of the Sorensen similarity indices among14 *P. syringae* strains and the 11 *S. aureus* and *S. marcescens* strains are shown in Appendix 4. As expected, the 14 *P. syringae* pv. *tomato* strains were much more similar to one another than they were to either the *S. aureus* strains or the *S. marcescens* strains. As mentioned above, the average similarity among all *P. syringae* pv. *tomato* strains was 0.81 (or 81%), whereas that for the comparison of *P. syringae* pv. *tomato* with *S. aureus* strains was 0.27 (or 27%), and that of *P. syringae* pv. *tomato* strains with *S. marcescens* strains, was only 0.23 (or 23%). It appears from this limited cross-species study, that a false inclusion of genetically distinct organisms is even less likely when organisms from different species are compared. Both other species produced much lower similarity indices than did the 14 *P. syringae* pv. *tomato* strains tested in the original study. These data indicate the potential value of the assay, in that it produced results that one would expect when looking at multiple, genetically different, organisms.

Results from our comparisons using AFLP analysis support previous studies done on *P. syringae*. One study, using the AFLP technique, found that strains within the same species were between 51-100% similar, while the average level of similarity between bacteria belonging to different species was only 25% (Clerc *et al.*, 1998). Another study, using a DNA:DNA hybridization technique, found that the *P. syringae* pv. *tomato* strains that were tested showed 86-100% homology (Denny *et al.*, 1988). The same technology applied to comparisons of strains belonging to *P. syringae* pv. *tomato* versus *P. syringae* pv. *syringae* showed were only 37-47% homology (Denny *et al.*, 1988). Our comparisons using AFLP analysis coupled with the Sorensen similarity index agree well with the

previous work of Denny *et al.* (1988) using DNA:DNA hybridization methods and especially the work reported by Clerc *et al.* (1998) using the AFLP technique.

Using AFLP analysis and capillary electrophoresis, we were able to discriminate among all 14 *P. syringae* pv. *tomato* strains tested, and showed that they were more genetically similar to one another than to either *S. aureus* or *S. marcescens*. These findings are in keeping with prior research, in that AFLP has been used to discriminate among multiple pathovars of *P. syringae* (Sisto *et al.*, 2007) as well as multiple strains within *P. syringae* pv. *tomato* (Clerc *et al.*, 1998). Our findings suggest that AFLP analysis can be a useful tool for applications in microbial forensics. Investigation of other bacterial strains could perhaps identify other critical aspects of the technique that can enhance reproducibility even further. As was shown in this study, even when the assay is less than optimal, results can still be somewhat informative concerning the attribution of a pathogen to a potential source.

Equally important to differences among the AFLP profiles produced from pathogenic bacteria are those elements with the profiles that do not differ among all strains within a given species. When investigating an outbreak of disease affecting a host, identification of a pathogen through matching elements of its AFLP profile against a database containing profiles from a wide range of bacterial pathogens could quickly and efficiently identify the pathogen in question, and perhaps relate it to known strains endemic to particular regions of the world. Ultimately, it may be possible to develop an AFLP database containing those elements of the AFLP profile that are shared among all members of a given pathogen species or pathovar. AFLP elements such as those shared among all strain of *P. syringae* pv. *tomato* tested here would be good candidates for

inclusion in such a database. All 14 *P. syringae* pv. *tomato* strains contained identical AFLP profile elements in the 2nd, 8th, 13th, 17th and 18th bin of the blue channel, the 7th and 8th bin of the green channel, and the 1st and 21st bin of the yellow channel. Similar AFLP elements were also found in *S. marcescens* profiles produced by Beauman and also among strains of *S. aureus* (Allen, unpublished), where all *S. marcescens* strains shared common AFLP fragments in the 12th bin of the blue channel and all *S. aureus* strains produced identical AFLP profiles in the 16th bin of the green channel. If all strains of bacteria have these elements in their profiles, AFLP analysis may not only be useful for attribution, but also useful for species identification as well.

REFERENCES

Baker, C. (2008). Strain typing and characterization of s^{54}-dependant transcriptional activator mutants in *Pseudomonas syringae* pv. *tomato*. Doctoral Dissertation, Oklahoma State University, Stillwater, Oklahoma, December 2008.

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

Bless, C., Palmeter, H., & Wallace, M. (2006). Identification of *Acer rubrum* using amplified fragment length polymorphism. *Journal of Forensic Science*. 51(1):31-8.

Butler, J. (2005). *Forensic DNA typing: Biology, technology and genetics of STR marker* (2nd ed.) Elsevier Science & Technology Books. 45-47, 58, 69.

Centers for Disease Control and Prevention (CDC). (2006) Update on multi-state outbreak of E. coli 0157:H7 infections fresh spinach, October 6, 2006. http://www.cdc.gov/foodborne/ecolispinach/100606.htm

Centers for Disease Control and Prevention (CDC). (2009) Investigation update: outbreak of *Salmonella* Typhimurium Infections, 2008-2009. April 29, 2009. Retrieved from: http://www.cdc.gov/salmonella/ typhimurium/update.html, on June 9, 2009.

Clarke, A. & Meudt, H. (2008). Almost forgotten or latest practices? AFLP applications, analyses and advances. *Trends in Plant Sciences*, 12(3), 106-117.

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

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

Denny, T., Gilmour, M. & Selander, R. (1988) Genetic Diversity and Relationships of Two Pathovars of *Pseudomonas syringae*. *Journal of General Microbiology*, 1988(134), 1949-1960.

Fletcher, J., Bender, C., Budowle, B., Cobb, W. T., Gold, S. E., Ishamaru, C. A., Luster, D., Melcher, U., Murch, R., Scherm, H., Seem, R. C., Sherwood, J. L., Sobral, B. W. & Tolin, S. A. (2006). Plant Pathogen Forensics: Capabilities, Needs and Recommendations. *Microbiology and Molecular Biology Reviews*, 70(2), 450-471.

Frazier, R. A., Ames, J. M., & Nursten, H. E. (2000). Background theory and principles of capillary electrophoresis. *Capillary electrophoresis for food analysis*. Burlington House London: RSC Publishing.

Groenewald, S., Van Den Berg, N., Marasas, W. F., & Viljoen, A. (2006). The application of high-throughput AFLP's in assessing genetic diversity in F*usarium oxysporum* f. sp. *cubense*. *Mycological Research*, 110(Pt 3), 297-305.

Hanley, J.A. & Lippman-Hand, A. "If nothing goes wrong is everything all right? Interpreting zero numerators". *Journal of American Medical Association*, 249(13): 1743-45, 1983.

Hirano, S. S. and C. D. Upper (1990) Population biology and epidemiology of *Pseudomonas syringae*. *Annual Reviews in Phytopathology* 28:155-177.

Holmes, A., Govan, J., & Goldstein, R. G. (1998) Agricultural Use of *Burkholderia* (*Pseudomonas*) *cepacia*: A Threat to Human Health? http://www.cdc.gov /ncidod/eid/vol4no2/holmes.htm.

Jain, S. K., Persaud, D., Perl, T. M., Pass, M. A., Murphy, K. M., Pisciotta, J. M., et al.(2005). Nosocomial malaria and saline flush. *Emerging Infectious Diseases, 11*(7), 1097-1099.

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

Mueller, U. G., & Wolfenbarger, L. L. (1999). AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution, 14*(10), 389-394.

Neale, Tod. (2009). Cost of *Salmonella* Outbreak Likely to Top $1 Billion. Retrieved from: http://medpagetoday.com/InfectiousDisease/PublicHealth/13269, on June 9, 2009.

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

Oberst, R. (2004). Agricultural biosecurity taking on increase importance. http://www.eurekalert.org/feathers/doe/2004=06/dnl-abt061. Retrieved on February 4, 2009.

Park, Michael. (2006). Foxnews.com. "*E. coli* Outbreak Hurts Spinach Farming Industry, Restaurants." Retrieved from: http://www.foxnews.com/story/0,2933, 215257,00.html, on June 9, 2009.

President's DNA Initiative. Retrieved on June 4, 2009, from" http://www.nfstc.org/pdi/ Subject05 /pdi_s05_m02_01_a.htm.

Rudin, N., & Inman, K. (2001). An introduction to forensic DNA analysis (2nd ed.) CRC Press.

Savelkoul, P., Melles, D., Buffing, N., Gorkink, R., Simons, G., & van Belkum, A. (2007, October). High density whole genome fingerprinting of methicillin-resistant and -susceptible strains of Staphylococcus aureus in search of phenotype-specific molecular determinants. *Journal of Microbiological Methods*, 71(1), 44-54.

Sorensen, Thorvald. (1948). "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons". Biologiske Skrifter. Kongelige Danske Videnskabernes Selskab, 5 (4): 1-34.

Shaik, R., Pillay, D., & Pillay, B. (2009). Amplified fragment length polymorphisms reveal genetic differentiation among strains of Xanthomonas albilineans. *Journal of Microbial Methods,* 76 (1) 43-51.

Sisto, A., Cipriana, M., Tegli, S., Cerboneschi, M., Stea, G. & Santilli, E. (2007). Genetic characterization by fluorescent AFLP of Pseudomonas savastasoi pv. savastasoi strains isolated form different host species. *Plant Pathology*, 56, 366-372.

Tamaki, K., & Jeffreys, A. J. (2005). Human tandem repeat sequences in forensic DNA typing. *Legal medicine (Tokyo, Japan), 7*(4), 244-250.

United States Department of Agriculture (USDA). (2009). Latest U.S. Agricultural Trade Data. Retrieved from: http://wwwers.usda.gov/Data/FATUS/monthlysummary .htm, on June 9, 2009.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T. Hornes, M, Frijters, A., Pot, J., Peleman, J., & Kuiper, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, 23(21), 4407-4414.

Wheelis, M. (2005). Agricultural biowarfare and bioterrorism: What are the special features of an attack on an agricultural sector. Retrieved from: http://www.fas.org/ bwc/agr/main.htm, on April 15, 2009.

Wittendorp, J. (2007). AFLP® - genetic fingerprinting. http://www.keygene.com/techs-apps/index.html.

Young, J., Dye, D., Bradbury, J., Panagopoulos, C., & Robbs, C. (1978). The use of the term “pathovar” in the classification of plant pathogenic bacteria. Plant Pathogenic Bacteria. Proceedings of the 4th International Conference on Plant Pathogenic Bacteria, 359-363.

APPENDICES

Appendix 1

Haplotype code for *P. syringae* pv. *tomato* strains (blue-FAM)

CPST-147	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	1	1	0	0	2	1	0	0	0	0	0
CPST-232	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	1	1	1	0	0	1	2	0	0	0	0	0
RG4	0	1	0	0	0	1	0	1	0	1	0	0	1	0	0	0	1	1	0	0	0	1	2	0	0	0	0	0
880	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0
T1	1	1	0	0	0	0	1	1	0	1	0	0	1	0	0	0	1	1	0	0	0	1	2	0	0	0	0	0
TF1	1	1	0	0	0	0	1	1	0	1	0	0	1	0	0	0	1	1	1	0	0	1	2	0	0	0	0	0
T4B1	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	1	1	1	0	1	0	2	0	0	0	0	0
1318	0	1	0	0	0	1	1	1	1	0	1	0	1	0	0	0	1	1	0	0	1	0	2	0	0	0	0	0
30555	1	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	1	0	0	0	1	1	1	0	0	0	0
B125	1	1	0	0	0	0	1	1	0	1	0	0	1	0	0	0	1	1	0	0	0	1	2	0	0	0	0	0
3357	0	1	0	0	0	1	1	1	1	0	0	0	1	0	0	1	1	1	0	0	1	0	2	0	0	0	0	0
PT17	1	1	0	0	0	0	0	1	0	1	0	0	1	0	0	0	1	1	0	0	0	1	2	0	0	0	0	0
1008	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	1	0	0	1	1	0	1	0	0	0	0
188B	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	1	1	1	0	1	0	2	0	0	0	0	0

Haplotype code for *P. syringae* pv. *tomato* strains (green-JOE)

CPST-147	3	1	0	1	2	1	1	1	2	0	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
CPST-232	4	2	0	0	2	0	1	1	1	1	1	0	1	0	0	0	0	2	2	0	0	0	0	0	1	0	0	0
RG4	4	2	0	0	1	0	1	1	2	1	1	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
880	3	1	0	0	3	1	1	1	1	0	1	0	0	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
T1	4	2	0	0	2	0	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
TF1	4	2	0	0	1	0	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
T4B1	4	1	0	0	2	0	1	1	2	2	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
1318	4	2	0	1	2	0	1	1	2	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
30555	4	2	0	0	2	0	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
B125	4	2	0	0	1	0	1	1	1	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0
3357	4	2	0	0	2	0	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
PT17	4	1	0	0	2	0	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
1008	3	1	0	1	2	1	1	1	1	0	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0
188B	4	2	0	0	2	0	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0

Haplotype code for *P. syringae* pv. *tomato* strains (yellow-NED)

CPST-147	1	1	1	0	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0	1	1	0	0	0	0	0	0	0	0
CPST-232	1	2	1	0	1	0	0	0	0	1	0	0	0	1	2	0	0	0	0	0	1	0	0	1	0	0	0	0	0
RG4	1	2	0	0	1	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
880	1	0	1	0	1	0	0	0	0	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0
T1	1	2	2	0	1	0	0	0	0	1	0	0	0	1	2	0	0	0	0	0	1	0	0	1	0	0	0	0	0
TF1	1	2	2	0	1	0	0	0	0	1	0	0	0	1	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0
T4B1	1	2	2	0	0	0	0	0	0	2	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0
1318	1	2	1	0	1	1	0	0	0	1	0	0	0	0	2	0	1	0	0	0	1	0	0	1	0	0	0	0	0
30555	1	2	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0
B125	1	2	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0
3357	1	2	2	0	2	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0
PT17	1	2	2	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	0
1008	1	0	1	0	1	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0
188B	1	2	2	0	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	1	0	0	1	0	0	0	0	0

Appendix 2

Sorensen similarity indices for *P. syringae* pv. *tomato* strains (blue-FAM)

	CPST 147	CPST 232	RG4	880	T1	TF1	T4B1	1318	30555	B125	3357	PT17	1008	188B
CPST 147	1	0.81	0.7	0.74	0.67	0.73	0.73	0.64	0.8	0.67	0.64	0.7	0.74	0.73
CPST 232	0.81	1	0.73	0.67	0.87	0.92	0.92	0.75	0.82	0.87	0.75	0.82	0.67	0.92
RG4	0.7	0.73	1	0.63	0.86	0.82	0.64	0.73	0.7	0.86	0.73	0.9	0.63	0.64
880	0.74	0.67	0.63	1	0.6	0.57	0.67	0.67	0.84	0.6	0.67	0.63	1	0.67
T1	0.67	0.87	0.86	0.6	1	0.96	0.78	0.7	0.76	1	0.7	0.95	0.6	0.78
TF1	0.73	0.92	0.82	0.57	0.96	1	0.83	0.67	0.73	0.96	0.67	0.91	0.57	0.83
T4B1	0.73	0.92	0.64	0.67	0.78	0.83	1	0.83	0.73	0.78	0.83	0.73	0.67	1
1318	0.64	0.75	0.73	0.67	0.7	0.67	0.83	1	0.64	0.7	0.92	0.64	0.67	0.83
30555	0.8	0.82	0.7	0.84	0.76	0.73	0.73	0.64	1	0.76	0.64	0.8	0.84	0.73
B125	0.67	0.87	0.86	0.6	1	0.96	0.78	0.7	0.76	1	0.7	0.86	0.6	0.78
3357	0.64	0.75	0.73	0.67	0.7	0.67	0.83	0.92	0.64	0.7	1	0.64	0.67	0.83
PT17	0.7	0.82	0.9	0.63	0.95	0.91	0.73	0.64	0.8	0.86	0.64	1	0.63	0.64
1008	0.74	0.67	0.63	1	0.6	0.57	0.67	0.67	0.84	0.6	0.67	0.63	1	0.67
188B	0.73	0.92	0.64	0.67	0.78	0.83	1	0.83	0.73	0.78	0.83	0.64	0.67	1

Sorensen similarity indices for *P. syringae* pv. *tomato* strains (green-JOE)

	CPST 147	CPST 232	RG4	880	T1	TF1	T4B1	1318	30555	B125	3357	PT17	1008	188B
CPST 147	1	0.74	0.78	0.84	0.81	0.76	0.87	0.9	0.79	0.8	0.79	0.86	0.97	0.84
CPST 232	0.74	1	0.84	0.72	0.86	0.86	0.82	0.85	0.89	0.86	0.89	0.86	0.76	0.89
RG4	0.78	0.84	1	0.71	0.91	0.91	0.86	0.89	0.89	0.91	0.89	0.86	0.74	0.89
880	0.84	0.72	0.71	1	0.8	0.8	0.81	0.79	0.83	0.69	0.83	0.86	0.91	0.83
T1	0.81	0.86	0.91	0.8	1	0.94	0.89	0.92	0.97	0.94	0.97	0.94	0.83	0.97
TF1	0.76	0.86	0.91	0.8	0.94	1	0.89	0.92	0.97	0.94	0.97	0.94	0.83	0.97
T4B1	0.87	0.82	0.86	0.81	0.89	0.89	1	0.93	0.92	0.83	0.92	0.95	0.84	0.92
1318	0.9	0.85	0.89	0.79	0.92	0.92	0.93	1	0.95	0.86	0.95	0.92	0.87	0.95
30555	0.79	0.89	0.89	0.83	0.97	0.97	0.92	0.95	1	0.91	1	0.97	0.86	1
B125	0.8	0.86	0.91	0.69	0.94	0.94	0.83	0.86	0.91	1	0.91	0.88	0.76	0.91
3357	0.79	0.89	0.89	0.83	0.97	0.97	0.92	0.95	1	0.91	1	0.97	0.86	1
PT17	0.86	0.86	0.86	0.86	0.94	0.94	0.95	0.92	0.97	0.88	0.97	1	0.89	0.97
1008	0.97	0.76	0.74	0.91	0.83	0.83	0.84	0.87	0.86	0.76	0.86	0.89	1	0.86
188B	0.84	0.89	0.89	0.83	0.97	0.97	0.92	0.95	1	0.91	1	0.97	0.86	1

Sorensen similarity indices for *P. syringae* pv. *tomato* strains (yellow-NED)

	CPST 147	CPST 232	RG4	880	T1	TF1	T4B1	1318	30555	B125	3357	PT17	1008	188B
CPST 147	1	0.74	0.53	0.75	0.7	0.6	0.63	0.7	0.71	0.71	0.6	0.67	0.8	0.78
CPST 232	0.74	1	0.97	0.63	0.96	0.87	0.73	0.87	0.8	0.8	0.78	0.76	0.78	0.86
RG4	0.53	0.67	1	0.67	0.63	0.74	0.67	0.63	0.75	0.75	0.74	0.71	0.57	0.59
880	0.75	0.63	0.67	1	0.6	0.7	0.63	0.6	0.71	0.71	0.7	0.67	0.8	0.56
T1	0.7	0.96	0.63	0.6	1	0.92	0.78	0.83	0.76	0.76	0.83	0.91	0.63	0.91
TF1	0.6	0.87	0.74	0.7	0.92	1	0.87	0.75	0.86	0.86	0.92	0.91	0.63	0.82
T4B1	0.63	0.73	0.67	0.63	0.78	0.87	1	0.7	0.9	0.9	0.87	0.95	0.56	0.86
1318	0.7	0.87	0.63	0.6	0.83	0.75	0.7	1	0.76	0.76	0.75	0.73	0.74	0.82
30555	0.71	0.8	0.75	0.71	0.76	0.86	0.9	0.76	1	1	0.86	0.95	0.63	0.84
B125	0.71	0.8	0.75	0.71	0.76	0.86	0.9	0.76	1	1	0.86	0.95	0.63	0.84
3357	0.6	0.78	0.74	0.7	0.83	0.92	0.87	0.75	0.86	0.86	1	0.91	0.63	0.82
PT17	0.67	0.76	0.71	0.67	0.91	0.91	0.95	0.73	0.95	0.95	0.91	1	0.59	0.9
1008	0.8	0.78	0.57	0.8	0.63	0.63	0.56	0.74	0.63	0.63	0.63	0.59	1	0.71
188B	0.78	0.86	0.59	0.56	0.91	0.82	0.86	0.82	0.84	0.84	0.82	0.9	0.71	1

Sorensen similarity indices for *P. syringae* pv. *tomato* strains (combined)

	CPST 147	CPST 232	RG4	880	T1	TF1	T4B1	1318	30555	B125	3357	PT17	1008	188B
CPST 147	1.00	0.76	0.67	0.78	0.73	0.70	0.74	0.75	0.77	0.73	0.68	0.74	0.84	0.78
CPST 232	0.76	1.00	0.85	0.67	0.90	0.88	0.82	0.82	0.84	0.84	0.81	0.81	0.74	0.89
RG4	0.67	0.75	1.00	0.67	0.80	0.82	0.72	0.75	0.78	0.84	0.79	0.82	0.65	0.71
880	0.78	0.67	0.67	1.00	0.67	0.69	0.70	0.69	0.79	0.67	0.73	0.72	0.90	0.69
T1	0.73	0.90	0.80	0.67	1.00	0.94	0.82	0.82	0.83	0.90	0.83	0.93	0.69	0.89
TF1	0.70	0.88	0.82	0.69	0.94	1.00	0.86	0.78	0.85	0.92	0.85	0.92	0.68	0.87
T4B1	0.74	0.82	0.72	0.70	0.82	0.86	1.00	0.82	0.85	0.84	0.87	0.88	0.69	0.93
1318	0.75	0.82	0.75	0.69	0.82	0.78	0.82	1.00	0.78	0.77	0.87	0.76	0.76	0.87
30555	0.77	0.84	0.78	0.79	0.83	0.85	0.85	0.78	1.00	0.89	0.83	0.91	0.78	0.86
B125	0.73	0.84	0.84	0.67	0.90	0.92	0.84	0.77	0.89	1.00	0.82	0.90	0.66	0.84
3357	0.68	0.81	0.79	0.73	0.83	0.85	0.87	0.87	0.83	0.82	1.00	0.84	0.72	0.88
PT17	0.74	0.81	0.82	0.72	0.93	0.92	0.88	0.76	0.91	0.90	0.84	1.00	0.70	0.84
1008	0.84	0.74	0.65	0.90	0.69	0.68	0.69	0.76	0.78	0.66	0.72	0.70	1.00	0.75
188B	0.78	0.89	0.71	0.69	0.89	0.87	0.93	0.87	0.86	0.84	0.88	0.84	0.75	1.00

Appendix 3

Haplotype code for *S. aureus* strains (blue-FAM)

ATCC	1	0	1	2	1	1	1	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0
4i2	3	1	2	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
27626-b	0	0	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Haplotype code for *S. aureus* strains (green-JOE)

ATCC	1	0	2	1	2	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
4i2	1	0	1	0	2	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
27626-b	0	1	0	1	1	3	2	0	0	1	0	0	0	1	0	1	2	0	0	0	0	1	0	0	0	0	0

Haplotype code for *S. aureus* strains (yellow-NED)

ATCC	0	1	0	2	0	0	1	2	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
4i2	1	1	0	3	0	0	1	1	0	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
27626-B	0	2	1	1	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0

Haplotype code for *S. marcescens* strains (blue-FAM)

731-17	0	0	0	1	1	0	1	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
ATCC-2987	0	0	1	3	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0
DB-11	0	0	0	0	2	0	1	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
HO1A	0	0	1	0	2	0	1	0	2	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0
PO1A	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
ROZ	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
ZO1A	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
WO1A	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

Haplotype code for *S. marcescens* strains (green-JOE)

731-17	0	0	0	1	1	0	1	0	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
ATCC-2987	0	0	1	2	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0
DB-11	0	0	0	1	0	0	1	0	1	0	2	1	0	1	0	1	0	2	0	0	0	0	1	0	0	0	1	0
HO1A	0	0	0	0	2	0	0	0	0	0	1	1	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	0
PO1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
ROZ	0	0	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
ZO1A	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1
WO1A	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1

Haplotype code for *S. marcescens* strains (yellow-NED)

731-17	0	0	0	4	0	2	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0
ATCC-2987	0	0	0	1	0	2	1	1	1	2	3	1	0	0	3	0	1	0	0	0	1	0	1	0	0	0	2	1
DB-11	0	0	1	1	0	2	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0
HO1A	0	0	0	2	0	2	0	0	0	0	1	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
PO1	0	0	0	3	0	1	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
ROZ	0	0	0	1	1	1	0	0	0	2	0	1	1	1	1	2	1	1	0	0	0	0	0	1	1	0	0	0
ZO1A	0	0	0	3	0	1	1	0	0	2	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1
WO1A	0	0	0	3	1	1	0	0	0	2	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0

Appendix 4

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

	ATCC	4i2	27626-b
CPST 147	0.1	0.19	0.28
CPST 232	0.27	0.35	0.25
RG4	0.1	0.19	0.14
880	0	0.1	0.31
T1	0.19	0.27	0.13
TF1	0.27	0.35	0.125
T4B1	0.27	0.35	0.25
1318	0.18	0.26	0.38
30555	0.1	0.19	0.29
B125	0.19	0.27	0.13
3357	0.1	0.19	0.14
PT17	0.1	0.19	0.14
1008	0	0.1	0.31
188B	0.27	0.35	0.25

Sorensen similarity indices for *S. aureus* and *P. syringae* pv. *tomato* strains (green-JOE)

	ATCC	4i2	27626-b
CPST 147	0.44	0.38	0.42
CPST 232	0.3	0.31	0.24
RG4	0.24	0.25	0.32
880	0.4	0.42	0.39
T1	0.38	0.4	0.38
TF1	0.31	0.32	0.38
T4B1	0.36	0.37	0.35
1318	0.41	0.36	0.4
30555	0.37	0.38	0.36
B125	0.25	0.26	0.33
3357	0.37	0.38	0.36
PT17	0.38	0.4	0.38
1008	0.46	0.4	0.44
188B	0.37	0.38	0.36

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

	ATCC	4i2	27626-b
CPST 147	0.24	0.32	0.24
CPST 232	0.2	0.27	0.32
RG4	0.25	0.33	0.27
880	0.12	0.21	0.13
T1	0.19	0.26	0.3
TF1	0.19	0.26	0.3
T4B1	0.2	0.27	0.32
1318	0.19	0.26	0.2
30555	0.22	0.3	0.35
B125	0.22	0.3	0.35
3357	0.19	0.26	0.3
PT17	0.21	0.29	0.33
1008	0.13	0.22	0.13
188B	0.21	0.29	0.33

Sorensen similarity indices for *S. aureus* and *P. syringae* pv. *tomato* strains (combined)

	ATCC	4i2	27626-b
CPST 147	0.26	0.30	0.31
CPST 232	0.26	0.31	0.27
RG4	0.20	0.26	0.24
880	0.17	0.24	0.28
T1	0.25	0.31	0.27
TF1	0.26	0.31	0.27
T4B1	0.28	0.33	0.31
1318	0.26	0.29	0.33
30555	0.23	0.29	0.33
B125	0.22	0.28	0.27
3357	0.22	0.28	0.27
PT17	0.23	0.29	0.28
1008	0.20	0.24	0.29
188B	0.28	0.34	0.31

Sorensen similarity indices for *S. marcescens* and *P. syringae* strains pv. *tomato* (blue-FAM)

	731-17	ATCC-2987	DB-11	HO1A	PO1	ROZ	ZO1A	WO1A
CPST 147	0.11	0	0.12	0.19	0.29	0.31	0.17	0.29
CPST 232	0.2	0	0.21	0.26	0.38	0.27	0.35	0.38
RG4	0	0.11	0.12	0.09	0.14	0.15	0.13	0.14
880	0.12	0	0.13	0.2	0.31	0.33	0.29	0.31
T1	0.11	0	0.22	0.18	0.27	0.14	0.25	0.27
TF1	0.1	0	0.21	0.17	0.25	0.13	0.24	0.25
T4B1	0.1	0	0.21	0.26	0.38	0.27	0.35	0.38
1318	0.1	0.09	0.32	0.35	0.38	0.27	0.35	0.38
30555	0.11	0	0.12	0.19	0.29	0.31	0.27	0.29
B125	0.11	0	0.22	0.18	0.27	0.14	0.25	0.27
3357	0.2	0.09	0.21	0.26	0.38	0.27	0.35	0.38
PT17	0	0	0.12	0.09	0.14	0.15	0.13	0.14
1008	0.12	0	0.13	0.2	0.31	0.33	0.29	0.31
188B	0.2	0	0.21	0.26	0.38	0.27	0.35	0.38

Sorensen similarity indices for *S. marcescens* and *P. syringae* strains pv. *tomato* (green-JOE)

	731-17	ATCC-2987	DB-11	HO1A	PO1	ROZ	ZO1A	WO1A
CPST 147	0.44	0.23	0.39	0.43	0.09	0.17	0.15	0.16
CPST 232	0.32	0	0.32	0.21	0.09	0.17	0.15	0.16
RG4	0.32	0.08	0.28	0.23	0.1	0.18	0.16	0.17
880	0.24	0.17	0.34	0.46	0.19	0.18	0.16	0.17
T1	0.38	0.08	0.33	0.37	0.18	0.26	0.23	0.25
TF1	0.38	0.08	0.33	0.37	0.18	0.26	0.23	0.25
T4B1	0.29	0.07	0.31	0.41	0.17	0.24	0.21	0.23
1318	0.41	0.14	0.39	0.4	0.16	0.23	0.21	0.22
30555	0.37	0.08	0.32	0.43	0.17	0.25	0.22	0.24
B125	0.33	0.09	0.21	0.24	0.1	0.19	0.16	0.18
3357	0.37	0.08	0.32	0.43	0.17	0.25	0.22	0.24
PT17	0.38	0.08	0.33	0.44	0.18	0.26	0.23	0.25
1008	0.46	0.24	0.4	0.44	0.09	0.17	0.15	0.17
188B	0.37	0.08	0.32	0.43	0.17	0.24	0.22	0.24

Sorensen similarity indices for *S. marcescens* and *P. syringae* pv. *tomato* strains (yellow-NED)

	731-17	ATCC-2987	DB-11	HO1A	PO1	ROZ	ZO1A	WO1A
CPST 147	0.11	0.28	0.25	0.11	0.25	0.17	0.2	0.21
CPST 232	0.09	0.25	0.32	0.19	0.11	0.38	0.26	0.36
RG4	0.22	0.14	0.13	0.24	0.27	0.36	0.32	0.44
880	0.21	0.21	0.38	0.22	0.25	0.35	0.3	0.42
T1	0.09	0.24	0.3	0.18	0.1	0.37	0.25	0.35
TF1	0.17	0.18	0.4	0.27	0.2	0.44	0.33	0.43
T4B1	0.18	0.25	0.32	0.19	0.21	0.38	0.35	0.36
1318	0.17	0.36	0.3	0.27	0.3	0.44	0.25	0.35
30555	0.2	0.2	0.35	0.21	0.24	0.33	0.29	0.3
B125	0.2	0.2	0.35	0.21	0.24	0.33	0.29	0.3
3357	0.17	0.18	0.3	0.18	0.2	0.37	0.25	0.35
PT17	0.19	0.19	0.33	0.1	0.22	0.32	0.27	0.29
1008	0.11	0.28	0.27	0.12	0.27	0.27	0.21	0.33
188B	0.1	0.25	0.22	0.1	0.22	0.24	0.18	0.19

Sorensen similarity indices for *S. marcescens* and *P. syringae* pv. *tomato* strains (combined)

	731-17	ATCC-2987	DB-11	HO1A	PO1	ROZ	ZO1A	WO1A
CPST 147	0.22	0.17	0.25	0.24	0.21	0.22	0.17	0.22
CPST 232	0.20	0.08	0.28	0.22	0.19	0.27	0.25	0.30
RG4	0.18	0.11	0.18	0.19	0.17	0.23	0.20	0.25
880	0.19	0.13	0.28	0.29	0.25	0.29	0.25	0.30
T1	0.19	0.11	0.28	0.24	0.18	0.26	0.24	0.29
TF1	0.22	0.09	0.31	0.27	0.21	0.28	0.27	0.31
T4B1	0.19	0.11	0.28	0.29	0.25	0.30	0.30	0.32
1318	0.23	0.20	0.34	0.34	0.28	0.31	0.27	0.32
30555	0.23	0.09	0.26	0.28	0.23	0.30	0.26	0.28
B125	0.21	0.10	0.26	0.21	0.20	0.22	0.23	0.25
3357	0.25	0.12	0.28	0.29	0.25	0.30	0.27	0.32
PT17	0.19	0.09	0.26	0.21	0.18	0.24	0.21	0.23
1008	0.23	0.17	0.27	0.25	0.22	0.26	0.22	0.27
188B	0.22	0.11	0.25	0.26	0.26	0.25	0.25	0.27

VITA

Andrew Kenneth Taylor

Candidate for the Degree of

Master of Science

Thesis: EVALUATING THE SUITABILTIY OF AFLP GENOTYPING TECHNOLOGY FOR DISCRIMINATING AMONG STRAINS OF *PSEUDOMONAS SYRINGAE*

Major Field: Forensic Science

Biographical:

Education:

Completed a Bachelor of Science in Zoology at The University of Oklahoma, Norman, Oklahoma in May, 2007.

Completed the requirements for a Master of Science in Forensic Science at Oklahoma State University – Center for Health Sciences, Tulsa, Oklahoma in December, 2009.

Experience:

Professional Memberships:

Name: Andrew Kenneth Taylor Date of Degree: December, 2009

Institution: Oklahoma State University Location: Tulsa, Oklahoma

Title of Study: EVALUATING THE SUITABILITY OF AFLP GENOTYPING TECHNOLOGY FOR DISCRIMINATING AMONG STRAINS OF *PSEUDOMONAS SYRINGAE*

Pages in Study: 73 Candidate for the Degree of Master of Science

Major Field: Forensic Science

Scope and Method of Study: This study was designed to assess the suitability of amplified fragment length polymorphism (AFLP) analysis for discriminating among strains of *Pseudomonas syringae* pathovar (pv.) *tomato*, a species of plant pathogenic bacteria. 14 strains of *P. syringae* pv. *tomato* were used as model plant pathogens to assess the reliability and discriminatory power of the assay. DNA was extracted using organic extraction and quantitated using spectrophotometry. The strains were processed using the AFLP Microbial Fingerprinting kit from Applied Biosystems. Capillary electrophoresis and GeneMapper ID software were used to analyze the data.

Findings and Conclusions: The AFLP profiles generated in this study exhibited over 95% reproducibility through three replicate assays done on each bacterial strain. Visual comparisons of electropherograms, as well as the use of a numerical haplotype code describing the AFLP profiles, showed that all *P. syringae* pv. *tomato* strains tested were distinguishable from on another. The discriminatory power of the AFLP assay was characterized through pairwise comparisons of haplotype codes among all 14 strains of *P. syringae* pv. *tomato*. Power was also assessed across species by comparing haplotype codes of *P. syringae* pv. *tomato* with the codes of eight strains of *Serratia marcescens* and three strains of *Staphylococcus aureus*. The assay discriminated 100% of the *P. syringae* pv. *tomato* strains tested, as well as 100% of the strains across the three species mentioned above. The technology provided a probability of exclusion of 100% in practice (with the lower interval of a 95% C.I. equal to 98.35%) within the *P. syringae* pv. *tomato* species and 100% (with the lower interval of a 95% C.I. equal to 99.5%) across the three species. AFLP technology is therefore reliable and sufficiently powerful for routine use in microbial forensics.

ADVISER'S APPROVAL: Dr. Robert Allen