Koehn, Alyssa. Identification of Unknown PCR Products Generated During STR Analysis of DNA from Bone Samples, Master of Science (Biomedical Science) Forensic Genetics. May, 2013, 34pp., 4 tables, 17 figures, references, 15 titles.

STR typing of DNA extracted from bone samples exhibited additional amplified products with the PowerPlex® 16 HS Kit (Promega Corp., Madison, WI) and the Identifiler® Plus Amplification kit (Life Technologies, Foster City, CA). Microbial DNA found in soil that is co-extracted and amplified with the human DNA from bones may be the source of the artifacts seen during the STR analysis of the DNA samples. This project tested the hypothesis that the additional PCR products found in the STR analysis of DNA derived from bone samples are microbial or fungal in origin. This study has demonstrated that amplification of soil samples with an STR kit can produce artifacts, similar to those seen in the bone samples. A protocol was developed to isolate the artifact products. Attempts were made to Sequence the amplified PCR product; however, no sequencing results have been obtained, further experiments continue to attempt to generate sequencing results.

# IDENTIFICATION OF UKNOWN PCR PRODUCTS GENERATED DURING STR ANALYSIS OF DNA FROM BONE SAMPLES

THESIS

Presented to the Graduate Council of the

Graduate School of Biomedical Sciences

University of North Texas

Health Science Center at Fort Worth

Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Alyssa Koehn, B.S.

Fort Worth, Texas

May 2013

I would like to thank the University of North Texas Center for Human Identification for allowing me to perform my research project in their laboratory. I would especially like to thank my Major Professor and Mentor Dr. Bruce Budowle for all of his guidance and patience throughout this project. I would also like to thank Dr. John Planz for his continued support and thoughtful input throughout the writing process. My sincere thanks go to Dr. Harlan Jones and Dr. Claire Kirchoff for being a part of my thesis committee.

I would like to thank Dixie Peters and Mark Ingram at the UNT Center for Human Identification for providing samples and materials in order for me to complete my project. My utmost gratitude goes to Jonathan King for always taking the time to brainstorm, I would not have been able to accomplish what I did without his advice and guidance in the laboratory.

I am grateful for all of my classmates and lab mates for all of their support, encouragement and laughs over the last two years. My sincere thanks goes to my family and friends for being the greatest support system I have. Their love and faith in me has allowed me to continue to pursue my educational goals.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

# CHAPTER I

## INTRODUCTION

Over 99.7% of DNA in the human genome is the same between individuals. In human identification it is important to focus on markers in the genome that potentially allow for individualization. The most widely used forensic genetic markers are short tandem repeat (STR) loci. STR markers make up about 3% of the human genome and occur on average about every 10,000 nucleotides (1-4). STRs are repeated DNA sequences that vary in the number of repeats at the specific loci among individuals. The variation at a STR locus is especially important in forensic DNA typing to distinguish among potential contributors of a forensic evidence sample. An advantage of the use of STR markers is their small size, between 100-400 base pairs. The small size of the PCR products produced from STRs allows them to be amplified efficiently and to be easily combined into a multiplex assay. Multiplexing allows many STR markers to be amplified at the same time in the same reaction. The combination of the STRs in the multiplex kit provides a high discriminating power among individuals. Additionally, a smaller amount of sample is consumed in the one multiplex reaction than would be by performing many amplification reactions, and the number of manipulations is reduced.

Each multiplex STR kit must undergo validation before introduction into a forensic laboratory. The purpose of a validation study is to determine the efficacy and reliability of a new technique or technology (QAS 2009) (5). Validation is a critical part of forensic DNA typing; it confirms that a method is acceptable for its intended use,

defines the limitations of the assay and offers confidence in the STR profile results. A developmental validation has many components, which include, but are not limited to characterization of the STR markers, species specificity, sensitivity studies, reproducibility studies, population studies, and evaluation of case type samples (SWGDAM 2004) (6). STR typing with the PowerPlex® 16 HS (Promega Corp., Madison, WI) and Identifiler® Plus Amplification kits (Life Technologies, Foster City, CA) has been validated showing that the methodologies supporting these kits are robust and reliable. There are a wide range of samples that can be analyzed ranging from bones, hair, nails, and biological fluids such as blood, semen, and saliva.

STR analysis on DNA derived from bones is particularly useful in missing persons and unidentified human remains cases. The small size of STR fragments allows analysis of degraded DNA samples. The STR profiles obtained from bone samples can be compared to STR profiles of close relatives by performing kinship analysis and potentially identify the human remains.

During a previous bone study by Blake Myers of the Institute of Applied Genetics, amplification products were observed that were not attributable to human STR alleles during analysis of DNA from several bone samples. Initially, these additional off ladder artifacts were seen when performing STR analysis with the PowerPlex® 16 HS System (Figure 1). The same DNA extracts were re-amplified using the AmpflSTR Identifiler® Plus PCR Amplification Kit to determine if these unexplained amplification products were detectable with another STR amplification kit. No artifacts were observed for any of the samples when amplified with the Identifiler® Plus PCR Amplification kit (Figure 2)

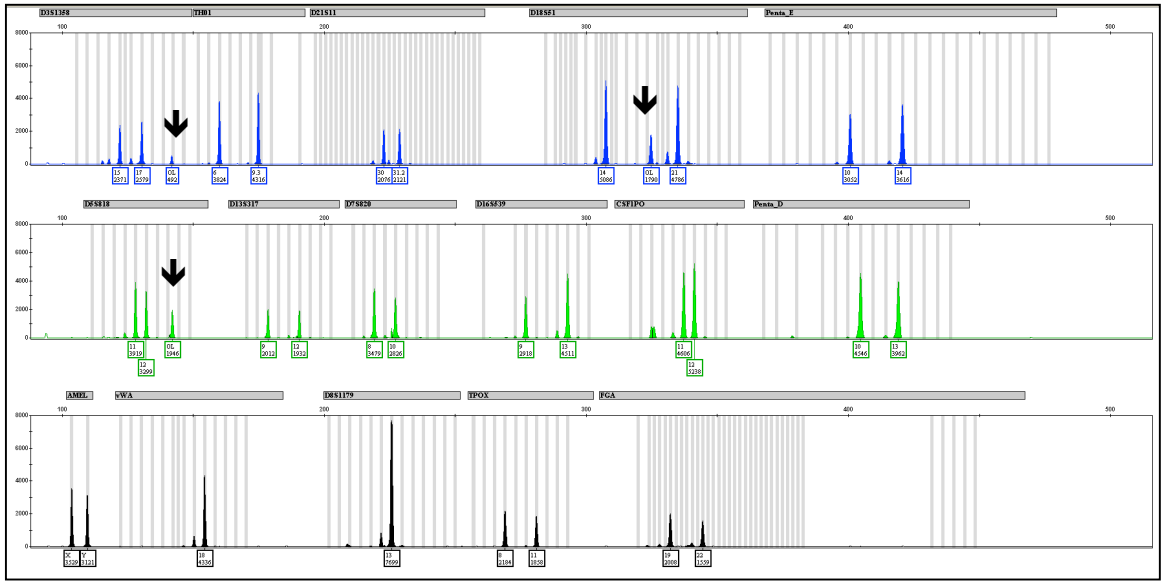

Figure 1: Sample 0019-12 electropherogram; off ladder alleles (indicated by the arrows) observed with PowerPlex® 16 HS; 1ng total input human DNA.

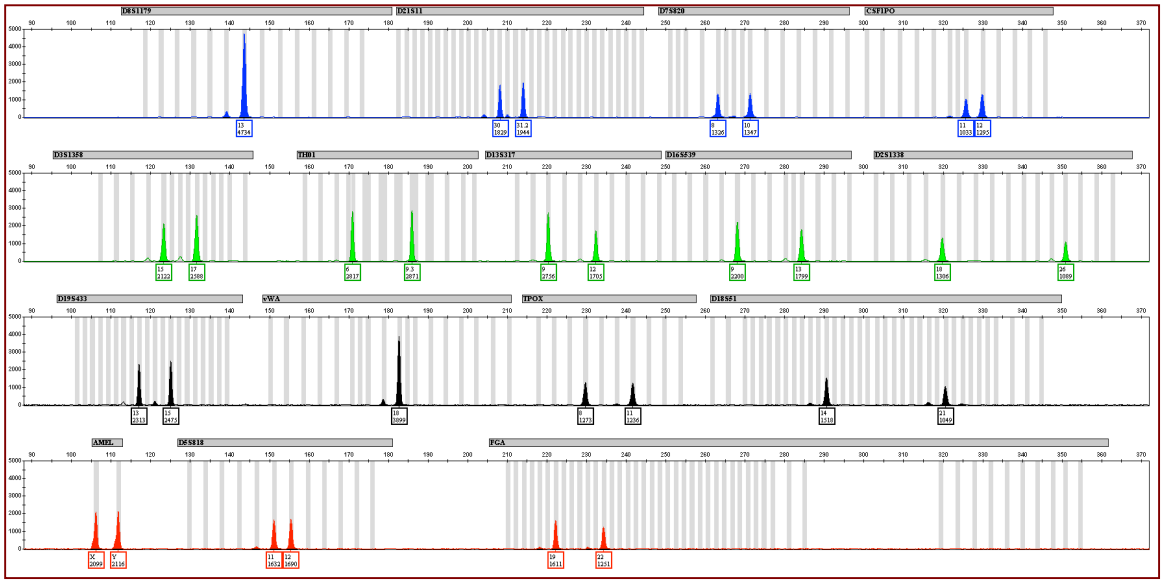

Figure 2: Sample 0019-12 electropherogram; no off ladder alleles with Identifiler® Plus; 1ng total input human DNA

Manipulations were carried out to reduce or eliminate the presence of these off ladder artifacts. A spectral calibration was performed to determine how much fluorescent signal overlap would be expected by the dyes in a particular instrument under its conditions of use. If the spectral calibration was not working properly to separate the color overlap of the dyes, pull-up or an uneven baseline signal can occur. This spectral overlap could contribute to the appearance of the additional peaks in the electropherograms. However, when a new four dye spectral calibration was performed there were no observed differences in the additional PCR products (Figure 3). In addition, the peaks did not migrate to the same position of peaks in other dye channels further supporting that pull up was not a viable explanation for the artifacts.

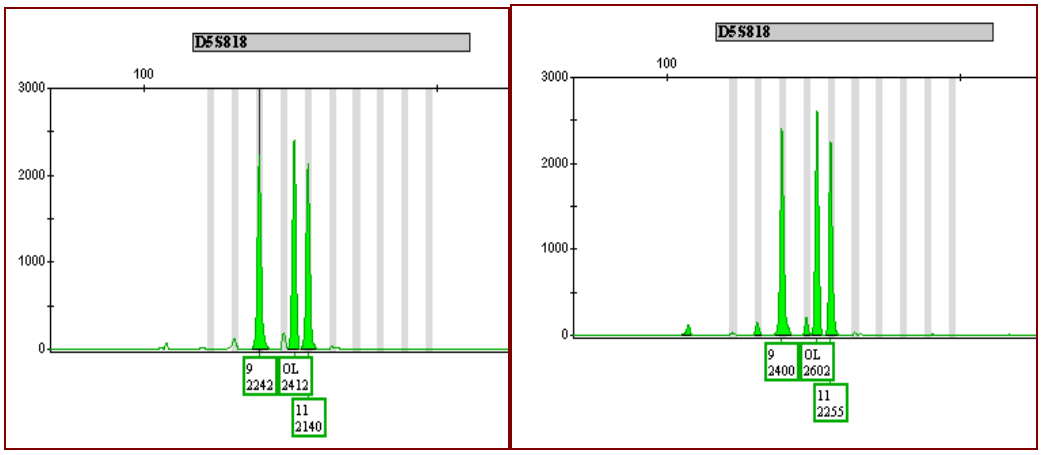


Figure 3: Left: the D5S818 locus generated with standard PowerPlex® 16 HS amplification protocol; Right: the D5S818 locus of same sample with the new four dye spectral calibration. The peak indicated OL is the artifact seen in this sample.

Additional off ladder artifacts also were observed by analysts at the UNTHSC Center for Human Identification Missing Person's Unit when performing STR analysis on bone samples using the Identifiler® Plus Amplification kit (Figure 4). The same

DNA extracts demonstrating artifacts in the Identifiler® Plus Amplification were run using the PowerPlex® 16 HS kit to determine if these products were present with the PowerPlex® 16 HS kit. The products that were seen during amplification with the Identifiler® Plus Amplification kit were not observed using the PowerPlex® 16 HS kit (Figure 5).

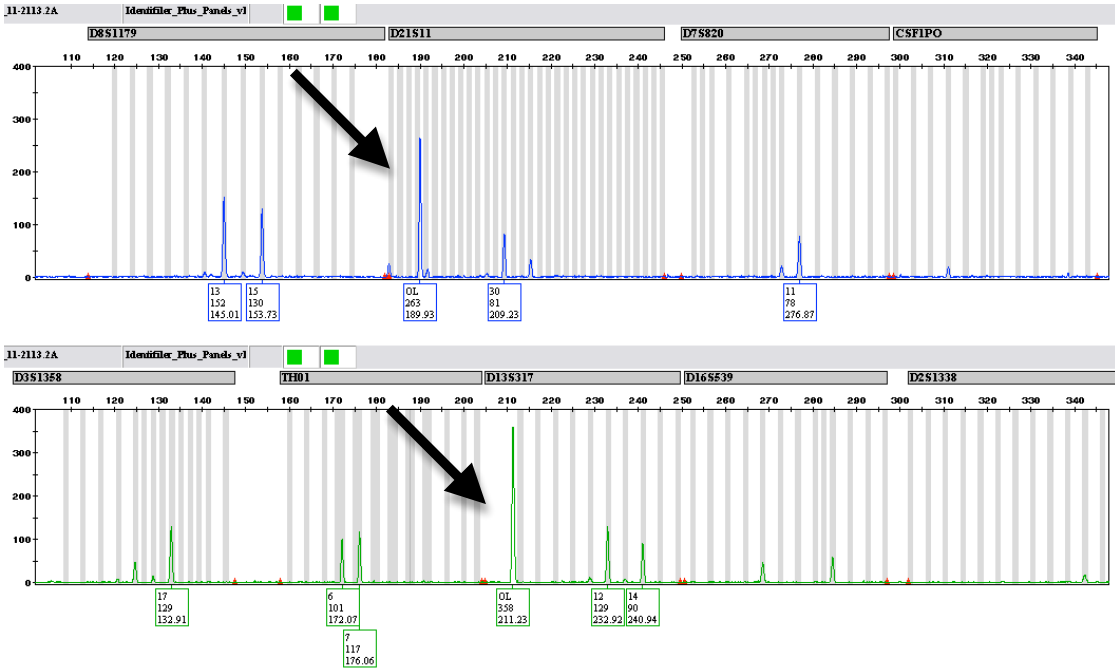


Figure 4: Sample A electropherogram displaying off ladder peaks (indicated by the arrows) with Identifiler® Plus Amplification Kit.

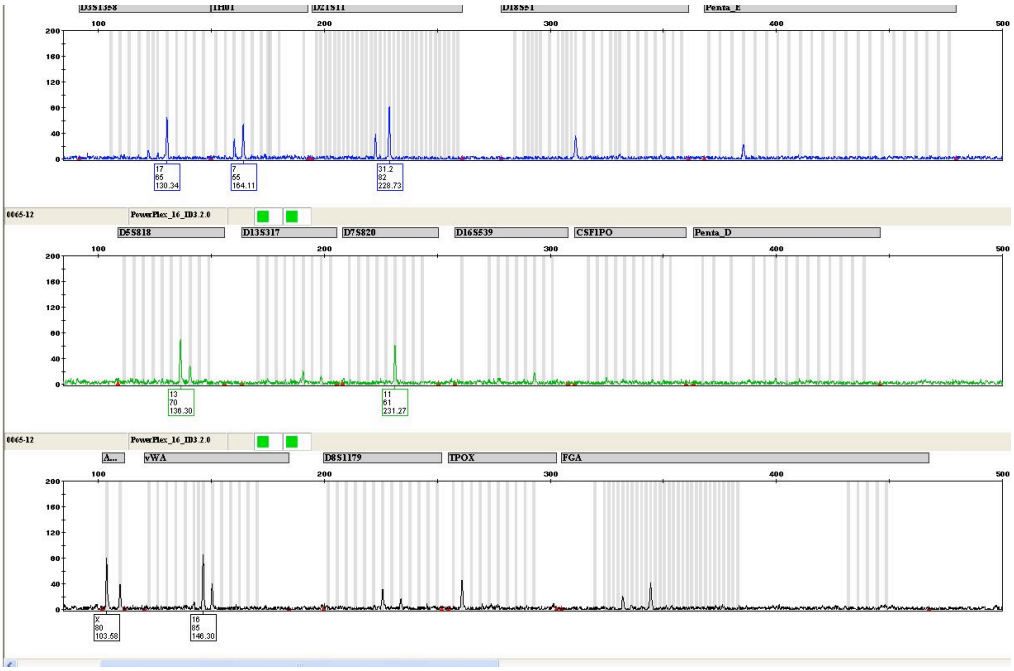

Figure 5: Sample A electropherogram displaying no off ladder peaks with PowerPlex® 16 HS

In forensic cases, bone samples often are collected from soil. The additional amplification products seen from the bone samples may originate from microbial or fungal DNA that resides in the surrounding soil. Although intended for human DNA some of the primers in the PowerPlex® 16 HS Amplification kit and Identifiler® Plus Amplification kit may cross-react with DNA from microbial or fungal species. Soil samples were collected by Blake Myers to determine if they too yielded products that were not consistent with human STR alleles. Soil samples from three different locations were collected and extracted in the same manner as the bone samples and amplified using both the PowerPlex® 16 HS and Identifiler® Plus kits. When the samples were amplified using PowerPlex® 16 HS kit artifact products were observed (Figure 6). However, when the samples were amplified using Identifiler® Plus no artifacts were observed (Figure 7). Although the peaks were in different positions on the

electropherogram, the results were similar with those from the bone analysis; i.e. the reagents in the PowerPlex® 16 HS kit generated artifacts with these samples and the Identifiler® Plus kit reagents did not. A plausible explanation for the artifacts is that microbial DNA, which is abundant in soil, is the origin of the additional PCR products generated using the kits.

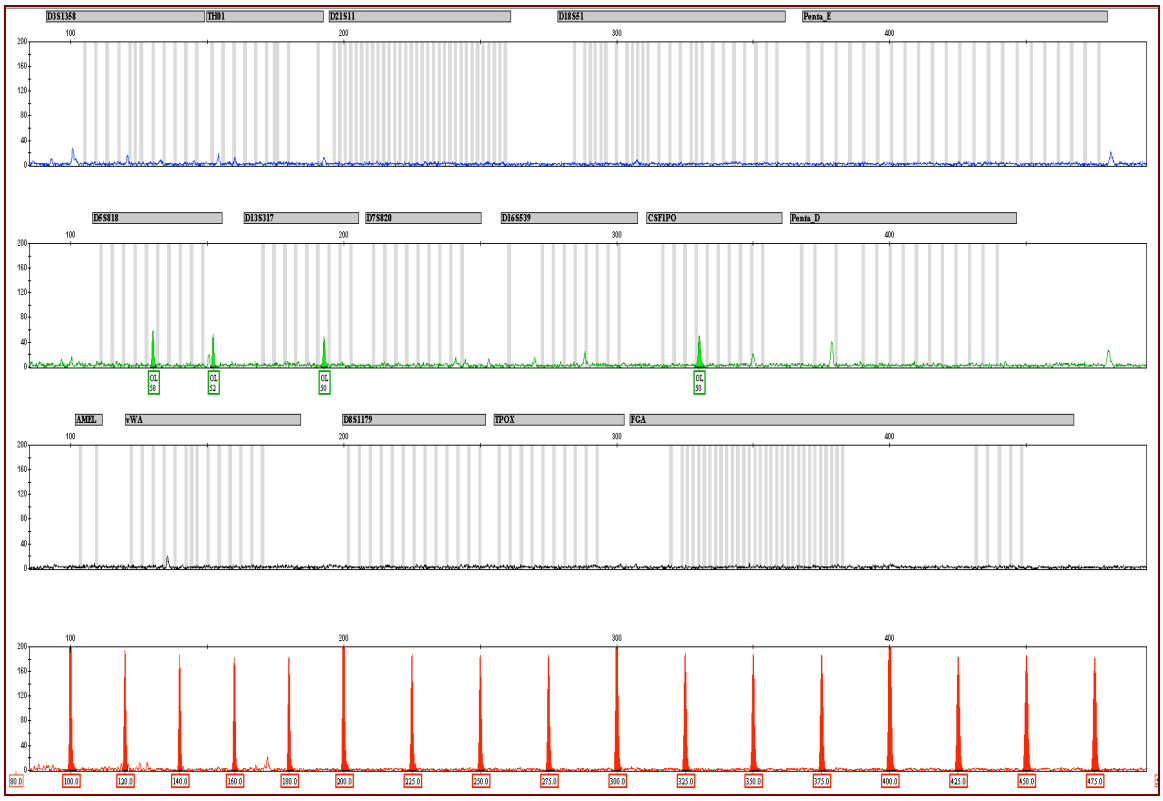


Figure 6: Electropherogram of extract from soil from duck pond, amplified with PowerPlex® 16 HS. Artifact products present.

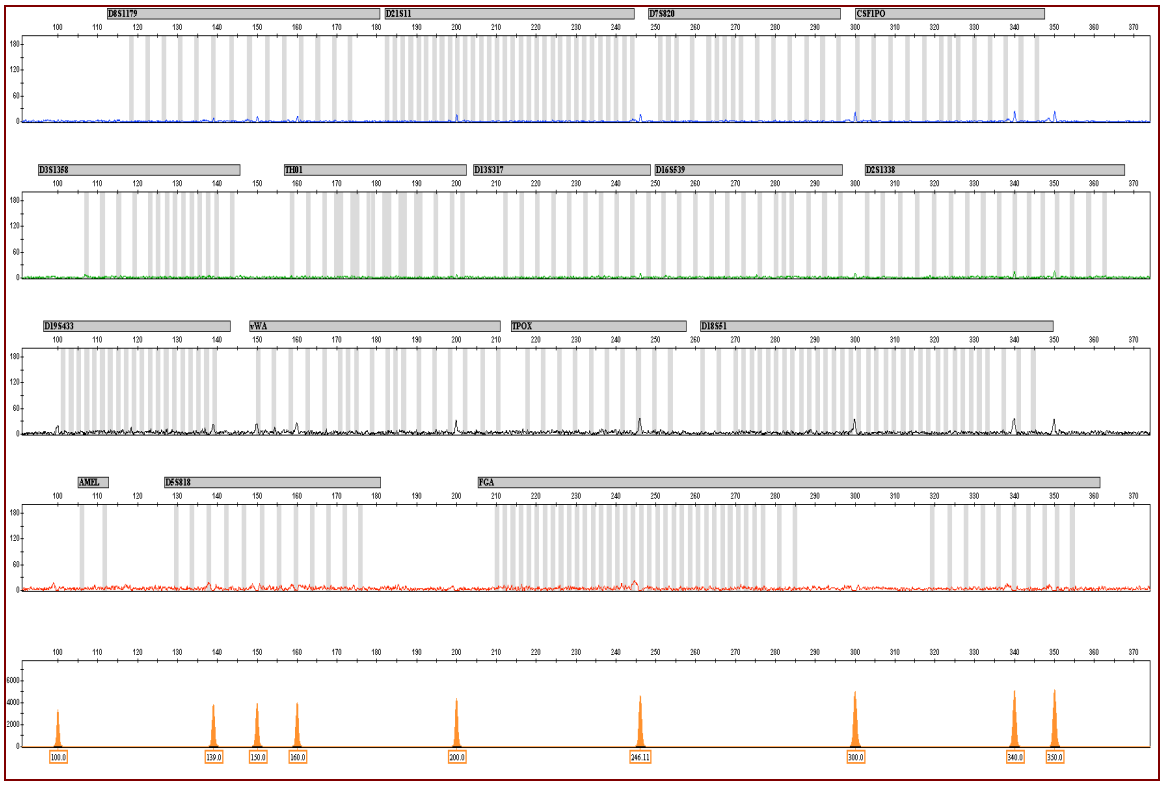


Figure 7: Electropherogram of extract from soil from duck pond, amplified with Identifiler® Plus. No artifact products observed.

Soil is a complex and diverse environment containing more than $10^9$ microbial cells per gram (7). Bone samples often are found buried in the soil and thus may be contaminated with DNA from microbial or fungal microorganisms. Standard human DNA extraction procedures cannot differentiate between human DNA and DNA from microorganisms and therefore these molecules will co-extract (8). Both PowerPlex® 16 HS System and Identifiler® Plus amplification kits were found to be only primate specific during developmental validation studies (9, 10). Neither kit was found to produce detectable profiles for a variety of microorganisms (9, 10). However, microbial cross-reaction studies typically test only a few homogeneous samples and cannot possibly be tested for the vast diversity of microorganisms.

Potential non-human artifacts were detected in STR profiles amplified with the PowerPlex® 16 HS kit (Promega Corp.) when DNA extracts from bone samples and independently collected soil samples were tested. The same DNA samples did not yield any artifacts using the Identifiler® Plus Amplification Kit (Life Technologies). In addition, other artifacts were detected in STR profiles of DNA samples amplified using the Identifiler® Plus Amplification kit (Life Technologies) that were not seen when the same samples were amplified using the PowerPlex® 16 HS kit (Promega Corp.). Microbial DNA that is co-extracted and amplified with the human DNA may be the source of the artifacts seen during the STR analysis of DNA from the bone and soil samples.

The successful STR typing of DNA from bone samples can be difficult due to degradation, contamination, and the limited quantity of nuclear DNA (11-13). Understanding the specificity of DNA typing of biological evidence in any forensic case is critical to ensure reliability of typing results and define the limitations of interpretation of the profiles. Bone and teeth DNA extract samples have been found to contain mixed amounts of Human DNA and microbial DNA (14). The presence of microbial or fungal DNA that can be amplified by PowerPlex® 16 HS (Promega Corp.) or Identifiler® Plus amplification kits (Life Technologies) may compete with the nuclear human DNA during the amplification reaction. The PCR conditions could lead to the preferential amplification of the microbial or fungal DNA template, producing a poor STR profile even if artifact peaks from the microbial or fungal DNA fragments do not reside at known STR allele positions. In some instances the presence of non-specific products may be incorrectly attributed as human PCR products and possibly confound interpretation. A

wrong profile or compromised profile may direct investigators to falsely associate human remains to a victim or be unable to associate the remains with a victim.

The hypothesis I will be testing is that the additional PCR products found in the STR analysis of DNA derived from bone samples using PowerPlex® 16 HS and Identifiler® Plus kits are microbial or fungal in origin. To test this hypothesis, the additional PCR products from the bone samples will be isolated. Then the isolated PCR products will be sequenced. The DNA sequences obtained will be compared with known DNA sequences in the NCBI database to determine the possible origin of the DNA from which these PCR products were generated.

# CHAPTER II

# MATERIALS AND METHODS

## *Sample Collection*

Bone fragments were provided by Dixie Peters, Technical leader, from the UNTHSC Center for Human Identification and Mark Ingram from the UNT Denton Anthropology Lab. DNA extracts and PCR product from bone samples also were provided by Dixie Peters. A total of twenty soil samples were collected. Soil samples were collected from several outdoor locations (a levee, duck pond and stairway) where additional PCR products had previsoulsy been observed. Soil samples were also collected in new locations along the Trinity River and Downtown areas of Fort Worth, Texas.

## *Sample Analysis*

DNA from Bone and soil samples was extracted using the Hi-Flow® DNA Purification Spin Columns, a silica-based device (Generon L.L.C., Maidenhead, UK). Two elutions were collected from each sample using the Hi-Flow® DNA Purification Spin Columns. Reagent blanks and negative controls were run with the soil and bone samples. The samples were purified further and concentrated using the DNA IQ™ Casework Pro kit (Promega Corp.) with the Maxwell® 16 (Promega Corp.). The quantity of human DNA was determined using the Quantifiler™ Human DNA Quantification Kit (Life Technologies). All DNA samples were amplified using the PowerPlex® 16 HS Amplification Kit and AmpflSTR Identifiler® Plus Amplification

Kit. Human DNA (1 ng) (or 10μL) was added into the PCR reaction for amplification. The reagent blank as well as the negative control were carried through the amplification process. A positive control also was amplified along with the samples. Amplification was performed on a GeneAmp® PCR System 9700 (Life Technologies) according to manufacturer's recommendations unless otherwise stated. STR typing was performed on a 3130*xl* Genetic Analyzer capillary electrophoresis system (Life Technologies) according to the manufacturer's recommendations. Data analysis was performed using GeneMapper™ ID Software (Life Technologies). Samples showing additional PCR products during STR analysis were carried through to further experiments.

*Analysis of Known DNA Sample*

Two known buccal swab DNA samples donated from the Institute of Applied Genetics, following the IRB protocol were sequenced at two homozygous loci. Sample 1 was homozygous at D13S317 and D8S1179. Sample 2 was homozygous at D3S1358 and TPOX. The DNA extracts were amplified using the Identifiler® Plus Amplification Kit to confirm the homozygous peaks. The DNA extracts were then sequenced using STR primers that were kindly provided by Life Technologies and using the Big Dye® Terminator Kit v3.1 Cycle Sequencing Kit (Life Technologies) according to the manufacturer's recommendations. Sequencing was performed on the 3130*xl* Genetic Analyzer capillary electrophoresis system. Positive and negative controls were used during the amplification and sequencing analysis. Data analysis was performed using Sequence Scanner Software (Life Technologies). The sequence data were then compared to the known STR sequences and repeats numbers from the STR data previously analyzed (15).

*Analysis of Additional PCR Products*

Samples with additional PCR products were amplified for a second time to produce more PCR product, with the Identifiler® Plus amplification kit according to the manufacturer's recommendations. The additional PCR products were isolated using the Pippin Prep™ DNA size selection system with the 2% agarose ethidium bromide cassette (Sage Science, Beverly, MA). The collection of the isolated PCR products was confirmed using the Agilent 2200 Tapestation System (Agilent Technologies, Santa Clara, CA). The isolated DNA fragments then were amplified once again with the Identifiler Plus® Amplification kit under standard conditions. The PCR products were run on a 3130*xl* Genetic Analyzer capillary electrophoresis system to visualize the single PCR product and produce more of the isolated DNA fragment.

DNA fragments were visualized on an agarose gel with GelRed™ Nucleic Acid Stain (Phoenix Research Products, Candler, NC). The stained GelRed™ bands produced by the DNA fragments were isolated by excising them from the agarose gel. The excised bands were purified using the PureLink® Quick Gel Extraction Kit (Life Technolgies). The isolated DNA fragments were amplified using the Identifiler® Plus amplification kit and run on a 3130*xl* Genetic Analyzer capillary electrophoresis system to visualize the single PCR product. Positive and negative controls were used for each amplification step.

*Sequencing DNA fragments*

The DNA fragments were sequenced using STR primers that were kindly provided by Life Technologies and using the Big Dye® Terminator Kit v3.1 Cycle Sequencing Kit according to the manufacturer's recommendations. Sequencing was performed on the 3130*xl* Genetic Analyzer capillary electrophoresis system. Positive and Negative

controls were used during the sequencing analysis. Data analysis was performed using Sequence Scanner Software. The sequences obtained were compared against known DNA sequences in the NCBI Database (http://www.ncbi.nlm.nih.gov/) using the Basic Local Alignment Search Tool (BLAST).

# CHAPTER III

# RESULTS

## *Soil Samples*

Soil samples were collected and analyzed to determine if similar PCR products could be detected as those seen in the soil samples previously tested. Soil samples were collected from the same locations as the soil samples originally tested (duck pond, levee, stairway) as well as other locations around Fort Worth, Texas including along the Trinity River trails and downtown areas. PCR product was found in eight of the twenty locations where soil samples were collected (Figure 8). However, PCR product was not found in the soil samples that were collected from the same areas as those that had previously been tested.


Figure 8: Electropherogram of soil sample S7E1 amplified with PowerPlex® 16 HS, artifact products present.

In each soil sample all the peaks from the PCR products were seen in the green dye channel when amplified with the PowerPlex 16 HS kit (Table 1). Some of the soil samples, S5, S10, S11, and S16, had PCR product present in both the first elution and second elution of the DNA extraction (Table 1). All of these soil samples were from different locations.

| Amplification of Soil Samples with PowerPlex® 16 HS | | | |
|---|---|---|---|
| **Sample** | **Fragment Size (bp)** | **Peak Height (RFUs)** | **Locus Area** |
| S5E1 | 150 | 50 | D5S818 |
| S7E1 | 150 | 64 | D5S818 |
| S7E1 | 250 | 112 | D7S820 |
| S7E1 | 375 | 2359 | Penta D |
| S10E1 | 200 | 182 | D13S317 |
| S11E1 | 150 | 875 | D5S818 |
| S12E1 | 150 | 284 | D5S818 |
| S12 E1 | 250 | 114 | D7S820 |
| S13E1 | 150 | 155 | D5S818 |
| S13E1 | 250 | 84 | D7S820 |
| S16E1 | 150 | 74 | D5S818 |
| S4E2 | 200 | 54 | D13S317 |
| S5E2 | 150 | 108 | D5S818 |
| S10E2 | 200 | 89 | D13S317 |
| S11E2 | 150 | 820 | D5S818 |
| S16E2 | 150 | 52 | D5S818 |

Table 1: Artifact products present in soil samples amplified with PowerPlex® 16HS. Sample nomenclature: S refers to sample number; each sample was collected from a different location. E refers to the elution number of that sample from the DNA extraction process.

Three of the soil samples extracts (S7E1, S10E1, and S11E1) that showed PCR products were amplified with the Identifiler® Plus Amplification Kit to determine if the PCR products would be generated. The samples were analyzed on the 3130xl for STR analysis. Two of the three soil samples did not show any PCR product (Figure 9). S10E1 showed one peak; however this observation may have been due to contamination from the positive control. No PCR products were observed in any soil samples tested, except for S10E1, when amplified with the Identifiler Plus Amplification Kit.

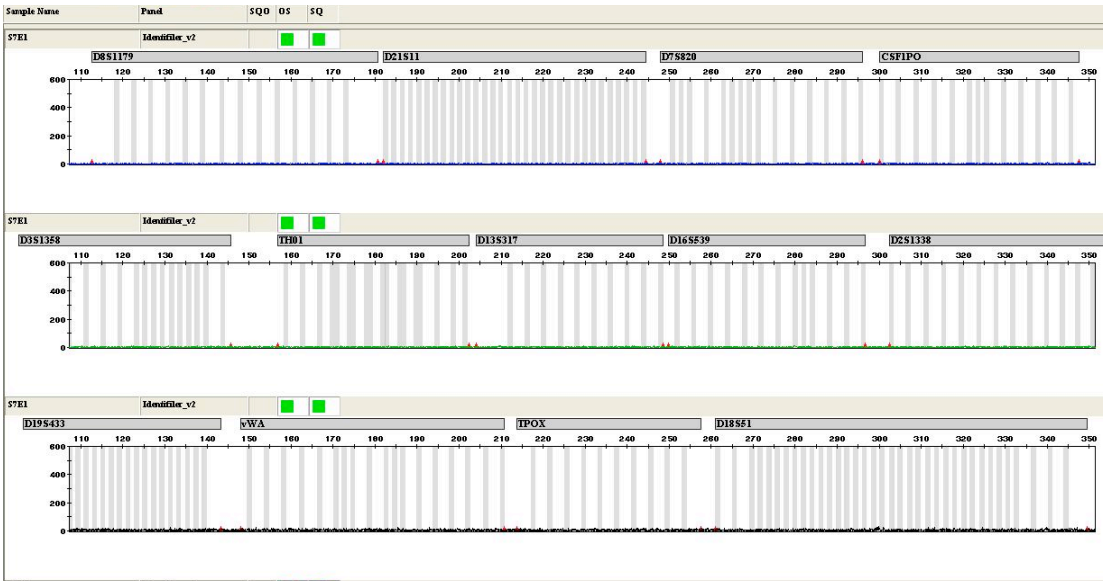


Figure 9: Electropherogram of soil sample S7E1 amplified with the Identifiler® Plus Amplification Kit, no artifact products present.

*Sequencing Known Samples*

Samples with known DNA profiles were used to ensure that a procedure was in place that could successfully sequence PCR products derived from the profile. Each of these known samples was homozygous at two specific loci when amplified with the Identifiler® Plus Amplification kit. Sample 1 was homozygous at the loci D13S317 and D8S1179. Sample 2 was homozygous at the loci D3S1358 and TPOX. Without any isolation, the PCR products were amplified for sequencing with the specific primer pairs for each locus. The amplification of only those DNA fragments in each reaction was confirmed using the Agilent 2200 Tapestation. Then the confirmed PCR products were sequenced. Sequencing results were obtained for these samples and analyzed using the Sequence Scanner software. The sequences obtained from the samples were concordant with the known sequences of the STR regions (15). The number of repeats in the

sequence data was concordant with the number of repeats obtained from the STR analysis with the Identifiler® Plus kit. For example, at the D13S317 locus, the STR results for known DNA sample 1 showed this sample to be homozygous with a 12 allele. The sequencing results for the D13S317 locus also showed 12 repeats (Table 2).

| | |
|---|---|
| Reference Sequence | ATcACAGAAGTCTGGGATGTGGAGGAgagttcatttctttagtgggcatccgtgactctctggac tctgacccatctaacgcctatctgtatttacaaatacattatctatctatctatctatctatctatctatct atctatcaatcaatcatctatctatctttctgTCTGTCTTTTTGGGCTGCC |
| Sequence from Forward Primer | TacaTTAtcTAtcTAtcTAtcTAtcTAtcTAtcTAtcTAtcTAtcTAtcAAtcAAtcAtcTAt cTAtcTTtcTGtcTGtcTTTTTGGGCTGCCTATGGCtcAACCCAAgTTGAAGgaGAGATT |
| Sequence from Reverse Primer | ATAGATAGATGATTGATTGATAGATAGATAGATAGATAGATAGATAGATAGATA GATAGATAGAtAATGTATTTGTAAATACAGATAGGCGTTAGATGGGtcAGAGtcCA GAGAGtcACGGATGCCCACTAAAGAAATGAACTC |

Table 2: Sequencing results for the D13S317 locus. The yellow area indicates where the reference sequence and obtained sequences were concordant. The green area indicates the repeat units.

The STR results for known DNA sample 2 at the TPOX locus showed this sample to be homozygous with an 8 allele. The sequencing results for the TPOX locus also showed 8 repeats for this sample (Table 3).

| | |
|---|---|
| Reference Sequence | GCACAGAACAGGCACTTAGGgaaccctcactgaatgaatgaatgaatgaatgaatgaatgaatgtttgggcaaataaacgctgacaaggacagaagggcctagcgggaagggaacaggagtaagaccagcgcacagcccgacttgtgttcagaagacctgggattggacctgaggagttcaatttttggatgaatctcttaattaacctgtggggttcccagttcctcccctgagcgcccaggacagtagagtCAACCTCACGTTTGAGCG |
| Sequence from Forward Primer | GAAGGGCCTAGCGGGAAGGGAacAGGAGTAAGACCAGCGCACAGCCCGACTTGTGTtcAGAAGACCTGGGATTGGACcTGAGGAGTtcAATTTTGGATGAAtcTCTTAATTAACCTGTGggGTtcCCAGTtcC |
| Sequence from Reverse Primer | TgTTccCTTccCgcTaGGCCTtCTgtCCTtGtcAgcGTTATTTCCCAACATtcATtCATtCATtcATtCATtCATtCATtCAGtgAGGGTtcCC |

Table 3: Sequencing results for the TPOX locus. The yellow area indicates where the reference sequence and obtained sequences were concordant. The green area indicates the repeat units.

*Isolation of Additional PCR Product in DNA samples from Bone*

The first step to identify the additional PCR products present in the DNA samples collected from bones was to isolate the DNA fragments. The PCR product from bone samples were re-amplified using the Identifiler® Plus Amplification Kit to generate more product. The samples were run on a 3130xl Genetic Analyzer capillary electrophoresis system to confirm the presence of the additional PCR product. The presence of the additional PCR product was consistent (by position on the electropherogram) with the previous STR analysis that was performed on these samples.

Amplification product from bone samples 66 and 70 were chosen to isolate the DNA fragments observed using the Pippin Prep™ DNA size selection system. Each sample was run twice to capture the peak in the blue dye channel as well as the peak in the green dye channel. In both DNA samples the peak of interest in the blue dye channel was approximately 188 bp and the peak of interest in the green dye channel was approximately 211 bp in length. These base pair measurements were relative lengths

based on STR analysis data on the 3130xl Genetic Analyzer. A "tight" range program was used to isolate the peaks of interest. The "tight" range program attempts to capture DNA fragments within a forty base pair range. The ranges were chosen based on the base pair size of each DNA fragment and the DNA fragments that were surrounding the DNA fragment of interest (Table 4).

| Sample | Dye channel | DNA fragment Size (bp) | tight range program |
|---|---|---|---|
| 66 | Blue | 188 bp | 166 bp – 194 bp |
| | Green | 211 bp | 181 bp – 213 bp |
| 70 | Blue | 188 bp | 166 bp - 194 bp |
| | Green | 211bp | 202 bp – 238 bp |

Table 4: Tight Programs used for the Pippin Prep™ isolation of DNA fragments.

After recovering the samples from the Pippin Prep™ size selection system the samples were run on the Agilent 2200 TapeStation to confirm that the correct size fragment was captured (Figure 10). The size fragment that was in the particular sample was captured. In lane B1, the lower and higher size markers ran further when compared to the other lanes, and therefore the band from the DNA fragment also ran further than expected.

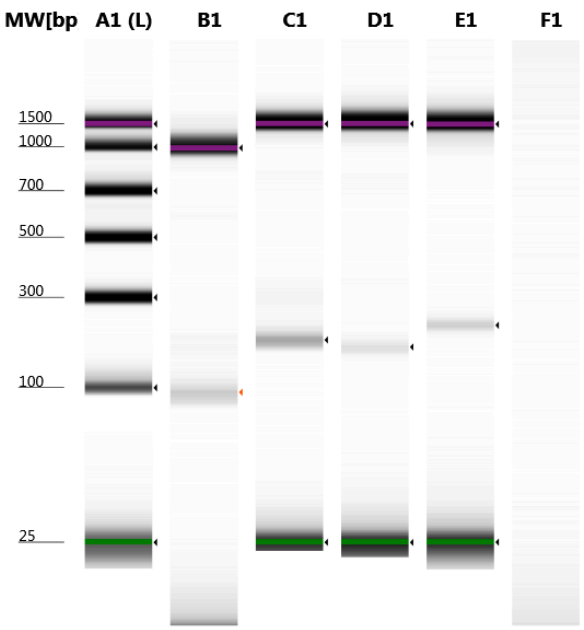


Figure 10: Agilent 2200 TapeStation (Agilent Technologies) results of DNA fragments captured from the Pippin Prep™ (Sage Science). Lane A1- Ladder, Lane B1- sample 66 blue dye channel tight range program, Lane C1- sample 66 green dye channel tight range program, Lane D1- sample 70 blue dye channel tight range program, Lane E1- sample 70 green dye channel tight range program, Lane F1- negative control.

The samples captured from the Pippin Prep™ then were amplified with the Identifiler® Plus Amplification Kit and run on a 3130xl Genetic Analyzer to confirm through STR analysis that only the peak of interest was captured. Although the tight range programs on the Pippin Prep™ should have elimanated any peaks that were outside the range of the program there were still peaks present in the sample outside that specified range (Figure 11 and Figure 12)

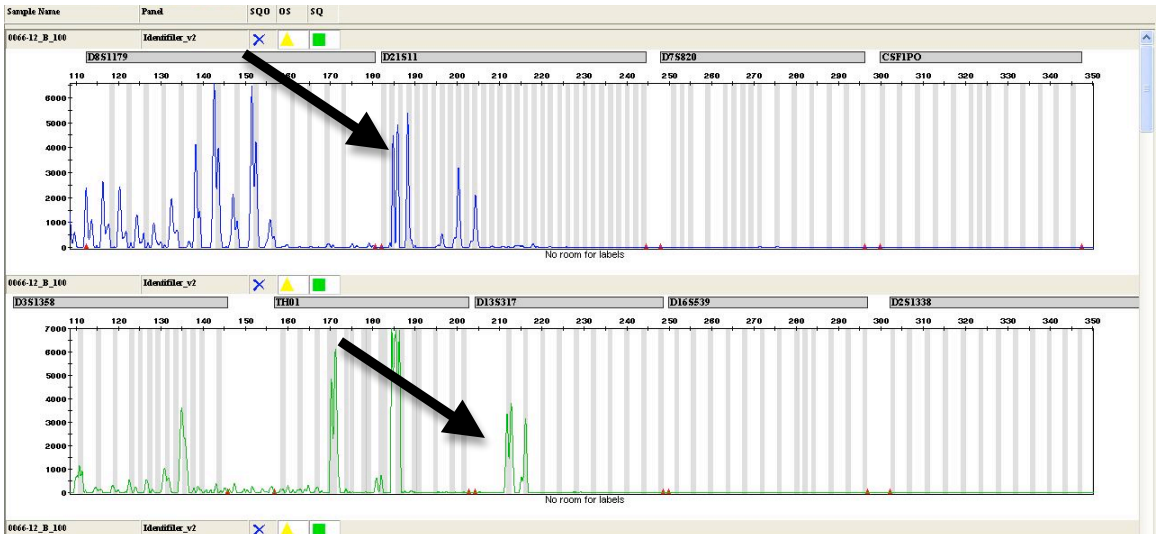


Figure 11: Sample 66 Amplified with Identifiler® Plus Amplification Kit after isolation with the Pippin Prep™. The arrows indicate the desired peaks in each dye channel.

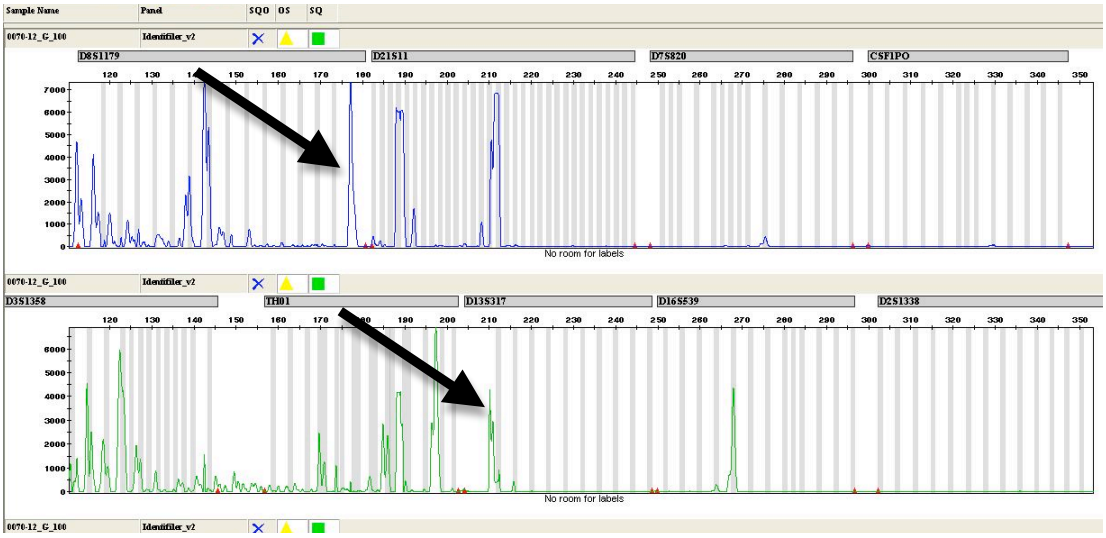


Figure 12: Sample 70 Amplified with Identifiler® Plus Amplification Kit after isolation with the Pippin Prep™. The arrows indicate the desired peaks in each dye channel.

To remove the DNA fragments outside the desired range, the samples were separated again on a 1.5% agarose gel, visualized and the bands were excised and purified. The purified DNA then was amplified with the Identifiler® Plus Amplification Kit and analyzed to determine if the DNA fragment of interest had been isolated and spurious products were eliminated. Although there were still some peaks present in addition to the peaks of interest, the samples did appear to be cleaner, with fewer peaks (particular higher molecular products) than solely by isolation with the Pippin Prep™ (Figure 13 and Figure 14). Only some DNA fragments smaller in size than the peaks of interest remained. The smaller DNA fragments resembled an allelic ladder with a peak at nearly each bin of the D8S1179 Locus in the blue dye channel and the D3S1358 locus in the green dye channel (Figure 13 and Figure 14).

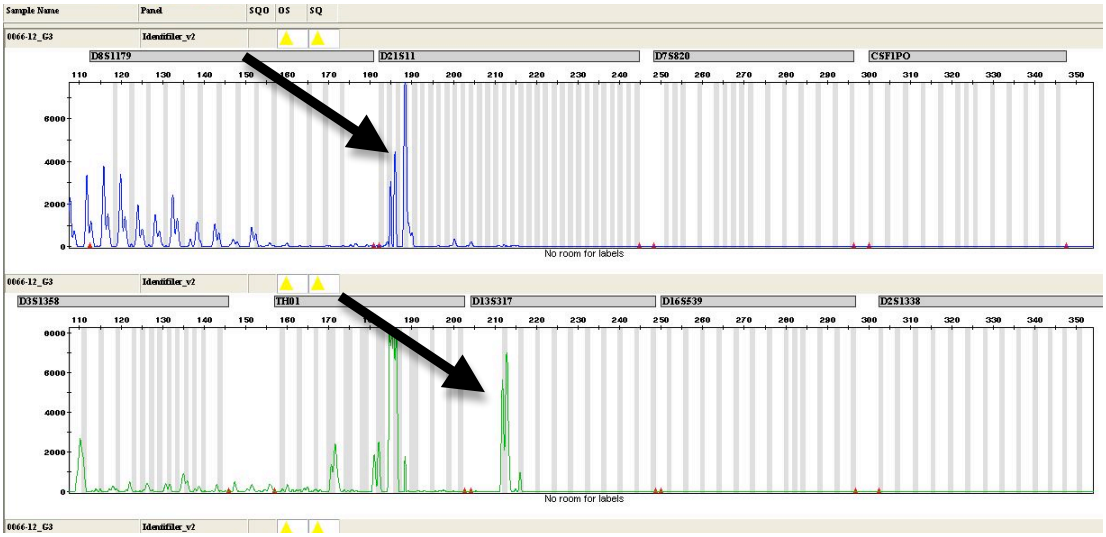


Figure 13: Sample 66 Amplified with Identifiler® Plus Amplification Kit after isolation with the Pippin Prep™ and Agarose Gel. The arrows indicate the desired peaks in each of the dye channels.

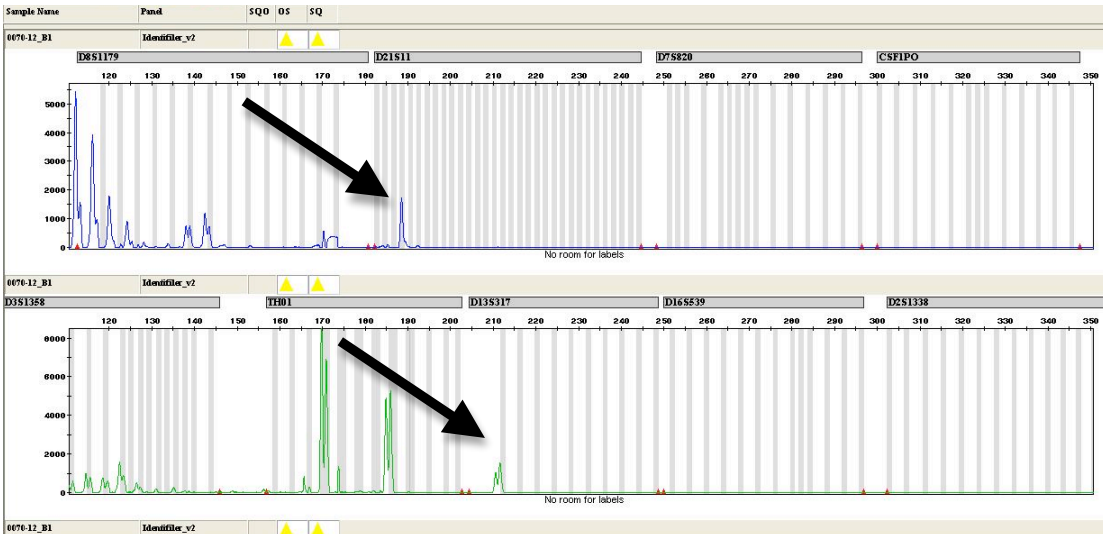


Figure 14: Sample 70 Amplified with Identifiler® Plus Amplification Kit after isolation with the Pippin Prep™ and Agarose Gel. The arrows indicated the desired peaks in each dye channel.

*Sequencing of Additional PCR Products in DNA Samples from Bones*

The next step to identify the additional PCR products was to sequence the isolated fragments. The product amplified after the agarose gel isolation was used for the sequencing reactions. Sequencing was performed using the Big Dye Terminator v3.1 Cycle Sequencing Kit. In the first run, sample 66 was amplified with each individual primer for the STR loci from the blue dye channel, to target the peak in the D21S11 locus region (Figure 15). Since it was unknown whether the forward or reverse primers were labeled with the fluorophore, each forward and reverse primer was used separately as the sequencing primer. The control was 9947A DNA, which is homozygous at the TPOX locus, and was sequenced using the TPOX primers. The control reactions produced results that were concordant with the reference sequence for the TPOX locus (Figure 16). However, no sequencing

results were obtained for any of the 66 sample reactions with primers from the blue dye channel (Figure 17).

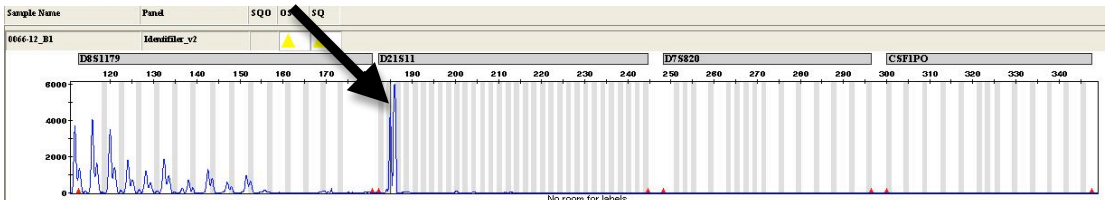


Figure 15: Sample 66 used for sequencing. Each primer in the blue channel was tested. The arrow indicates the peak that was targeted for sequencing.

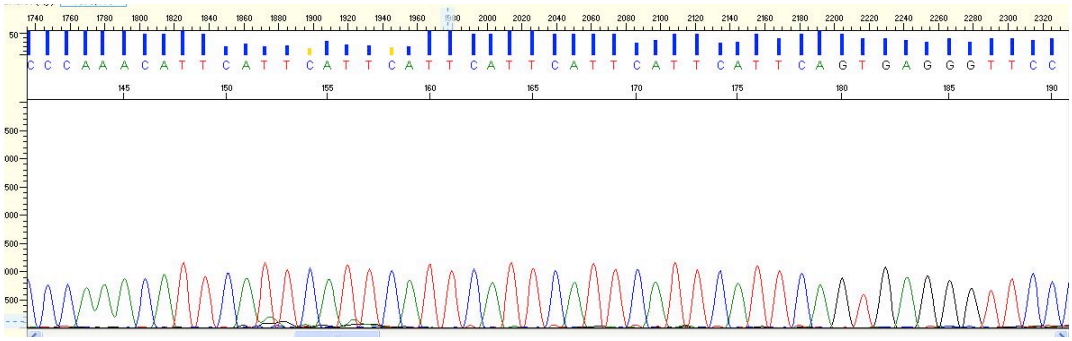

Figure 16: Sequencing Control; TPOX Reverse Primer with 9947A Control DNA

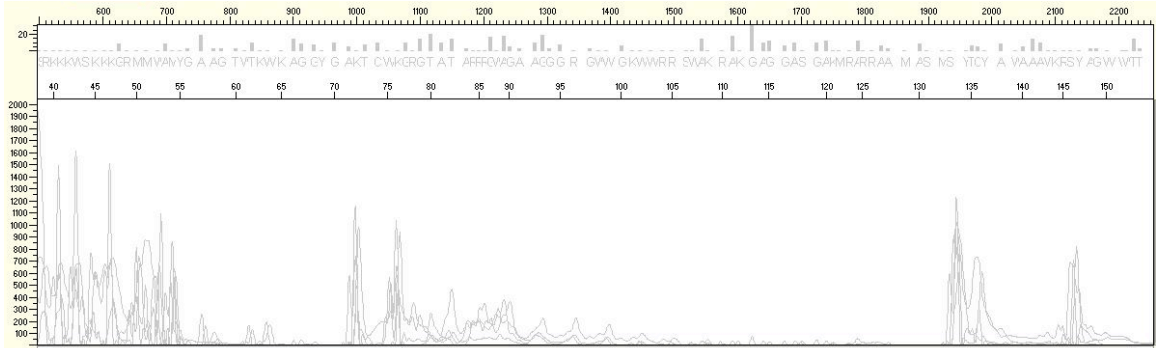

Figure 17: Sample 0066 Sequencing Results; D21S11 Forward Primer.

# CHAPTER IV

# CONCLUSIONS

STR analysis of DNA derived from bone samples is a common practice in many Forensic Laboratories and is particularly useful in missing persons or unidentified human remains cases. Successful STR typing of DNA from bone samples can be difficult due to degradation, contamination, and the limited quantity of nuclear DNA (11-13).

Bone samples may be collected from soil at a crime scene and have been shown to contain mixed amounts of human DNA and microbial DNA (14). The microbial DNA present in bone samples may be the source of the additional artifacts observed during STR analysis on the bone samples. The presence of non-specific PCR products in bone samples can confound interpretation. If these non-specific products are mistaken for human STR alleles, DNA analysts may report an inaccurate profile leading investigators to falsely associate human remains to a victim or be unable to associate the remains of a victim with a true family. The goal of this study was to determine if the additional PCR products found in the STR analysis of DNA derived from bone samples were microbial or fungal in origin.

Artifact products were found to be present in bone samples amplified with both the PowerPlex® 16 HS and Identifiler® Plus Amplification Kits. However, bone samples where artifact alleles were observed with the PowerPlex® 16 HS System were not observed when those same bone samples were amplified with the

Identifiler® Plus Amplification Kit. Additionally, Bone samples in which artifact alleles were observed when amplified with the Identifiler® Plus Amplification Kit were not observed when those same bone samples were re-amplified with the PowerPlex® 16 HS System. Non-Specific PCR products also were observed in soil samples amplified with the PowerPlex® 16 HS System. No PCR products were observed when the same soil samples were amplified with the Identifiler® Plus Amplification Kit. The presence and absence of the additional artifacts with each kit suggests that there may be specificity differences. This variation in specificity may be due to differences in primer design. The different primers in each kit may select for different targets; so this observation was not unexpected.

Experiments were performed to determine if non-specific PCR products could be detected in soil samples similar to those previously observed. Samples were collected from the same areas as previously tested soil samples as well as locations around the Fort Worth area. No PCR products were observed in the soil samples from the areas where samples had previously been tested (duckpond, levee, and stairway) with either the PowerPlex® 16 HS system or the Identifiler® Plus Amplification Kit. This may be due to the environmental conditions, the samples were collected at different times of the year and the exact locations of the collections may have differed as well. The moisture of the soil also may influence the extraction process; DNA may be more difficult to extract from dry soil samples than from wet samples. However, many samples from other areas of Fort Worth, including along the trinity trails did show non-specific PCR products when amplified with the PowerPlex® 16 HS System. Furthermore, the non-specific PCR products

were observed in multiple elutions of the DNA extract, showing that the results were reproducible. No PCR products were observed in any soil samples when amplified with the Identifiler® Plus Amplification Kit.

Approaches that were considered for determining the origin of the non-specific PCR product included cloning and sequencing. After further research, and some preliminary testing, cloning proved to be difficult. The DNA fragments present in the bone and soil DNA samples that were targeted for cloning and sequencing had a fluorescent tag that was attached via primer elongation during amplification with the STR kits so that the fragment can be visualized on the capillary electrophoresis instrument. The fluorescent tag prevented the DNA fragment from being incorporated into the vector for cloning. Therefore, an alternate protocol was developed to isolate the DNA fragment of interest and sequence the DNA fragment directly without cloning.

Samples of known DNA profiles were used to ensure that an effective procedure was in place to sequence the additional PCR products observed in the DNA isolated from bone samples. Homozygous peaks at the D13S317, D8S1358, D3S1358, and TPOX STR loci were tested to determine if the protocol was performing appropriately that would be used for sequencing the off ladder PCR product. After confirming the successful amplification of those specific loci, the known DNA samples were sequenced and analyzed. The known DNA samples were successfully sequenced and were concordant with the reference sequences and contained the same number of repeats obtained by standard STR typing using the Identifiler® Plus STR Amplification Kit (Table 2 and Table 3).

A protocol for the isolation of the non-specific PCR products found in the DNA from bone samples was developed utilizing two different isolation methods. The first method was using the Pippin Prep™ DNA Size Selection System. A "tight" range program was used to isolate the fragments of interest. This program attempts to capture DNA fragments within a 40 base pair range. The recovered samples were analyzed on the Agilent 2200 TapeStation and confirmed that the correct size fragment was isolated (Figure 10). However, after amplifying the recovered DNA with the Identifiler® Plus Amplification Kit for STR analysis there were still peaks present in the sample outside that specified range (Figure 11 and Figure 12). The size that was selected to be the target selection size was determined using the STR data that was produced before isolation. The differences between the target size range and the sizes that were actually captured may be due to the differences between the separation medium in which the DNA fragments migrated in the 3130xl Genetic Analyzer and the Pippin Prep™ DNA Size Selection System. The DNA fragments may travel differently in 2% agarose gel in the Pippin Prep™ and Pop-4® Polymer in the 3130xl Genetic Analyzer.

To attempt to remove the DNA fragments that were outside of the target range, the samples were further separated on a 1.5% agarose gel. The DNA fragments were visualized by staining the gel with GelRed™ Nucleic Acid Stain. The bands that were visible were excised and purified. The purified DNA was amplified with the Identifiler® Plus Amplification Kit and analyzed. The results from STR amplification after gel isolation and purification show that the target DNA fragments were isolated. However, some smaller sized fragments were captured as well. The

gel isolation and purification did produce a cleaner product, eliminating all of the non-targeted DNA fragments that were larger in size than the desired DNA fragment. There were also fewer DNA fragments present that were smaller in size than the desired DNA fragment. Although, there were still some peaks present that were not targeted in the isolation procedure, the targeted peaks have been reproducible after each isolation step and throughout each amplification step, supporting that these artifacts are PCR products from DNA present in the sample.

After the non-specific PCR product was isolated the next step was to sequence these products. At this time, no sequence data was obtained from the artifact products. This part of the project is ongoing. There could be several reasons why no sequencing results were obtained. One reason may be that there was not enough template DNA for the sequencing reaction. Experiments will continue to try to obtain sequence data from the non-specific PCR products found in the bone and soil samples.

The artifacts observed in the DNA from bone samples can confound the interpretation of the STR profile. Identification of these non-specific PCR products will better define the limitations and provide guidelines for the interpretation of STR typing results from samples in which such artifacts are observed. The results of this study have demonstrated that the additional artifact peaks in the bone samples are DNA-based. The artifact peaks were reproducible throughout each step of the study even after dilution, gel separation, and multiple amplifications. In addition, a method has been developed for isolating non-specific PCR products, such as those that may be observed in bone and soil samples. This study also has demonstrated

that amplification of soil samples with an STR kit can produce artifacts that appear similar to those seen in the bone samples. The artifacts observed in the soil samples are seen in soil samples from a range of locations and are reproducible. The results from this study support that non-human DNA found in soil may be the source of the additional artifact peaks seen in the DNA samples from bone and soil.

# REFERENCES


1. Ellegren H. Microsatellites: Simple sequences with complex evolution. Nature Reviews Genetics. 2004 06;5(6):435-4.

2. Edwards A, Civitello A, Hammond HA, Caskey CT. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet. 1991 Oct;49(4):746-5.

3. Collins JR, Stephens RM, Gold B, Long B, Dean M, Burt SK. An exhaustive DNA micro-satellite map of the human genome using high performance computing. Genomics. 2003 7;82(1):10-9.

4. Subramanian S, Mishra RK, Singh L. Genome-wide analysis of microsatellite repeats in humans: Their abundance and density in specific genome regions. Genome Biol. 2003;4(2):R13.

5. FBI — QAS for DNA testing labs [Internet].; cited 7/24/2012]. Available from: http://www.fbi.gov/about-us/lab/codis/qas_testlabs.

6. FBI — forensic science communications - july 2004: Revised validation guidelines-scientific working group on DNA analysis methods (SWGDAM) [Internet].; 2004; cited 8/6/2012]. Available from: http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/july2004/standards/2004_03_standards02.htm/.

7. Principal methods for isolation and ... [folia microbiol (praha). 2012] - PubMed - NCBI [Internet].; cited 7/20/2012]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22791233.

8. Milos A, Selmanovic A, Smajlovic L, Huel RL, Katzmarzyk C, Rizvic A, et al. Success rates of nuclear short tandem repeat typing from different skeletal elements. Croat Med J. 2007 Aug;48(4):486-93.

9. Wang DY, Chang C, Lagacé RE, Calandro LM, Hennessy LK. Developmental validation of the AmpF?STR® identifiler® plus PCR amplification kit: An established multiplex assay with improved performance. J Forensic Sci. 2012;57(2):453-65.

10. Ensenberger MG, Thompson J, Hill B, Homick K, Kearney V, Mayntz-Press KA, et al. Developmental validation of the PowerPlex® 16 HS system: An improved 16-locus fluorescent STR multiplex. Forensic Science International: Genetics. 2010 7;4(4):257-64.

11. Anderung C, Persson P, Bouwman A, Elburg R, Götherström A. Fishing for ancient DNA. Forensic Science International: Genetics. 2008 3;2(2):104-7.

12. Lee HY, Park MJ, Kim NY, Sim JE, Yang WI, Shin K. Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. Forensic Science International: Genetics. 2010 10;4(5):275-80.

13. Zupanic Pajnic I, Gornjak Pogorelc B, Balazic J, Zupanc T, Stefanic B. Highly efficient nuclear DNA typing of the world war II skeletal remains using three new autosomal short tandem repeat amplification kits with the extended european standard set of loci. Croat Med J. 2012 Feb 15;53(1):17-23.

14. Alonso A, Andelinovic S, Martin P, Sutlovic D, Erceg I, Huffine E, et al. DNA typing from skeletal remains: Evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples. Croat Med J. 2001 Jun;42(3):260-6.

15. Short tandem repeat DNA internet DataBase [Internet]. 10/2/1997; 2013 [updated 2/25/2013; cited 3/5/2013]. Available from: http://www.cstl.nist.gov/strbase/.