This is to certify that the

thesis entitled

THE EVALUATION OF INDIVIDUAL LOCUS PERFORMANCE USING THE PROMEGA POWERPLEX 16 SYSTEM FOR USE WITH SINGLE SOURCE CODIS SAMPLES

presented by

Teri Lynn Lawton

has been accepted towards fulfillment
of the requirements for

M.S. degree in Forensic Science

Major professor

Date 11/2/01

THE EVALUATION OF INDIVIDUAL LOCUS PERFORMANCE USING THE PROMEGA POWERPLEX™16 SYSTEM FOR USE WITH SINGLE SOURCE CODIS SAMPLES

By

Teri Lynn Lawton

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

School of Criminal Justice

2002

ABSTRACT

THE EVALUATION OF INDIVIDUAL LOCUS PERFORMANCE USING THE PROMEGA POWERPLEX 16 SYSTEM FOR USE WITH SINGLE SOURCE CODIS SAMPLES

By

Teri Lynn Lawton

In the field of forensic DNA analysis, a genotyping kit (the Promega PowerPlex™16 System) has been developed which identifies genotypes at sixteen different loci. This kit would be ideal for the genotyping of samples that must be entered into the Combined DNA Index System (CODIS), a DNA database for convicted offenders of violent crimes. It would be an improvement over the current method of analysis, because it only requires one reaction (as compared to two)—but it must first prove to be an effective and reliable method for the analysis of single source samples. This project will evaluate the performance of this kit at each locus and determine whether it would be suitable for single source samples. This project will evaluate locus performance by comparing values such as peak height ratios and relative fluorescence units between and within each locus. One hundred and fourteen individuals will be tested to provide data needed to compare these values and determine whether the PowerPlex™16 System would be a reliable method for genotyping single source samples in the forensic DNA laboratory.

ACKNOWLEDGEMENTS

I would like to acknowledge the Michigan State Police Forensics Laboratory (DNA Unit) in Lansing, Michigan for the opportunity to perform research at their facility. I would especially like to thank DNA Laboratory Supervisor Charlie Barna for giving me the opportunity to work with the PowerPlex™ 16 System, Forensic Scientists Don Yet and Glen Hall for their assistance throughout the project, and the rest of the MSP DNA Unit Staff who were always there to answer endless questions. In addition, I would also like to thank the Promega Corporation for providing the PowerPlex™ 16 Kits and DNA extracts that were utilized in this evaluation.

I would also like to acknowledge Dr. Jay Siegel for his encouragement, support, and persistence throughout my time as a graduate student at Michigan State University. Lastly, I would like to thank my parents and family for always encouraging me to finish my projects, and for all of the support they have given me the past few years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

I. The Value of DNA Evidence

The application of deoxyribonucleic acid (DNA) analysis to the field of forensic science has broadened the horizons of criminal investigation procedures. Evidence obtained from a crime scene that contains DNA can be the prime incriminating or exonerating factor in a case. DNA is present in most biological fluids such as blood, semen, vaginal fluid, saliva, and can be occasionally found in urine or feces. Other samples from which DNA can be extracted include hair, bone, and tooth pulp. Blood samples tend to produce the best DNA yield (twenty to forty thousand ng/ml), while urine and bone the least (one to twenty ng/ml) [1]. From their introduction to forensic science applications, many types of DNA analysis have evolved through time, with the most recent type termed short tandem repeat (STR) analysis. This process utilizes short repeating sequences that occur throughout the genome to calculate a frequency of a particular genetic makeup (genotype).

The Federal Bureau of Investigation (FBI) has established the COmbined DNA Index System (CODIS)—a database containing genetic profiles of persons convicted of sexual offenses and other violent crimes. Now, investigators have a useful tool with which they can compare DNA evidence found at the crime scene with a database of known convicted offenders. The ultimate goal for CODIS is for every state to have a database of DNA profiles collected from the scene of the crime, or from the criminals themselves. This data can then be centralized, allowing each state the ability to search and compare their data with all of the other states.

The FBI has determined the thirteen STR loci which must be included in every genetic profile of an individual convicted of a number of offenses, which are determined by each of the member states in the National CODIS database (the loci being: D3S1358, THO1, D21S11, D18S51, vWA, D8S1179, TPOX, FGA, D5S818, D13S317, D7S820, D16S539, and CSF1PO). A match with these STR loci between two samples can produce random match probabilities in the quadrillions.

Currently, most forensic DNA analysts are utilizing the Applied Biosystems (Foster City, CA) AmpF*l*STR® Profiler Plus™ and CoFiler™ genotyping kits; these kits test multiple genetic STR loci in one reaction, and therefore are termed multiplex kits. In this procedure, two separate runs must be set up to obtain results for all thirteen CODIS loci. Setting up two separate analyses depletes the original sample, consumes valuable human and monetary resources, and most importantly—these additional testing steps can increase the chance of inadvertent sample transfers and other types of sample integrity concerns.

With the concerns of the forensic scientist in mind, the Promega Corporation (Madison, WI) has developed the GenePrint® Power Plex™16 multiplex kit which determines the genetic profile of an individual at sixteen different loci, thirteen of which are the required CODIS loci. This genotyping system will save the analyst (and the agency for which he/she works) time, money, and sample—if it is proven to be an effective, reliable tool for the typing of CODIS samples. The PowerPlex™16 system would allow for faster processing of convicted offender samples, whose DNA profiles could be entered into the CODIS database quicker, thus providing the states with another DNA profile of which to search.

This thesis project is the evaluation of individual locus performance using Promega's PowerPlex™ 16 multiplex system for use with single source CODIS samples. In this thesis, the following issues will be addressed:

1. The optimal target quantity of DNA per reaction
2. The level of performance at each locus
3. The viability of this kit for database samples

This thesis project was part of a larger project by the Promega Corporation and a number of other forensic DNA testing facilities. The data generated in this evaluation was included in the "STR primer concordance study" [2], which compared DNA profiles obtained with the PowerPlex™16 typing kit with their corresponding DNA profiles using Applied Biosystem's Profiler Plus™ and CoFiler™ typing systems. Results in the STR primer concordance study indicated that the primers used in the PowerPlex™16 Kit, Profiler Plus, and CoFiler Kits produced reliable, consistent DNA typing results obtained on reference samples.

To understand the use of STRs with CODIS and how a genetic profile is obtained, some background material is presented on the structure and function of DNA, and the history of various DNA analysis methods leading up to the most recent application of DNA analysis—the development of STR multiplexing.

II. The structure of DNA

Deoxyribonucleic acid was first isolated in 1869 by a Swiss chemist Johann Friedrich Miescher, but it was not until the 1950's when James Watson and Francis Crick combined all of the data and created a working model of DNA. DNA is composed of subunits, with each subunit containing a nitrogenous base, a pentose sugar, and a phosphate group. The nitrogenous bases fall into two categories: the pyrimidines (a six

member ring) and the purines (structures composed of two rings). There are four types of nitrogenous bases that make up DNA—two pyrimidines (cytosine and thymine) and two purines (guanine and adenine).

There are two types of nucleic acids—deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). What differentiates RNA from DNA is that the sugars contained in the molecule RNA are riboses (not deoxyriboses, as in DNA). In addition, the bases that make up RNA include guanine, adenine, cytosine, and uracil (instead of thymine). Uracil base pairs with adenine (like thymine) but is structurally different than thymine (lacks a methyl group). After formation of the DNA template in a cell, it is transcribed (or read) into a specific kind of RNA molecule, which is then translated into a specific protein. DNA is the hereditary material found in all living organisms, while RNA (or DNA) can be the hereditary material found in viruses.

When a nitrogenous base is bonded to a pentose sugar and a phosphate group, the structure is termed a nucleotide. Chains of nucleotides are what make up the structure of nucleic acids. In 1953, Watson and Crick proposed that the DNA molecule was made up of two polynucleotide chains and that each of the nucleotides was paired such that a pyrimidine is always opposite a purine. They also determined that the proportion of cytosine to guanine was always 1:1, as was the proportion of adenine to thymine. This suggested that in the double helix, the cytosine was always paired with guanine (a purine with a pyrimidine) and thymine with adenine. This is referred to as complimentary base pairing. Watson and Crick determined that the two polynucleotide chains stay associated in the double helix structure by hydrogen bonding of the nitrogenous bases, with three hydrogen bonds holding cytosine and guanine together and two hydrogen bonds between

the adenine and thymine bases. The structure of the DNA double helix has often been compared to a ladder, with the rungs of the ladder representing the nitrogenous bases held together by hydrogen bonds.

The organization of the bases in the DNA molecule form the alphabet for the genes contained in a particular sequence. The DNA is transcribed into an mRNA molecule, which is then translated into proteins that are expressed by the cell. Different forms of the same gene can occur and are termed alleles. An example of a gene with multiple alleles is that for eye color. It is often useful to measure the variations of an allele in a population, as it is used in forensic science to determine a frequency of an individual's genotype within a distinct population.

III. DNA Inheritance

In 1944, O.T. Avery and colleagues performed laboratory experiments in which they observed the effect of injecting virulent and avirulent strains of the bacterium *Pneumococcus* into mice [3]. They found that by isolating DNA from one strain of bacteria, its traits could be transferred to a new colony of bacteria, thus proving DNA to be the hereditary material for which genetic information was transferred. Not only is DNA the hereditary material of bacteria, but for all living organisms, and it resides in the nucleus of cells, organized into structures termed chromosomes. Humans have forty-six chromosomes in every cell that contains a nucleus—two sets of twenty-three (one set is paternally derived, and the other maternally). The only exception to this is that there are only twenty-three chromosomes in the sex cells (gametes). Upon fertilization of an egg and a sperm, the cell then contains the full forty-six chromosomes. In addition to nuclear DNA, the cells containing mitochondria (organelles in the cytoplasm for cellular

respiration) have their own circular DNA template termed mitochondrial DNA (mtDNA). Mitochondrial DNA, like nuclear DNA, is inherited, though only maternally. Due to limited DNA typing capabilities, mtDNA is generally used for the tracking of relatives and/or when nuclear DNA typing has failed due to minimal quality or quantity.

IV. DNA Replication

A robust method for the amplification of DNA sequences is termed the polymerase chain reaction (PCR). The principles underlying the polymerase chain reaction involves the *in vivo* process of DNA replication and how scientists were able to mimic this process *in vitro* to produce an invaluable technique for the analysis of DNA sequences. Since DNA is base-paired in its double helix—each strand can serve as a template for the synthesis of a complimentary daughter strand, and each strand of the original double helix is paired with a new strand. In other words, the parental DNA helix replicates to form two identical daughter duplexes. This type of replication is called semi-conservative, since one strand of the "original" is conserved in the next round of replication. This is very similar to the process of PCR amplification—which also utilizes the original strands of DNA to replicate the desired sequence millions of times. One strand of the double-stranded DNA has all the information needed to build its complimentary strand, but requires an enzyme to catalyze the addition of these complimentary nucleotides. In vivo, this enzyme is DNA polymerase. DNA polymerase does not act alone; it requires a primer (a short sequence of single-stranded DNA) to first bind the DNA template, which is what initiates the elongation of the complimentary strand by the polymerase. The in vitro process utilizes a similar DNA polymerase and will be discussed in the section on PCR analysis.

V. The History of Forensic DNA Analysis

The first type of forensic DNA analysis was Restriction Fragment Length Polymorphism (RFLP), developed by Alec Jeffries in 1985. In this process, after DNA is isolated, it is digested with a restriction enzyme that recognizes specific sequences in the DNA and cleaves at these sites. The fragmented DNA is then separated by size using agarose gel electrophoresis. The DNA fragments are then transferred to a nitrocellulose or nylon membrane (Southern blot). The membrane is then hybridized with a radioactive or chemiluminescent probe, that identifies the alleles at one locus by complimentary base-pair binding of the probe to the fragmented DNA. The probe can be stripped from the membrane and a new locus can be examined. Using multiple probes on a sample decreases the probability of another person having those alleles. RFLP analysis is excellent for samples that are not degraded, since the fragments obtained with this type of assay are in the order of five hundred base-pairs and larger. RFLP analysis for the genotyping of degraded samples can be difficult, since the detection of larger alleles might be missed.

By 1986, Kary Mullis had invented a new typing technique, capable of analyzing limited or degraded samples and was termed the polymerase chain reaction (PCR). This process exponentially multiplies the amount of DNA present, therefore making it easier to analyze. PCR amplification will be described in depth in the discussion of mulitiplexing.

In addition to RFLP and PCR, DNA sequencing is another method which can be used by the forensic scientist to determine the genetic profile of an individual. Unlike the previous methods, there is no comparison of alleles in this method; the process

works by fluorescent based detection of the DNA sequence products. Though this method provides the examiner with an exact genetic makeup of the individual, the process is very time consuming and requires costly equipment. Until the day when the equipment becomes less expensive and the process less cumbersome, other methods of analysis will generally be employed.

VI. Short Tandem Repeat (STR)

One of the fastest growing methods of DNA analysis in the forensic science community is the analysis of short tandem repeats. STRs are sequences in the genome made up of approximately three to seven base pairs that repeat as a sequence a variable number of times and occur repeatedly throughout portions of the genome. Currently, there are an estimated thirteen hundred different STR loci [1], but less than two percent of those are currently used in the forensic science laboratory. STRs are a type of Variable Number Tandem Repeat (VNTR), meaning they contain a tandemly repeated sequence and their fragment size depends on how many repeats are in the sequence. An example of an STR is the sequence "AATG", which can be found at the THO1 locus. This sequence occurs in tandem repeats, where the sequence is repeated as a unit (AATG-AATG). The number of repeats is responsible for the genetic variation at that particular locus. For example, a person may have two alleles—one with six repeats, and one with ten, each derived from a parent. In addition, STRs fall into non-coding regions of the genome, but they are still inherited, just like alleles coding for a particular gene.

The most recent method of DNA fragment analysis is based on fluorescence detection of PCR products. Like DNA sequencing, this process also utilizes DNA fragments labeled with fluorescent dyes (incorporated during PCR amplification); but

instead of obtaining the actual DNA sequence, one obtains the alleles present for specific STR loci. One of the first instruments utilized for fluorescent detection of STRs was a flat bed laser scanning instrument (Hitachi FMBIO), but because of the multiple problems associated with this cumbersome method, it was soon replaced by capillary electrophoresis (CE). As where the FMBIO required a polyacrylimide gel for separation, CE incorporates a polymer-filled capillary (or column) for separation of fluorescently labeled DNA. Like the FMBIO, CE utilizes a laser to excite the fluorescently labeled DNA products (which in turn produce a spectrum of light). The different fluorescent signals are separated according to their wavelength and are displayed by a CCD camera. A filter allows for separation of the signals, and a matrix is incorporated to normalize the fluorescent intensities for each dye. The ABI PRISM® 310 Genetic Analyzer (Applied Biosystems) is an instrument that uses a capillary based separation scheme to fluorescently detect tagged DNA fragments. The ABI 310 allows the analyst to test multiple loci in one reaction and allows for examination of alleles with overlapping size ranges (because of the different dyes employed). In addition, an internal lane standard (ILS) is incorporated with every sample; this enables automated sizing of fragments based on a curve derived from the specified ILS fragments. The ILS normalizes any differences in electrophoretic mobility and is labeled with a different dye than all of the other loci. This is to ensure that there is no overlap of DNA fragments with the ILS itself. There are three different software programs associated with the 310 Genetic Analyzer: the 310 Data Collection® software (ensures proper operation of the instrument); Genescan® software (applies the matrix and determines the size of the DNA fragments using the ILS; and lastly, Genotyper®/PowerTyper™ software determines the

genotypes based on an allelic ladder from the same run. An allelic ladder, sequenced to verify fragment lengths and repeat structure, is included with every PowerPlex™16 System. It contains the most common alleles for each locus and by comparisons of sizes obtained for samples with the known sizes in the allelic ladder, a genotype can be assigned. The size ranges of each locus are the actual base pair sizes of the sequenced alleles. Within each locus, each complete four base pair repeat unit is designated by a whole number, and alleles that contain a partial repeat are the whole number designation, followed by a decimal point and the number of base pairs in the partial repeat. For example, the D21S11 33.2 allele contains 33 complete four base pair repeats and a partial repeat of two base pairs. The exception to this method is for loci with five base pair repeats (such as Penta D and Penta E), where the whole number designation must be a five base pair unit.

The Polymerase Chain Reaction and the Development of STR Multiplexing

I. The Polymerase Chain Reaction

The polymerase chain reaction (PCR) is a technique that copies a specific sequence of DNA and amplifies that segment exponentially so that millions of copies are present at the end of the replication procedure. Since PCR needs only a small amount (0.5-1ng) of template DNA to start the reaction process—this technique is ideal for the analysis of limited or degraded samples, especially when compared to RFLP analysis, which can consume up to five hundred nanograms of high molecular weight DNA to generate a genetic profile.

II. PCR operation

PCR has three main steps: denaturation, annealing, and extension of the primers. Denaturation is the first step and takes place at approximately 95° C for one minute. This cleaves the double stranded DNA molecule into two complementary daughter strands. This allows the DNA to be accessible for the next step in this reaction. The second step in the PCR process is the annealing of the primer sequences to the original DNA template strands. Primers are designed so they are complimentary to the flanking regions of the desired sequence, allowing them to bind specifically at those locations under thermal condition where the exact complementary pairing of the primer to the template DNA strand is favored. The annealing step generally occurs at 55° C to 60° C for 30-45 seconds. A delicate balance must be found to optimize a PCR reaction, a temperature too low will allow non-specific PCR products (miss-pairing) to occur, and an annealing temperature too high will inhibit the primers from binding to the template DNA strand. The final step in PCR is the extension of the primers by the use of a DNA polymerase.

Taq polymerase (derived from the bacterium *Thermus aquaticus*) is a commonly used thermostable polymerase which facilitates the addition of nucleotides to the extending molecule, while maintaining its viability in temperature ranges required for PCR. The Taq polymerase attaches to the DNA template at the primers and begins the additions of the complimentary nucleotides to extend the primers. This step takes place at approximately 72° C for 30 seconds. This increase in temperature will also aid in the dissolution of miss-paired primers from the template DNA. At the end of this step, two identical copies have been made from one original double-stranded sequence. In order to achieve millions of copies of the particular sequence, the three steps are repeated for about 30 cycles. This results in an exponential amplification of the original sequence (the first cycle produces two copies from one molecule; the next cycle produces four copies from the two present, and so on.).

III. Reagents/instruments needed for PCR

There are many items that work together to optimize the PCR reaction, and these reagents are combined together in what is termed a "PCR cocktail". A standard cocktail contains KCl (a salt), Tris-HCl (a buffer), deoxynucleotide triphosphates (dNTPs; these are free nucleotides that are incorporated in the extension of the molecule), $MgCl_2$ (the Mg ions work with Taq polymerase and the dNTPs), BSA (bovine serum albumin helps prevent PCR inhibitors by binding to them), and Taq polymerase. All of these reagents must be balanced within the reaction for optimal results, usually by altering one reagent at a time until an optimal point is reached. With commercial STR typing kits, this optimization has already been performed and the reagents arrive pre-mixed in a reaction buffer cocktail—all that is needed is the Taq polymerase and the template DNA. Once

the PCR cocktail (including the Taq) has been aliquoted and the DNA templates added, the samples (usually 25ul reactions) are amplified in a thermal cycler. This instrument allows for the alternate heating/cooling cycles that PCR reactions must undertake. An important consideration in the setup of a PCR reaction is the amount of template DNA required for optimal results. Normally, between approximately 0.5ng and 2.5ng of template DNA produce the best results, but each kit must be tested with a range of DNA quantities to determine the amount which produces the best results. If too much DNA template is added, then an overabundance of PCR product is generated. In fluorescent detection of PCR products, the ramifications of an excess of PCR product can include data containing "off-scale" peaks (there is too much fluorescent intensity for detection by the instrument). Additionally, excess PCR product can cause "pull-up" (an artifact caused by poor separation of the fluorescent dyes due to the excessive amount of fluorescently labeled PCR product). On the other hand, if too little DNA is added to the reaction, unbalanced amplification can occur at heterozygous loci (loci with two different alleles).

IV. The first types of PCR tests developed for forensic use

Tests that identified sequence polymorphisms were the first forensic tests which utilized PCR. Sequence polymorphisms occur when there is a mutation in one base-pairing of a particular sequence. These polymorphisms can be identified through a test called a reverse dot blot. In a reverse dot blot, PCR product is added to a nylon membrane which has DNA probes attached to it. The probes and the PCR product are complimentary sequences of the same locus—the probes are commercially available and the PCR product is obtained by the analyst via extraction and amplification. Before PCR

product is added to the probe-coated membrane strips, the strips are white—upon addition of the PCR product, complimentary sequences of the probe and the product turn blue, thus identifying the alleles in the PCR product. The most common forensic DNA tests developed which identified sequence polymorphisms were DQα (Applied Biosytems) and the AmpliType®PM system (Roche Molecular Systems).

In addition to tests that identified sequence polymorphisms, PCR was utilized to amplify length polymorphisms (VNTRs), such as those occurring at the locus D1S80. As described previously, length polymorphisms contain a core repeating sequence; the number of times the core repeats is determined by its fragment size on a polyacrylamide gel. Fragments are compared to a molecular ladder and the alleles determined (the alleles are numbered according to the number of repeats in the sequence). For example, an individual could be a 10, 12 at this locus, meaning they have an allele with ten repeating core elements, and an allele with 12 repeating core elements. Unfortunately, this method of analysis is not highly discriminative, since it only analyzes one locus.

The most recent technology in forensic DNA analysis that applies PCR technology is in the typing of STRs. STRs, a type of VNTR, are length-based polymorphisms, but the core repeat is much smaller (three to seven base-pairs) compared with the sixteen base-pair core repeat in D1S80. Though one individual STR locus does not have a significant amount of variation (usually about five to twenty alleles), STR loci can be combined in a multiplex reaction—which simultaneously types all of the STR loci, thereby greatly increasing the variability. Mulitplexing uses the same concept of PCR as if it were a single locus amplification, but instead of one set of primers in the

PCR cocktail, there is a set of primers for each locus (the PowerPlex ™ 16 kit thus has sixteen sets of primers).

V. Previous Multiplex PCR Genotyping Kits

Wallin and coworkers [4] developed one of the first commercially available multiplex kits available for forensic identity casework. It genotyped three STR loci—D3S1358, vWA, and FGA, and was called the AmpFlSTR® Blue PCR Amplification Kit. Blue referred to the color of the dye (5-FAM-5-carboxyfluorescein) used to fluorescently label the primers. Instead of visualizing the PCR product on an acrylamide gel, capillary electrophoresis was utilized to detect the alleles present. As described earlier, capillary electrophoresis is a separation technique in which the DNA sample is carried through a capillary and the fluorescently labeled fragments are excited by a laser at the end of the capillary journey. The fluorescence is collected and focused by mirrors onto a CCD camera and visualized as an electropherogram generated by the instrument. Two instruments used for the detection of fluorescent PCR products are the ABI PRISM® 310 Genetic Analyzer (Applied Biosystems), a capillary electrophoresis instrument, and the ABI PRISM® 377 DNA Sequencer (Applied Biosystems), an instrument which utilizes slab gel electrophoresis. GeneScan® software is utilized to analyze the electropherograms; it applies a mathematical matrix model to separate emissions from different dyes in the sample being analyzed, and it creates and applies a size curve to the data generated from the sample being analyzed. The alleles from the DNA sample are represented as peaks on a graph, with the size of the peak being related to how much DNA is present in the amplified sample. The information that can be derived from an electropherogram includes: the alleles in the sample (peaks), relative

amount (measured by relative fluorescence units (r.f.u.)), and the time with which the laser detected them (or that they emerged through the capillary). As in other separation techniques, the smaller fragments elute quicker than the larger fragments. In addition to the DNA samples run on the 310 or 377, an internal lane standard (ILS) is also run with each DNA sample. This standard contains labeled fragments of known lengths and thus provides a manner in which all of the peaks in the run can be determined. Lastly, an allelic ladder, containing all of the common allele types for the loci is run in a separate injection, so that the alleles in the DNA sample can be compared and determined, using either manual determinations or automation software. Genotyper™ (Applied Biosystems) is the software program available to perform this task; it uses the allelic ladder to create allelic matching windows, which are applied to all of the samples in the run. In Wallin's validation of the AmpF*l*STR® Blue Kit, the 377 DNA Sequencer and the ABI PRISM® 310 were utilized and Genescan® software was utilized for analysis. The loci chosen for the AmpF*l*STR® Blue kit were chosen because the overall size ranges (per locus) were very small, hence, the primers produced short PCR products (one hundred to three hundred-fifty basepairs). This is ideal for samples containing degraded DNA, and their sizes minimized the occurrence of preferential amplification. Preferential amplification (as defined by Wallin) occurs when there is a difference in amplification of two alleles within the same locus (evident upon comparison of peak heights of heterozygotes). Preferential amplification can lead to a loss of information regarding that locus, and a mistyping could possibly occur. To avoid preferential amplification, STRs with small size ranges are generally employed (as with the Blue kit). In their validation studies, Wallin and coworkers noted differential amplification with increasingly degraded DNA,

evident by the drop-out of the largest locus (FGA). Wallin described differential amplification as the occurrence of one or more loci amplifying less than the other loci—a difference in amplification between loci. Differential amplification may also result in a loss of information regarding a particular locus, though no mistyping can occur. Differential and preferential amplification must be considered when developing and validating multiplex PCR kits. Though a kit may be optimized for minimal preferential and differential amplification, other factors can promote the same effects, such as degraded DNA and the presence of inhibitors (soil (which contains metal ions, enzymes, and other proteins), bleach, dyes, etc.). There is a way to determine if something is inhibited versus being degraded. In degraded DNA samples, one gets more signal (higher r.f.u.) when more DNA is added. In the case of an inhibited sample, if it is diluted, it may promote a more efficient amplification. It is important to determine if the kit itself promotes preferential/differential amplification, since many of the samples forensic analysts encounter are either degraded or inhibited. Overall, the AmpF*l*STR® Blue kit had minimal preferential/differential amplification. In addition to testing the occurrence of differential/preferential amplification, Wallin also compared single locus to multilocus performance. For the single locus amplifications, individual primers were used and only one locus per reaction was employed. They obtained the same genotyping results for both systems and no significant difference in peak heights were noted. They concluded that the single locus reaction had no benefit over the multiplex reaction.

During this same time, Micka and coworkers [5] had also been validating two multiplex kits which each typed three STRs. Like Wallin, they determined that the performance of a single monoplex reaction had no advantage over the use of the

multiplex kits. This led researchers to develop larger multiplex kits—such as the one developed by Lins [6]. Lins and coworkers developed an eight-locus, two-color STR multiplex system for human identification. This eight-plex combined two four-plex systems: the CTTv and GammaSTR™ multiplex system. The CTTv multiplex genotyped STRs at the following loci: CSF1PO, TPOX, THO1, and vWA, while the GammaSTR™ system genotyped D16S539, D7S820, D13S317, and D5S818. When combined together, the multiplex was termed the PowerPlex™ System. The Power Plex™ System combined the primers for CTTv and GammaSTR™ loci. Originally, all primers were labeled with fluorescein (FL—recognized as a blue dye), but to achieve better separation, the CTTv multiplex was changed to contain primers labeled with carboxy-tetramethyl rhodamine (TMR—recognized as a yellow dye). This provided better resolution of alleles within the one hundred to four hundred base-pair range. Though labeled with different dyes, the eight sets of primers were combined into one reaction mix. In addition, allelic ladders were developed, which contained all of the alleles for both systems (both the FL and TMR labeled products), to aid in the determination of allele calls. The primers were modified to eliminate artifactual bands due to incomplete terminal addition of adenine, also referred to as –A or +A. With incomplete terminal nucleotide addition, two peaks are present, usually of unequal peak height. One peak represents the allele without the terminal adenine, and the other peak represents the STR with the terminal nucleotide addition. In this case, the primers were modified to help minimize this occurrence. In addition, a thirty minute extension at sixty degrees Celcius following the PCR cycling process reduces –A/+A. This process promotes full terminal nucleotide addition to one hundred percent—therefore only one

peak is present on the electropherogram. In the development of this system, Lins chose STRs with low mutation rates and with few microvariants. These factors contribute greatly to the reliability of the genotyping of the alleles in the system. If an STR has a high mutation rate, or has many variants of the common alleles, the allelic ladder cannot be used by the computer software to accurately determine the allele types. Also, the STRs that were chosen by Lins had relatively low stutter. Stutter is the occurrence of a smaller band, or peak, one repeat smaller (or larger) than the primary band. A peak is deemed stutter if it is approximately ten to fifteen percent of the primary peak (this percentage differs between loci). One theory, supported by Walsh [7], explains that the occurrence of stutter is due to slipped strand mispairing. During PCR, Taq polymerase may fall off, giving one strand a chance to loop out before they bind again. Once together again, one of the strands is shorter by one repeat unit. Stutter is a reproducible artifact and does not interfere with genotyping of a particular sample, unless the sample is a mixture. That is why it is important to choose STRs with low stutter percentages, since many samples encountered in forensic science analysis are mixtures and deciphering the genotypes would be difficult if stutter were high. Lastly, Lins combined the eight STR multiplex system with the primers for the Amelogenin locus to obtain gender information. Amelogenin was labeled with TMR and produced specific fragments for the X and Y alleles.

After development of an STR multiplex, such as the PowerPlex™ System, it must undergo validation studies to be used for forensic science casework. Micka [8] validated the GenePrint® PowerPlex™ 1.1/Amelogenin System developed by Lins and coworkers (for the Hitachi FMBIO Fluorescent Scanner). As previously described, this system

contained two groups of four STRs—each group labeled with different fluorescent dyes. Additionally, the locus for gender identification was also incorporated. Micka's results indicated no differential amplification (allelic drop-out), or any other artifactual bands. Other artifactual bands may include primer-dimers (when the primers bind to themselves) or partial binding of the primer to a complimentary sequence in the DNA. In addition, Micka did single locus versus multi-locus amplification studies. In both cases, the systems produced the same genetic typing results.

The development of STR multiplexing kits has made life for the forensic DNA analyst a little bit easier. Previous kits have shown that multiplexing produces the same genotyping results as single locus amplifications—meaning one can obtain genetic typing results for many loci in one reaction, rather than setting up individual reactions for each locus. This saves sample amount, decreases the chance for contamination, and allows for quicker genotyping results. Like the multiplexing kits that came before it, the PowerPlex™ 16 kit would be another useful tool for the forensic scientist. It allows the genotyping of the thirteen required CODIS loci to be analyzed in one reaction, instead of setting up two different reactions, as is the current method. To determine whether the PowerPlex™ 16 system is suitable for single source samples, such as those submitted for CODIS, evaluations on locus performance must first be made.

Materials and Methods

Samples: Promega Corporation (Madison, WI) provided three of the DNA extracts used in this project: B15 (1ng/ul), H9 (1ng/ul), and standard DNA template 9947A (10 ng/ul). They had previously been quantified by Promega.

The Michigan State Police DNA/Forensic Biology Unit (MSP) provided a blood stain of known origin and genotype (MBILC). This blood stain sample required an organic extraction. To perform the extraction, small cuttings of the bloodstain were placed in a Spin-Ease (Gibco BRL) extraction tube. Approximately 400ul stain extraction buffer (SDS, EDTA, NaCl, and Tris) and 20ul of Proteinase K (BRL) were added. After an overnight incubation at 56° C, the cuttings were removed and approximately 400ul phenol/chloroform/isoamyl alcohol (24:1:1) was added. After vigorous vortexing, the sample was then spun down at 14,000 x g to facilitate the separation of the aqueous and organic layers. The upper layer (the aqueous layer containing the DNA) was transferred into a Centricon® 100 Concentrator (Millipore Corp.). The Centricon 100 employs a size exclusion membrane which retained the large molecules (the DNA) and allowed for passage of the smaller ones (such as salt ions, detergents, and fragmented proteins). The sample was then rinsed with approximately 400ul TE^{-4} buffer and spun at 500 x g for 15 minutes. This step was repeated three times. Lastly, the DNA that was retained by the membrane was captured by inverting the sample reservoir and spinning the sample at 1000 x g for 3 minutes. The DNA recovered was then transferred to a new 1.5ml microcentrifuge tube for long term storage at –20 C.

Quantitation: Upon completion of the extraction and purification of the DNA, the sample was assessed for quality and quantity using a yield gel (an ethiduium bromide

fortified 1% agarose gel). The yield gel contains a visual marker (lambda HINDIII/ECOR1) and human DNA quantification standards (BRL) to which one can compare extracts to determine the relative amount of DNA present. In addition, the yield gel will identify whether the sample is degraded (observable by a smear instead of a band when the gel is subjected to short wave UV radiation). Following detection using the yield gel, slot blot quantitation using the Applied Biosystems QuantiBlot® Kit was performed to determine a more accurate representation of the DNA present (this method utilizes the knowledge obtained from the yield gel to make an approximate dilution for the slot-blot). This quantitation method entailed the immobilization of denatured target DNA on to a charged nylon membrane (Pall-Biodyne B). In addition, this method is species-specific by utilizing a primate specific DNA probe (D17Z1) in the hybridization process. Once again, this kit also contains human DNA standards, which are compared to the unknown samples to determine an approximate quantity of DNA in the sample. Chemiluminescent detection utilizing ECL (Amersham Pharmacia) and Kodak XLS film was utilized to observe the results of the QuantiBlot.

MSP also provided DNA extracts of approximately 114 individuals for analysis using the Geneprint® PowerPlex™ 16 System (Promega Corp.) The extracts were quantified using the Applied Biosystems Quantiblot® Kit prior to analysis using the PowerPlex™ 16 kit.

PCR Amplification: Dilutions: Upon extraction of MBILC, dilutions of 2ng/10ul, 1ng/10ul, 0.5ng/10ul, 0.25ng/10ul, 0.125ng/10ul, and 0.0625ng/10ul were set up for each of the samples (except for 9947A at 0.0625ng/10ul). Dilutions were made with 0.2um filtered high purity ($18M\Omega cm^{-1}$) water and placed in new microcentrifuge tubes. The

dilutions were based on Promega's quantitations of B15, H9, 9947A, and the quantitation of MBILC performed in this study.

Amplification Set-up: Promega Corp. provided the genotyping kits (GenePrint® PowerPlex™ 16 System) utilized in this project. For pre-PCR, the kit includes a buffer (Gold ST*R 10X Buffer), fluorescently labeled primers (PowerPlex™ 16 10X Primer Pair Mix), and a standard DNA template (9947A). A master mix was set-up containing nuclease-free water, buffer, primers, and a thermostable DNA polymerase, AmpliTaq Gold® (Applied Biosystems). 10ul of sample were added to 15ul of master mix. Negative and positive controls were set up with each amplification run.

Reactions of 25ul were set up in MicroAmp®(Applied Biosystems) reaction tubes and amplified using the Applied Biosystems GeneAmp® PCR System 2400 Thermal Cycler. The cycling protocol (from the Geneprint®PowerPlex™16 Technical Manual [9]) was as follows:

95° C for 11 minutes
96° for 1 minute
Ramp 100% to 94° C for 30 seconds
Ramp 100% to 60° C for 30 seconds
Ramp 23% to 70° C for 45 seconds
Repeat this for ten cycles, then:

Ramp 100% to 90° C for 30 seconds
Ramp 100% to 60° C for 30 seconds
Ramp 23% to 70° C for 45 seconds
Repeat this for 22 cycles, then:

60°C for 30 minutes
4°C soak

Upon completion of amplification, the samples were placed in a –20° C freezer in the post-amplification room.

Capillary Electrophoresis: Instrument Preparation: Before every run on the ABI PRISM® 310 Genetic Analyzer (Applied Biosytems), the instrument was thoroughly cleaned using distilled water and dried using lint free wipes and compressed air. This was to prevent the occurrence of fluorescent spikes (which may be due to crystallized polymer) and to keep the polymer and buffer fresh throughout the run. In addition, a matrix was generated on the Genetic Analyzer in order to analyze samples run on that specific instrument. Promega provided the GenePrint® Matrix FL-JOE-TMR-CXR for matrix standardization.

Sample Preparation: Once removed from the –20° C freezer, the samples were thawed, vortexed, and centrifuged. A loading cocktail was set up containing 24ul/sample of deionized formamide (Ultra Pure Grade, Amresco) and 1ul/sample of Internal Lane Standard 600 (ILS 600). 1ul of DNA sample was then combined with 25ul cocktail. In addition, PowerPlex™ 16 Allelic Ladders were set up (two per run). The samples were then denatured (heated for 3 minutes at 95° C and immediately cooled in an ice bath for approximately 3 minutes). The samples were then loaded onto the 310 Genetic Analyzer. The following parameters, found in the ABI 310 Collection Software, were set with accordance to the parameters specified in the PowerPlex™16 Technical Manual.

Injection time:	3 seconds
Injection kV:	15.0
Run kV:	15.0
Run °C:	60° C
Run Time:	30 minutes

In addition, the "GS STR POP4 (1ml)A" module was employed. Additional analysis parameters were specified by the PowerPlex™16 Technical Manual:

Analysis Range	Start: 3200 Stop: 10000
Data Processing	Baseline: Checked MultiComponent: Checked Smooth Options: Light
Peak Detection	Peak Amplitude Thresholds: Blue: 50-150 Yellow: 50 Green: 50 Red: 50 Min. Peak Half Width:
Size Call Range	Min: 60 Max:600
Size Calling Method	Local Southern Method
Split Peak Correction	None

To assign a new size standard for the ILS 600, the peaks were labeled according to Figure 1 Panel D of Section VIII.D in Promega's Technical Manual for the GenePrint® PowerPlex™ 16 System. Both the PowerPlex™ 16 matrix and size standard were saved on the hard drive of the MacIntosh computer for use throughout the project.

Data Analysis: Samples were analyzed using GeneScan® /Genotyper® software (Applied Biosystems) using a Macintosh G3 233 MHz (for 310 Collection Software) and a Macintosh G3 350 MHz (for analysis software). These software programs were used in conjunction with the PowerTyper™ 16 macro (Promega Corp.). Allele tables were generated containing the allele call, relative fluorescent units, and the fragment size (base pairs) and exported out to files in Microsoft® Excel (contained in a DELL OptiPlex GX1 400MHz P2 system).

Results and Discussion

I. Determination of an optimal target quantity of DNA

The first goal of this project was to determine the optimal target quantity of DNA per reaction, based on a sensitivity assay using the samples 9947A, MBILC, B19, and H9. Six target quantities of DNA (0.0625ng, 0.125ng, 0.25ng, 0.5ng, 1.0ng, and 2.0ng) were amplified using the PowerPlex™16 System. To determine the optimal target quantity of DNA per reaction, the peak heights (R.F.U.) obtained at each target DNA amount were compared. A peak height threshold of 150 R.F.U. was utilized; this is the minimum threshold that the Michigan State Police (MSP) utilize for reporting alleles. Anything below 150 R.F.U. would not be reportable (this threshold is represented by a horizontal line in Figure 1). The Technical Manual for the PowerPlex™16 System recommends optimal peak heights less than 2000 R.F.U. Table 1 (Average Peak Heights (R.F.U.) for 9947A, MBILC, B15, and H9 at various DNA template amounts) reveals that for the target DNA quantity of 0.0625ng, the average peak height was 165 R.F.U.—only 15 R.F.U. greater than the threshold of 150 R.F.U. Similarly, the average peak height at 0.125ng is still only 345 R.F.U. The average peak height for 0.25ng of target DNA was 598 R.F.U., which was a little more reasonable and well above the threshold. At 0.5ng target DNA, the average peak height was 1208 R.F.U., considerably higher than the value at 0.25ng, but still under the optimal limit of 2000 R.F.U. recommended by Promega. It is important to note, though, that the homozygote alleles were halved and counted twice, so that the average R.F.U. for homozygotes at 0.5ng might be closer to 2400 R.F.U., which is still only 400 R.F.U. greater than Promega's suggested optimal R.F.U. The average peak heights obtained for 1.0ng (2003 R.F.U.) and 2.0ng (2810

R.F.U.) were above the recommended limit of 2000 R.F.U., and those would only be the heterozygote heights—the homozygote heights would be approximately twice the R.F.U. For this reason, the target DNA amounts for 1.0ng and 2.0ng are not shown in Figure 1: R.F.U. vs Target DNA. By analyzing peak height data alone, the DNA template amounts of 0.25ng and 0.5ng provided the best results. In addition to observing peak heights, heterozygote peak height ratios (PHR) were calculated to evaluate the amplification of heterozygote alleles within a locus. The PHRs were compared to a threshold PHR of 0.7 (a value MSP utilizes in the comparison of heterozygote alleles). Table 2 (Peak Height Ratios of Heterozygotes at Different Template Amounts) shows the average PHR obtained for 0.0625ng was only 0.599504—clearly there was not enough DNA template available to effectively amplify each of the alleles. Average PHRs of 0.76975 and 0.774011 were obtained for template amounts of 0.125ng and 0.25ng. The PHRs were above 0.70, indicating the template amounts were sufficient in amplifying heterozygote alleles, but the closer the ratio is to one, the more efficient the amplification. At 0.5ng template DNA, the PHR calculated was 0.847903—considerably greater than the average PHRs obtained for 0.125ng and 0.25ng. For the 1.0ng and 2.0ng template amount, PHRs of approximately 0.88 were obtained, which indicates the most efficient amplification of heterozygote alleles were at these template amounts. When determining the optimal quantity of DNA per reaction, both the average peak heights and the peak height ratios were considered. Based on having an average peak height of 1208 R.F.U. (~2400 for homozygotes) and a peak height ratio of 0.847903, 0.5ng template DNA was chosen to be the optimal quantity of DNA per reaction.

Table 1: Average Peak Heights (R.F.U.) for 9947A, MBILC, B15, and H9 at various DNA template amounts.

Locus		0.0625ng	0.125ng	0.250ng	0.50ng	1.0ng	2.0ng
D3S1358		136	182	385	838	1562	2691
THO1		115	182	337	771	1900	2777
D21S11		195	355	519	1150	2143	2959
D18S51		148	377	613	1269	1691	2228
Penta E		177	353	588	1120	1373	1605
D5S818		106	238	427	881	1636	2635
D13S317		105	274	459	944	1909	2984
D7S820		168	359	522	1075	1845	2563
D16S539		178	405	723	1408	2214	3151
CSF1PO		232	509	733	1380	2199	3053
Penta D		203	442	863	1391	1907	2125
vWA		122	272	499	1033	1887	3213
D8S1179		134	343	627	1633	2881	3952
TPOX		300	432	840	1546	2510	3181
FGA		163	454	828	1683	2393	3026
Ave Peak Height Per DNA Template Amount		165	345	598	1208	2003	2810
Standard Deviation		53	98	169	292	394	551

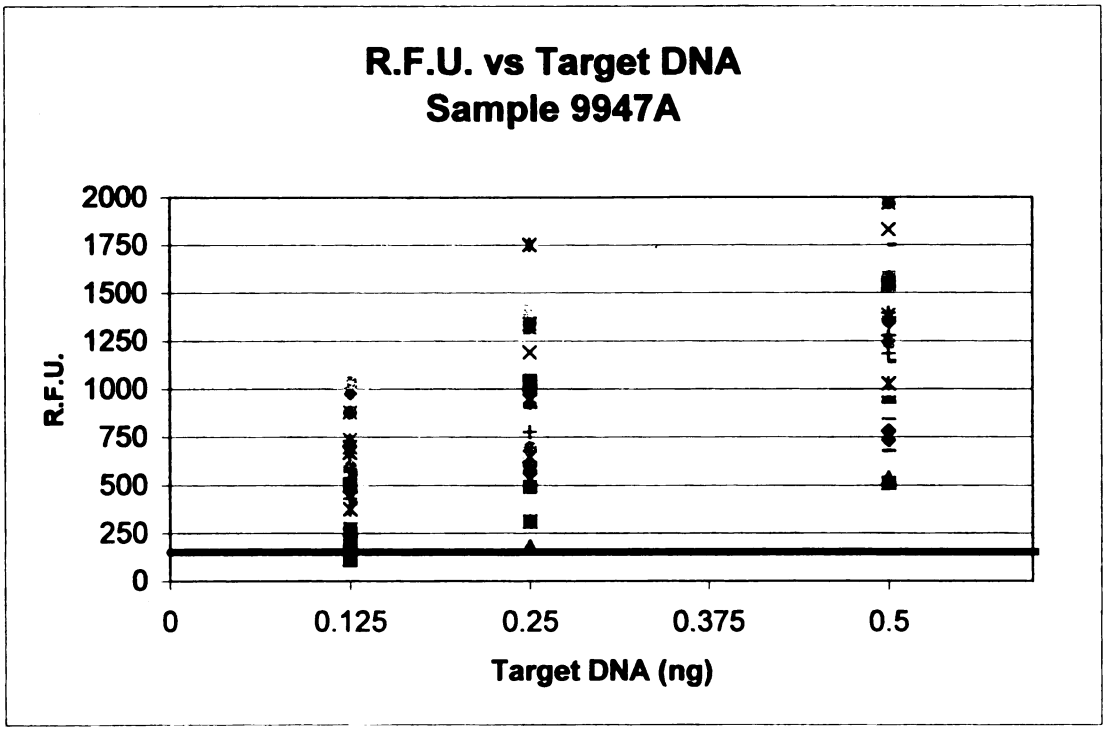
Note: Homozygote peak heights were halved and counted twice.
Peak heights have been rounded to whole numbers.

Table 2: Peak Height Ratios of Heterozygotes at Different Template Amounts (9947, MBLIC,B15, & H9)

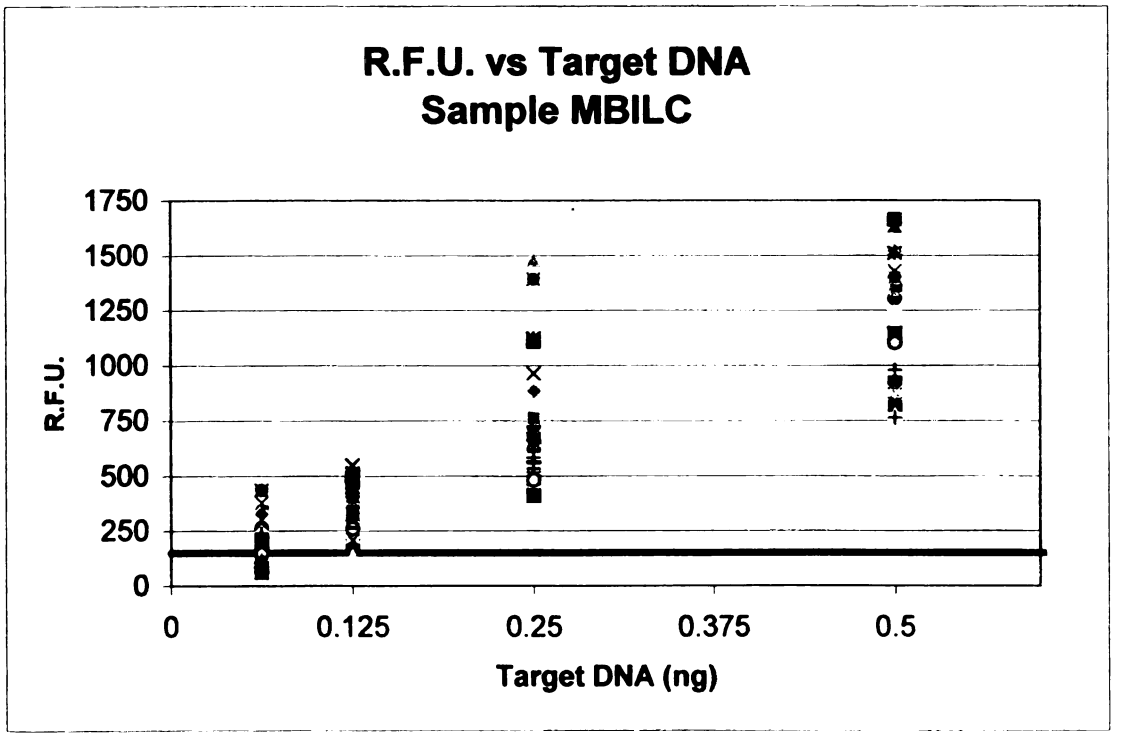
Locus	0.0625ng	0.125ng	0.25ng	0.5ng	1.0ng	2.0ng
D3S1358	0.61869	0.727	0.851667	0.843133	0.906	0.8222
THO1	0.670033	0.757475	0.6432	0.957225	0.804025	0.945675
D21S11	0.536567	0.839133	0.8977	0.815533	0.911633	0.8427
D18S51	0.505467	0.896525	0.873225	0.89955	0.89535	0.827275
Penta E	0.645867	0.638525	0.744675	0.8658	0.83725	0.7833
D5S818	0.5327	0.81115	0.6761	0.652	0.9226	0.96835
D13S317	0.834233	0.768933	0.892567	0.8533	0.872867	0.958233
D7S820	0.646033	0.861475	0.87225	0.75435	0.809525	0.83665
D16S539	0.688133	0.765075	0.746975	0.794775	0.917775	0.85485
CSF1PO	0.32945	0.82165	0.652825	0.923075	0.82575	0.8988
Penta D	0.631533	0.554367	0.664333	0.830567	0.903367	0.8856
vWA	0.77765	0.946167	0.788567	0.951467	0.951033	0.916767
D8S1179	0.5055	0.77	0.90945	0.8802	0.9414	0.9277
TPOX	0.567	0.6181	0.72685	0.87195	0.91425	0.9333
FGA	0.5037	0.770675	0.669775	0.825625	0.905725	0.873225
Average Peak Height Ratio	0.599504	0.76975	0.774011	0.847903	0.887903	0.884975
Standard Deviation	0.12382	0.10469	0.100545	0.077821	0.047036	0.056151

Note: Homozygote alleles were omitted.

Figure 1: R.F.U. vs Target DNA, (a) Sample 9947A; (b) Sample MBILC; (c) Sample B15; and (d) Sample H9



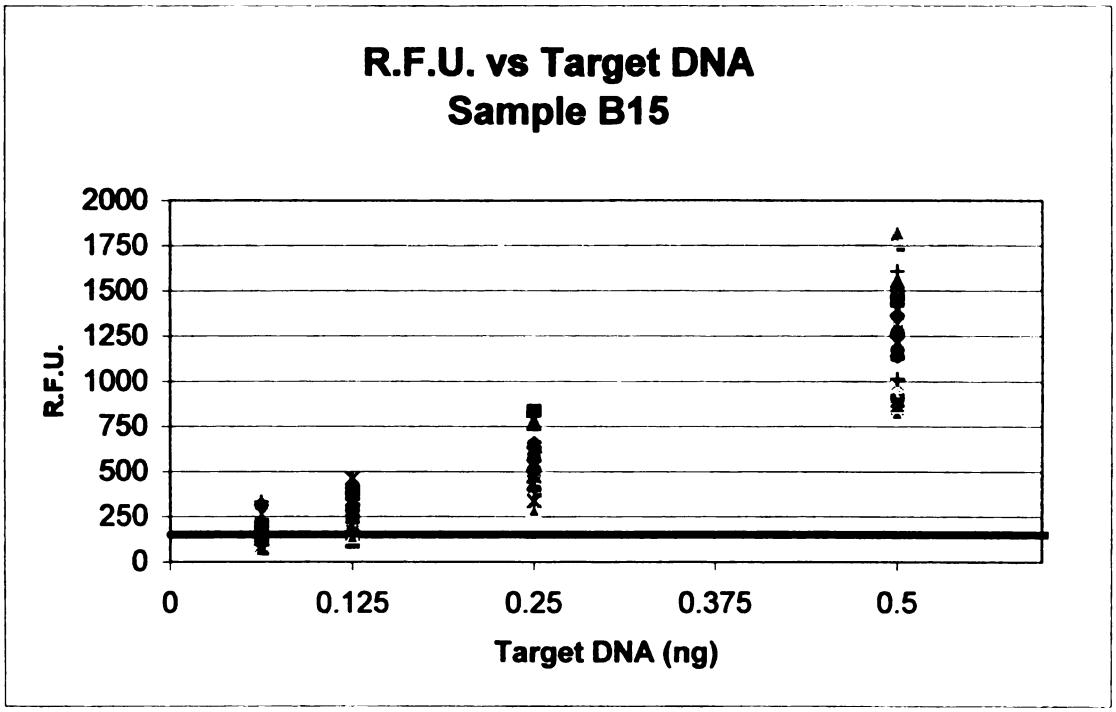
(a)



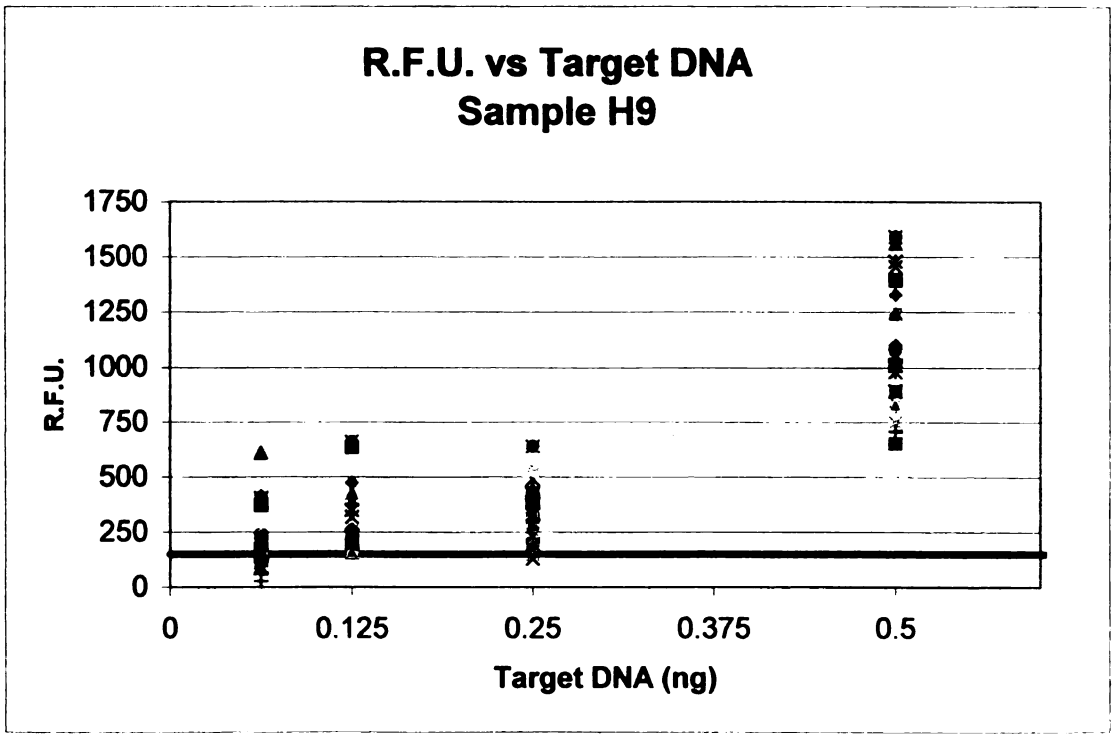
(b)

Note: Homozygote alleles were halved and counted twice.

Figure 1: (cont.)



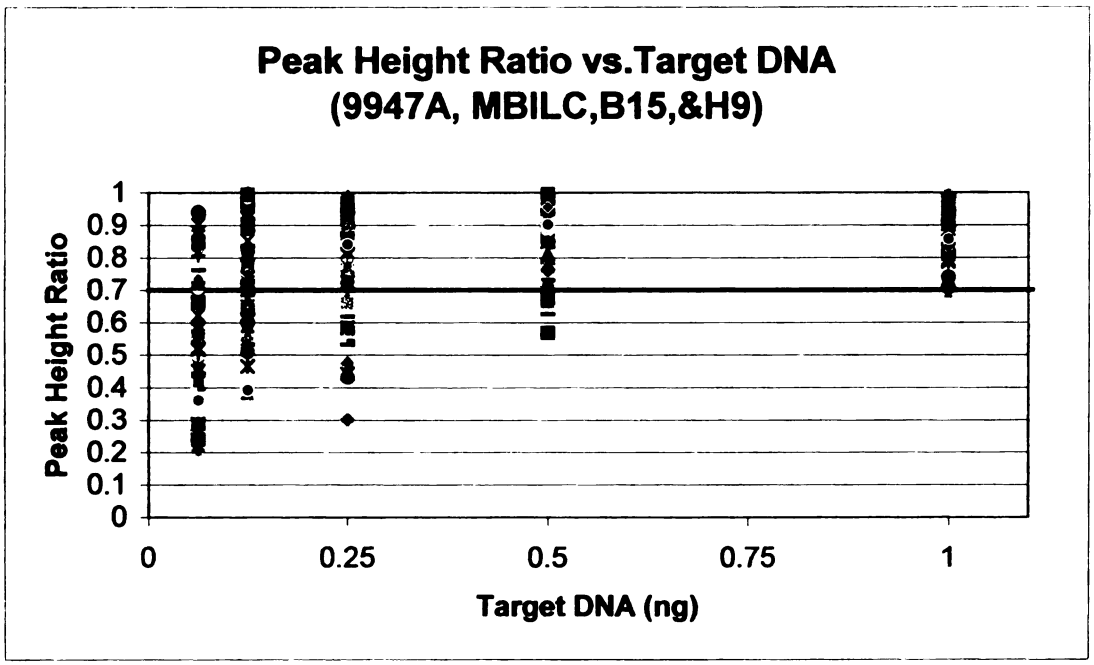
(c)



(d)

Note: Homozygote alleles were halved and counted twice.

Figure 2: Peak Height Ratio vs Target DNA (9947A, MBILC, B15, & H9)



Note: Peak height ratios were calculated with the exclusion of homozygote alleles.

II. Observation of performance at each locus

After the optimal target quantity of DNA was determined to be 0.5ng, a set of one hundred and fourteen samples were run (with 0.5ng template DNA) with the PowerPlex™16 Amplification System. The results are shown in Figure 3: Heterozygote Peak Height Ratio vs Locus, Table 3: Heterozygote Peak Height Ratios, Table 4: Average Peak Heights (R.F.U.) Per Locus, and Figure 4: R.F.U. vs Fragment Size (Per Locus). The second goal of this project was to evaluate the level of performance at each of the PowerPlex™16 loci (excluding Amelogenin). To observe locus performance, one can determine heterozygote peak height ratios (PHR) for each locus and compare them to an optimum value (in this project, heterozygote peak height ratios were compared to a value of 0.70). The MSP DNA Unit utilized this value in validation studies of previous genotyping kits. Evaluating PHRs helps identify whether a DNA genotyping kit has the tendency to preferentially amplify alleles (this occurs when there is a difference in amplification of two alleles within the same locus). Preferential amplification is evident when two alleles within a locus differ greatly by peak heights—hence, by PHR. An ideal PHR is equal to one (or 100%), meaning the peaks are exactly the same height. The smaller the values are from 100%, the greater the difference in heterozygote peaks. MSP has determined in previous studies that an acceptable PHR for heterozygotes is 0.70 (70%) or greater. It is useful to evaluate PHRs with a new genotyping kit, especially if the kit will be used in analyzing casework samples (which are often degraded and contain mixtures). Since there should not be any mixtures or degradation with single-source database samples, it may be a little easier to evaluate the viability of this genotyping kit. From Table 3: Heterozygote Peak Height Ratios, it is evident that the average PHRs for

the fifteen loci range from 0.800437 (Penta E) to 0.876938 (TPOX). This range of average PHRs is well above 0.70, indicating that there is minimal preferential amplification occurring at each locus. In addition, the percentages of PHRs that fell under 0.70 were also calculated. Percentages ranged from 2.2% (THO1) to 23% at Penta E. These percentages reveal that though the average PHRs are above 0.70, there is some preferential amplification occurring. This also shows that a threshold of 0.70 for heterozygote PHRs may be too high for this particular genotyping kit, and that a value slightly lower than 0.70 may be more suitable.

Another method to evaluate locus performance is to compare peak heights (R.F.U.) and fragment sizes of the alleles at each locus. Graphs showing the relationship of peak height and fragment size were generated (Figure 4: R.F.U. vs Fragment Size (Per Locus)). From these graphs, one can determine whether differential amplification is evident (the occurrence of larger sized alleles amplifying less than the smaller sized alleles). Often, in multiplex PCR kits, the larger sized alleles (or loci) get amplified less than the smaller fragments—which results in lower peak heights for the larger alleles. In this project, minimal differential amplification was observed at each locus and between loci, as can be observed in Figure 4—the graphs which compare peak height and fragment size, and Table 4: Average Peak Heights Per Locus. Some of the graphs did produce a slight downward trend, indicating a reduction of peak height with an increase in size (D18S51, Penta E, and FGA). To better evaluate the relationship of peak height and fragment size, more samples should be run so that an equal representation of all alleles would be present.

Lastly, Table 4: Average Peak Heights (R.F.U.) Per Locus compares the average peak heights at each locus. Peak heights ranged from 962 R.F.U. (Penta E) to a height of 1819 R.F.U. at D8S1179. As noted previously, Promega suggests an optimal peak height under 2000 R.F.U. The average peak heights obtained for this project fell within that optimal range, though it is important to note that the homozygote alleles were halved and counted twice. Therefore the range for homozygotes may be closer to 1800-3600 R.F.U. (a little higher than recommended by Promega).

Figure 3: Heterozygote Peak Height Ratios vs Locus

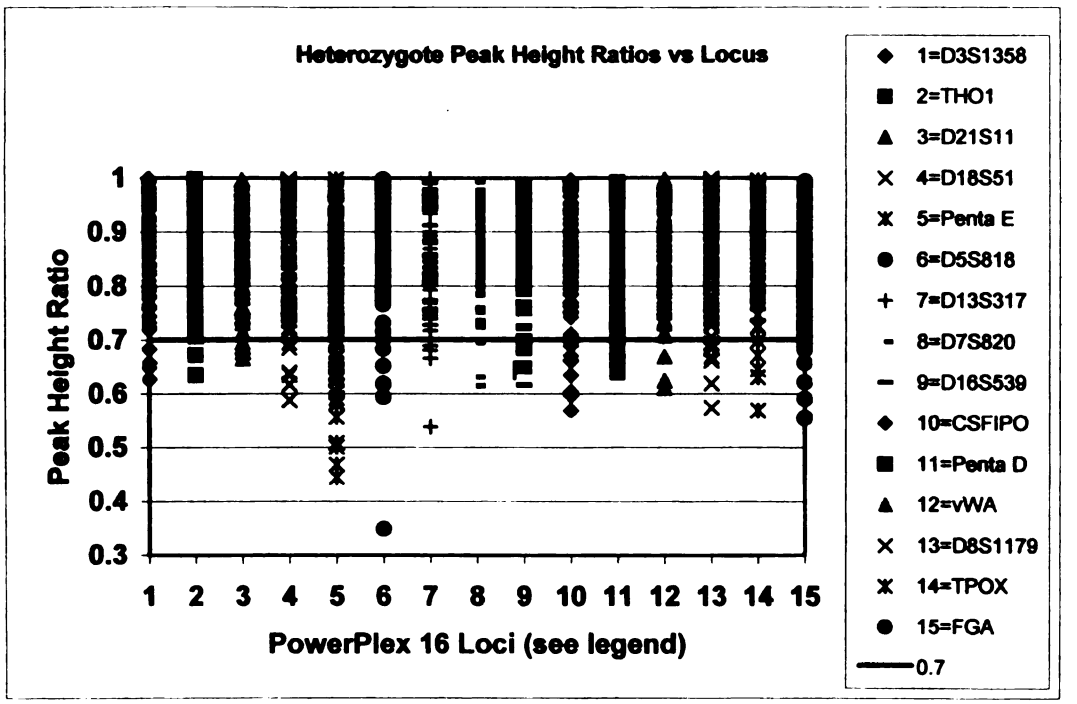


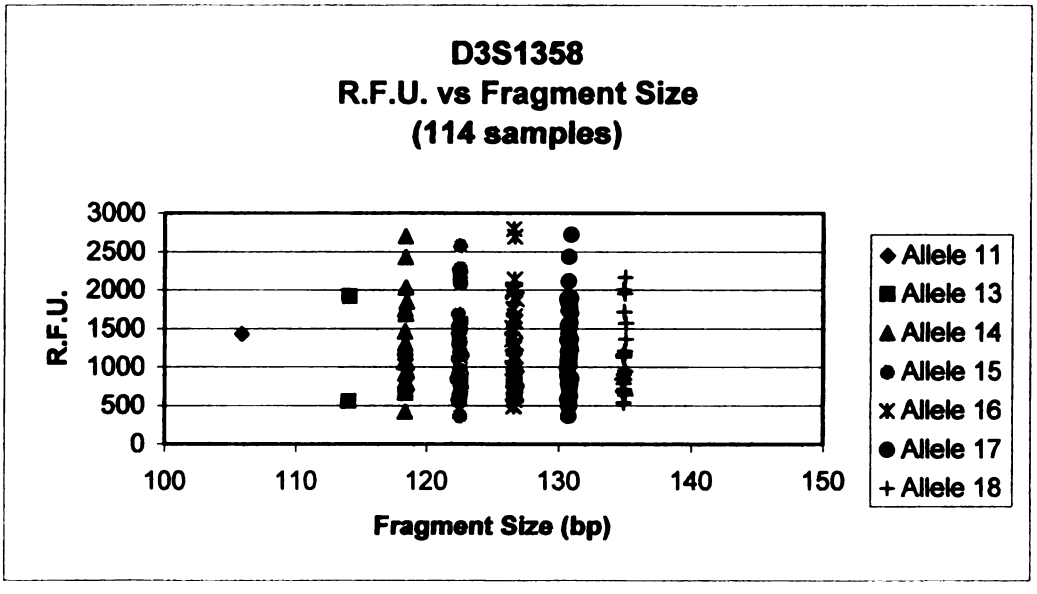
Table 3: Heterozygote Peak Height Ratios

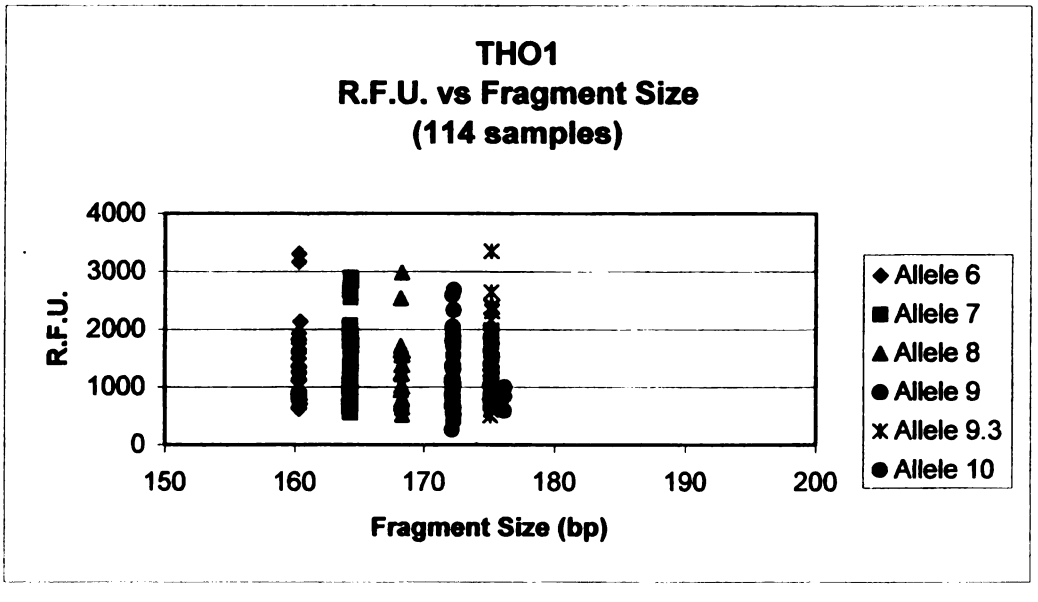
Locus	D3S1358	THO1	D21S11	D18S51	Penta E	D5S818	D13S317
Samples	88	91	90	97	100	90	88
Average	0.871856	0.875807	0.868036	0.848737	0.800437	0.857274	0.864429
Median	0.891664	0.882656	0.877778	0.863177	0.821655	0.880224	0.863921
Std Dev	0.093666	0.081407	0.079097	0.110684	0.133053	0.10576	0.092161
< 70%	5.68%	2.20%	4.44%	10.31%	23%	6.67%	4.55%
D7S820	D16S539	CSF1PO	Penta D	vWA	D8S1179	TPOX	FGA
97	93	82	100	96	88	73	95
0.87198	0.853188	0.851628	0.859954	0.872133	0.866916	0.876938	0.845391
0.882129	0.865478	0.862699	0.8744	0.878005	0.882867	0.891164	0.859031
0.08237	0.097068	0.108027	0.090788	0.087882	0.097309	0.097608	0.099755
4.12%	10.75%	10.98%	9%	3.13%	7.95%	6.85%	8.42%

Note: Values have been calculated with the exclusion of homozygotes.

Average: The average value of peak height ratios for that locus.
Median: The number in the set of values which has one half of the values higher and one half lower.
Std. Dev: The standard deviation measures how widely values are dispersed from the average value.
<70%: The percentage of peak height ratios that fell under 0.70.

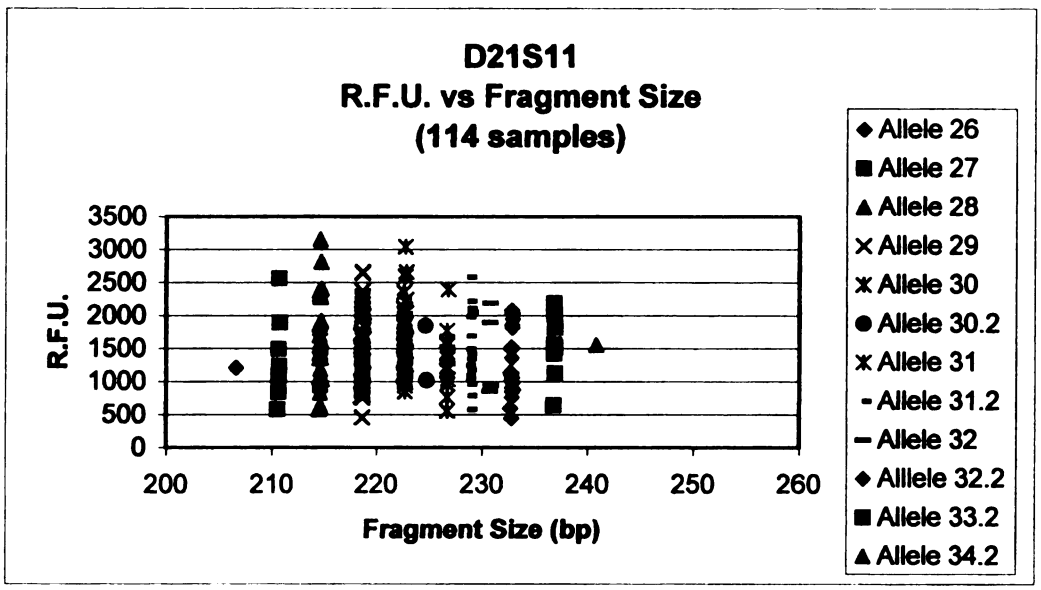
Figure 4: R.F.U. vs Fragment Size (Per Locus)

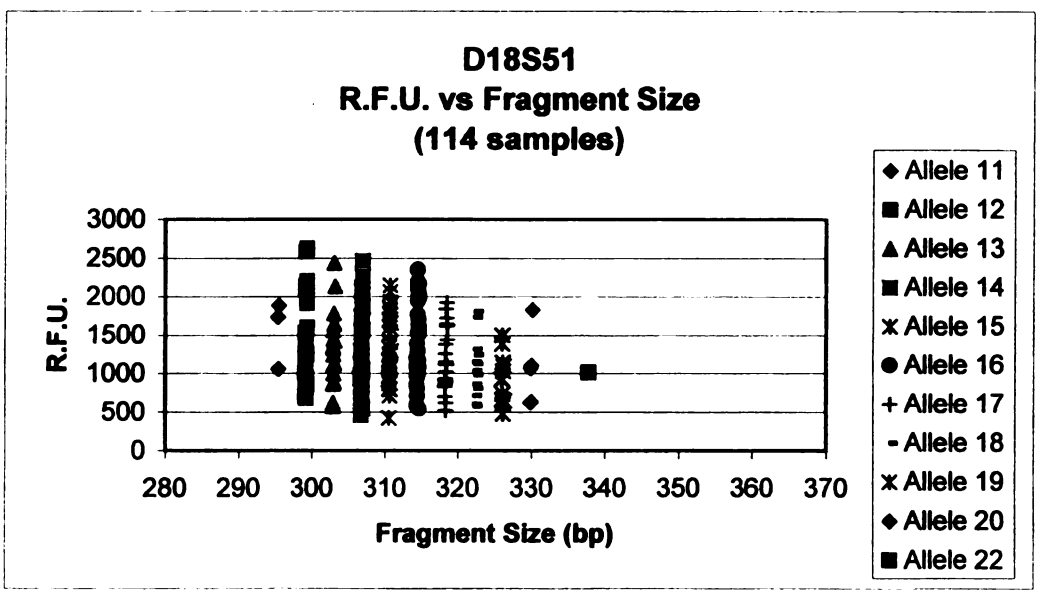




Note: Homozygote alleles were halved and counted twice.

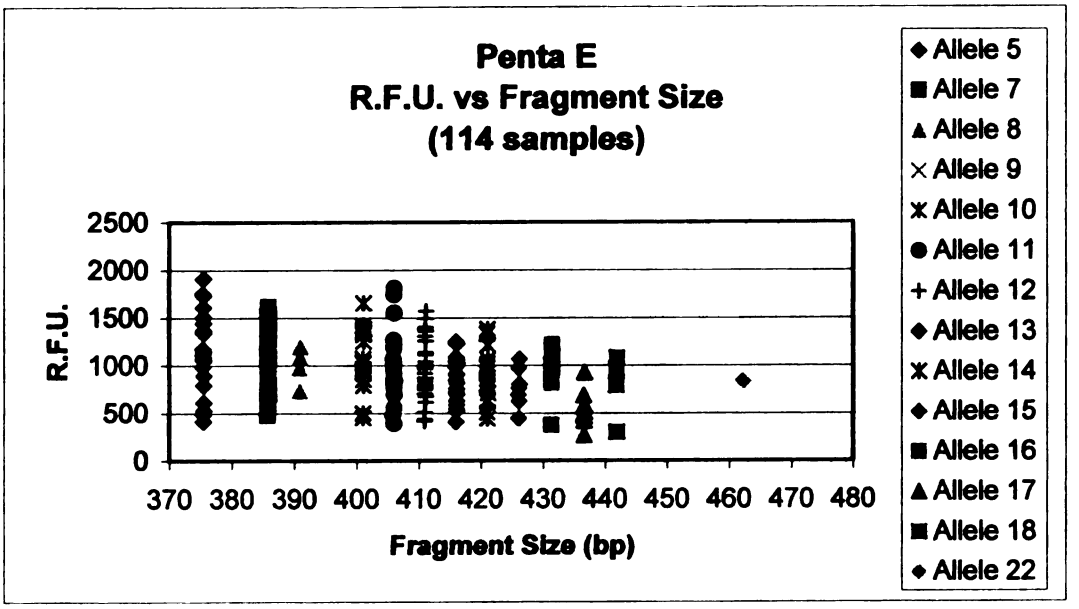
Figure 4: R.F.U. vs Fragment Size (Per Locus)

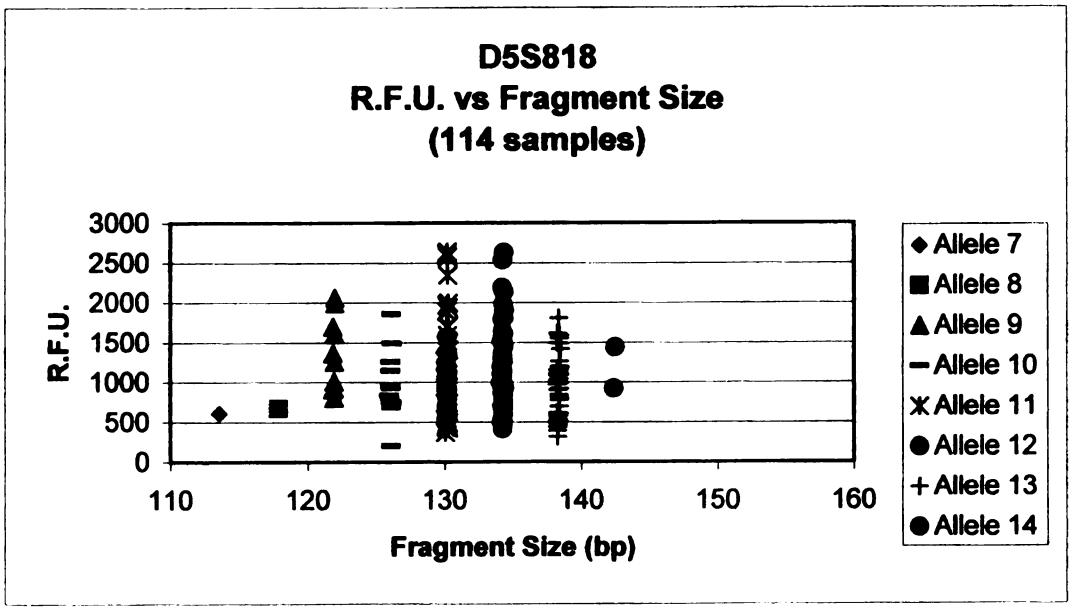




Note: Homozygote alleles were halved and counted twice.

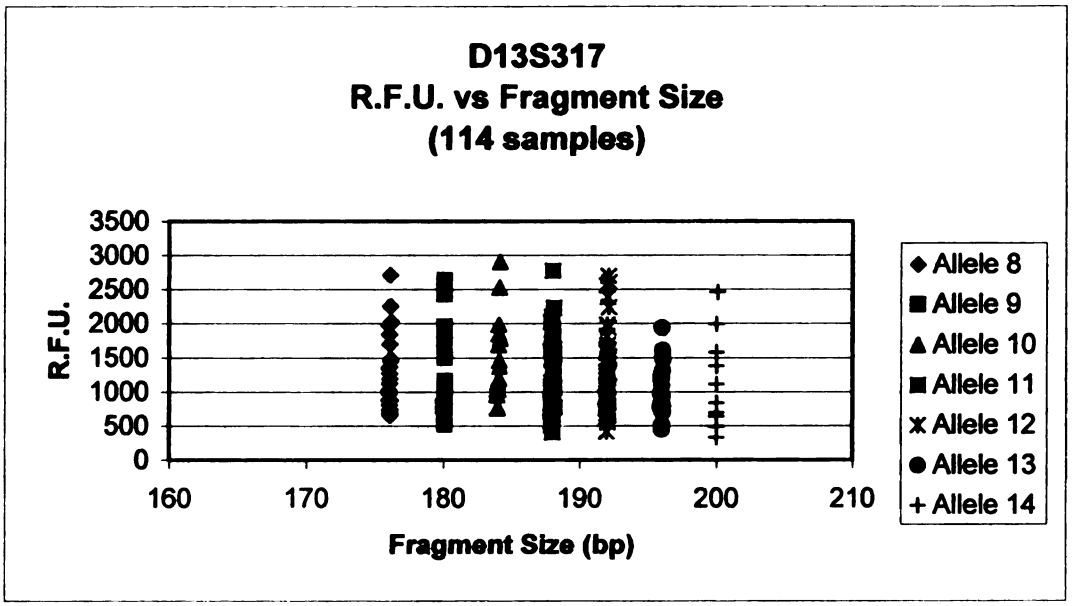
Figure 4: R.F.U. vs Fragment Size (Per Locus)

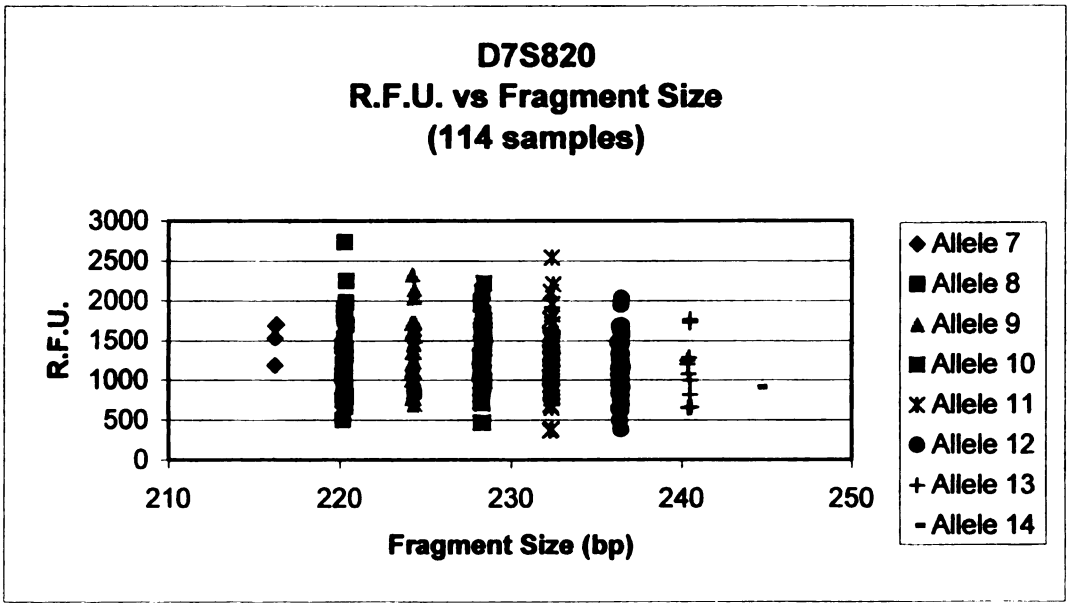




Note: Homozygote alleles were halved and counted twice.

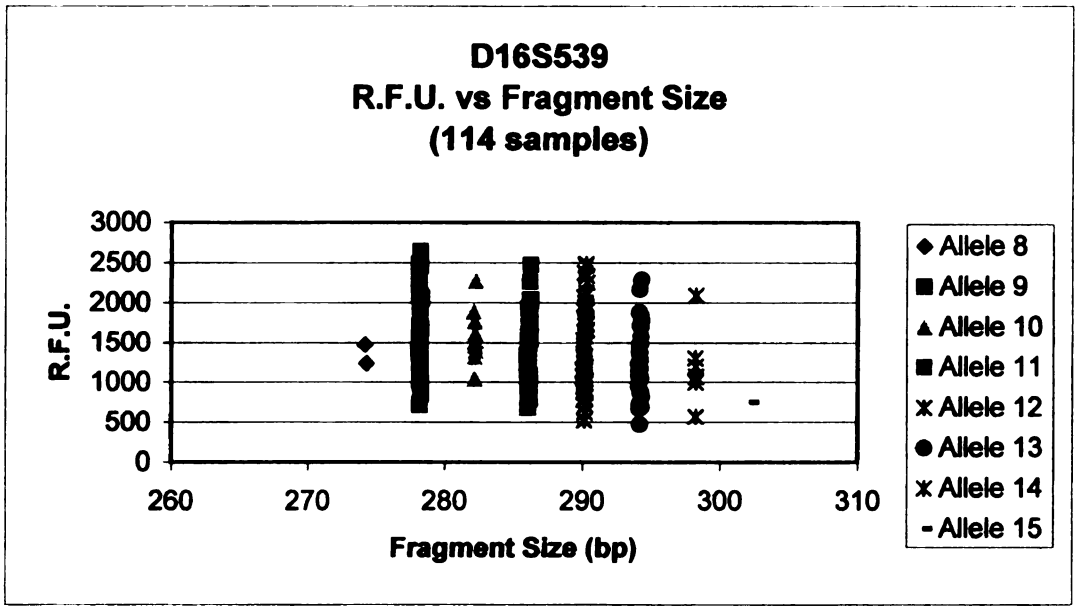
Figure 4: R.F.U. vs Fragment Size (Per Locus)

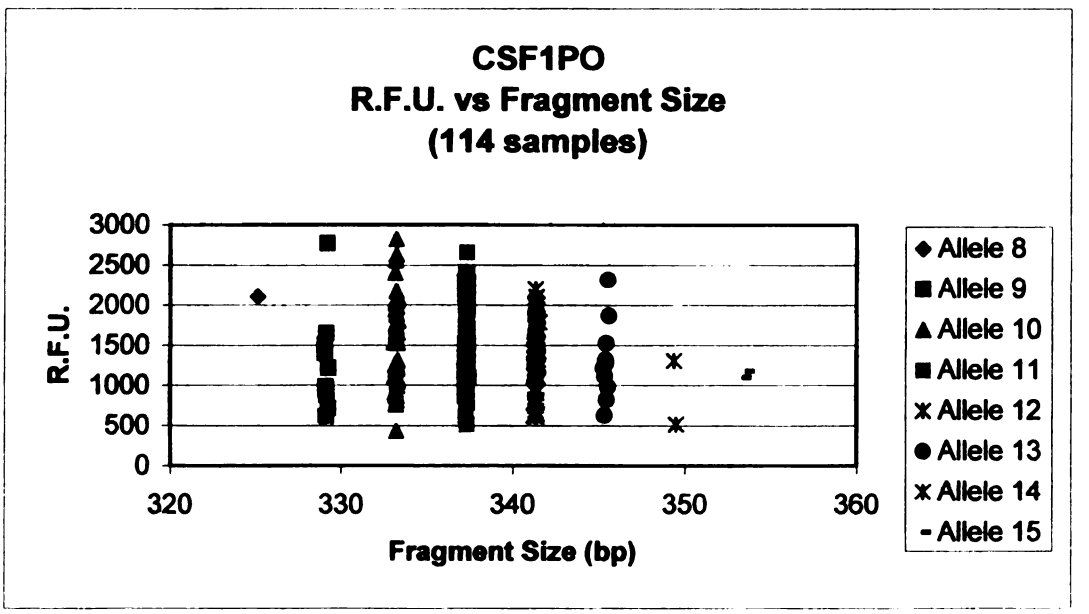




Note: Homozygote alleles were halved and counted twice.

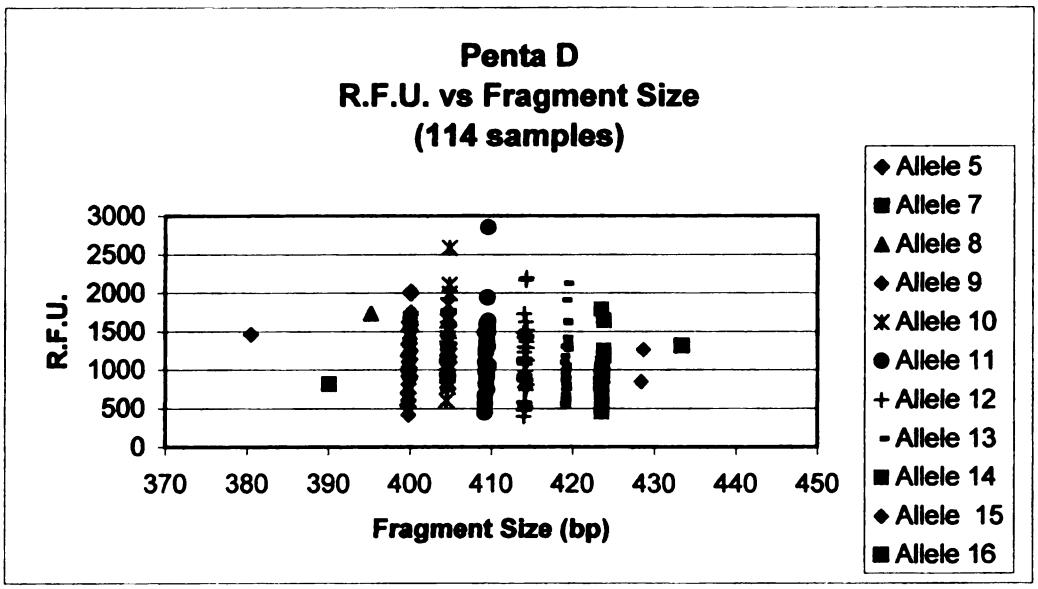
Figure 4: R.F.U. vs Fragment Size (Per Locus)

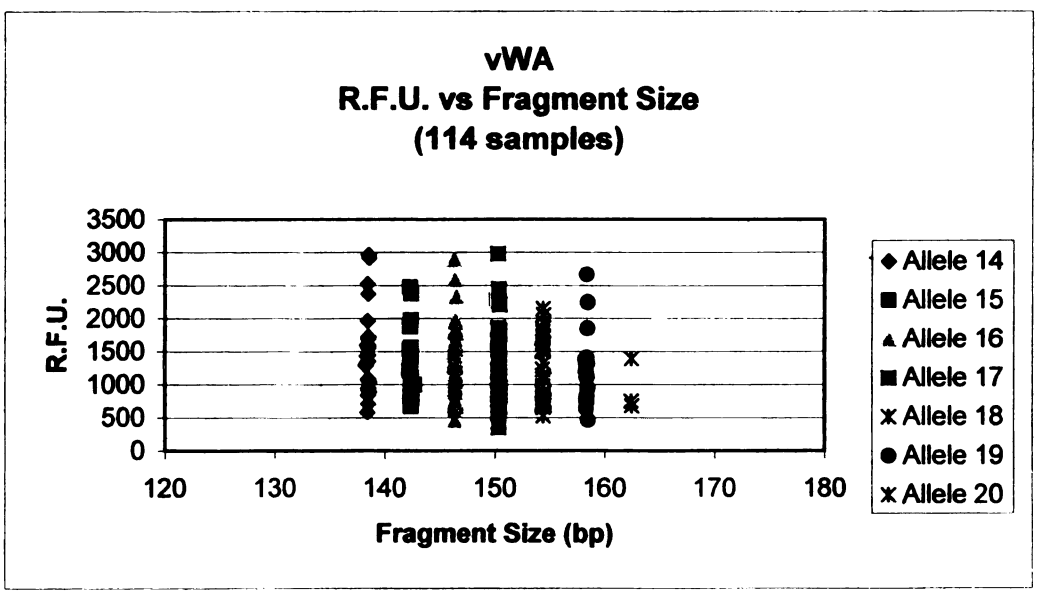




Note: Homozygote alleles were halved and counted twice.

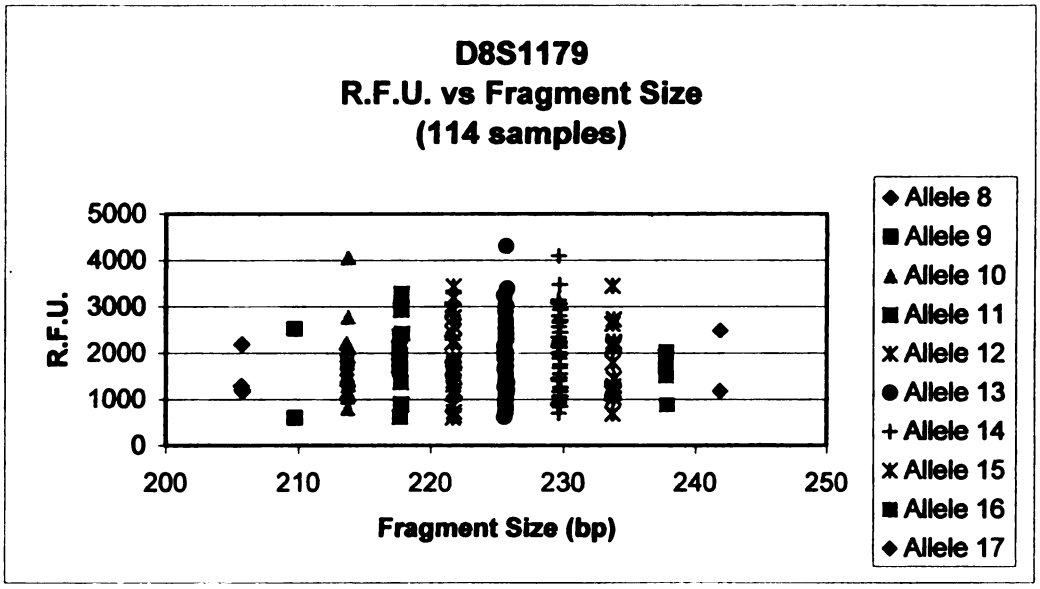
Figure 4: R.F.U. vs Fragment Size (Per Locus)

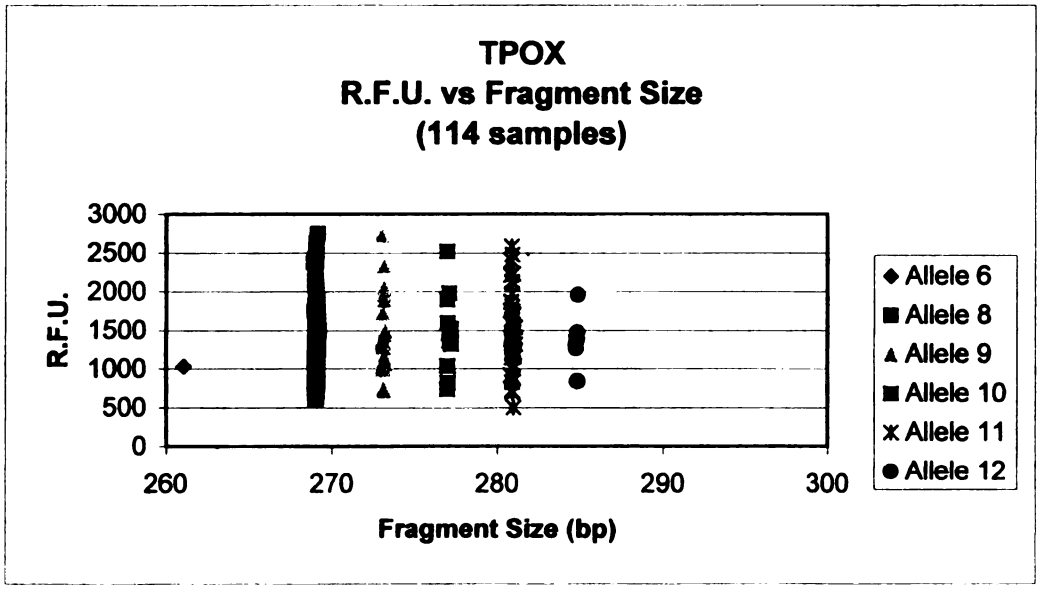




Note: Homozygote alleles were halved and counted twice.

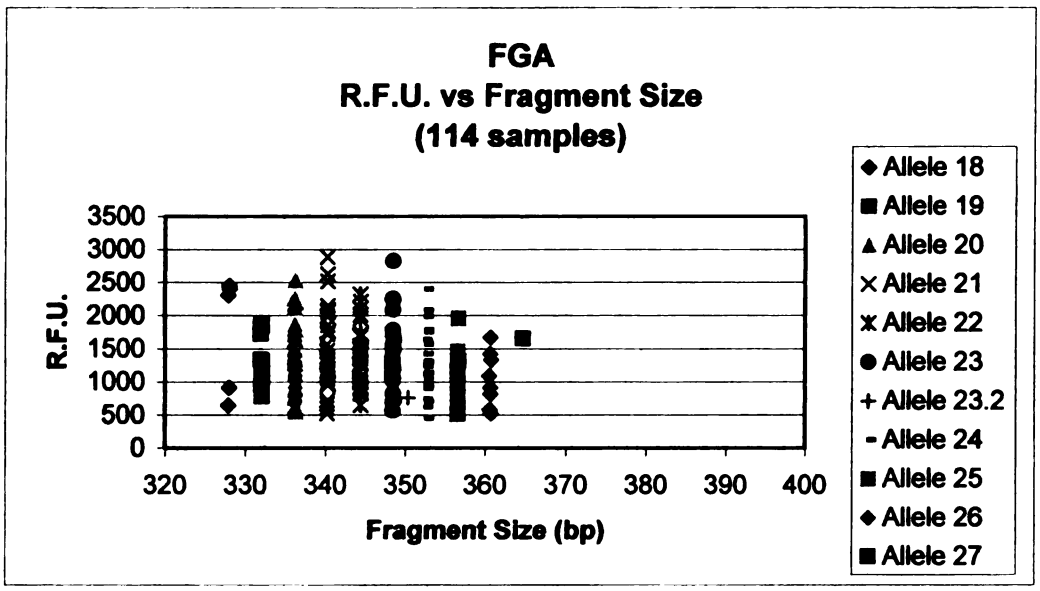
Figure 4: R.F.U. vs Fragment Size (Per Locus)





Note: Homozygote alleles were halved and counted twice.

Figure 4: R.F.U. vs Fragment Size (Per Locus)



Note: Homozygote alleles were halved and counted twice.

Table 4: Average Peak Heights (R.F.U.) Per Locus

Locus		Average Peak Height	Median	Standard Deviation
D3S1358		1164	1070	509
THO1		1291	1130	602
D21S11		1481	1398	535
D18S51		1277	1151	471
Penta E		962	962	296
D5S818		1108	1011	479
D13S317		1299	1202	535
D7S820		1243	1163	425
D16S539		1409	1308	464
CSF1PO		1451	1362	473
Penta D		1151	1117	403
vWA		1255	1198	530
D8S1179		1819	1746	739
TPOX		1488	1411	478
FGA		1340	1300	465

Note: Homozygote alleles were halved and counted twice.

Conclusion and Future Research

The third goal of this project was to evaluate the viability of the PowerPlex™ 16 Amplification Kit for database samples. As stated previously, database samples (as those used for CODIS) should all be single source samples. No degraded samples or mixtures would be included in a database such as CODIS, therefore lessening the chance for preferential amplification. In addition, by observation of the heterozygote PHRs (approximately 0.8 to 0.86), preferential amplification was minimal. Peak heights were also compared with fragment size of alleles at each locus; this would reveal any differential amplification trends (within loci). Differential amplification would be evident by decreasing peak height with increasing size at each locus, or between loci. There was minimal differential amplification within loci (as can be noted in Figure 4), meaning the PowerPlex™16 Kit amplified all of the alleles at each locus equally. Also, the average peak heights per loci did not display any signs of decreasing peak height with increasing locus size, meaning the kit had amplified each of the loci equally. In addition, the average heterozygote peak heights obtained per locus were under the recommended peak height limit recommended by Promega, with the height of the homozygotes a little higher than the suggested optimal height.

It should be noted that overall, the kit performed equally within and between loci, with the exception occurring at the locus Penta E. The Penta E locus was slightly problematic, being such a large locus and containing the largest sized alleles in the PowerPlex™16 allelic ladder. From the results, preferential amplification is observed at the Penta E locus, with 23% of the peak height ratios falling under the 70% threshold. This may be due to the overall size and base pair lengths of this locus. The occurrence of

preferential amplification indicates that the heterozygote alleles are not amplifying equally, which again could be due to the size of this locus, meaning the larger sized alleles would be amplifying less than the smaller alleles. In addition, a downward trend was noted in the graph of Penta E (Figure 4), indicating slight differential amplification (a reduction of peak height with an increase in fragment size). Though some anomalies were observed in the Penta E locus, it should be stressed that it is not one of the thirteen required CODIS loci, that it was added simply to increase specificity of the DNA typing system.

Future research for the PowerPlex™16 System include developing an appropriate threshold for acceptable peak height ratios. The value of 0.70 was utilized in this project, but that value was generated from validation studies of a genotyping kit of only ten loci, whereas the PowerPlex™16 kit has sixteen. Research using single source samples should be done to observe amplification trends per locus and to develop a suitable cutoff for determining sister alleles. Additional research should also be done to observe any trends in differential amplification within and between loci. This should be done using a large number of single source samples so that all of the PowerPlex™16 alleles are represented.

Overall, the PowerPlex™16 System is a robust method for obtaining the DNA profiles for single source samples to be entered into the CODIS database. Though this project did not include a concordance study to verify the DNA profiles of the one hundred and fourteen samples, it should be noted that the known samples used in this evaluation (9947A, MBILC, B19, and H9) all typed as expected, and that the allelic ladders utilized in this project produced the expected results as well (as compared to the

allelic ladder in the PowerPlex™16 Technical Manual). As stated previously, this evaluation of locus performance was part of a larger project by Promega and a number of other forensic DNA testing facilities. The data generated in this evaluation was included in a concordance study with the Applied Biosystem's Profiler Plus™ and CoFiler™ typing systems, which evaluated the concordance of the primers used in each kit. Results in the "STR primer concordance study" indicated that the primers used in the PowerPlex™16 System and the Profiler Plus™/CoFiler™ systems produced reliable results on reference samples.

The PowerPlex™16 system is a powerful method of obtaining the DNA profile of an individual at sixteen different loci in one reaction. Not only does this system save the analyst time, money, and resources, but also the agency for which he/she works. This typing system would allow for faster processing of convicted offender samples, whose DNA profiles could be entered into the CODIS database quicker, thus providing investigators with another DNA profile of which to search.

WORKS CITED

WORKS CITED

1. Lee, et al. "DNA Typing in Forensic Science, I. Theory and Background." *The American Journal of Forensic Medicine and Pathology* 1994; 15(4), 2699-282.

2. Budowle, et al. "STR primer concordance study." *Forensic Science International.* 2001; in press.

3. Avery, MacLeod, and McCarty. "Studies on the chemical nature of the substance inducing transformation of *pneumococcal* types." *Journal of Experimental Medicine* 1944; 98, 451-460.

4. Wallin, et al. "TWGDAM Validation of the AmpFLSTR™ Blue PCR Amplification Kit for Forensic Casework Analysis." *Journal of Forensic Science* 1998; 43(4), 117-133.

5. Micka, et al. "Validation of Multiplex Polymorphic STR Amplification Sets Developed for Personal Identification Applications." *Journal of Forensic Sciences* 1996; 41(4), 582-590.

6. Lins, et al. "Development and Population Study of an Eight-Locus Short Tandem Repeat (STR) Multiplex System." *Journal of Forensic Science* 1998; 43(6), 1168-1180.

7. Walsh, Fildes, and Reynolds. "Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA." *Nucleic Acid Research* 1996; 24(14), 2807-2812.

8. Micka, et al. "TWGDAM Validation of a Nine-Locus and a Four-Locus Fluorescent STR Multiplex System." *Journal of Forensic Science* 1999; 44(6), 1243-1257.

9. Promega Corporation. "Technical Manual, GenePrint® PowerPlex™16 System." Copyright © 2000 Promega Corporation, Madison, Wisconsin.

ADDITIONAL WORKS CONSULTED

Balding and Donnelly. "Evaluating DNA Profile Evidence When the Suspect Is Identified Through a Database Search." *Journal of Forensic Science* 1996; 41(4), 603-607.

De Pancorbo, et al. "Population Genetics and Forensic Applications Using Multiplex PCR (CSF1PO, TPOX, and THO1) Loci in the Basque Country." *Journal of Forensic Science* 1998; 43(6), 1181-1187.

Dictionary of Scientists. Johann Friedrich Miescher. University Press © Market House Books Ltd 1999.

Dieffenbach and Dveksler. "PCR Primer, A Laboratory Manual." Copyright © 1995 by Cold Spring Harbor Laboratory Press, Plainview, New York.

Erlich. "PCR Technology, Principles and Applications for DNA Amplification." Cpyright © 1992 by Oxford University Press, Inc., New York, New York.

Gill, et al. "Automated short tandem repeat (STR) analysis in forensic casework—a strategy for the future." *Electrophoresis* 1995: 16, 1543-1552.

Gill, Jeffreys, and Werrett. "Forensic application of DNA 'fingerprints'." *Nature* 1985; 318, 577-579.

Inman and Rudin. "An Introduction to Forensic DNA Analysis." Copyright © 1992 by CRC Press, Boca Raton, Florida.

Lazaruk, et al. "Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument." *Electrophoresis* 1998: 19, 86-93.

Lewin. "Genes VI." Copyright © 1997 by Oxford University Press, Inc., New York, New York.

Lorente, et al. "Sequential Multiplex Amplification: Utility in Forensic Casework with Minimal Amounts of DNA and Partially Degraded Samples." *Journal of Forensic Science* 1997; 42(5), 923-925.

Mullis. "The unusual origin of the polymerase chain reaction." *Scientific American* 1990; 262(4), 56-65.

PE Applied Biosystems. "AmpF*l*STR Profiler Plus™ PCR Amplification Kit User's Manual." Copyright © 1998 by the Perkin-Elmer Corporation, Foster City, California.

Schmitt and Benecke. "Five cases of forensic short tandem repeat DNA typing." *Electrophoresis* 1997: 18, 690-694.

Watson and Crick. "A structure for deoxyribose nucleic acid." *Nature* 1953; 171, 737-738.