2023

Development and Validation of a Multiplex STR Amplification Reaction for Academic and Research Purposes

Ahana Chatterjee
University Of Alabama At Birmingham

DEVELOPMENT AND VALIDATION OF A MULTIPLEX STR AMPLIFICATION REACTION FOR ACADEMIC AND RESEARCH PURPOSES

by

AHANA CHATTERJEE

JASON G LINVILLE, COMMITTEE CHAIR
ELIZABETH A GARDNER
M TINO UNLAP

A THESIS

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Master of Science

BIRMINGHAM, ALABAMA

2023

DEVELOPMENT AND VALIDATION OF A MULTIPLEX STR AMPLIFICATION REACTION FOR ACADEMIC AND RESEARCH PURPOSES

AHANA CHATTERJEE

FORENSIC SCIENCE

ABSTRACT

Short Tandem Repeat (STR) typing has been accepted globally as a gold standard for human identity testing. In 1997, the Federal Bureau of Investigation (FBI) Laboratory's newly developed Combined DNA Index System (CODIS) consisted of 13 core loci to be targeted while DNA typing. In this context, leading biotechnology companies like Promega (Madison WI), Applied Biosystem/Life Technologies (Foster City, CA) etc. offer commercial STR typing kits for crime laboratories. However, these expensive kits are not the most suitable for academic and research laboratories. Students learning the method of STR typing and researchers showing proof of concept do not require identifying 13 loci. Instead, a smaller number of loci can be targeted, reducing the cost for laboratories operating on a low budget.

For the past 10 years a custom multiplex STR amplification reaction has been used in the Forensic Science program at the University of Alabama at Birmingham (UAB). This cost-effective reaction has worked well, but supplementing its use with a second reaction would increase the significance of the results and provide a tool for troubleshooting contamination issues in the lab. Additionally, the process for designing and validating the first reaction was not published, so publishing the details of this reaction in a thesis provides a resource for other labs. Hence, in this project an additional reaction targeting 3 new loci was developed and validated. The cost for each 25 µL reaction was

reduced to $0.74, much lower than the commercial kits that cost as much as $30 per reaction.

To develop the reaction, the three loci D5S818, D16S539 and D18S51 and their corresponding primer sequences were selected. Then, a monoplex reactions for each locus was completed followed by a multiplex reaction. Finally, a validation study (sensitivity study, thermal cycling parameter selection, precision study, accuracy, and concordance study) was performed. The project has resulted in a new, inexpensive, validated multiplex STR amplification reaction that is efficient, and can be used by academicians and researchers in the field of forensic science.

Keywords: STR, DNA, Custom mix

DEDICATION

I dedicate this to my Maa, Mrs. Kalyani Chatterjee and my Baba, Mr. Alok Chatterjee. Thank you for letting me follow my dreams.

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to my advisor and committee chair, Dr. Linville, who guided me at every step of this research project. None of it would have been possible without him. I am extremely grateful to my committee members, Dr. Gardner, and Dr. Unlap for making my defense an enjoyable experience and helping with important suggestions in my thesis.

I shall always be grateful to Dr. Tribedi who introduced me to research and saw my potential when no one else did. I am thankful to my parents who have always believed in me. A special thanks to Rushil for looking out for me and being a constant support. I have immense gratitude for my girls, Sejal, Shruti, Gracie and Emma who cheered for me throughout my journey. My friends, Prathamesh, Amit, Varun, and Abigail, thank you for always being ready to help.

Finally, I want to thank my family and friends from India who have been by my side through thick and thin. I hope I can keep making them proud.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

DNA in Forensics: How Did it all Start?

In 1886 in the novel *A Study in Scarlet* written by Arthur Conan Doyle, Sherlock Holmes discusses a reagent he had discovered which is only precipitated by hemoglobin and he named the test the Sherlock Holmes test.[i] A hundred years later, at the University of Leicester, Sir Alec Jeffreys visualized reproducible patterns of DNA and started working on developing a method towards understanding the uniqueness of these patterns for everyone. Unlike Holmes, instead of naming the test the Jefferys test, he called it DNA fingerprinting.[ii] This discovery brought forth an era of new discoveries in science which would be used in different areas of biology from anthropological studies to understanding diversity. However, it was not until the first application of DNA fingerprinting in civil and criminal cases was published that the true potential of fingerprinting was understood outside of academic circles.

Forensic DNA fingerprinting describes the early processes of comparing the DNA obtained from a person's biological fluid to biological evidence found at a crime scene or to another person's DNA for the purpose of identification or exclusion. Though using these techniques provided a new and interesting tool for criminal investigations, the first case of DNA fingerprinting was not used for a criminal case, instead the technique of comparing DNA was used for an immigration case. In this regard, in March of 1985, DNA fingerprinting was applied to prove familial relations to save a young boy from deportation.

This first case caught attention from public.[iii] In the University of Leicester Bulletin Supplement of August of 2004, Alec Jefferys said "'If our first case had been forensic, I believe it would have been challenged and the process may well have been damaged in the courts". [iv]

Finally, as advances in science were used to improve the process, DNA fingerprinting was used in a forensic case for the first time in 1987. In this case, in an English village, two young girls were raped and murdered in 1983 and 1986. The police suspected Richard Buckland, as he had already confessed to murdering the teenage girls. Jefferys compared the DNA profiles obtained from the semen samples collected from the crime scene to that of Buckland, but there was no match. In search for the real culprit, around the Leicestershire area the police took blood and saliva samples from men between the ages of 17-34 to analyze their DNA. Eventually, a match was found between the crime scene sample and a man named Colin Pitchfork. Interestingly, Pitchfork had tried to evade the DNA dragnet and had asked a friend to provide samples on his behalf. On September 19, 1987, Pitchfork was arrested and in the following January he was sentenced to a life sentence.[v]

Initially, the Lister Institute of Leicester was the only institute performing DNA fingerprinting as it was Jefferys home institution. Later, around 1987 companies like Cellmark, the national police, and other law enforcement agencies started analyzing DNA for purposes involving paternity testing and human identification. The 1990s were the golden age for DNA fingerprinting as there was continuous research and advances in the field. During that time Jefferys' initial techniques of DNA analysis started to evolve as scientists started using PCR instead of Southern Blots, using fluorescent labels instead of

radioactive ones and using capillary electrophoresis instead of slab gels. Later, as observations of the gel patterns produced by DNA fingerprinting were replaced with measurements of the number of nucleotides repeat units, the technique was more appropriately called DNA profiling. The processes became automated, simple, and involved straightforward statistics. This is when scientists all over the world started using these methods.

Short Tandem Repeats: May Your Bands Be Variable!

According to Kirby, DNA typing, fingerprinting, profiling, testing, or analysis can be defined as the characterization of the unique patterns in an individual's genetic makeup or genome.[vi] A person's DNA is identical no matter if it is extracted from blood, semen samples or hair roots. This uniqueness of DNA within all tissues of an individual provides the foundation of DNA typing. The concept of DNA typing was first introduced in 1980 by Wyman and White with the polymorphic DNA locus which is characterized by variable-length restriction fragments named as restriction fragment length polymorphisms (RFLPs).[vii] Wyman and White suggested that the variation was not simply due to mutations in the restriction site but was due to the large number of alleles observed among individuals. These alleles indicated there were length differences within the locus that resulted in the variable alleles.

In 1985, shortly before applying their technique to the immigration case, Alec Jefferys and his coworkers published "Hypervariable Minisatellite Regions in Human DNA" where they discussed DNA fingerprinting.[viii] They discovered a region on human myoglobin gene with a region of 33-base pairs (bp) which was repeated 4 times between

an intervening sequence (IVS). The tandem repeat was called a minisatellite and repeats with a similar core sequence were found to occur at other loci. The region was said to be "hypervariable" as the number of repeats varied within and in between loci. The core repeat units consisted of 16 bp repeats. A complex ladder of DNA fragments was visualized when DNA was isolated and cleaved by enzymes and hybridized with a probe consisting of the repeat sequence under low stringency conditions. This pattern of DNA fragments turned out to be unique to every individual.

To advance this method, in 1986 Lifecodes investigators used two probes which targeted two polymorphic regions at a specific locus.[vi] This method was sensitive and effective compared to the prior DNA analyzing methods. In 1987, Nakamura et al. coined the term variable number of tandem repeats (VNTR) in reference to DNA typing which indicates the allele composition of different tandem repeats at individual loci which vary in the number of core units.[ix] Three years later in 1990 Nakamura along with Kasai and White published another article about using polymerase chain reaction (PCR) to successfully amplify a small quantity of DNA having a locus (D1S58) with VNTR. This resulted in good reproducible resolution of the individual alleles which differed by a single repeat unit. [x] This paved the way for a better understanding of PCR and how it could be used in forensic science practices.

One major limitation of RFLP, VNTR processes were related to the quantity and quality of DNA. Thus, in the early 1990s these methods of DNA fingerprinting started to get replaced by PCR based methods as PCR offered reduced time of analysis, enhanced precision, and better sensitivity.[xi] During this time, microsatellites referred to as short tandem repeats (STR) got introduced which became the ideal target for forensic

applications. STR typing was more sensitive than RFLP and less prone to allelic dropout in comparison to VNTR. The STR typing also proved to be more discriminating than the early PCR based typing methods like major histocompatibility complex, class II, DQ alpha 1 (HLA-DQA1).[xii,xiii]

Later, as more research went on to be performed in this area, more than 2000 articles got published describing the new tools in STR typing like miniSTRs, studying a variety of population groups, developing standards, and validating standard protocols in laboratories all over the world.[xiv] Currently, DNA typing has become much less complicated and easy to perform. It uses a panel of multi-allelic STR markers which are structurally like the initial minisatellites but has much shorter repeat units of 4 or 5 base pairs, which simplifies the amplification and analysis process.

Commercial STR Amplification Kits and the CODIS STR Amplification Loci

The PCR of short tandem repeat loci quickly replaced RFLP analysis of VNTR loci as the standard method for identifying biological evidence. STR typing was the most effective, sensitive, and precise technique for human identification.[xv,xvi,xvii] STR typing increased from a single locus reaction to a complex multiplex reaction involving the amplification of as many as 30 loci in one reaction. Detection of STR loci has progressed from using radioactivity and silver stain methods to using fluorescent instrumentation.

Multiplex STR amplification reactions go back to 1996 when the Forensic Science Services in Birmingham UK published an article validating an STR amplification kit targeting 7 loci along with amelogenin for identifying the sex of an individual. [xviii] A similar study was conducted around the same time by Sparkes et al. (1996) where they focused on

7 loci including the amelogenin.[xix] The first companies to manufacture multiplex STR amplification systems were Promega (Madison, WI) and Applied Biosystems (Foster City, CA). Lins et al. (1998) gives a detailed description of the multiplex STR kits that Promega manufactured during that time. They mention that the array of loci including CSF1PO, TPOX, TH01, vWA, D16S539, D7S820, D13S317, D5S818, F13A01, FESFPS, F13B, LPL were incorporated in multiple multiplex systems that were developed by Promega. Initially there were triplex systems and five multiplex systems. The triplex systems included CTT Multiplex, FFv Multiplex, and SilverSTR III Multiplex which were manufactured for silver stain detection.[xx,xxi,xxii] Their five multiplex systems (CTTv Multiplex, FFFL Multiplex, GammaSTR Multiplex, PowerPlex 1.1 System, and PowerPlex 1.2 System) used dyes attached to primers for fluorescent detection.

Around the same time, Applied Biosystems manufactured their first STR amplification kit named the AmpFISTR Blue PCR Amplification Kit which targeted 3 specific loci namely D3S1358, vWA and FGA.[xxiii] The dye 5-FAM was used in this kit which looked blue in the electropherogram, hence the name. Applied Biosystems then produced AmpFISTR Profiler and AmpFISTR COfiler, which when combined could amplify the newly established 13 loci to be included in the US database.

It was in the same year that the Federal Bureau of Investigation (FBI) published these 13 loci to be included while performing human identification. The 13 loci consisted of TPX, D3S1359, FGA, CSF1PO, D5S818, D7S820, D8S1179, TH01, VWA, D13S317, D16S539, D18S531, D21S11. Shortly after the database was established, kits like AmpFISTR Identifiler and Powerplex 16 were manufactured that targeted the 13 CODIS core loci in one reaction and were adopted by crime laboratories all over the world.[xxiv,xxv]

As time progressed the biotechnology companies continually manufactured new kits or updated the old kits. In 2008, Applied Biosystems released a new kit designed for degraded DNA which was specific to 8 of the largest size loci (D7S820, D13S317, D16S539, D21S11, D2S1338, D18S51, CSF1PO, and FGA).[xxvi] In the following year Promega upgraded their Powerplex 16 to Powerplex 16 HS where 2 more pentanucleotide loci (Penta D and Penta E) were included along with an improved buffer system.[xxvii] Thereafter, in 2012 Promega manufactured the Powerplex 18D which targeted the 13 CODIS loci, the amelogenin, Penta D, Penta E along with 2 other loci D2S1338 and D19S433. This commercial kit was specific to direct amplification of buccal swabs and blood stains directly from FTA Cards.[xxviii] Next in 2014, the Powerplex 21 was manufactured by Promega targeting D1S1656, D2S1338, D3S1358, D5S818, D6S1043, D7S820, D8S1179, D12S391, D13S317, D16S539, D18S51, D19S433, D21S11, Amelogenin, CSF1PO, FGA, Penta D, Penta E, TH01, TPOX, and vWA. The kit included the 13 CODIS loci along with some loci commonly used in Europe and Asia.[xxix] Following the footsteps of coming up with a kit which is compatible with DNA databases of different nations Applied Biosystems manufactured the Globalfiler Kit in 2014 which targeted 24 loci including the expanded CODIS core loci, the ESS (European Standard Set) and 3 other commonly used loci in other STR kits (D19S433, D2S1338, and SE33).[xxx] In 2014 Promega too came up with a similar commercial kit targeting the 24 loci and called it Powerplex Fusion.[xxxi]

In January of 2017, FBI officially published an expansion of the original 13 CODIS loci to include a total of 20 core loci to increase the strength of identification even further. The commercial kits have been actively in use in crime labs, paternity testing laboratories,

academic and research institutes all over the world. These kits have made human identification a sensitive, and reproducible process that takes less than one day. However, the expense involved in using these kits is quite high, mostly because of the cost involved in research and development along with customer support. For instance, as of January of 2023, the cost/reaction of AmpFISTR Identifiler is $36.20/reaction[xxxii] , Powerplex 16 is $33.38/reaction[xxxiii] , while that of AmpFISTR GlobalFiler is $28.35/reaction[xxxiv] , PowerPlex Fusion is $26.03/reaction[xxxv] .

In this regard, the academic institutes whose aim of performing STR amplification reactions is to learn the process of STR amplification, and the need is to keep repeating and practicing so that eventually when these students are working as forensic scientist in crime labs the rate of human error is low. Hence, the aim of this research project was to further develop a multiplex STR amplification reaction for students and researchers in the field which would be inexpensive but ideal for academicians.

Validation Studies: The Beginning, the Middle, the Now

From the 20th to the 21st century as the forensic laboratories kept evolving, the necessary rules and regulations that the laboratories needed to be followed also kept changing. While DNA evidence was gaining more importance in solving crimes, new working groups were being created to make the process reliable and legitimate. In this regard, accreditation of laboratories played a significant role in making sure the same standards were being followed all over the nation. An important part of the accreditation process was performance of validation studies to make sure the same criteria of standardization were being maintained.

In the 1980s, the FBI (Federal Bureau of Investigation), DEA (Drug Enforcement Administration) and other federal agencies sponsored technical working groups to assist the forensic science community in developing and applying laboratory methods and standards. As questions about the validity of DNA evidence in the courts appeared, the FBI established the Technical Working Group on DNA Analysis Methods (TWGDAM). TWGDAM was established in 1988 to improve communication between the different scientific disciplines and the federal, state, and local forensic communities. At the end of 1989, TWGDAM first published certain quality guidelines in the Crime Laboratory Digest for forensic biology laboratories.[xxxvi] These guidelines were revised and expanded in 1991.[xxxvii] The 1991 guidelines mostly focused on quality control protocols and quality assurance for DNA typing methods using RFLP/VNTR technologies. In this version of the guidelines, validation was not highlighted or discussed. A similar pattern was followed in the Crime Laboratory Digest published in 1993.[xxxviii]

The 1995 Crime Laboratory Digest discussed the creation of a Quality Assurance Unit in the Forensic Science Research and Training Center in the FBI academy.[xxxix] This was an attempt to help maintain proper laboratory functions in forensic laboratories as they submitted revised guidelines to the Technical Working Group on DNA Analysis Methods (TWGDAM) Quality Assurance (QA) Subcommittee. Eventually, certain changes were incorporated among those years' published guidelines. A section of 1995 publication discussed Validation Studies where Validation was defined as the process which is employed by the scientific communities to obtain the required information to determine the ability of a procedure to produce reliable results, acquire knowledge about the conditions under which such results are obtained and know what the limitations of the

procedure are. This section involved multiple rules that were to be followed while performing validation studies. Since this research primarily focuses on amplification of STR (Short Tandem Repeats) loci, the 4.4.1 section in the Crime Laboratory Digest is relevant as this section discusses the rules to be followed while performing validation of amplification reactions. The following are the factors to be considered:

- The sequence of the PCR primers used in the amplification reaction should be known.
- Specific conditions and measurements should be determined so that the pre-PCR samples are not contaminated by the post PCR products.
- The reaction conditions including thermal cycle parameters and concentration of the different reagents, needed to maintain the specificity of the reaction should be determined.
- The number of cycles needed to produce reliable results should be determined.
- An assessment of differential amplification should be performed.
- If in one sample mixture, more than one locus is amplified, its effect on each system should be documented.

Promega collaborated with the FBI, Palm Beach Sheriff's Crime Laboratory, The Blood Center of Southeastern Wisconsin, and DNA Berkeley Laboratory and validated the CTT Multiplex. The validation studies of the same was published in 1996 issue of the Journal of Forensic Science.[xl] The article followed the guidelines recommended by TWGDAM.

Alongside all these developments, in 1994 the DNA Advisory Board was established under the DNA Advisory Act. This Board was administered by the FBI and

was a distinct and separate entity whose aim was to manage and upgrade the forensic DNA standards of the nation. This board emphasized the importance of accreditation of crime laboratories, and the ability for labs to maintain the same standard protocols and follow quality control guidelines. Finally in 1998 the "Quality Assurance Standards for Forensic DNA Testing Laboratories" took effect.[xli] These standards would guide laboratories in setting up Quality Assurance Programs and included information on staffing, facilities, evidence control, and equipment maintenance. While the 1998 QAS document did have a section on validation, the section lacked the specific details for STR validation protocols that were contained in the 1995 TWGDAM publication.

The DNA Advisory Act of 1994 also established the national Combined DNA Index System (CODIS). It was in 1990 that CODIS began as a pilot project for building a database, but after the DNA Advisory Act the FBI authorized the formation of the National DNA Index System (NDIS) which eventually began operation in 1998.[xlii]

In 1999, TWGDAM was re-named as the Scientific Working Group on DNA Analysis Methods (SWGDAM) and added more guidelines towards betterment of the forensic community. In 2003, due to the increased use of STR typing throughout the nation, revised validation guidelines were published separating the validation section from the 1995 QAS.[xliii] These new 2003 validation guidelines were written specifically for the now established STR multiplex amplification and analysis of biological evidence. Since 2003, commercial companies follow these validation guidelines as new STR multiplex kits are developed and released and when new protocols are adopted. As the years passed, the guidelines for performing a validation study were updated, and the most recent version was published in December of 2016. These current guidelines describe two types of validation,

namely developmental validation, and internal validation. These validation studies consist of some studies which are mentioned below:

- Characterization of genetic marker: The fundamental characteristics of a genetic marker including mapping, inheritance, detection, and polymorphism should be tested and documented.
- Species specificity: The ability to detect non targeted species should be considered.
- Sensitivity study: The limit of detection along with dynamic range, ideal target range, signal to noise ratio associated with the assay should be determined.
- Stability studies: The ability to obtain DNA from compromised samples.
- Precision and accuracy: The repeatability and reproducibility of the assays should be determined.
- Case type samples: The ability to produce results for samples which are representative of typically encountered samples in case work should be examined.
- Population studies: In relevant population groups, the distribution of genetic markers should be checked.
- Mixture studies: The ability to obtain DNA profiles from mixture samples should be determined.
- PCR based studies: Different reaction conditions including thermal cycling parameters should be determined. Primer sequences used for the reaction need not be published but should be available if needed.

For laboratories wishing to design their own amplification reactions, these validation guidelines highlighted the areas that should be addressed to assure the reaction results are accurate and reproducible.

Custom Kit: Why Is It Required?

The commercial STR Amplification kits available in the market are expensive as seen in the case of the most popularly used AmpFISTR Identifiler or the most recent GlobalFiler PCR Amplification Kits. As a result of which it becomes very difficult for academic institutes running on a low budget to let students use the kits for the purpose of training. Multiple research centers running on a low budget face similar problems where the entry level researchers or employees in the field of forensic biology may produce more human errors because of lack of practice. For these reasons, beginning in 2013, UAB developed and has been using a custom made multiplex STR amplification reaction targeting 3 loci (D3S1358, Th01, and D13S317). The reaction has served the purpose of allowing students to perform STR analysis on samples used for teaching and research but has limitations. First, having only 3 loci the kit limits the ability to identification in comparison to a kit with 6 loci. Second, in the recent past the reactions performed using the old kit have been showing contamination issues. Third, the first kit was not scientifically published to authenticate it's working.

In this project an additional 3 STR loci were targeted for a new multiplex reaction, and a validation study was performed. The aim was to have two multiplex STR reactions available in our lab, having a total of 6 loci to increase the significance of the resultant DNA profile. Moreover, this would also mean in case of any contamination issues there would always be a second kit to switch to in order to troubleshoot the problem. Another aim was to publish our process of developing and validating a method which could be followed by any other academic or non-academic laboratories wanting to create their own STR amplification kits.

CHAPTER 2

MATERIALS AND METHODS

Biological Samples and Control DNA

Cheek cells were collected at the University of Alabama at Birmingham from participants following protocols approved by the IRB (IRB 300003083). Cheek cells were collected via self-swabbing using a foam tipped swab. Blood and semen samples used in this study were purchased from BioIVT (Westbury, NY). Blood and semen stains were prepared by applying 50 µL of biological fluid to 100% cotton fabric and allowing samples to dry at least 24 hours before extraction. For this study two different control DNA samples were used. One sample of 2800M Control DNA (0.25ng/uL) was purchased from Promega (Madison, WI) and Applied Biosystems TaqMan Control Human DNA Male (10 ng/uL) was purchased from Thermo Fisher Scientific (Waltham, MA).

Quantification and Dilution of DNA

The extractions from participants were quantitated using the Quantifiler Duo Human DNA Quantification Kit (Thermo Fisher Scientific, Waltham, MA) following the manufacturer's recommended protocol for a 25 µL reaction. For each reaction, to 12.5 µL of MasterMix, 10.5 µL of primer mix and 2 µL of DNA were added. Samples were amplified using an Applied Biosystems 7500 real time PCR instrument (Thermo Fisher Scientific) and analyzed using SDS software version 1.2.3 (Thermo Fisher Scientific).

Dilutions of participant's DNA and control DNA were prepared based on the quantitation of the participant DNA or based on the given DNA concentrations of the 2800M (0.25 ng/L) and Human DNA Male (10 ng/µL) control DNA samples. Dilutions of all the participants' DNA were prepared at a concentration of 0.5 ng/µL. These dilutions were used for making the allelic ladder.

The dilutions of 250 pg/µL and 125 pg/µL that were prepared for the sensitivity study (see below) were also quantitated after the sensitivity study was completed. These were dilutions from the 2800M standard, the Human Male DNA standard, and from one of the participant extractions. These dilutions were quantitated using the Quantifiler Duo Human DNA Quantification Kit at the parameters previously mentioned. These quantitation results were used to help determine the accuracy of the dilution concentrations when interpreting the data from the sensitivity study.

Selection and Preparation of Primers

The 3 loci selected for this project were D5S818, D18S51 and D16S539. Hereafter, D5S818 shall be abbreviated as D5, D18S51 as D18 and D16S539 as D16. The selected primer set for D5 had been previously published by Promega,[xliv] D18 was published by Urquhart et al.[xlv] and D16 was published as part of the NIH Cooperative Human Linkage Center (CHLC), a human genome mapping project with data that is no longer available online.[xlvi] For D5 and D18 6-FAM was selected as the dye as the common alleles from these loci had amplicon sizes of 119-155 bp and 266-342 bp respectively, and the alleles would not overlap. While for D16, VIC was selected as the dye as the common alleles from this locus had amplicon sizes of 133-173 bp. By using a different dye, these allele sizes

could not overlap with those of D5. Table 1 shows the selected primer sequences and the respective dyes and amplicon sizes. Forward primers (fluorescently labelled) were purchased from Thermo Fisher Scientific (Waltham, MA) while the reverse primers were purchased from Eurofins Genomics LLC (Louisville, KY). For monoplex reactions, a 5 pmol/µL working solution of each primer was prepared. For multiplex reactions, a custom primer mix was prepared by first preparing 100 µL of each forward and reverse primer at a concentration of 5 pmol/µL, then mixing all the solutions resulting in a total volume of 600 µL with a concentration of 0.83 pmol/µL for each primer. Unless otherwise noted, 6 µL of this primer mix was used for the multiplex amplification reactions, resulting in 5 pmol of each primer in the reactions.

Table 1: The Loci Selected for the Multiplex STR amplification Reaction and the Corresponding Primer Sequences

LOCI	PRIMER SEQUENCE	AMPLICON SIZE (bp)
D5S818	5'-GGTGATTTTCCTCTTTGGTATCC-3' 5'-[6FAM]-AGCCACAGTTTACAACATTTGTATCT-3	119-155
D18S51	5'-[6FAM]-CAA ACC CGA CTA CCA GCA AC-3' 5'-GAG CCA TGT TCA TGC CAC TG-3'	266-342
D16S539	5'-[VIC]-GATCCCAAGCTCTTCCTCTT-3' 5'-ACGTTTGTGTGTGCATCTGT-3'	133-173

Monoplex Testing of Primer Sets

At first, monoplex reactions for each of the 3 loci (D5, D18 and D16) were performed using the primer sets described above. For performing monoplex reactions, the total reaction volume was 25 µl and included 12.5 µL of Promega GoTaq Hot Start Polymerase: Colorless Master Mix (Madison, WI), 1 µl of forward primer (5 pmol/µL) and 1 µl of reverse primer (5 pmol/µL) and 1 ng of template DNA, with the remaining volume consisting of nuclease free water (Madison, WI). The prepared samples were then amplified in an Applied Biosystem's thermal cycler using a program of a 95°C hold for 11 minutes, a 96°C hold for 1 minute, 10 cycles of 94°C for 30 seconds, 58°C for 30 seconds, 70°C for 45 seconds, then 20 cycles at 90°C for 30 seconds, 58°C for 30 seconds, 70°C for 45 seconds, and a final hold at 60°C for 30 minutes.

Multiplex Amplification of DNA

Unless otherwise noted, the following multiplex reaction components and amplification conditions were used for all reactions. Each 25 µL amplification reaction contained 12.5 µL of Promega GoTaq Hot Start Polymerase: Colorless Master Mix (Madison, WI), 6 µL of a the custom D5, D18, D16 primer mix described above resulting in 5 pmol of each primer in each reaction, 1-6 µL of DNA extraction, and an amount of nuclease free water to bring the total volume to 25 µL. The prepared samples were then amplified in an Applied Biosystem's thermal cycler using a program based on the recommendation in the GenePrint Powerplex 1.2 System Technical Manual.[xlvii] The annealing temperature was changed from the recommended 60°C to 58°C based on the results of the annealing temperature optimization study described below. The adjusted parameters were a 95°C hold for 11 minutes, a 96°C hold for 1 minute, 10 cycles of 94°C

for 30 seconds, 58°C for 30 seconds, 70°C for 45 seconds, then 20 cycles at 90°C for 30 seconds, 58°C for 30 seconds, 70°C for 45 seconds, and a final hold at 60°C for 30 minutes.

Separation and Analysis of Amplified Product

Unless otherwise noted, the amplified product was separated and analyzed using the following process. Capillary electrophoresis was performed using an Applied Biosystems 3130 Genetic Analyzer (Life Technologies) and GeneMapper ID (Life Technologies) software to analyze the DNA profiles. The amplified product was prepared by adding 1 µl of the amplified DNA to 8.7 µl of Applied Biosystems Hi Di Formamide (Life Technologies) and 0.3 µl of Applied Biosystems GeneScan 500 LIZ Size Standard (Thermo Fisher Scientific). Amplified product was separated using the instrument conditions recommended for separation of samples using the Applied Biosystems AmpFISTR Identifiler PCR Amplification Kit, including an injection at 3 kV for 5 seconds, a run voltage of 15 kV, and an oven temperature of 60°C. Nano POP4 (MCLAB, South San Francisco, CA) was used as the polymer along with a 36 cm 4 capillary array. Alleles were called with peak heights greater than the threshold of 100 Relative Fluorescent Unit (RFU).

Making the Allelic Ladder

To make the allelic ladder, the 0.5 ng/µL dilutions from 8 participants DNA samples were amplified separately using the multiplex amplification reaction described above. In the amplification reactions, 4 µL (2ng) of DNA were added. Twenty (20) µl of the amplified product from each volunteer were mixed together. This mixture of amplified

products would be used as the allelic ladder throughout the remainder of the research study. From this mixture of amplified product, 1 µL was prepared and separated on the 3130 Genetic Analyzer following the procedure described above. The remaining portion of the mixture served as an allelic ladder, which was run alongside all amplified product for all validation samples. One (1) ng of the 2800M Control DNA was also amplified following the same protocol and separated on the 3130 Genetic Analyzer alongside the newly developed allelic ladder and the volunteer DNA samples. The resulting allele calls of the ladder were compared with the already known alleles of the 2800M Control DNA. The alleles were also compared to a population database to determine whether the alleles in the ladder were common alleles.

Validation of the Multiplex Reaction

Thermal Cycling Parameters – Annealing Temperature

To determine the optimum annealing temperature to produce peaks with adequate height while maintaining a low amount of noise, four different temperatures 56°C, 58°C, 60°C and 62°C were tested. The programs were based on the program described above. The program for 58°C included a 95°C hold for 11 minutes, a 96°C hold for 1 minute, 10 cycles of 94°C for 30 seconds, 58°C for 30 seconds, 70°C for 45 seconds, then 20 cycles at 90°C for 30 seconds, 58°C for 30 seconds, 70°C for 45 seconds, and a final hold at 60°C for 30 minutes. The other programs changed the annealing temperature for all 30 cycles. For these reactions, 1 ng of 2800M DNA was amplified in triplicate for each annealing temperature.

Thermal Cycle Parameters – Cycle Number

A similar approach was taken to determine the optimum number of cycles. Four different PCR thermal cycling programs were tried, including a total of either 28, 30, 32 or 34 cycles. The programs were based on the program described above. The program for 28 cycles included a 95°C hold for 11 minutes, a 96°C hold for 1 minute, 10 cycles of 94°C for 30 seconds, 58°C for 30 seconds, 70°C for 45 seconds, then 18 cycles at 90°C for 30 seconds, 58°C for 30 seconds, 70°C for 45 seconds, and a final hold at 60°C for 30 minutes. The other programs also included an initial 10 cycles with the 94°C denaturing temperature but increased the number of cycles using the 90°C denaturing temperature to 20, 22, or 24. For these reactions, 1 ng of 2800M DNA was amplified in triplicate for each cycle number.

Sensitivity Study

The sensitivity of the multiplex reaction was tested to determine how a decreasing amount of total input DNA would affect the resulting profile. For this purpose, three serial dilutions were prepared from the 2 control DNA (2800M and Human DNA Male) along with one participant's DNA sample (volunteer 7). The concentrations of dilutions were 250 pg/µL, 125 pg/µL, 62.5 pg/µL, 31.25 pg/µL, 15.6 pg/µL, 7.8 pg/µL and 3.9 pg/µL. For the amplification reactions, 4 µL of each dilution were added to the 25 µL reactions resulting in 1 ng, 0.5 ng, 0.25 ng, 0.125 ng, 62.4 pg, 31.25 pg or 15.6 pg of DNA in each reaction. Amplification reactions were performed in duplicate using the 30-cycle program with an annealing temperature of 58°C. Preparation of amplified products and capillary electrophoresis were conducted following the above-mentioned process.

Precision Study

To determine the variation in the sizing of alleles, a precision study was performed. For the precision study, eight samples of prepared allelic ladder were injected two times, for a total of 16 injections. The injections used the separation and analysis parameters for the 3130 Genetic Analyzer described above.

Accuracy and Concordance Study

The accuracy of the allele calls was tested through a concordance study comparing the DNA profiles obtained from the custom reaction with that of the AmpFISTR Identifiler kit. For this study, DNA from blood and semen were first amplified using the custom reaction described above, and then using the Applied Biosystems AmpFISTR Identifiler PCR Amplification Kit (Thermo Fisher Scientific). The Identifiler reactions were prepared according to the manufacturer's recommended protocol for a 25 µL reaction and the manufacturer's recommended thermal cycling conditions. The amplified product was prepared for analysis and analyzed using the parameters recommended for the Identifiler Kit.

CHAPTER 3

RESULTS AND DISCUSSION

Selection and Preparation of Primers

The original 13 STR loci chosen by the FBI for inclusion in the Combined DNA Index System (CODIS) database were considered due to the availability of published primer sequences for those loci.[xlviii,xlix] The loci D3S1358, Th01, and D13S317 were excluded as candidates as these three loci had already been chosen for a separate multiplex reaction currently used in the UAB forensic laboratory. Two important factors, a small amplification product size and avoiding overlapping alleles of the 3 loci were considered while selecting the appropriate STR loci and the respective primer sequences. An organized list of previously published primer sequences can be found on the STRbase website.[l] This list was used to select primers based on the amplicon size. In choosing the dyes, Applied Biosystems dyes were selected due to their compatibility with the dye matrix used for analysis of AmpFISTR Identifiler samples.

Monoplex Testing of Primers

All of the primer sets successfully amplified DNA in the expected size range when 5 pmol of the forward and reverse primer were added to a 25 µL amplification reaction. Figure 1 shows a representation of the electropherograms obtained when separate monoplex amplification reactions were performed for D5S818 (D5), D18S51 (D18) and

D16S539 (D16). The expected two allele peaks were observed at the respective loci. Some stutter peaks, which are one repeat unit shorter than allele peaks and are a small percentage of the peak height, were observed at D18. Stutter peaks are a common phenomenon when amplifying short tandem repeat (STR) loci and have been documented in previous validation studies.[li,lii] Stutter peaks are generally observed because of slipped strand mispairing (SSM) during the Polymerase Chain Reaction which occurs due to insertion or deletion of a repeat unit on the newly formed strand.[liii] Overall, the alleles at each loci demonstrated that the primers for each locus were amplifying the alleles present at the expected location of DNA.

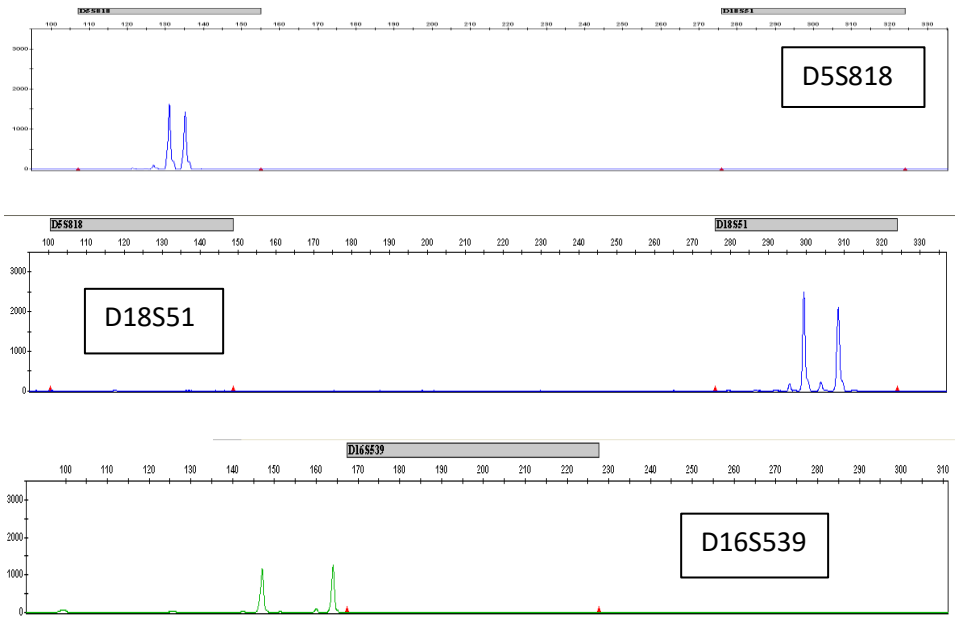

Figure 1: Representative electropherograms of 1 ng 2800M Control DNA amplified in three separate monoplex reactions for D5S818, D18S51 and D16S539. The top and middle panels (blue) and the bottom panel (green) correspond to 6-FAM and VIC labeled peaks respectively. The reactions were specific to amplifying only the targeted alleles and proved successful working of the forward and reverse primers at 5 pmol of each primer in all reactions.

Making the Allelic Ladder

The combined 20 μL of amplified DNA from 8 research participants produced at least 3 alleles at each locus when using a multiplex reaction of the D5, D18, and D16 loci. There were 3 peaks at D5, 7 peaks at D18 and 3 peaks at D16. The sizes of these peaks (in base pairs) were compared to that of the 2800M Control DNA as shown in Table 2. The 2800M Control DNA was amplified and separated on the 3130 the same day as the ladder. Since the alleles of the control DNA were known, the remaining alleles of the ladder could be easily assigned. Next, the amplified alleles were compared to a population database,[liv] to confirm that the allelic frequency was 0.04 or higher, further supporting that the amplified DNA represented the expected alleles from the participants. The tallest allele peaks were selected to be included in the panel file used for the GeneMapper analysis of the custom multiplex samples. The panel file is a text file that can be edited, which instructs the software to identify certain allele peaks and create bins around those peaks. The panel file can also instruct the software to create virtual bins for alleles which may be present in future samples but are not present in the ladder. The tallest alleles included in the panel were 11,12,13 at D5, at 12,13,14,16,17,18,19 at D18 and 9,11,12 at D16. Figure 2 shows the electropherogram of the allelic ladder which was used throughout the research study.

Table 2: Comparing the peak sizes of the known alleles of the 2800M Control DNA to the peak sizes of the unknown peaks obtained from the mixture of 20 µl amplified DNA from 8 volunteers.

Labels	D5		D18		D16	
Mixture	Peak size (bp)	Allele	Peak size (bp)	Allele	Peak size (bp)	Allele
	130.91	11	282.31	12	147.09	9
	135.00	12	286.44	13	155.79	11
	139.21	13	290.64	14	160.00	12
			298.98	16		
			303.60	17		
			308.24	18		
			312.70	19		
2800M Control DNA	135.18	12	299.44	16	147.31	9
			308.77	18	164.26	13

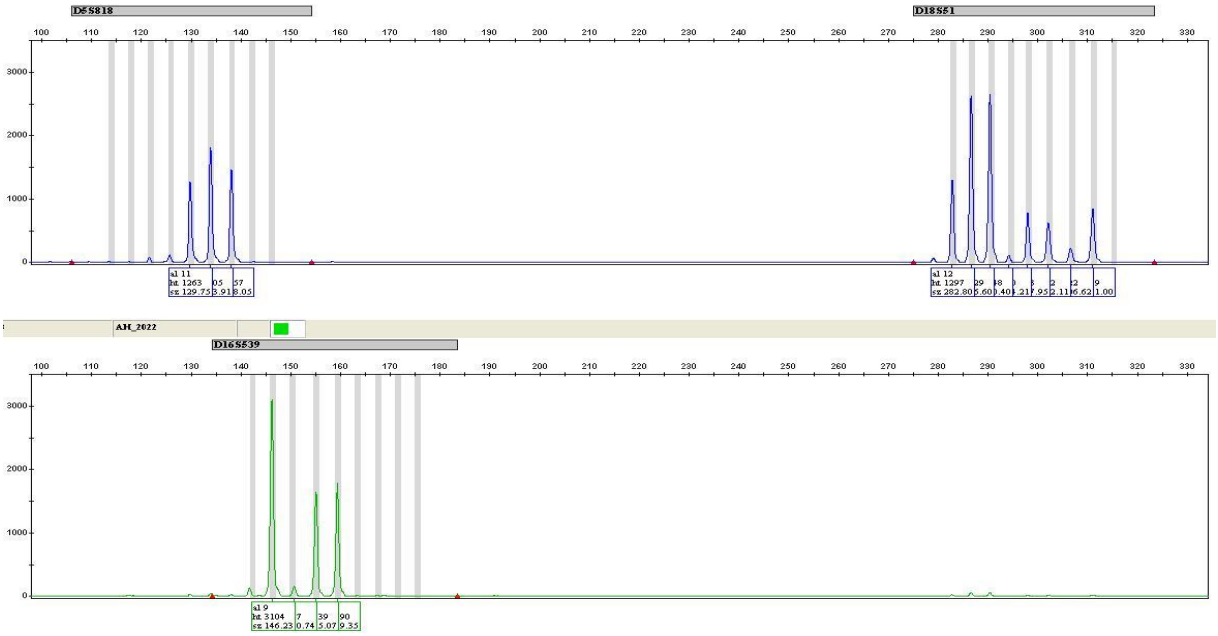

Figure 2: Representative electropherogram of the allelic ladder. The allelic ladder was developed by individually amplifying 2ng DNA from 8 volunteers using a multiplex reaction for D5, D18 and D16. The top panel (blue) and the bottom panel(green) correspond to 6-FAM and VIC labeled peaks respectively. The multiplex reaction used 5 pmol of each primer.

Multiplex Validation

Thermal Cycling Parameters – Annealing Temperature

One of the key features for a successful STR amplification reaction is the use of the correct annealing temperature. In this study, 56°C, 58°C, 60°C and 62°C were tested. The samples that were amplified at an annealing temperature of 58°C showed the best results with the highest peak heights and a limited amount of noise. Thus, 58°C was selected as the optimum annealing temperature for this STR amplification reaction. The average peak heights obtained at the different temperatures where it is clearly visible that 58°C showed highest average peak height is shown in Figure 3. Similar results are reported by validation studies for commercial kits. When validating the AmpFISTR NGM Select PCR Amplification Kit in 2013, Green et el. determined 59°C to be the optimal temperature when the range of temperatures checked were 55°C, 57°C, 59°C, 61°C, and 63°C. [lv] Similar results were observed by Collins et al. (2004) while validating the AmpFISTR Identifiler and by Ensenberger et al. (2014) while performing Powerplex 21 system where the optimum annealing temperature was 59°C.[lvi,lvii] Since that time, other validation studies have not reported the results of testing an optimal annealing temperature, instead reported that the recommended protocol uses 59°C, the same annealing temperature as previously validated kits. This is true for the recently validated Globalfiler[lviii] and Powerplex Fusion[lix] kits too.

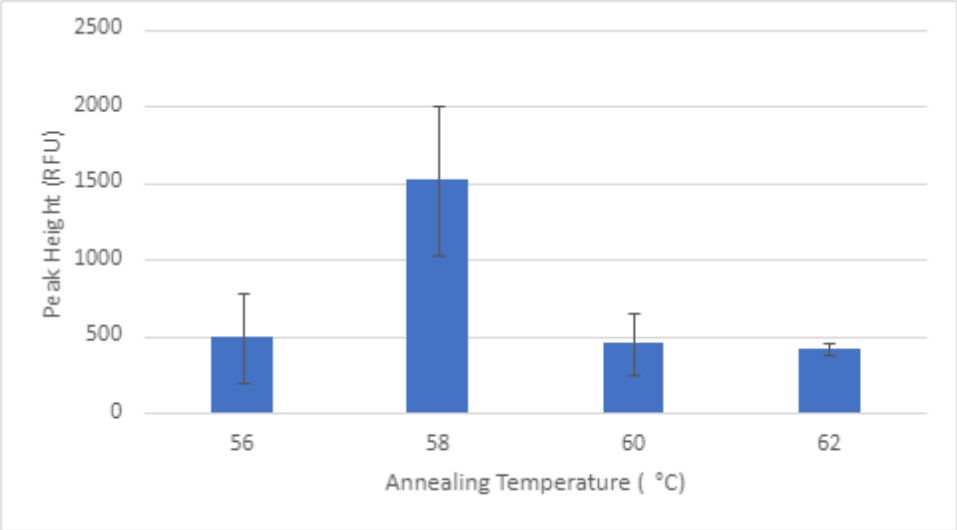

Figure 3: Effect of different annealing temperatures on the peak height average. One ng of 2800M Control DNA was amplified in triplicate using the multiplex reaction. The annealing temperature of 58 °C showed the highest average peak height.
N: 3; Error bars: ± 1SD

Thermal Cycling Parameters – Cycle Numbers

A range of amplification cycles were selected to be included in this study, based on the number of cycles used for commercial kits (28, 30, 32 and 34). The allele peaks at 30 cycles and 32 cycles showed the best results in terms of high peak heights and a low amount of stutter and background noise. Figure 4 shows the average peak heights obtained at the different cycle numbers. Amplified alleles failed to cross the threshold when amplified using 28 cycles. When samples were amplified using 34 cycles, stutter was increased, and other non-allele peaks were present (Figure 5B). Though 32 cycles showed slightly better results in terms of peak heights, 30 cycles were selected as the chosen cycle number for several reasons. First, the combination of an annealing temperature of 58°C and a cycle number of 30 gave consistently acceptable results throughout the early validation process

(Figure 5A). In this context, instead of using 1 ng of template DNA if more template DNA is used, 30 cycles would result in less stutter and non-allelic peaks in comparison to 32 cycles. Second, using 30 cycles takes less time for the thermal cycling to be completed when compared to 32 cycles. Additionally, the other multiplex reaction that is already in use was previously validated at 30 cycles with an annealing temperature of 58°C. Using the same program as the other multiplex reaction would be more efficient and the new multiplex reaction could be completed on the same thermal cycler at the same time.

When compared to other multiplex validations, use of a similar number of cycles was observed, although commercial kit recommendations may vary based on the type of sample being amplified. In the same study for the AmpFISTR NGM Select PCR Amplification Kit mentioned above, Greens et al. tested 27, 28, 29, 30 and 31 cycles.[lv] The article recommended the cycle number for 1ng DNA should be 29 cycles. The recent Globalfiler kit also recommends 29 cycles.[lviii] Ensenberger et al. (2014) recommended a 30 cycle protocol while validating the Powerplex 21 System.[lvi] The Powerplex Fusion kit recommends 30 cycles, but gives different recommendations for solid samples where the amount of input DNA may be higher.[lix] For example, FTA punches were only amplified for 27 cycles in the Powerplex Fusion validation study.

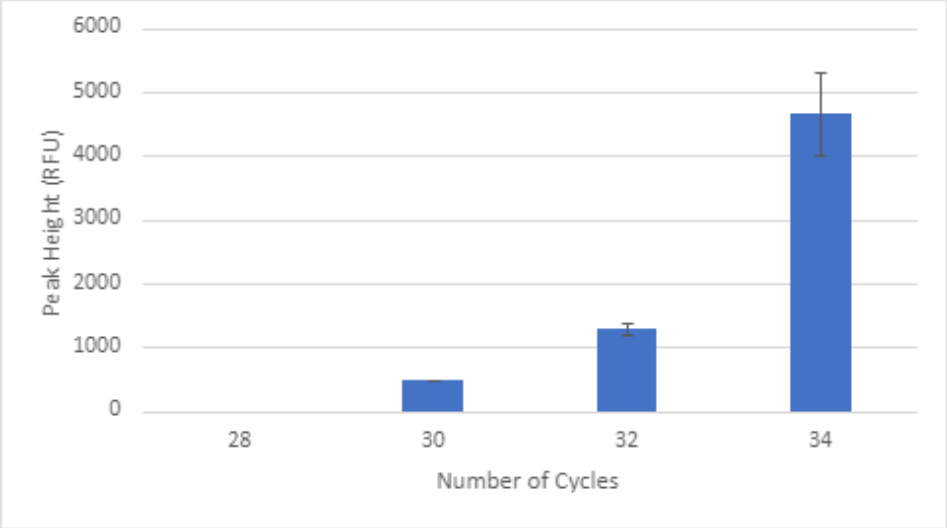

Figure 4: Effect of different PCR cycle numbers on the peak height average. One ng of 2800M Control DNA was amplified in triplicate for 28, 30 and 32 cycles in a multiplex reaction. For 34 cycles 1 ng of 2800M Control DNA was amplified in duplicate in a multiplex reaction.
N: 3 for 28, 30 and 32 cycles; N: 2 for 34 cycles; Error bars: ± 1SD

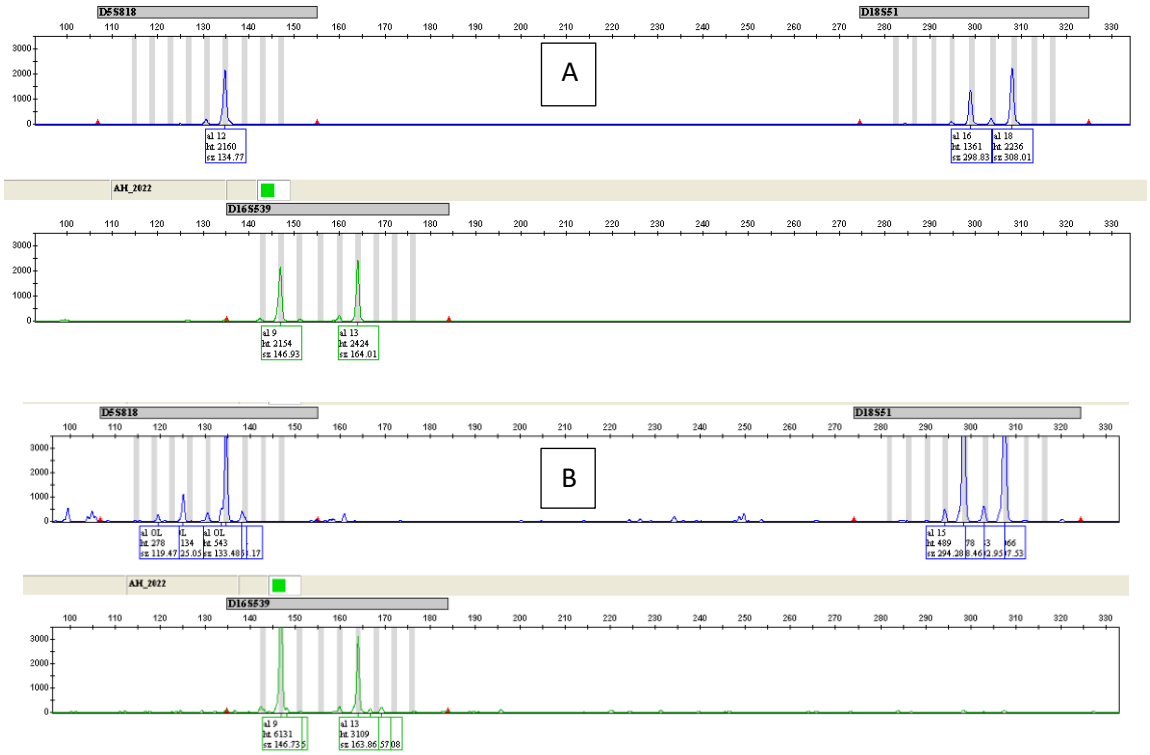

Figure 5: Representative electropherogram of 1ng 2800M Control DNA amplified with the new STR amplification reaction in Applied Biosystems Thermal Cycler. (A) The combination of thermal cycle parameters of 30 cycle numbers and a 58°C Annealing Temperature. (B) The cycle number used for the PCR was 34. The obtained results showed high peaks with stutter and background noise at the targeted loci of D5, D18 and D16. The top panel (blue) and the bottom panel(green) correspond to 6-FAM and VIC labeled peaks respectively. The multiplex reaction used 5 pmol of each primer.

Sensitivity Study

The sensitivity study was performed to determine how the quantity of input DNA affects the amplification of alleles from all three loci. As expected, there was a decrease in peak heights as the amount of DNA in each reaction decreased from 1 ng to 15.6 pg in the buccal swab DNA, 2800M Control DNA and Human DNA Male. For the DNA extracted from the buccal swab of a participant, full DNA profiles were produced at 0.25 ng of DNA, while partial profiles were observed at 0.125 ng and 62.5 pg, and alleles failed to amplify

at 31.25 pg and below. For the 2800M Control DNA, full profiles were produced with 0.25 ng of DNA, followed by partial profiles at 0.125 ng, 62.5 pg and 31.25 pg, and a failure to amplify alleles at 15.6 pg. In the case of Human DNA Male, a complete profile was observed when 1 ng, 0.5 ng, 0.25 ng, 0.125 ng, 62.5 pg and 31.25 pg were amplified, followed by one full and one partial profile at 15.6 pg.

These results are similar to the reported sensitivity of commercial kits, where many more loci are amplified. The Globalfiler reaction resulted in alleles dropping out when 62.5 pg of DNA were amplified, and successful amplification of some alleles when 15.6 pg of DNA were amplified.[lviii] The Powerplex Fusion reaction resulted in some dropout of alleles when 50 pg of DNA were amplified.[lix] A similar study conducted by Lucy et al (2009) performed a sensitivity study using the AmpFISTR Minifiler kit, a reaction specifically designed for amplifying small amounts of DNA by only targeting 9 loci and using primer sets that produce an amplified product of a reduced size when compared to other commercial kits. A full profile was obtained when 50 pg of control DNA were amplified, while for reference samples a full profile required 62.5 pg of DNA in the reaction.[lx]

Overall, the results indicated that the new multiplex STR amplification reaction would give full DNA profiles for all 3 loci till a DNA concentration of 0.25ng. It will result in partial profiles at a DNA concentration of 0.125 ng - 62.5 pg. However, in case of Human DNA Male full profiles were visible at 31.2 pg and 15.6 pg too. The probable cause of this increased sensitivity was a higher concentration of DNA in the dilutions than what was theoretically calculated. This hypothesis was supported after quantifying the 250 pg/µL dilution of Human DNA Male, which resulted in a measured DNA concentration of 570 pg/µL. The Human DNA Male is sold as positive control DNA at a reported concentration

of 10 ng/µL. The certificate of analysis states the DNA concentration is determined by UV spectroscopy, which lacks the accuracy of quantitation by real-time PCR. Hence, the amplification of DNA resulting in a full profile at the concentrations of 31.5 pg and 15.6 pg of Human DNA Male is reasonable. The results from the participant DNA and 2800M are likely more accurate representation of the newly developed STR amplification reaction's sensitivity.

Precision Study

Consistency in the sizing of alleles was checked by calculating the average peak size and standard deviation for each allele present in multiple injections of the allelic ladder. The reproducibility and consistency were maintained among all the injections with an average SD of 0.09 while the highest SD was 0.16 at the 19th allele in the D18 locus. These results are consistent with the precision reported for commercial kits, where the highest SD for the Powerplex Fusion kit was 0.16 when run on the 3130 Genetic Analyzer.[lix] The Globalfiler kit precision study was completed on an Applied Biosystems 3500 Genetic Analyzer and the highest standard deviation for an allele was reported to be 0.06.[lviiilviii]

Accuracy and Concordance Study

The DNA profiles obtained for the extracted blood and semen samples when the new multiplex STR amplification reaction was used were consistent with the profiles obtained using the AmpFISTR Identifiler kit. A comparison of the DNA profiles obtained using the 2 different kits is shown in Table 3.

Table 3: Alleles obtained when DNA samples were analyzed using the new multiplex STR amplification reaction compared to the Identifiler kit.

Loci	Alleles obtained with AmpFISTR Identifiler PCR Amplification Kit	Alleles obtained with the custom made STR amplification reaction
Blood		
D5	13, 13	13, 13
D18	13, 17	13, 17
D16	9, 10	9, 10
Semen		
D5	12, 12	12, 12
D18	14, 20	14, 20
D16	9, 13	9, 13

Pricing

Another aim of this research was to provide a multiplex STR amplification reaction that is inexpensive when compared to commercial kit reactions. In that regard, the price of all the products used to make the kit was evaluated. As of March 5, 2023, the three labelled forward primers bought from Applied Biosystems were $343.77, and the three reverse primers from Eurofins were $10.08. These primers can be diluted and used to perform 2000 reactions at a reaction volume of 25 µL. The cost of the Promega Go Taq Hot Start Polymerase Colorless Mix, which includes the nuclease free water was $1120.00 for 2000 reactions. A total of 2000 reactions could be performed by using all these products at a total cost of $1473.85, which makes the cost/reaction approximately $0.74.

For commercial kits, cost per reaction is much higher. The current listed price for a 200 reaction Globalfiler kit is $5670.00 resulting in a cost/reaction of $28.35. The current listed price for a 200 reaction Powerplex Fusion kit is $5206.00 resulting in a cost/reaction

of $26.03. Although these commercial kits do come with positive control DNA and with a prepared allelic ladder, research labs would be able to use previously extracted DNA as a positive control and would be able to make their own allelic ladder using the process previously described for a minimal cost. This 30 times reduction in the cost/reaction for the custom multiplex kit is the most obvious advantage for using these reactions for teaching or research.

Future Work

This research study has resulted in a new multiplex STR amplification reaction which is both efficient and inexpensive. Other academic or research laboratories can use the guidelines provided in this study to manufacture their own STR amplification kit by modifying certain parameters like the allelic ladders, primers, number of selected loci, etc. The allelic ladder can be improved by including more participants to result in a wider range of alleles Additionally, a greater number of loci can be used in one reaction to increase the significance of the profile, while only moderately increasing the cost. In this study the primers were selected from the already published primer sets, future research work can also include the designing of new primers for the chosen loci. This project holds a lot of potential that can be undertaken by other scientists.

CONCLUSION

The commercial STR amplification kits are widely used in forensic DNA labs and in academic institutes. The crime labs' use of the commercial kits is necessary in spite of the expense since the aim is human identification that requires highly sensitive and efficient reactions. However, the situation is slightly different for academic institutes and research centers where these kits are used in forensic education and research to teach the concepts of STR amplification. In the best-case scenario, availability of an inexpensive STR amplification kit would allow the DNA amplification techniques to be repeated as many times as possible. This could result in reduced training requirements once these students are hired by crime labs. In this regard, a new multiplex STR amplification reaction was developed and validated in this project. The cost of this custom STR reaction is $0.74/reaction, thereby giving the opportunity for students to repeat STR typing until they are confident with their performance.

In this new reaction, D5, D18 and D16 were selected as the targeted loci. A monoplex reaction was performed using a set of primers whose sequences had been published by other companies and researchers. Once each forward and reverse primer set was established as working, a validation study was performed following the SWGDAM guidelines after studying multiple validation studies performed by other scientists. Following the guidelines a validation study consisting of sensitivity study, precision study,

accuracy study, concordance study along with checking the thermal cycling parameters including cycle numbers and annealing temperature were performed.

In the sensitivity study, full profiles were obtained to a DNA concentration of 0.25 ng. The size precision study produced allele sizes with standard deviations between 0.04 and 0.16. PCR based studies showed the best results at an annealing temperature of 58°C. In the case of the cycle numbers, 30 and 32 gave good DNA profiles. Though 32 cycles gave an average peak height which was higher than the 30 cycles, the combination of 30 cycles and an annealing temperature of 58 °C resulted in consistent adequate peak heights throughout the rest of the study. Additionally, these conditions are known to work well with the other reaction which has already been in use at UAB. For the concordance study, the results obtained from the new STR amplification reaction and the AmpFISTR Identifiler Kit produced same allelic profile at each of the 3 loci. This showed that the developed reaction could work as an alternative to the Identifiler kit in academic institutes when the purpose is not identifying an individual with a high level of certainty, but to show proof of concept. Overall, a successful multiplex STR amplification reaction was developed and validated in this project holding potential for future scientists to use the reaction or create their own STR amplification kit following the process described in this research work.

REFERENCES

[i] Doyle AC. A study in scarlet, Beeton's Christmas Annual. 1887

[ii] Jeffreys AJ, Wilson V, Thein SL. Individual-specific "fingerprints" of Human DNA. Nature. 1985;314:67.-74

[iii] Jeffreys AJ, Brookfield JF, Semeonoff R. Positive identification of an immigration test-case using human DNA fingerprints. Nature. 1985;317:818.-819

[iv] University of Leicester Bulletin Supplement August/. 2004, September

[v] https://cen.acs.org/analytical-chemistry/Thirty-years-DNA-forensics-DNA/95/i37 Reviewed January 24, 2023

[vi] Kirby LT, editor. DNA fingerprinting: an introduction. Oxford University Press; 1993 Sep 23.

[vii] Wyman AR, White R. A highly polymorphic locus in human DNA. Proceedings of the National Academy of Sciences. 1980 Nov;77(11):6754-8.

[viii] Jeffreys AJ, Wilson V, Thein SL. Hypervariable 'minisatellite'regions in human DNA. Nature. 1985 Mar;314(6006):67-73.

[ix] Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science. 1987 Mar 27;235(4796):1616-22.

[x] Kasai, K., Nakamura, Y., & White, R. (1990). Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science. *Journal of forensic sciences*, *35*(5), 1196-1200.

[xi] Edwards A, Civitello A, Hammond HA, Caskey CT: DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet. 1991, 49: 746-756.

[xii] Budowle B, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC: Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. Am J Hum Genet. 1991, 48: 137-144

[xiii] Saiki RK, Bugawan TL, Horn GT, Mullis KB, Erlich HA: Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. Nature. 1986, 324: 163-166. 10.1038/324163a0

[xiv] Coble MD, Butler JM: Characterization of new miniSTR loci to aid analysis of degraded DNA. J Forensic Sci. 2005, 50: 43-53

[xv] Edwards, A. et al. (1991) "DNA Typing and Genetic Mapping with Trimeric and Tetrameric Tandem Repeats," Am. J. Hum. Genet. 49, 746

[xvi] Edwards, A. et al. (1992) "Genetic Variation at Five Trimeric and Tetrameric Tandem Repeat Loci in Four Human Population Groups," Genomics 12, 241.

[xvii] Lins, A.M. et al. (1996) "Multiplex Sets for the Amplification of Polymorphic Short Tandem Repeat Loci – Silver Stain and Fluorescent Detection," BioTechniques 20, 882.

[xviii] Kimpton, C. P., Oldroyd, N. J., Watson, S. K., Frazier, R. R., Johnson, P. E., Millican, E. S., ... & Gill, P. (1996). Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. *Electrophoresis*, *17*(8), 1283-1293.

[xix] Sparkes, R., Kimpton, C., Watson, S., Oldroyd, N., Barnett, L., Arnold, J., ... & Clayton, T. (1996). The validation of a 7-locus multiplex STIR test for use in forensic casework: (I) Mixtures, ageing, degradation and species studies. *International Journal of Legal Medicine*, *109*, 186-194.

[xx] Budowle, B., Moretti, T. R., Keys, K. M., Koons, B. W., & Smerick, J. B. (1997). Validation studies of the CTT STR multiplex system. *Journal of forensic sciences*, *42*(4), 701-707.

[xxi] Entrala, C., Lorente, J. A., Lorente, M., Alvarez, J. C., Budowle, B., & Villanueva, E. (1998). Population Studies and Casework Application with the New GenePrint™ SilverSTR™ III Multiplex (D16S539, D7S820, D13S317). *Forensic Sci*, *42*, 701-707.

[xxii] Góes, A. C. D. S., Silva, D. A. D., Domingues, C. S., Marreiro Sobrinho, J., & Carvalho, E. F. D. (2002). Identification of a criminal by DNA typing in a rape case in Rio de Janeiro, Brazil. *Sao Paulo Medical Journal*, *120*, 77-79.

[xxiii] Wallin, J. M., Buoncristiani, M. R., Lazaruk, K. D., Fildes, N., Holt, C. L., & Walsh, P. S. (1998). TWGDAM validation of the AmpFISTR™ Blue PCR amplification kit for forensic casework analysis. *Journal of Forensic Science*, *43*(4), 854-870.

[xxiv] Krenke, B. E., Tereba, A., Anderson, S. J., Buel, E., Culhane, S., Finis, C. J., ... & Sprecher, C. J. (2002). Validation of a 16-locus fluorescent multiplex system. *Journal of Forensic Sciences*, *47*(4), 773-785.

[xxv] Collins, P. J., Hennessy, L. K., Leibelt, C. S., Roby, R. K., Reeder, D. J., & Foxall, P. A. (2004). Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFℓSTR® Identifiler® PCR amplification kit. *Journal of Forensic Science*, *49*(6), JFS2002195-13.

[xxvi] Mulero, J. J., Chang, C. W., Lagace, R. E., Wang, D. Y., Bas, J. L., McMahon, T. P., & Hennessy, L. K. (2008). Development and validation of the AmpFℓSTR® MiniFilerTM PCR Amplification Kit: a miniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. *Journal of forensic sciences*, *53*(4), 838-852.

[xxvii] Ensenberger, M. G., Thompson, J., Hill, B., Homick, K., Kearney, V., Mayntz-Press, K. A., ... & Krenke, B. E. (2010). Developmental validation of the PowerPlex® 16 HS System: an improved 16-locus fluorescent STR multiplex. *Forensic Science International: Genetics*, *4*(4), 257-264.

[xxviii] Oostdik, K., French, J., Yet, D., Smalling, B., Nolde, C., Vallone, P. M., ... & Sprecher, C. (2013). Developmental validation of the PowerPlex® 18D System, a rapid STR

multiplex for analysis of reference samples. *Forensic Science International: Genetics*, *7*(1), 129-135.

[xxix] Ensenberger, M. G., Hill, C. R., McLaren, R. S., Sprecher, C. J., & Storts, D. R. (2014). Developmental validation of the PowerPlex® 21 System. *Forensic Science International: Genetics*, *9*, 169-178.

[xxx] Martín, P., de Simón, L. F., Luque, G., Farfán, M. J., & Alonso, A. (2014). Improving DNA data exchange: validation studies on a single 6 dye STR kit with 24 loci. *Forensic Science International: Genetics*, *13*, 68-78.

[xxxi] Oostdik, K., Lenz, K., Nye, J., Schelling, K., Yet, D., Bruski, S., ... & Storts, D. R. (2014). Developmental validation of the PowerPlex® Fusion System for analysis of casework and reference samples: a 24-locus multiplex for new database standards. *Forensic Science International: Genetics*, *12*, 69-76.

[xxxii] https://www.thermofisher.com/order/catalog/product/4322288
Reviewed January 20, 2023

[xxxiii]https://www.promega.com/products/forensic-dna-analysis-ce/str-amplification/powerplex-16-system/?catNum=DC6531
Reviewed January 20, 2023

[xxxiv] https://www.thermofisher.com/order/catalog/product/4476135
Reviewed January 20, 2023

[xxxv]https://www.promega.com/products/forensic-dna-analysis-ce/str-amplification/powerplex-fusion-system/?catNum=DC2402
Reviewed January 20, 2023

[xxxvi] DOC, J. (1989). Crime Laboratory Digest. *Crime Laboratory Digest*, *16*, 27.
Reviewed March 20, 2023

[xxxvii] https://www.ojp.gov/pdffiles1/Digitization/132230NCJRS.pdf Reviewed March 20, 2023

[xxxviii] https://www.ojp.gov/pdffiles1/Digitization/143726NCJRS.pdf Reviewed March 20 2023

[xxxix] Technical Working Group on DNA Analysis Methods. Guidelines for a quality assurance program for DNA analysis. Crime Lab Digest 1995; 22(2):21–43

[xl] Micka, K. et al. (1996) "Validation of Multiplex Polymorphic STR Amplification Sets Developed for Personal Identification Applications," J. Forensic Sci. 41, 582.

[xli] https://scv10mr-cdnpre-p-cus-00.azureedge.net/-/media/files/resources/conference-proceedings/ishi-10/oral-presentations/04eisenberg.pdf?rev=46a12894f05d4a69928ae22481d6065b&sc_lang=en Reviewed March 20 2023

[xlii] https://oig.justice.gov/reports/FBI/a0632/intro.htm#:~:text=CODIS%20began%20as%20a%20pilot,NDIS%20became%20operational%20in%201998.
Reviewed March 20 2023

[xliii] Guideline, R. V. Scientific Working Group on DNA Analysis Methods (SWGDAM). *Forensic Science Communications*, *6*(3).
Reviewed March 20 2023

[xliv] Krenke, B. E., Tereba, A., Anderson, S. J., Buel, E., Culhane, S., Finis, C. J., ... & Sprecher, C. J. (2002). Validation of a 16-locus fluorescent multiplex system. Journal of Forensic Sciences, 47(4), 773-785.

[xlv] Urquhart, A., Oldroyd, N.J., Kimpton, C.P. and Gill, P. (1995) Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. BioTechniques 18: 116-121.

[xlvi] Murray, J. C., Buetow, K. H., Weber, J. L., Ludwigsen, S., Scherpbier-Heddema, T., Manion, F., ... & Cann, H. (1994). A comprehensive human linkage map with centimorgan density. Science, 265(5181), 2049-2054.

[xlvii] Promega Corporation. GenePrint PowerPlex 1.2 System Technical Manual, Part #TMD009 (revised 6/07). Madison, WI, 1998.

[xlviii] Moretti, T., & Budowle, B. (1998, February). The CODIS STR project: evaluation of fluorescent multiplex STR systems. In Proceedings of the 50th Annual Meeting of the American Academy of Forensic Sciences (pp. 9-14). San Francisco, CA: American Academy of Forensic Sciences.

[xlix] Budowle, B., & Moretti, T. R. (1998, October). Examples of STR population databases for CODIS and for casework. In 9th International Symposium on Human Identification (Vol. 1, pp. 64-73).

[l] National Institute of Standards and Technology. (2023, March 19). Strbase. Retrieved March 19, 2023, from https://strbase.nist.gov/index.htm

[li] Greenspoon, S. A., Ban, J. D., Pablo, L., Crouse, C. A., Kist, F. G., Tomsey, C. S., ... & Schumm, J. W. (2004). Validation and implementation of the PowerPlex® 16 BIO system STR multiplex for forensic casework. *Journal of forensic sciences*, *49*(1), 71-80.

[lii] Krenke, B. E., Tereba, A., Anderson, S. J., Buel, E., Culhane, S., Finis, C. J., ... & Sprecher, C. J. (2002). Validation of a 16-locus fluorescent multiplex system. *Journal of Forensic Sciences*, *47*(4), 773-785.

[liii] Brookes, C., Bright, J. A., Harbison, S., & Buckleton, J. (2012). Characterising stutter in forensic STR multiplexes. *Forensic Science International: Genetics*, *6*(1), 58-63.

[liv] Hill, C. R., Duewer, D. L., Kline, M. C., Coble, M. D., & Butler, J. M. (2013). US population data for 29 autosomal STR loci. Forensic Science International: Genetics, 7(3), e82-e83.

[lv] Green, R. L., Lagacé, R. E., Oldroyd, N. J., Hennessy, L. K., & Mulero, J. J. (2013). Developmental validation of the AmpFℓSTR® NGM SElect™ PCR Amplification Kit: A next-generation STR multiplex with the SE33 locus. *Forensic Science International: Genetics*, *7*(1), 41-51.

[lvi] Ensenberger, M. G., Hill, C. R., McLaren, R. S., Sprecher, C. J., & Storts, D. R. (2014). Developmental validation of the PowerPlex® 21 System. *Forensic Science International: Genetics*, *9*, 169-178.

[lvii] Collins, P. J., Hennessy, L. K., Leibelt, C. S., Roby, R. K., Reeder, D. J., & Foxall, P. A. (2004). Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFℓSTR® Identifiler® PCR amplification kit. *Journal of Forensic Science*, *49*(6), JFS2002195-13.

[lviii] Martín, P., de Simón, L. F., Luque, G., Farfán, M. J., & Alonso, A. (2014). Improving DNA data exchange: validation studies on a single 6 dye STR kit with 24 loci. Forensic Science International: Genetics, 13, 68-78.

[lix] Oostdik, K., Lenz, K., Nye, J., Schelling, K., Yet, D., Bruski, S., ... & Storts, D. R. (2014). Developmental validation of the PowerPlex® Fusion System for analysis of casework and reference samples: a 24-locus multiplex for new database standards. Forensic Science International: Genetics, 12, 69-76.

[lx] Luce, C., Montpetit, S., Gangitano, D., & O'Donnell, P. (2009). Validation of the AMPFℓSTR® MiniFilerTM PCR Amplification Kit for use in forensic casework. *Journal of forensic sciences*, *54*(5), 1046-1054.

APPENDIX

IRB APPROVAL FORM

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM
Office of the Institutional Review Board for Human Use

470 Administration Building
701 20th Street South
Birmingham, AL 35294-0104
205.934.3789 | Fax 205.934.1301 |
irb@uab.edu

APPROVAL LETTER

TO: Linville, Jason G

FROM: University of Alabama at Birmingham Institutional Review Board
Federalwide Assurance # FWA00005960
IORG Registration # IRB00000196 (IRB 01)
IORG Registration # IRB00000726 (IRB 02)
IORG Registration # IRB00012550 (IRB 03)

DATE: 16-May-2022

RE: IRB-300003083
IRB-300003083-003
Development and Validation of a STR Amplification Kit for Teaching and Research Purposes

---

The IRB reviewed and approved the Personnel Amendment submitted on 13-May-2022 for the above referenced project. The review was conducted in accordance with UAB's Assurance of Compliance approved by the Department of Health and Human Services.

Type of Review:	Expedited
Expedited Categories:	3,
Determination:	Approved
Approval Date:	16-May-2022
Expiration Date:	15-May-2025

Although annual continuing review is not required for this project, the principal investigator is still responsible for (1) obtaining IRB approval for any modifications before implementing those changes except when necessary to eliminate apparent immediate hazards to the subject, and (2) submitting reportable problems to the IRB. Please see the IRB Guidebook for more information on these topics.