An Enzymatic Means for the Rehabilitation of Low-Copy Number and Degraded DNA

Nicole Sambol

University of Central Oklahoma

Forensic Science Institute

An Enzymatic Means for the Rehabilitation of Low-Copy Number and Degraded DNA

By: Nicole Sambol

A THESIS

APPROVED FOR THE W. ROGER WEBB FORENSIC SCIENCE INSTITUTE

May 2014

By _____

Dr. Dwight Adams Committee Chairperson

_____

James Creecy Committee Member

_____

Dr. Lillian Chooback Committee Member

Abstract

In spite of the immense success forensic DNA analysis has obtained over the last twenty-five years, a substantive challenge within the field of forensic DNA analysis is amplification and interpretation of degraded and low-copy number (LCN) DNA obtained from minimal and poor quality biological evidence. It has been well established that DNA profiles obtained from degraded samples are often of limited value due to the frequent occurrence of preferential amplification during polymerase chain reaction (PCR). The by-products of preferential PCR amplification are often observed as inter- and intra-locus peak imbalance, allelic dropout, and/or locus dropout. Inspired by advances in next-generation sequencing techniques, I propose a methodology for simultaneously normalizing the abundance of PCR products across all short tandem repeat (STR) loci using the DNA exonuclease, duplex-specific nuclease (DSN). DSN is a novel DNA nuclease that was isolated from the hepatopancreas of Red King (Kamchatka) crab and it possesses a strong affinity for digesting double stranded DNA (dsDNA) and has limited activity toward single stranded DNA (ssDNA). Degraded DNA known to display peak imbalance and allele dropout was amplified using AmpFISTR® Identifiler® Plus for 28 cycles. Following amplification, samples were denatured at 99.9 °C for 5 min and incubated with one unit of DSN at 62 °C in a 28 µl volume for 1 min. Nuclease activity was terminated through the addition of equal volume of 10 mM EDTA and 95 °C incubation for 2 min. The findings obtained support the potential use of DSN treatment as a method for the normalization of STR profiles, but further analysis must be done before the protocol is casework ready.

Keywords: Degraded DNA, Low-Copy Number DNA, Duplex-Specific Nuclease

Table of Contents

Table of Figures

Table of Tables

Introduction

Forensic DNA analysis is an applied biological science that seeks to identify individuals based on the genetic material that remains after a crime has been committed. Often, the genetic material obtained at a crime scene is harvested from biological materials like blood, semen, and various other tissues. Due to the complex nature of crime scenes and the number of environmental and man-made variables that can affect the quality of the DNA sample, DNA evidence recovered at crime scenes can be degraded and/or present in very low quantities. Over the last fifteen years, numerous technological advances have contributed to the successful analysis of degraded and low quantity DNA evidence. Due to these advances, the extraction, amplification, and analysis processes of degraded and low quantity DNA evidence has become common practice in the forensic science community (Gill, 2001). Despite technological advances, a DNA quality threshold remains and there is only so much damage a biological sample can endure before the DNA template in that sample declines and poor forensic DNA profiles are produced. The analysis of poor quality and quantity DNA evidence has become so frequent that a sub-discipline within forensic DNA analysis has emerged. Low-copy number (LCN) DNA analysis focuses on the analysis of DNA evidence that contains minimal genetic material and is the focus of much of the research effort within the forensic DNA community. Evidentiary DNA profiles generated using LCN DNA often include artifacts such as inter- and intra-locus peak imbalance, high stutter, allelic dropout, and complete locus dropout. Each of these artifacts alone can make the interpretation of evidentiary DNA profiles challenging, while combination of such artifacts can make interpretation impossible.

The challenge associated with interpreting both degraded and LCN DNA profiles is the result of the dramatic increase in observed artifacts caused by a polymerase chain reaction (PCR)

phenomena known as preferential amplification by stochastic effect. In principle the preferential amplification is best described as follows: samples containing small amounts of degraded DNA have very few copies of each genetic locus, but even fewer copies of the high-molecular weight genetic loci. This bias toward high-molecular weight loci is due to the fact DNA degradation is more likely to affect the larger loci because there is a greater area that may be affected by double stranded break events. During PCR, *Taq* DNA polymerase will more efficiently replicate the smaller, low-molecular weight loci because there are more intact copies of this genetic material. With each new round of DNA amplification, more and more copies of the genetic marker are produced, and each time the smaller loci are amplified in greater number in relation to the larger loci. At the conclusion of the PCR process, a genetic profile is biased toward the lower-molecular weight loci over those with higher molecular weights. This is commonly referred to as preferential amplification by stochastic effect. The number of artifacts resulting from preferential amplification can lead to many problems during the interpretation phase of forensic DNA analysis, which could result in the possible false inclusion or exclusion of an individual if the proper cautions are not taken (Taberlet et al., 1996).

The development of a method for enzymatically normalizing LCN and degraded DNA in order to avoid the detrimental effects of preferential amplification is of specific interest to this research project. The discovery and characterization of an enzyme known as duplex-specific nuclease (DSN), which cleaves double stranded DNA (Shagin et al., 2002), has provided the forensic community with the opportunity to repair the imbalance of genetic material prior to DNA interpretation. DSN has been successfully used in research settings to normalize samples with unequal representations of template DNA, but never for a forensic application. Previous experimental approaches have successfully normalized the amount of DNA in a given sample by

taking advantage of the differential reassociation rates of denatured DNA. Simply stated, DNA sequences that are more abundant in a sample are more likely to anneal to their complementary strands, and do so at a significantly faster rate than those that have fewer copies. Using the kinetics associated with the reannealing of denatured DNA, DSN can then be employed to differentially cleave DNA sequences in higher concentration as they reanneal, thereby reducing the amount of overrepresented DNA sequences. The removal of the overrepresented DNA sequences from the total population allows the underrepresented DNA sequences in the sample to become equalized (Shagina et al., 2010).

The ability to generate a full profile from degraded or minimal amounts of sample would prove to be an invaluable resource when attempting to associate an individual to a crime scene during a forensic investigation. Therefore, the application of DSN to normalize LCN and degraded DNA would be useful when applied to forensic DNA profiling. I hypothesize that DSN normalization prior to genetic analysis will alleviate the preferential amplification of lower-molecular weight loci, thereby allowing for equal representation of all genetic material and a more complete forensic DNA profile.

Literature Review

Molecular Structure of DNA

DNA is the biological macromolecule that contains all the genetic information needed to bring about life. In addition, the genetic information contained within DNA is responsible for the inheritance of traits from parent to offspring. The forensic analysis of DNA for the purpose of identification is fundamentally based on the structure of DNA and the method by which DNA replicates. The alternating sugar-phosphate backbone of the DNA macromolecule depicted in Figure 1 consists of nucleotides that are linked to one another by phosphodiester bonds. A single DNA strand is formed through the addition of nucleotides to the 3' end of the molecule. The fundamental unit of DNA is the nucleotide, seen in Figure 2, which consists of three distinct parts; a 5-carbon deoxyribose sugar, a phosphate group, and a nitrogenous base (Watson & Crick, 1953). The phosphate group is located on the backbone of the DNA molecule and provides DNA with an overall negative charge, an attribute that is ultimately utilized during DNA analysis. Attached to the phosphate group is the 2'-deoxyribose sugar. In addition to contributing to the overall backbone of the DNA molecule, the sugar's lack of a hydroxyl group provides the DNA molecule with an increased resistance to hydrolysis (Lindahl, 1993). Attached to it at the 1' carbon of the deoxyribose sugar is the third component, the nitrogenous base. There are four different nitrogenous bases that may be attached to the deoxyribose sugar, and they are adenine, thymine, cytosine, and guanine. Hydrogen bonds bind two complimentary, anti-parallel strands together at the nitrogenous bases, ultimately providing the DNA with its classic double-helix formation (Watson & Crick, 1953). Hydrogen bonding provides the DNA molecule with flexibility, which allows the bases to be moved, damaged bases to be removed during synthesis, and the strands to be methylated (Iyer, Pluciennik, Burdett, & Modrick, 2006).

Figure 1

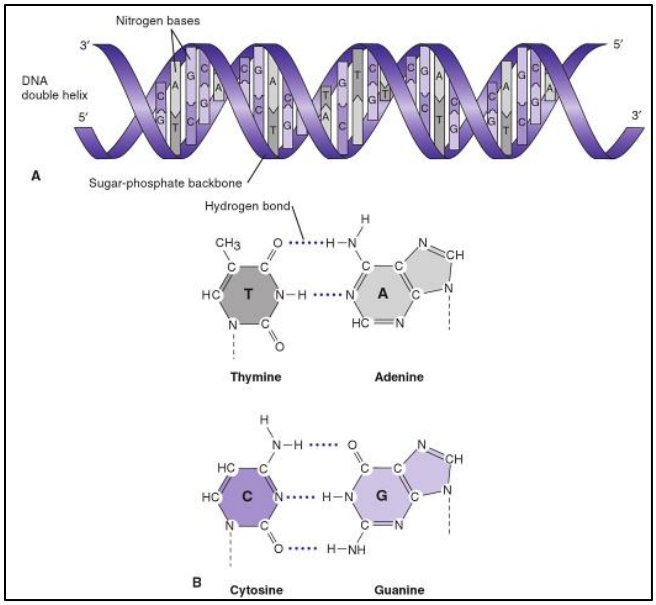

Figure 1: DNA double helix and nitrogenous bases
The anti-parallel strands are bound at the nucleotides through hydrogen bonding. Below are the four nitrogenous bases: adenine, thymine, cytosine, and guanine. Adenine and thymine are bound together by two hydrogen bonds, while cytosine and guanine are bound by three hydrogen bonds. (Buckingham, 2011)

Figure 2

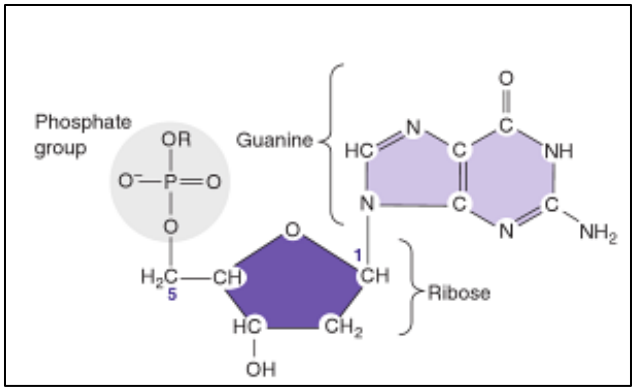

Figure 2: Nucleotide Structure
A nucleotide is the functional unit of DNA and consists of three parts: the phosphate group, a 2' deoxyribose sugar, and a nitrogenous base. (Buckingham, 2011)

The entirety of the human genome consists of two distinct regions, the "coding" region and the "noncoding" region. The coding region contains genes that control the expression of proteins, which could be useful for phenotypic identification (Kayser & Schneider, 2009). However, the frequencies of these alleles are constantly changing due to non-random mating, constant migration, and natural selection, and as a result Hardy-Weinberg equilibrium is not met (Hartl & Clark, 1997). Ultimately, this means that the coding region is of little use to the forensic science community for human identification purposes since Hardy-Weinberg statistical calculations cannot be performed. Noncoding regions of DNA do not provide genetic information, but have the potential to be used for human identification because of the high variability between individuals (Lander et al., 2001). Population studies have examined the non-

coding alleles in numerous populations, including the Caucasian, African-American, and Hispanic populations, and have discovered that they are inherited in predictable patterns. The frequencies of the alleles of interest in the forensic community have been determined to be consistent within each ethnic subgroup, which allows statistical analysis to be performed (Cohen, Lynch, & Tayplor 1991).

Variable Number Tandem Repeats for Human Identification

Modern forensic DNA analysis is the byproduct of decades of research and method development that culminated in the analytical technique employed today. The first instances of forensic DNA testing were performed using restriction fragment length polymorphism (RFLP) tests to detect variable number tandem repeats (VNTRs). Also referred to as minisatellites, these repetitive sequences are located within the noncoding region of the genome and are highly variable between individuals making them optimal targets for human identification (Jeffreys, Wilson, & Thein, 1985). RFLP testing required that the extracted DNA be digested with various restriction enzymes to reproducibly cut DNA samples into smaller fragments, which were then separated based on their size using agarose gel electrophoresis. A Southern hybridization technique was frequently used to transfer the DNA from the agarose gel onto a nylon membrane (Jeffreys et al., 1985). Each sample was labeled with a radioactive probe complementary to a specific VNTR sequence, which could be visualized using an X-ray film (Budowle, Smith, Moretti, & DiZinno, 2000). Despite its high variability and its capability for human identification, VNTR testing had several disadvantages. The environmental contaminants and insults often present at a crime scene made it difficult to obtain the 50 ng of intact DNA for analysis. RFLP analysis was also time-consuming, requiring up to a week to develop a single probe. In addition, the complexities associated with reading the results made interpretation of

alleles difficult, which limited their use in identification techniques (Edwards, Hammond, Jin, Caskey, & Chakraborty, 1992). As a result, and with the advent of new technologies, more sensitive and advanced methodologies were pursued for forensic DNA analysis.

Due to the nature of crime scenes, it is highly unlikely for forensic biological evidence to be recovered in pristine condition. Consequently, the validation of forensic RFLP analysis for human identification also included numerous tests to determine how contaminants and environmental insults altered a DNA profile. Various types of samples that could provide a DNA profile were subject to both biological and nonbiological insults, and it was determined that contaminants did not alter the DNA profile, but could ultimately result in sample degradation. Testing did not result in any false positives or negatives, and none of the sample profiles differed from the control profiles by more than 5%. These data showed that RFLP analysis could be successfully and reproducibly performed on all types of forensic samples under varying conditions (Adams et al., 1991), and was a solid foundation for the protocols used in forensic DNA analysis today.

Short Tandem Repeat Units for Human Identification

Short tandem repeats (STRs), used in forensic science analysis are tetranucleotide repeat units found within the noncoding region of the genome. These units may be repeated anywhere from 6 to over 30 times, depending on the locus. The high variability associated with STRs allows for the individualization of a person at the genetic level, which makes them extremely advantageous not only in identifying individuals in forensic casework, but for missing person's identification and mass disaster identification (Clayton, Whitaker, & Maguire, 1995).

Figure 3

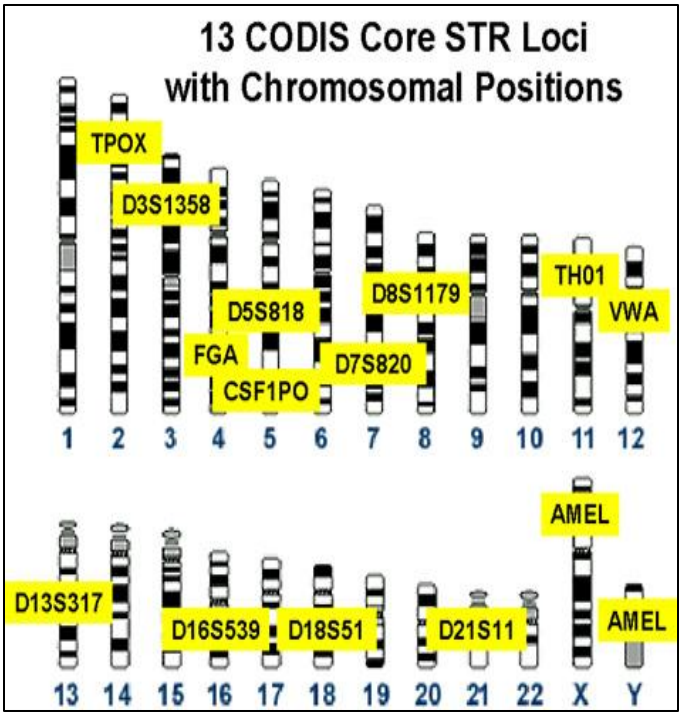

Figure 3: The 13 CODIS Core STR Loci with Chromosomal Positions

The FBI designated 13 CODIS genetic loci required for a complete forensic profile. Also depicted is the relative location of each locus on the chromosome (Butler & Reeder, 2011)

Thousands of STRs have been discovered in the human genome, and the FBI has designated 13 core loci that are required in order to build a complete forensic profile. They are: TPOX, D3S1358, FGA, D5S818, CSF1PO, D7S820, D8S1179, TH01, vWA, D13S317, D16S539, D18S51, and D21S11, and can be seen in Figure 3. The amelogenin locus is also included in analysis to determine the donor. The STRs analyzed are genetically unlinked from one another, so it is expected that during replication they will segregate independently. CSF1PO and D5S818 are the only STRs found on the same chromosome, 5, which is one of the largest human chromosomes containing 177 million base pairs (Schmutz et al., 2004). Since these loci separated by 26 million base pairs, they do not show any signs of linkage and are thus independently inherited (Butler & Reeder, 2012). Studies have also determined that the loci are highly variable among individuals of similar race and ethnic background (Budowle, Moretti, Baumstark, Defenbaugh, & Keys, 1999).

The majority of the STR loci analyzed are tetranucleotide repeats, but these sequences can be one of several different pattern types. The most basic pattern type, the simple repeat, contains repeat units that have both identical sequence and length. Compound repeats are made

of at least two adjacent repeats, while repeat motifs of a complex nature have the potential to contain multiple repeat sequences of different lengths (Urquhart, Kimpton, Downes, & Gill, 1994). Some STR locations are classified as microvariants, meaning they contain incomplete repeat units. A common microvariant is the 9.3 allele found at the TH01 locus. This allele contains nine complete repeats and has one unit with only three nucleotides. The partial repeat unit is caused by one of two mechanisms; a missing adenine in the seventh repeat unit, or a missing thymine in one of the TCAT repeats (Puers, Hammond, Jin, Caskey, & Schumm, 1993). Differences in pattern type and repeat number indicate that not all loci behave in the same manner under the same conditions, a factor that was analyzed in this research.

STR analysis by PCR is much more sensitive than the RFLP methodology, and as a result only 1 ng of DNA is required for the reaction (Applied Biosystems, 2012a). Since STR analysis can be performed on many loci at once, an individual's entire forensic profile can be determined in a matter of hours, which is a much faster turnaround time than those provided by VNTR analysis. Since STR testing does not require a Southern Blot to visualize the results, it is more cost effective than VNTR analysis and a much more practical method for forensic human identification (Edwards et al.,1992). The advantages provided by STRs make this method ideal for human individualization, and therefore the methodology of choice for forensic identification.

DNA Extraction using DNeasy® Blood and Tissue Kit

Extraction of DNA from the sample is often thought of as the rate-limiting step in DNA analysis, because how well the procedure is performed determines how much usable DNA there is to analyze. Multiple methodologies for extraction exist, but this research utilized a solid-phase extraction protocol designed by QIAGEN for the DNeasy® Blood and Tissue Kit. This approach takes advantage of DNA's negatively charged phosphate backbone previously

described, and is advantageous for DNA extraction in forensic settings due to the lack of extremely hazardous chemicals and high-throughput capabilities.

Extraction begins through the lysis of cellular material to liberate the DNA found on evidence. Cellular material is soaked in a solution that contains sodium dodecyl sulfate (SDS) and proteinase K to break open the cells and digest protein contaminants respectively (QIAGEN, 2006). Following lysis, the solution is placed on a silica-based membrane that contains a glass bead support. When high concentrations of chaotropic salts like sodium iodide (Vogelstein & Gillespie, 1979), guanidine hydrochloride, guanidine isothiocyanate, and sodium perchlorate are present, the negatively charged DNA backbone selectively binds to the glass beads incorporated into silica membrane (Boom et al., 1990). The addition of high-salt wash buffers effectively moves contaminants and impurities through the membrane while the DNA continues to be bound to the glass beads. Once wash steps have alleviated unwanted contaminates from the sample, the DNA is eluted from the beads through the addition of a low-salt wash buffer. The resulting purified product can then be used for STR analysis.

DNA Quantitation using UV Spectroscopy and Real-Time Polymerase Chain Reaction

The quantitation of DNA within an evidentiary sample is an essential step in the forensic analysis of DNA. Guidelines implemented by the DNA Advisory Board mandate that all samples be measured to determine the amount of amplifiable human DNA in the extract (DNA Advisory Board, 1998). One of the oldest and most reliable methods for quantifying DNA in solution is the use of UV spectrophotometry. In principle, UV light of a known intensity is directed at the DNA sample and the absorption of the UV light by the nitrogenous bases in DNA is measured. The absorption is directly proportional to the concentration of the DNA in the sample. A light with a 260 nm wavelength, and one with a 280 nm wavelength, moves through

the sample, and the absorbance levels are taken. The absorbance allows the user to determine the concentration of the DNA using the 260 nm value, while contaminants such as proteins and organic solvents are measured by the 280 nm value. Using the absorbance values, a 260/280 ratio can be calculated to determine the overall purity of the DNA. Samples with increased 260/280 values are most likely contaminated with RNA, while lower than expected numbers could be the result of protein contamination (Nicklas & Buel, 2003). While the UV spectroscopy is useful, fast, and easy to use, it cannot be employed in forensic casework because it is not human specific. In this experiment, both the Quantifiler® protocol and UV spectroscopy were used to quantify the sample DNA, and the results were compared.

Currently, the majority of DNA forensic laboratories in the United States use real-time PCR (qPCR) to obtain the concentration of amplifiable DNA. This technology evaluates the change in fluorescence as a result of the amplification of the target sequences after each cycle of PCR (Higuchi, Dollinger, Walsh, & Griffith, 1992). Since the instrument monitors the fluorescence inside the reaction tube, there is little potential for contamination (Higuchi, Fockler, Dollinger, & Watson, 1993). This research utilized the Quantifiler® Human DNA Quantification Kit, which employs TaqMan® probes to quantify the DNA. This methodology is depicted in Figure 4. These probes are specific to the Human Telomerase Reverse Transcriptase (hTERT) gene, and contain both a "reporter" and a "quencher" dye. During the annealing phase of PCR, if human DNA is present, the probe hybridizes to the specific DNA sequence within the hTERT gene. *Taq* DNA polymerase then begins to incorporate nucleotides into the growing DNA strand, and simultaneously utilizes its 5' exonuclease activity to break up the probe into single nucleotides. Ultimately, this separates the reporter from the quencher, and when the two are no longer in close proximity to one another, the reporter emits a fluorescent signal that is

measured (Applied Biosystems, 2010). The amount of fluorescence seen is directly proportional to the amount of DNA in the sample. A fluorescence threshold is set, and when the fluorescence level crosses it, the C_T value, or the number of cycles it took to cross that threshold, can be compared to a standard DNA concentration curve. This curve shows the C_T values for known concentration of DNA. By comparing the cycle numbers, it is possible to determine the concentration of DNA in the unknown sample (Applied Biosystems, 2010). The DNA concentrations are then utilized for calculations in further stages of forensic DNA analysis.

Figure 4

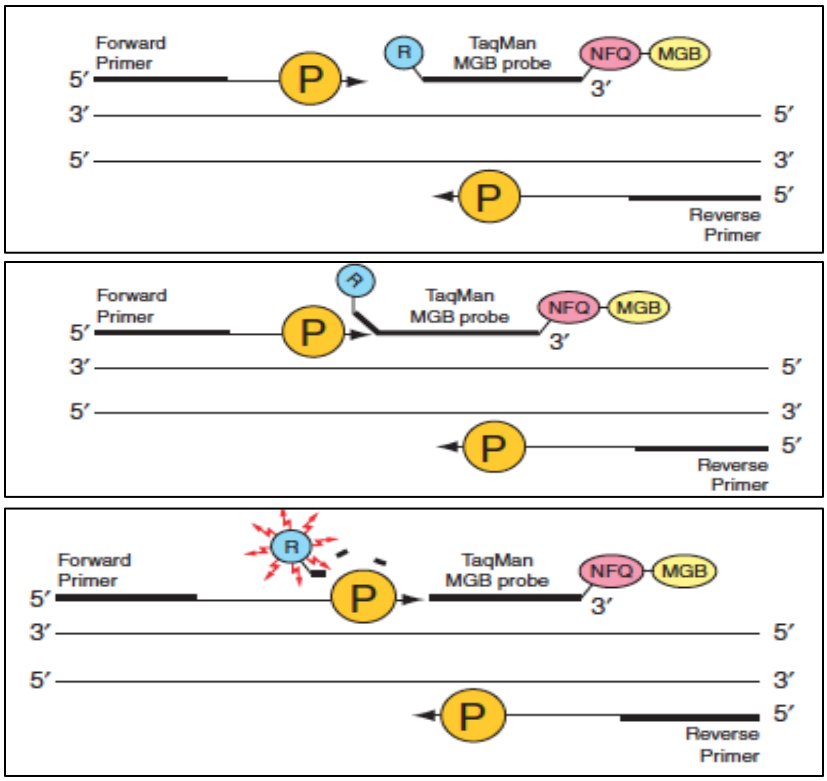

Figure 4: Quantification through the use of Taqman® Probes

Figure 4 demonstrates the TaqMan® methodology for the quantification of DNA. The first image depicts primer and probe binding to the target DNA strands. The second image shows the polymerase adding nucleotides to the growing strand. As the polymerase approaches the probe in the third image, the 5' exonuclease activity liberates the reporter from the probe, and a fluorescent signal is emitted (Applied Biosystems, 2010)

This test also contains an IPC, or an internal PCR control, that monitors for inhibition during amplification. If the IPC is amplified, but the sample is not, it can be said that the sample does not contain human DNA. If both the IPC and sample show no amplification, the

experiment was most likely inhibited. In order to distinguish between the IPC and human DNA, two different reporter dyes are used, FAM for human-specific DNA, and VIC for the IPC (Applied Biosystems, 2010).

DNA Amplification using PCR

PCR is a molecular biology technique developed by Kary Mullis (Mullis & Faloona, 1987) to increase a specific portion of a DNA sequence by a factor of 10^6 (Saiki et al., 1988), using only a minute amount of starting sample. PCR is performed by mixing the extracted template DNA obtained from an evidence item with a "master mix," which contains buffer, oligonucleotide primers, $MgCl_2$, deoxynucleotides (dNTPs), and a DNA polymerase, a thermostable enzyme most commonly derived from a bacterium *Thermus aquaticus.* The DNA polymerase acts to replicate DNA from the nucleotide sequence on the complementary DNA strand. This is achieved by binding to the region specific primers and moving down the template DNA strand and incorporating free dNTPs found in the reaction mixture into the new DNA strand (Saiki et al., 1988).

PCR is used by the forensic community to amplify the 13 core STR loci and to obtain a forensic DNA profile. There are three steps to the PCR process: denaturing, annealing, and extension (Saiki et al., 1988). The denaturing step heats the sample to approximately 94 °C for 1 minute, effectively separating the complementary DNA strands. PCR of forensic samples is often performed on multiple loci at once, meaning that multiple primers are being utilized during the reaction. Primers used must flank regions of different sizes and are incorporated with a unique fluorophore so that regions of similar size can be differentiated (Edwards & Gibbs, 1994). These primers all have different optimal annealing temperatures based on base GC composition, length, and the individual sequences (Henegariu, Heerema, Dlouhy, Vance, & Vogt, 1997). As a

result, there is a range of effective annealing temperatures that can be utilized during the reaction. During the annealing phase of the Identifiler® Plus protocol, the sample is cooled to approximately 59 °C for one minute to effectively bind all primers to the regions flanking the regions of the complementary target DNA. *Taq* DNA polymerase has the capabilities to be active over a broad range of temperatures, with maximum nucleotide addition around 75 °C (Innis, Myambo, Gelfand, & Brow, 1988). In order for the extension of the growing strand to occur in the Identifiler® Plus protocol, the sample temperature is then heated to approximately 72 °C for 1 minute (Applied Biosystems, 2012a). The three steps are repeated 28 times, and, once completed, the sample contains millions of copies of the target DNA sequence (Applied Biosystems, 2012a).

Commercial kits available for forensic use currently have the capabilities of amplifying 16 loci at once, but there are ongoing studies examining the potential for the analysis of even more loci. The FBI has designated a CODIS Core Loci Working group, and these individuals have determined that expanding the CODIS core loci in the United States would have three major advantages. The first would be the reduction of chance matches within the ever-growing database, since laboratories are becoming more efficient at processing samples. An increase in the number of core CODIS loci would also increase compatibility with international law enforcement and further expand the discrimination power of the DNA profiles (Hares, 2012).

The expansion of the core loci is still being researched, meaning current forensic protocols use commercial kits that amplify 15 autosomal STR markers and the amelogenin sex-determining locus to individualize forensic DNA samples. This experiment utilized the Identifiler® Plus PCR Amplification System, seen in Figure 5, that is commonly used in forensic laboratories in the United States. Before a particular PCR system protocol is implemented in the

laboratory, it must be validated to ensure it provides accurate and reproducible results (DNA Advisory Board, 1998). Previous validation studies of the Identifiler® Plus kit have demonstrated its effectiveness in the amplification of low quality and quantity biological samples (Wang, Chang, Lagacé, Calandro, & Hennessy, 2012), as demonstrated in sensitivity studies preformed which exhibit the success of Identifiler® Plus obtaining complete profiles from biological samples containing as little as 0.125 ng of DNA per reaction, and partial profiles with as little as 0.0062 ng of DNA per reaction. Further studies preformed using degraded samples have proven Identifiler® Plus has the capabilities to amplify complete STR profiles after exposure to 3 units of DNase I, while samples exposed to more than 3 units exhibited some dropout on the larger loci. In addition to the increased sensitivity of the amplification of poor and low-quality samples, the Identifiler® Plus kit chemistry demonstrated an increased balance of between intralocus peak height ratios (Wang et al., 2012). Due to the balanced nature of the Identifiler® Plus chemistry and its proven success in the amplification of degraded and low-quantity samples, this experiment will utilize the AmpFISTR® Identifiler® Plus PCR reaction kit for the amplification of the CODIS 13 core STR loci.

Figure 5

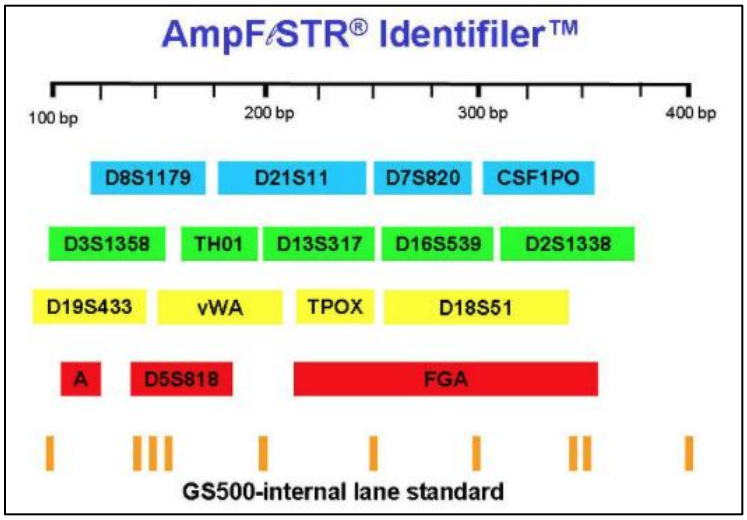

Figure 5: The Identifiler® Plus 16 Loci and Five Dye Channels

The five dye channels analyzed by the Identifiler® Plus system. The loci are shown in the position they will be found on a genetic profile. (Butler & Reeder, 2011)

DNA Separation using Capillary Electrophoresis

Amplification of the genetic material not only serves to provide more copies of the DNA fragments of interest, but to incorporate fluorescent dyes to the loci for analysis. The negative charge provided by the phosphate backbone of the DNA strand allows PCR products formed during amplification to be separated and analyzed based on size and fluorescent label using capillary gel electrophoresis. Every basic capillary electrophoresis instrument consists of a narrow glass capillary, two buffer reservoirs, and a high-voltage power supply connected to two electrodes. The capillary is filled with a polymer while the buffer acts to solubilize and stabilize the DNA while simultaneously acting as a carrier of the electrical charge (Jorgenson & Lukacs, 1981). Once the voltage is applied by the electrodes, the current pushes the DNA through the polymer. Since the DNA fragments being analyzed are of different sizes, they migrate at different rates through the capillary, with the smaller fragments moving faster than the larger ones (Grossman & Colburn, 1992).

Advanced capillary electrophoresis instruments contain specialized cameras that are able to detect the fluorescent labels attached to the DNA fragments. As the DNA migrates, the attached fluorophore is in a ground energy state and no fluorescence is given off. Once the DNA passes by a laser within the instrument, a photon from the laser acts to excite a fluorophore electron to a transition state, causing the electron to undergo a conformational change that results in a relaxed singlet excitation state. As the excited electron falls back to its ground state, a photon is released (Lakowicz, 2007). The fluorophores used in forensic DNA analysis have known light absorption and emission patterns, and a charge-coupled device (CCD) camera within the instrument is able to detect and analyze which fluorescent dye is being emitted. The emitted light is measured in relative fluorescent units (RFUs), and is depicted as a peak on the

electropherogram. Higher RFU values indicate a high concentration of a particular DNA fragment. During electrophoresis, the data analysis software records all peaks above a standard 50 RFU baseline threshold, and a second threshold called the interpretation, or the stochastic threshold, is set to detect the target loci (Butler, Buel, Crivellente, & McCord, 2004).

Each sample is run through capillary electrophoresis with an internal size standard that contains DNA fragments of various known sizes. The internal size standard is ultimately utilized to build a standard curve to determine the relationship between the number of base pairs the known fragment contains and the time it takes for that fragment to pass by the laser. The unknown DNA fragment sizes are then compared to the standard curve in order to determine the number of base pairs present in those fragments. Another standard, the allelic ladder, is also analyzed through capillary electrophoresis. This standard contains all alleles commonly observed in the human population, so, on the resulting electropherogram, the allelic repeat units are spaced one unit apart from one another (Butler et al., 1995). The allelic ladder provides alleles that are the same size as those in the unknown samples. The analysis software has the ability to convert the size of the unknown samples in base pairs to the size of the alleles in repeat number based off the allelic ladder (Applied Biosystems. 2004).

Technological advances have made it possible for capillary electrophoresis to be automated, which means numerous samples can be analyzed simultaneously. Many crime scenes provide only small amounts of biological sample, so the small capillary size is ideal, as only small sample quantities need to be consumed for results to be obtained. Unlike with agarose gels, the heat generated from the ionic movement is more evenly dispersed during capillary electrophoresis, which prevents distortion of the bands, and contamination risks are decreased since the samples remain within their own capillary. The automated process also eliminates the

need to estimate band size, as the computer analyzes the fragments as they are moved through the polymer, which makes the results highly reproducible (Butler et al., 2004). The accuracy, reproducibility, and high-throughput capabilities associated with capillary electrophoresis makes it an ideal methodology for forensic DNA analysis, and as such, is the standard analysis methodology for the forensic community.

Low Copy Number DNA Analysis

The advent of PCR analysis and the use of multiplexing techniques has allowed for the testing and examination of minuscule amounts of forensically-relevant biological samples (van Oorschot & Jones, 1997). Most commercial kit protocols state that optimum multiplexing is performed with approximately 1 ng of DNA, but much lower concentrations can also be analyzed (Sparkes et al., 1996). LCN DNA is typically defined in the literature as DNA samples containing less than 100 pg of DNA, or approximately 15 diploid cell copies, but new kit chemistries have demonstrated the ability to amplify reactions containing only 0.0062 ng of DNA (Wang et al., 2012). In theory, because one cell contains the entirety of the genome, an individual's genetic profile could be obtained from a single cell. However, when such a small amount of template is available, primer binding may occur more often at some loci over others during the first few rounds of amplification. The resulting profile will ultimately demonstrate in an imbalance among the alleles represented in the sample. This phenomenon is known as preferential amplification, and is most likely to benefit the lower molecular weight STRs.

The amplification of low concentrations of DNA can also result in other artifacts that are much more difficult to analyze, some of which are depicted in Figure 6. Allelic drop-out, for example, is when certain alleles that should be in the DNA profile may not be amplified during PCR, leading a heterozygous locus to falsely appear homozygous. In the most extreme cases of

preferential amplification, the entire locus can disappear from the profile (Gill, Whitaker, Flaxman, Brown, & Buckleton 2000). Allelic drop-in, which is often a result of contaminants, may also be present as a result of LCN DNA analysis. The peaks seen on the electropherogram as a result of allelic drop-in are usually not reproducible because this process occurs randomly, and is not based on the size of the repeat unit (Gill et al., 2000). Both allelic drop-out and drop-in are ultimately PCR phenomenon, so to ensure that the alleles being called in LCN DNA profiles are accurate and are not the result of PCR analysis, the samples are typically amplified two to three times (Gill et al., 2000). If the allele is consistently observed throughout all of the amplifications it may be included in the final profile, but if it appears inconsistently during analysis it must be excluded because it is unknown if the peak on the electropherogram is the individual's true allele (Gill et al., 2000). This type of analysis is not feasible for a forensic science protocol because forensic samples are often small, so there may not be enough sample available to perform repeat tests. Additionally, having to run triplicates of each sample would require increased amounts of both time and money.

Figure 6

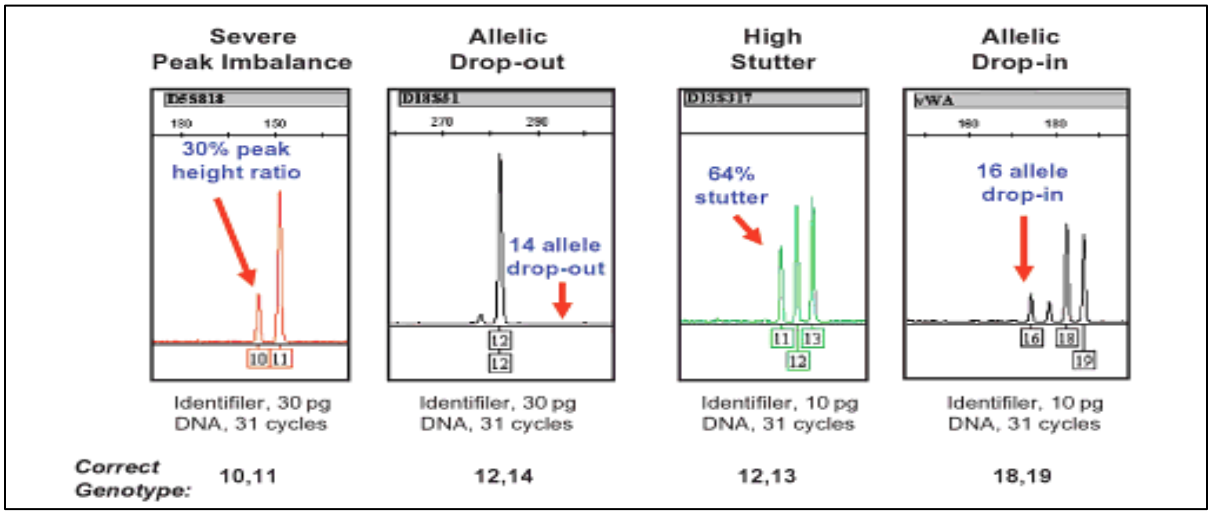

Figure 6: Stochastic effects of PCR

The different types of stochastic PCR effects commonly observed in forensic analysis. Peak height imbalance, allelic drop-out, high stutter, and allelic drop-in are all artifacts that may occur in a forensic profile (Butler, 2011)

Another common artifact seen in LCN DNA analysis is a phenomenon commonly referred to as stutter. During the process of DNA replication strand slippage may occur resulting in DNA fragments that are one repeat unit shorter, or in rare cases, one repeat unit longer than the true allele (SWGDAM, 2010). When the data analysis software interprets the fragment during capillary electrophoresis, it is seen as a smaller peak next to the true peak on a particular locus, and at optimal DNA concentrations are typically anywhere from 10-15% the size of the true allele. The amplification of low copy number samples tends to produce stutter peaks that are much higher than average and can exceed the height of the true peak, making it difficult to determine which allele is actually present in a forensic DNA profile (Gill et al., 2000).

The United Kingdom's Forensic Science service began a project in the late 1990s that utilized increased PCR cycle number to improve the detection of LCN samples (Gill et al, 2000). Several experiments were and have since been performed in which 34 amplification cycles were used instead of the standard 28, which increased the number of copies for each target sequence from 67 million to 4.3 billion (Butler, 2011). In some instances, increasing the number of amplification cycles led to the resolution of a more complete STR profile, and complete profiles were obtained using only 25-50 pg of DNA (Gill et al., 2000). Additionally, using an increased number of amplification cycles had no effect on the amount of stutter or the size of the peaks seen on the electropherogram (Kloosterman & Kersbergen, 2003). This protocol is the most common methodology used to analyze LCN samples in forensic laboratories today, and the potential results can be seen in Figure 7.

Despite the potential for 34 PCR cycles to generate a more complete profile, the protocol for an increase in PCR cycle number in unlikely to be used in a forensic setting. *Taq* DNA polymerase activity plateaus at higher cycle numbers, which results in a decrease of exponential

amplification (Lawyer et al., 1993). Due to the extreme sensitivity of the reaction, increased PCR cycles also leads to an increase in laboratory-based contamination, which must be considered during interpretation (Gill, 2001). Low-level amounts of DNA have been documented in reagents from laboratory personnel and from cross-contamination from other specimens. Contaminant DNA will amplify simultaneously with the target DNA, making it difficult to determine which peaks are contamination and which are true alleles. The negative control used in current DNA protocols acts to detect contamination within a reagent, but due to the sensitive nature of LCN DNA analysis, even the negative control has been shown to produce contaminant alleles, meaning the negative control cannot act in the same way for LCN protocols as it does for traditional protocols (Gill et al., 2000). The Quality Assurance Standards for Forensic DNA Testing Laboratories mandates that negative controls be concurrently run with all samples (Federal Bureau of Investigation, 2011). Since the negative control in LCN analysis is subject to contamination and therefore does not meet the quality assurance standards, other methodologies for LCN analysis must be pursued.

Figure 7

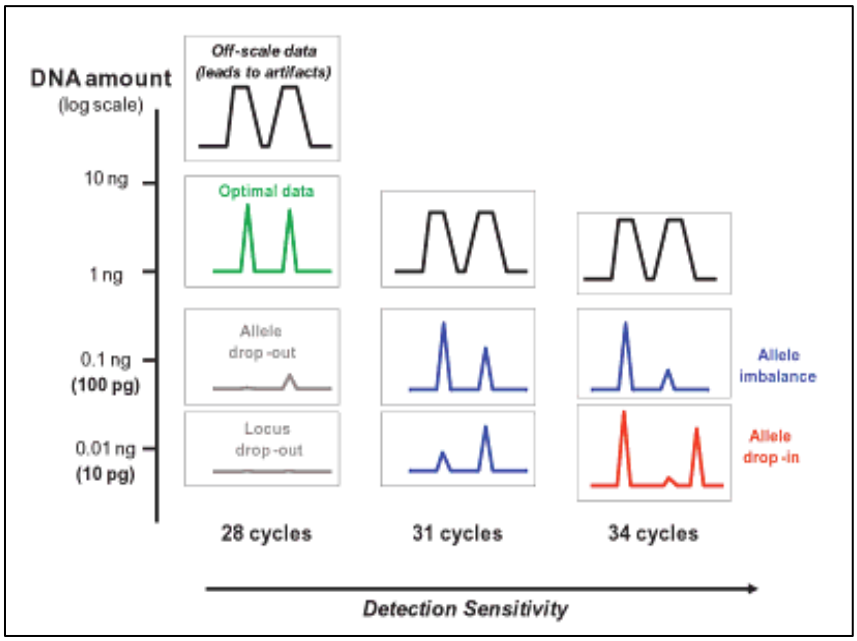

Figure 7: Hypothetical outcome of increased PCR cycle number and various amount of DNA
The potential outcome at a heterozygous locus after amplification and increased cycle number. At the standard 28 cycles, 1 ng of DNA provides optimal data and allelic and locus dropout are observed at the pictogram level. As cycle number increases, imbalance begins to occur and allelic drop-in is present (Butler, 2011)

Degraded DNA Analysis

Crime scenes are often exposed to both environmental and man-made insults which ultimately damage any recoverable DNA from biological samples. Degradation can occur in a variety of ways including oxidative damage, the crosslinking of purines, and through hydrolysis; all of which prevent the complete analysis of a forensic sample (Lindahl, 1993).

A predominate cause of forensic DNA degradation is UV irradiation, which alters the basic structure of the DNA (McNally et al., 1989). In nature, DNA is exposed to a highly energetic form of solar radiation known as UVB radiation. UVB radiation is readily absorbed by the thymine and cytosine nitrogenous bases, and causes photochemical reactions within the cell. These reactions ultimately result in the formation of cyclobutane pyrimidine dimers (CPDs), depicted in Figure 8, through the cycloaddition of C5-C6 double bonds of neighboring pyrimidines and (6-4) photoproducts (6-4PPs) by a cycloaddition of C5-C6 double bonds between a 5' pyrimidine and a 3' thymine (Revanat, Douki, & Cadet 2001). A second type of UV light, known as UVC was used to model degraded samples for this research. This type of UV radiation is absorbed by the atmosphere in nature, but has been shown in laboratory settings to replicate the effects of UVB radiation (Hall & Ballantyne, 2004). These artifacts prevent the addition of nucleotides by *Taq* DNA polymerase during PCR amplification, resulting in truncated products (Wellinger & Thoma, 1996). When these samples are analyzed through capillary electrophoresis, allelic dropout is observed, especially in the higher molecular weight loci (Hall & Ballantyne, 2004). The reduction in allele size and allelic dropout prevent the complete interpretation of a forensic DNA profile.

Figure 8

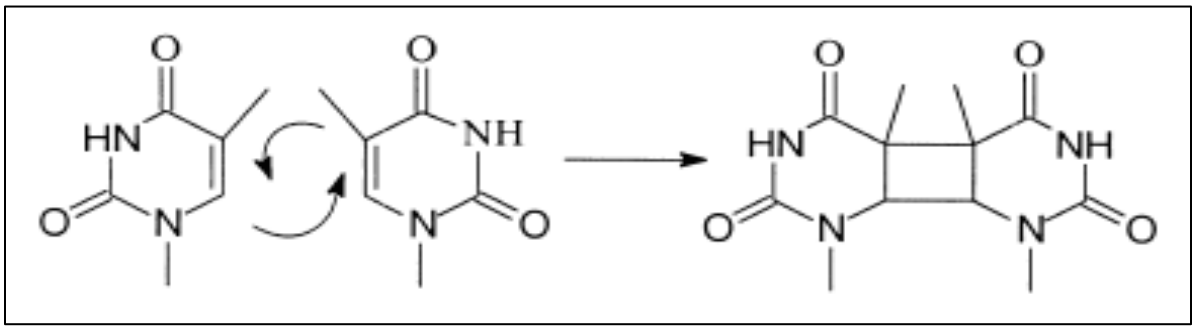

Figure 8: Thymine dimer

A visual on the formation of thymine dimers as a result of UV radiation (Revanat, Douki, & Cadet 2001)

Degradation has been a constant issues in forensic DNA analysis, and as such, several methodologies have been employed to alleviate its effects so an adequate profile can be produced. Some laboratories have turned to the analysis of the hypervariable regions of mitochondrial DNA (mtDNA) (Holland & Parsons, 1999). In itself, mtDNA testing is time-consuming, expensive, and unavailable in several laboratories. Additionally, the sequences analyzed are maternally inherited, and thus are not suitable for the identification of an individual to the exclusion of all others (Holland & Parsons, 1999). Other laboratories utilize commercially available kits like the AmpFISTR® MiniFiler® for the analysis of the core loci when the samples are highly degraded. These kits contain primers that anneal to the nuclear DNA much closer to the target repeat regions, effectively amplifying a smaller region (Mulero et al., 2008). However, these kits only analyze eight of the core loci and therefore do not provide a complete forensic profile (Applied Biosystems, 2012b).

The importance of degradation of DNA for the purpose of this research is that it can be used to model samples that contain low quantities of DNA. Previous studies have demonstrated that UV irradiation effectively models the inter- and intra- locus peak imbalance, allelic dropout, and locus dropout commonly observed in LCN samples. These studies have also shown that

preferential amplification was observed, implying that UV irradiation effectively reduces the concentration of the high-molecular weight loci in the sample (Pang & Cheung, 2007). For these reasons, UV irradiation was applied to known DNA samples in this research to model LCN DNA samples.

Normalization using Duplex-Specific Nuclease

DSN normalization of low-copy number and degraded DNA samples is a logical alternative to increasing the number of amplification cycles in order to obtain usable STR profiles. DSN is classified as a restriction endonuclease, a type of highly specific enzyme that is utilized by the cell to hydrolyze nucleic acids. Naturally occurring restriction endonucleases evolved to allow for changes in the genetic material to occur, such as DNA recombination, repair, and replication, and for cellular defense against viruses and digestion. Restriction endonucleases are routinely utilized for molecular biology research due to their predictable mechanisms and specificity (Nathans & Smith, 1975). DSN was first isolated from the hepatopancreas of the Kamchatka crab (*Paralithodes camtschaticus*) through the use of anti-DSN antibodies. DSN acts to selectively degrade the DNA strand of a DNA-RNA hybrid molecule, or the double-stranded DNA strands containing nine or more base pairs (Anisimova et al., 2008). As seen in Figure 9, this enzyme has been found to be highly thermostable, functioning over a temperature range of 20-80 °C, with optimal digestion occurring at 60 °C. The presence of a divalent cation such as Mg^{2+} is required for enzyme activity. Analysis has determined DSN activity is at 100% when cation concentration is approximately 20 mM. DSN has the capability to function over an extremely wide pH range, with maximum digestion efficiency occurring at a pH of 6.6. This enzyme displays a strong resistance to treatment with proteinase K, and is still active after incubation with chemicals such as 1% SDS and 10 mM β-

mercaptoethanol at 37 °C (Evrogen©, Product insert, #EA002). DSN is effectively inhibited through the use of EDTA, which acts to chelate divalent cations in the system (Garvan, 1964). These parameters have successfully been implemented in laboratory settings to incorporate DSN into research protocols.

Figure 9

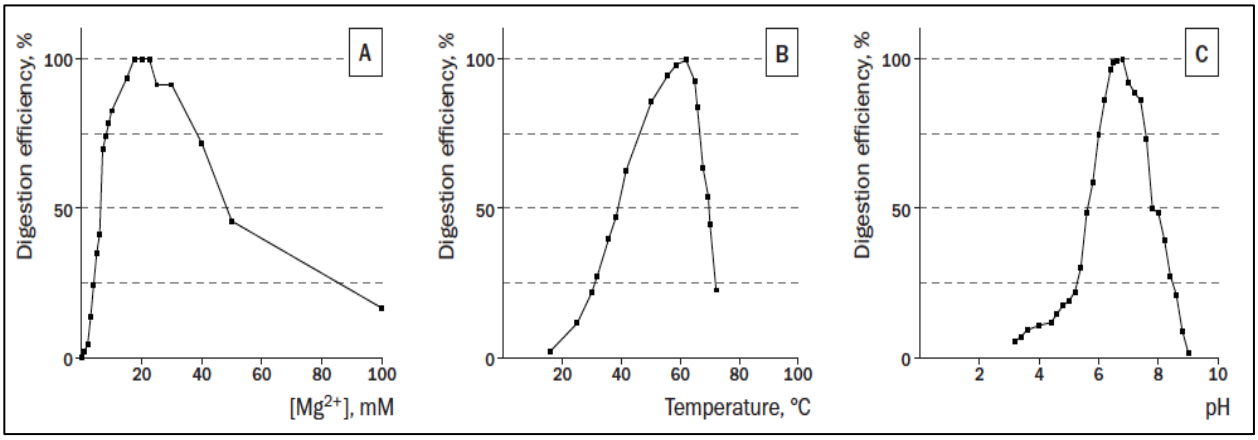

Figure 9: The effects of Mg^{2+}, Temperature, and pH on DSN activity

The wide range with which DSN is able to efficiently digest double-stranded DNA molecules. Optimum cation concentration is over 20 mM, optimal temperature is at 60 °C, and optimum pH is 6.6 (Evrogen©, Product insert, #EA002)

Research laboratories have used the enzymatic properties of DSN to cleave double stranded DNA to successfully normalize DNA samples during DNA sequencing (Shagina et al., 2010; Gijavanekar, Strych, Fofanov, Fox, & Willson, 2012; Yin, Liu, & Ye, 2012), for the detection of single nucleotide polymorphisms (SNPs), and for quantitative telomeric overhang determination (Anisimova et al., 2008). In the genome, there is a high concentration of highly conserved repetitive sequences and a fraction of low-copy number sequences (Shagina et al., 2010). DSN treatment is used to reduce the number of repetitive sequences and allows for the low-copy number sequences to be more equally represented (Shagina et al., 2010). Once DSN has normalized the strands, the DNA sample can be amplified though PCR to obtain adequate

concentrations for sequencing (Zhulidov et al., 2004). Research has shown that DSN can reduce the amount of the highly conserved sequences without altering the representation of the unique genes (Shagina et al., 2010). This same principle can be applied to previously amplified forensic DNA samples, in that DSN can act to selectively cleave the preferentially amplified and overly abundant low-molecular weight loci.

Figure 10

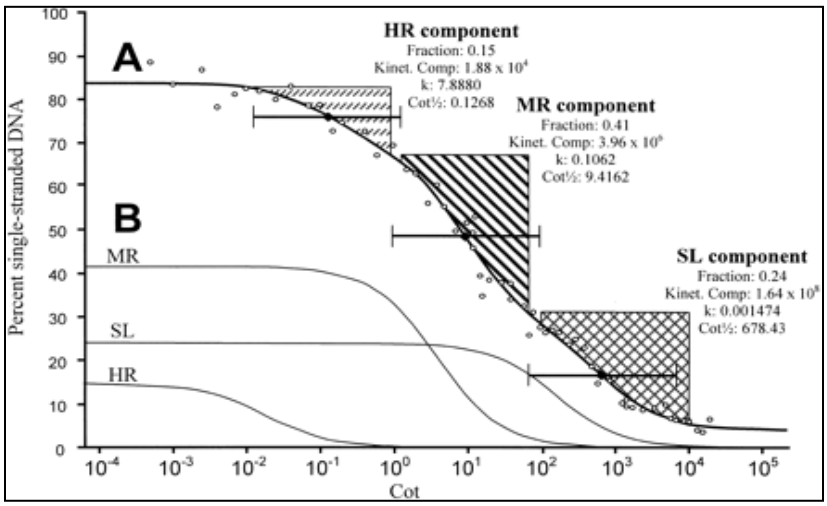

Figure 10: C_0T Curve for Eukaryotic DNA Reassociation

A C_0T curve for the reassociation of single stranded DNA fragments. Highly repetitive (HR) fragments find their complement quickly during reassociation, while single/low-copy fragments take much longer to anneal to their complementary strand (Peterson et al., 2002)

Normalization of DNA by DSN utilizes the disassociation and reassociation of denatured DNA during the first step of PCR amplification (Zhulidov et al., 2004). It is well known that some genetic loci are amplified preferentially during multiplex PCR, and therefore become overrepresented in the resulting PCR product. The overabundance of some genetic loci and amplification artifacts caused by this imbalance has been shown to be problematic for a forensic scientist during the interpretation of forensic profiles.

DSN can be used to help equalize the concentrations of each amplicon in the sample (Zhulidov et al., 2004). When there are several copies of an amplicon present, the denatured DNA strand is able to find its complementary strand and anneal to it quickly, as seen in Figure

10. The fewer copies present in the sample, the longer it will take for the reannealing of the complementary strands to occur. DSN specifically cleaves double stranded DNA (Shagin et al., 2002), and, therefore, when it is added to the PCR product following denaturing, it will only cleave DNA that was able to reanneal with a complementary strand. Since the more abundant copies will anneal first, the DSN activity will selectively degrade these overabundant amplicons (Zhulidov et al., 2004). In this manner, the number of copies of each amplified allele in the PCR reaction will be equalized. For forensic purposes, it is hypothetically possible to re-amplify the normalized sample to achieve DNA profiles free of preferential amplification of low-molecular weight loci. This hypothesis is depicted in Figure 11.

Figure 11

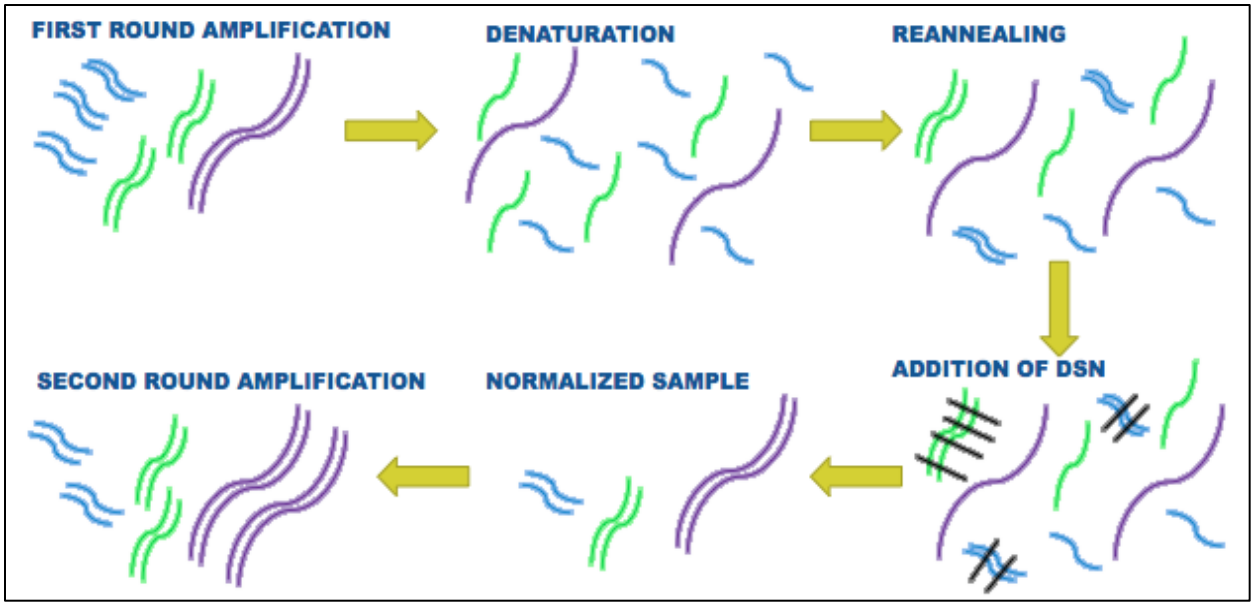

Figure 11: Normalization methodology

The hypothetical experimental design for this project. Sample DNA will be amplified and then denatured. Preferentially amplified low-molecular weight loci will find their sister strands first, and be selectively cleaved by DSN, resulting in a normalized sample. This sample will be re-amplified to produce a profile free of preferential amplification

This research will utilize the enzymatic capabilities of DSN to selectively degrade alleles that have been preferentially amplified as a result of multiplex PCR. As a result, the samples will be normalized prior to interpretation thus alleviating evidence of preferential amplification of the low-molecular weight loci. This will allow for equal representation of all loci analyzed and for a more complete forensic DNA profile to be generated.

Materials and Methods

This project was preformed at the University of Central Oklahoma in Edmond, Oklahoma. The university's Institutional Review Board (IRB) granted approval for the use of human subjects, and two consenting individuals were selected to provide buccal samples for analysis.

DNeasy® Blood and Tissue Kit

DNA from known donors was extracted from buccal swabs using the QIAGEN DNeasy® Blood and Tissue Kit (QIAGEN, Venlo, Limburg, Netherlands, #69504). Buccal swabs were placed in 180 μL of Buffer ATL and 20 μL of Proteinase K and incubated at 56 °C for 15 minutes. The swabs were placed in a spin-basket (Thermo Fisher Scientific, Waltham, MA, #14-223-062) and centrifuged on the Eppendorf 5430 Centrifuge (Eppendorf, Hamburg, Germany, Serial Number: 54272G011551) at 13,000 RPM for 1 minute to collect DNA from the swab. Equal concentrations of Buffer AL and ethanol were combined to make a stock solution, and 400 μL of solution was added to each sample. The entirety of the resulting solution was placed in a MinElute (QIAGEN, #28204) spin column and eluted into a 2 mL collection tube through centrifugation at 8000 RPM for 1 minute. The spin column was placed in a new collection tube and 500 μL of Buffer AW1 was added to the sample and was centrifuged at 8000 RPM for 1 minute and flow-through was discarded. The spin column was nested into a new collection tube with 500 μL of AW2 and centrifuged at 13200 RPM for 3 minutes and flow-through was discarded. Finally, the spin column was placed in a clean 2 mL collection tube and 200 μL of Buffer AE was allowed to soak on the membrane for 1 minute. Following the incubation, the samples were centrifuged for 1 minute at 8000 RPM to elute the DNA from the membrane. All samples were stored at 4 °C.

DNA Quantification

All samples were quantified using UV spectroscopy on a NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA, Serial Number: 9590) according to the manufacturer's recommended protocol. A sample volume of 1 μL of sample was placed on the stand and absorption was measured at 260 nm and 280 nm wavelengths. DNA concentration was measured and purity was determined by examining the 260/280 ratios provided by the instrument's NanoDrop 2000/2000c Software.

Each sample was also subjected to the Quantifiler® Human DNA Quantification Systems (Applied Biosystems, Carlsbad, CA, USA, #4343895) as a method for quantification. In order to build a standard curve, eight standards were run simultaneously with the unknown samples. Known standard DNA was created using 10 μL of 200 ng/μL of stock and 30μL of molecular-grade water. The stock solution was serially diluted to the following concentrations: 50.000 ng/μL, 16.700 ng/μL, 5.560 ng/μL, 1.850 ng/μL, 0.0620 ng/μL, 0.210 ng/μL, 0.068 ng/μL, and 0.023 ng/μL according to the Applied Biosystems protocol. Each reaction required 10.5 μL of Quantifiler Human Primer Mix and 12.5 μL of Quantifiler PCR Reaction Mix. Using these concentrations, enough master mix was created to analyze 30 samples, including the standards and reagent blank. The master mix was centrifuged briefly and 23 μL of the mix was aliquoted into each well of the MicroAmp® Optical 96-well Reaction Plate (Applied Biosystems, #N801-0560) and 2 μL of sample or standard was added to the appropriate well. Following addition of the master mix and sample, the reaction plate was covered with the MicroAmp® Optical Adhesive Film (Applied Biosystems, #4311971). The plate was placed on the Applied Biosystems 7500 Real-Time PCR System (Serial Number: 275001373).

Degraded Model Preparation

Extracted DNA was concentrated through the use of an Eppendorf Vacufuge Plus centrifuge (Serial Number: 5305AL010427), and UV spectroscopy was used to quantitate the samples. In order to model the effects of degradation on DNA, 100 μL volume samples were exposed to increasing durations of 254 nm UV light using the Spectro Linker XL-1500 UV Crosslinker (Spectrolinker, Westbury, NY, Serial Number: 1843457) ranging from 0 s, 300 s, 500 s, 700 s, and 900 s. The DNA concentration at each time point was once again determined using UV spectroscopy, and comparisons between pre and post exposure measures were made.

AmpFISTR® Identifiler® Plus PCR Amplification

Amplification Set Up

Amplification of all DNA samples was preformed using the AmpFISTR® Identifiler® Plus PCR Amplification Kit (Applied Biosystems, #4427368). Each reaction required 10μL of AmpFISTR® Identifiler® Plus Master Mix and 5 μL of the AmpFISTR® Identifiler® Plus Primer Set. The number of reactions, including the reagent blank, positive control, and negative control were determined to make the appropriate amount of master mix. Each sample was placed into a standard PCR reaction tube, and each sample amplification reaction volume consisted of 15 μL of master mix, 9 μL of molecular-grade water, and 1 μL of 1ng/μL DNA. A positive control was made using 15 μL of master mix and 10 μL of 9947A DNA. The negative control contained 15μL of master mix and 10 μL of molecular-grade water. The PCR tubes were placed on the GeneAmp® PCR System 9700 thermal cycler (Applied Biosystems, Serial Number: 80558201803).

Thermal Cycling Parameters

Proper cycling conditions required that the GeneAmp® PCR System 9700 be set to the 9600 Emulation Mode, and was done so for all samples. Applied Biosystems recommended that the standard 28-PCR-cycle protocol be run on samples containing the optimal 1 ng of DNA. The initial incubation step called for a hold at 95 °C for 11 minutes. The next 28 cycles consisted of a denaturation of 94 °C for 20 seconds, and an annealing/extension step at 59 °C for 3 minutes. A final extension for 10 min at 60 °C was preformed and the samples were stored at 4 °C until analysis.

AmpFISTR® Identifiler® Plus Genetic Analyzer

Analysis of the STR profiles was performed using the Applied Biosystems 3130 Genetic Analyzer (Serial Number: 21364-025) through capillary electrophoresis. A master mix was created using 8.7 μL of Hi-Di ™ Formamide (Applied Biosystems, #4311320) and 0.3 μL of GeneScan™ 500 LIZ™ (Applied Biosystems, #4322682) for each sample that needed analysis. The master mix was aliquoted in 9 μL volumes into a MicroAmp® Optical 96-well Reaction Plate, and 1 μL of amplified sample was added to each well according to the plate map. The plate was centrifuged briefly to ensure all sample was at the bottom of the well, and then loaded onto the Applied Biosystems 3130 Genetic Analyzer.

STR Data Analysis

GeneMapper® ID software version 3.2 was employed for fragment and allele analysis. Allele calls were based off the migration rate through the capillary relative to the internal lane standard and allelic ladder. An analytical threshold of 25 RFUs was set in order to visualize loci during all stages of analysis. Electropherograms were produced for all samples and analyzed for the presence of the 15 STR loci and the amelogenin sex-determining locus. A positive control

profile was analyzed with each run, and unknown samples were only analyzed if all control results were as expected.

DSN Treatment

Normalization of degraded DNA was concentrated on the samples exposed to 300 s of UV treatment as these best represented DNA commonly found at a crime scene. Lyophilized DSN enzyme (Evrogen©, Moscow, Russia, #EA002) was rehydrated through the addition of 50 μL DSN storage buffer (50 mM Tris-HCl, pH 8.0) and then briefly centrifuging the tube. The tube was incubated at room temperature for 5 minutes and an additional 50 μL of glycerol was added to the rehydrated enzyme. The entire mixture was spun briefly in a centrifuge and the reconstituted DSN was stored at -20 °C until use.

Following PCR amplification of the degraded DNA samples, 3 μL of 10x DSN master buffer (500 mM Tris-HCl, pH 8.0, 50 mM $MgCl_2$, 10 mM DTT) was mixed thoroughly with 24 μL of amplified product. The samples were incubated on the GeneAmp® PCR System 9700 for 5 minutes at 99.9 °C to ensure complete denaturation of all DNA in the samples. The samples were then allowed to cool to 80 °C in order to allow the highly repetitive sequences to begin reannealing to one another. Once the temperature reached 80 °C, 1 μL of DSN enzyme was added to each sample. The samples were allowed to further cool to 62 °C, the optimal temperature for nuclease activity, and the enzyme was allowed to selectively cleave those strands that had reannealed for 1 minute. The samples were then heated to 95 °C for 2 minutes to decrease the efficiency of the enzyme's activity, and 28 μL of 2x DSN stop solution (10 mM EDTA) was added to prevent further digestion of sample DNA. Following normalization, all samples were placed on ice, and 2 μL of each sample was analyzed on the Applied Biosystems 3130 Genetic Analyzer.

MinElute® Reaction Cleanup Kit and Second PCR Amplification

The DSN reaction was stopped via the addition of EDTA, which acts to chelate divalent cations necessary for PCR reactions. As a result, each sample underwent purification to remove both enzyme and stop solution before the second amplification could occur. The MinElute® Reaction Cleanup Kit Protocol (QIAGEN, #28204) was used to purify the double-stranded DNA remaining after the DSN digestion.

In order to purify the samples, 54 μL of the remaining sample was mixed with 300 μL of Buffer ERC and placed into a MinElute column and 2 mL microcentrifuge tube. The sample was centrifuged at 13,200 RPM for 1 minute to bind the DNA to the MinElute column. All flow-through was discarded and the MinElute was placed back into the microcentrifuge tube. The DNA was washed with 750 μL Buffer PE to the MinElute column and centrifuging the samples for 1 minute at 13,200 RPM. Flow-through was discarded and the column was centrifuged for an additional minute at 13,200 RPM to remove any remaining wash buffer. The MinElute was placed into a clean 2 mL microcentrifuge tube and 10 μL of Bufer EB (10 mM Tris-Cl, pH 8.5) was allowed to soak on the MinElute membrane for 1 minute. The MinElute column was centrifuged at 13,200 RPM for 1 minute to elute the DNA. Analysis of the purified product was performed using 2 μL of sample on the Applied Biosystems 3130 Genetic Analyzer. Following purification, second round of PCR was preformed. This protocol used 15 μL of master mix, 2 μL of molecular-grade water, and 8 μL of purified sample, and the thermal cycler parameters were set as previously described.

Split-Amplification Protocol

In order to stay within the standard 28 amplification cycles commonly employed in forensic laboratories, a protocol was created in which normalization occurred midway through

the first round of PCR. Sample extraction was preformed as previously described. Amplification setup using the Identifiler® Plus PCR Amplification system remained the same, as did all thermal cycler parameters. After the initial denaturation and the first 20 cycles, amplification was paused and the samples were treated using the DSN protocol previously described. All normalized samples were purified using the MinElute® Reaction Cleanup Kit Protocol detailed above, and all results were analyzed on the Applied Biosystems 3130 Genetic Analyzer.

Results

Artificially degraded genomic DNA was analyzed by STR analysis to determine if the PCR artifacts observed in degraded and LCN evidentiary DNA samples were similar to those generated under laboratory conditions. Five artificially degraded DNA samples were then exposed to the normalization treatment by DSN and peak height ratios from resulting electropherograms were analyzed to determine the effectiveness of the protocol.

DNA Quantification

Extracted DNA from two known donors was quantified to ensure adequate concentration of genetic material was present for all subsequent analysis. The extracted DNA samples were subject to both UV Spectroscopy and Quantifiler® Human DNA Quantification. Theoretically, the quantification values should not be affected by UV exposure. UV spectroscopy measures the 260 nm absorption by the bases, not the backbone of the DNA. The DNA sequence analyzed by the Quantifiler® methodology is only 94 base pairs long, which means it is unlikely to be significantly affected by degradation. As expected, the quantitation results for both methodologies were consistent with one another, as displayed in Table 1.

Table 1

	Male UV Spectroscopy (ng/μL)	Male Quantifiler® (ng/μL)	Female UV Spectroscopy (ng/μL)	Female Quantifiler® (ng/μL)
Measurement 1	2.2	N/A	1.8	N/A
Measurement 2	2.9	N/A	1.2	N/A
Measurement 3	2.4	N/A	1.6	N/A
Average:	2.5	2.9	1.5	1.7

Table 1: DNA Quantification Using UV Spectroscopy and Quantifiler®

Table 1 shows the concentration of extracted male and female DNA prior to degradation. DNA concentration measured through UV spectroscopy was done in triplicate, and the average was used for subsequent calculations. The Quantifiler® methodology determines the DNA concentration by comparison to a standard curve, so analysis was performed only once.

Due to the low template measured after extraction, the DNA was concentrated by vacuum centrifugation. The DNA was then exposed to UV light to model degradation and the concentration was again determined by UV spectroscopy after each exposure time. Analysis was run in triplicate, and the average concentration was determined, depicted in Table 2. Average concentration values were used in all subsequent calculations.

Table 2

UV Exposure (seconds)	Male Before UV Light Exposure (ng/μL)	Male After UV Light Exposure (ng/μL)	Female Before UV Light Exposure (ng/μL)	Female After UV Light Exposure (ng/μL)
0	6.7	N/A	5.2	N/A
300	6.5	6.3	6.0	4.6
500	7.0	5.6	5.6	4.6
700	6.9	6.3	6.3	5.3
900	7.0	7.4	5.7	6.2

Table 2: Quantification of concentrated DNA before and after UV irradiation

Table 2 depicts the concentration of the DNA measured by UV spectroscopy before and after exposure to UV irradiation. Values shown are the average of three measurements, and demonstrate that DNA concentration does not change after exposure to UV irradiation. The quantitation values were consistent across all UV exposure times and were not considered statistically different.

Modeling Degraded DNA Evidence

All five DNA samples degraded through UV irradiation were amplified using the Identifiler® Plus PCR Amplification and analyzed on the Applied Biosystems 3130 Genetic Analyzer. Figure 12 illustrates the effects of UV light exposure and its ability to degrade DNA after 0 s, 300 s, 500 s, 700 s, and 900 s. Manual analysis of the electropherogram obtained from the 300 seconds exposed DNA samples demonstrated allelic imbalance, a significant reduction in peak heights, and dropout of the high-molecular weight loci. These traits, seen in Figure 13, best modeled the effects of degraded and LCN DNA amplification documented in literature, and therefore the remainder of the experiment was performed using sample exposed to 300 seconds of UV light.

Figure 12

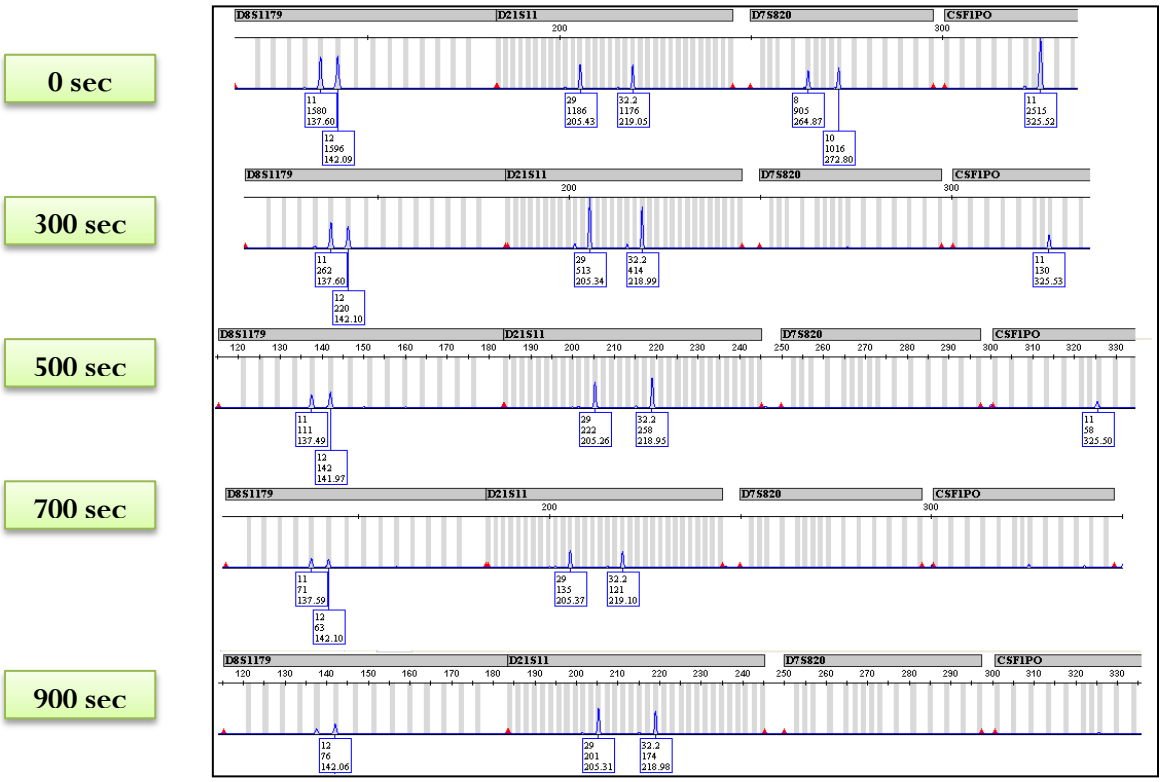

Figure 12: Effects of UV irradiation on sample DNA

The effects of UV irradiation on DNA. As exposure time is increased, peak heights are significantly reduced and high-molecular weight loci begin to drop out of the profile.

Five male and five female samples were subject to degradation through 300 seconds of UV irradiation and the subsequent profiles were analyzed to determine if the insult affected the samples consistently and reproducibly. DNA from both donors exhibited dropout at the D7 locus first, and, with the exception of the blue channel, ski slopes were constantly present throughout the various dye channels. The reproducibility of the degraded profiles indicated that UV irradiation was a suitable methodology for modeling degraded and LCN DNA profiles.

Figure 13

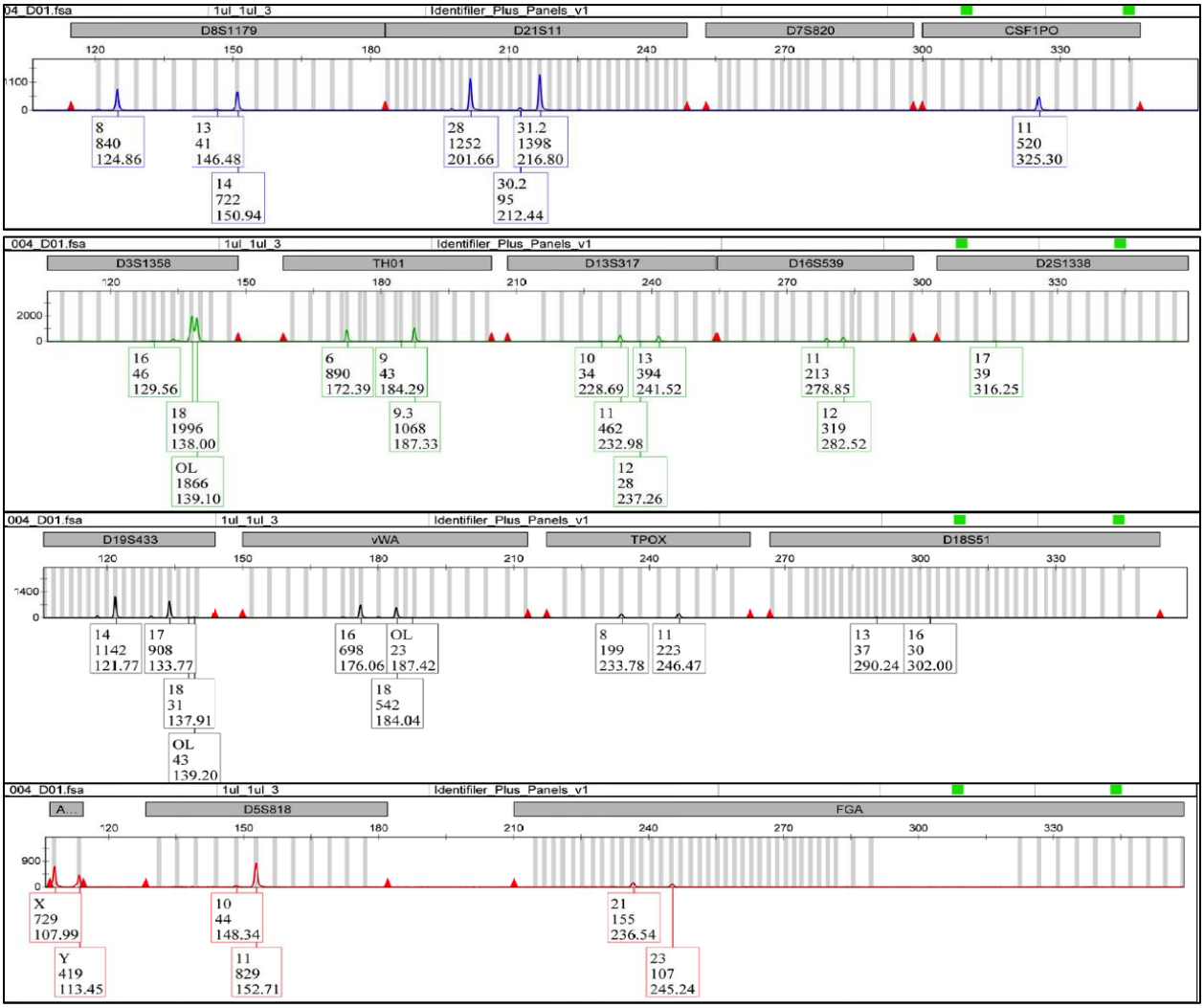

Figure 13: Sample DNA exposed to 300 seconds of UV irradiation

The effects of 300 seconds of UV irradiation on sample DNA across the four dye channels. Dropout of the high-molecular weight loci is observed across the entire profile.

DSN Treatment

Samples exposed to 300 seconds of UV light treatment were subjected to DSN normalization. Following sample normalization, 2 μL was removed for analysis on the Applied Biosystems 3130 Genetic Analyzer. Figure 14 demonstrates the effects and average outcome of

DSN treatment on DNA. The analytical threshold was set to 25 RFUs in order to visualize the majority of the alleles and to determine the extent of degradation across the profiles.

Figure 14

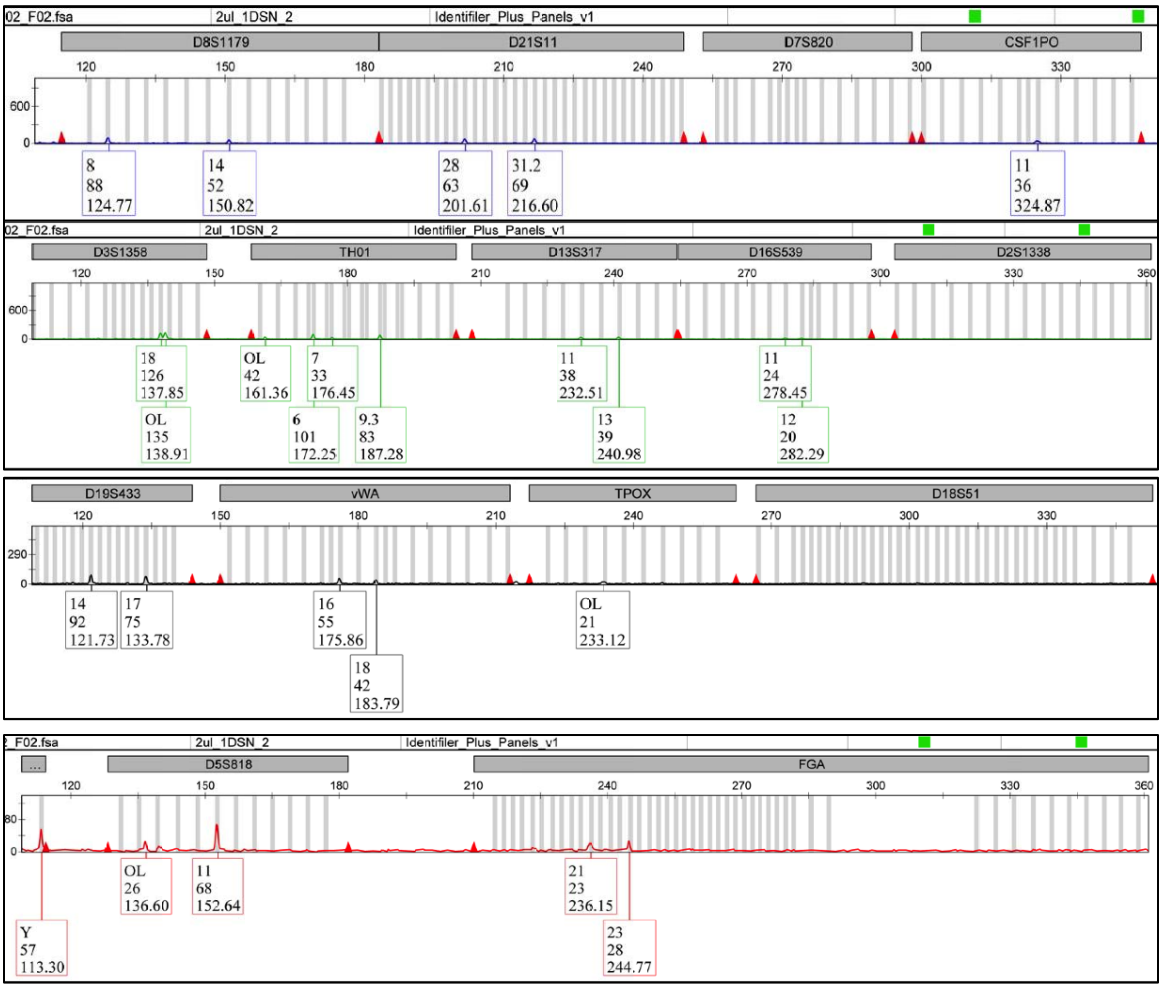

Figure 14: DNA after DSN Treatment

The samples depicted in Figure 14 show a severe reduction in peak heights and significant dropout of the loci. The majority of the loci are not above the previously validated standard analysis threshold of 95 RFUs

Purification

The DSN enzyme is able to function under denaturing conditions and EDTA acts to inhibit PCR, so all samples underwent spin-column purification prior to the second amplification step. Purified samples were analyzed on the Applied Biosystems 3130 Genetic Analyzer. The results are shown in Figures 15.

Figure 15

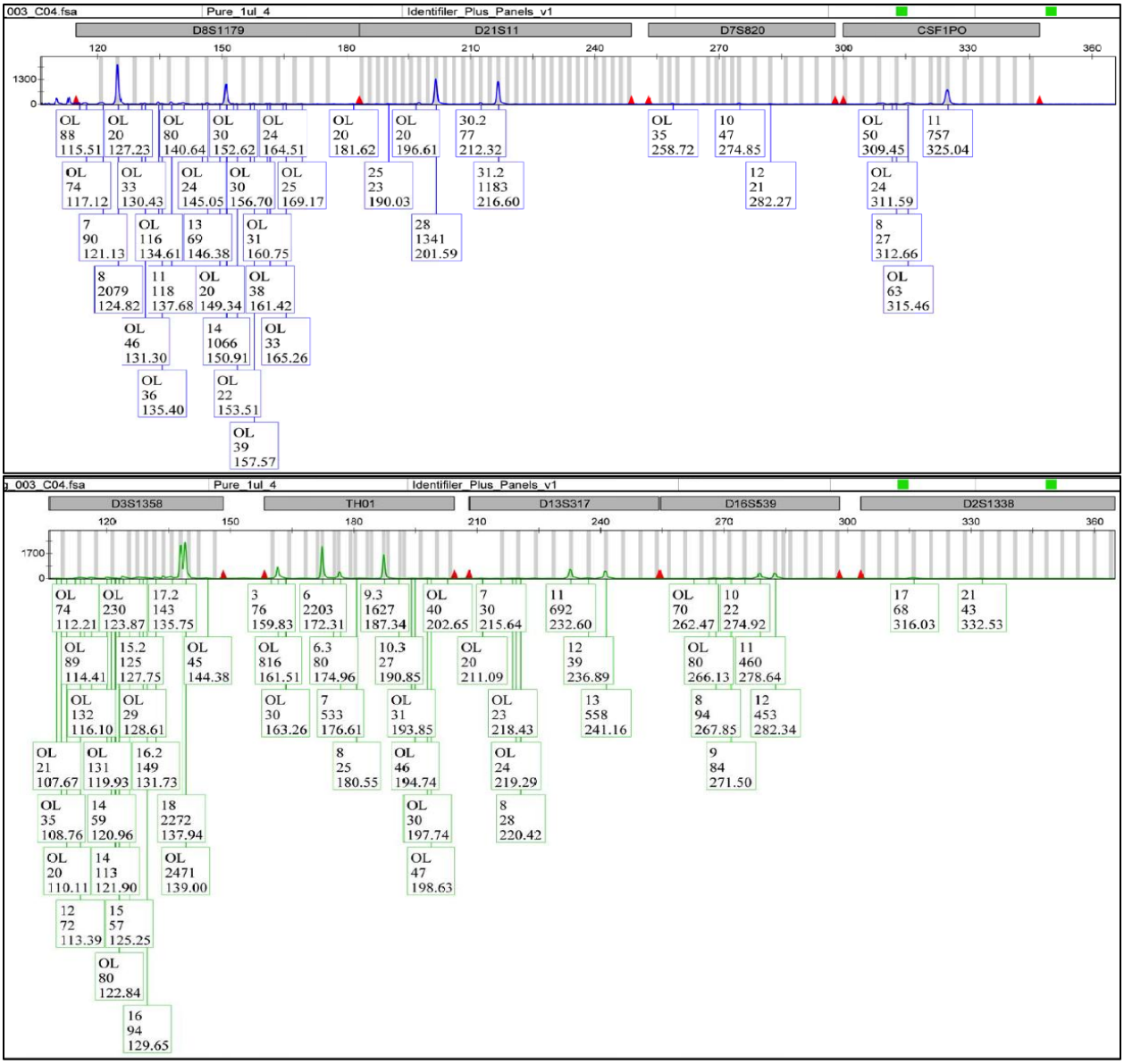

Figure 15: Blue and green channels post-purification

The blue and green channels shown in Figure 15 depict DNA samples following the purification process. A substantial increase in peak height ratios is observed compared to the samples analyzed after DSN treatment. Also depicted is an increased amount of artifacts, primarily off-ladder alleles.

The yellow and red channels of the purified DNA samples exhibited similar artifacts to those depicted above. The number of off-ladder alleles made analysis difficult, but the true peaks have much greater peak heights than the surrounding artifacts.

Second-Round Amplification

Purified samples underwent a second round of 28-cycle amplification and were analyzed by capillary electrophoresis. The resulting electropherogram is shown in Figure 16. The majority of peaks in all channels were off-ladder, and no allelic determinations were made.

Figure 16

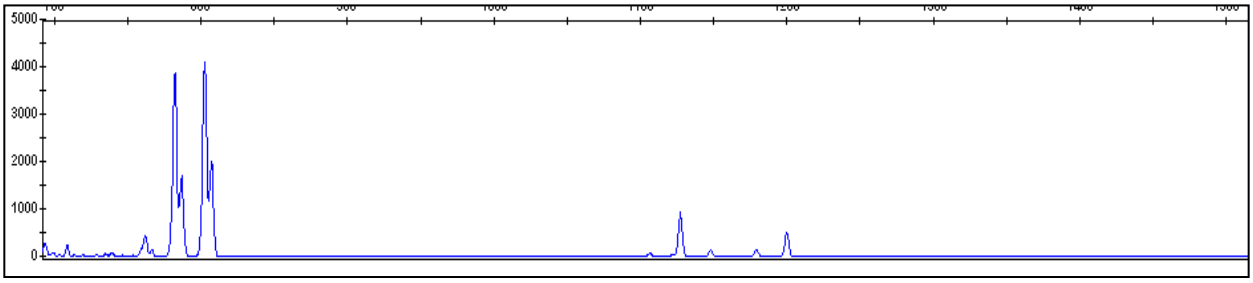

Figure 16: DNA samples after a second round of amplification

Results after the amplification of the purified product. No true loci are present on the electropherogram.

Split-Amplification

In order to maintain the standard 28-cycle forensic PCR methodology, DNA samples were allowed to amplify for 20 cycles, were normalized using DSN, and then amplified for another 8 cycles. The electropherogram results after the complete 28-cycle protocol are shown in Figure 17. The severe dropout prevented analysis and this methodology was not pursued.

Figure 17

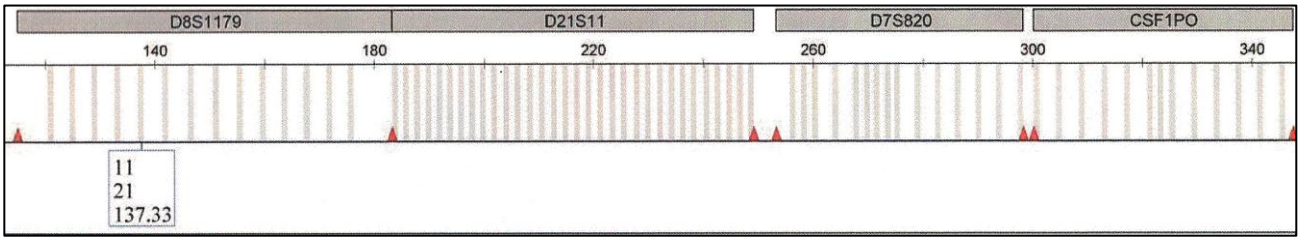

Figure 17: Samples after the split-amplification protocol

The electropherogram after the split-amplification protocol. The majority of the loci had dropped out of the profile, and those that were present were well below the defined threshold.

Analysis of Purified Samples

Before DSN treatment, the peak heights of each allele were divided by the tallest allele in the profile in order to establish a pre-normalization baseline. This allele was chosen in order to demonstrate normalization across the entire profile and not just in each channel. All five samples were analyzed and the values averaged. Post-DSN treatment peak height ratios were calculated in the same manner and compared to the pre-DSN treatment ratios. The percentage change was determined, as shown in Table 3. Alleles with the greatest percentage change demonstrated the most normalization relative to the tallest peak height.

Table 3

Locus	Before DSN Treatment		After DSN Treatment		Percentage Change	
	1st Allele	2nd Allele	1st Allele	2nd Allele	1st Allele	2nd Allele
D8S1179	41.3 +/- 2.5	32.5 +/- 4.9	83.4 +/-17.3	41.6 +/-4.7	42.1	9.1
D21S11	54.2 +/- 7.1	57.4 +/- 8.1	58.8 +/-6.5	58.6 +/-3.6	4.6	1.2
D7S820	1.4 +/- 0.1	1.1 +/- 0.7	N/A	N/A		
CSF1PO	21.4 +/- 4.7	Homozygous	40.4 +/-10.5	Homozygous	19	
D3S1358	96.7 +/- 5.2	96.3 +/- 2.3	86.7 +/-8.1	90.9 +/-5.0	-10	-5.4
TH01	49.3 +/- 9.1	46.1 +/- 5.0	79.8 +/-12.0	74.0 +/-13.5	30.5	27.9
D13S317	20.4 +/- 1.7	19.7 +/- 3.2	33.2 +/-4.5	25.4 +/-3.4	12.8	5.7
D16S539	12.2 +/- 2.4	11.2 +/- 1.8	26.9 +/-7.7	21.8 +/-7.5	14.7	10.6
D21S1338	2.3 +/- 0.6	2.2 +/- 0	N/A	N/A		
D19S433	54.1 +/- 3.6	49.6 +/-2.2	68.8 +/-9.8	60.3 +/-5.9	14.7	10.7
vWA	33.6 +/-3.3	28.0 +/-1.5	44.1 +/-5.9	35.1 +/-3.1	10.5	7.1
TPOX	12.1 +/-1.5	10.3 +/- 1.5	15.4 +/-0.7	15.6 +/-1.2	3.3	5.3
D18S51	2.7 +/-0.6	1.4 +/-0.2	N/A	N/A		
AMEL	39.5 +/-4.8	22.3 +/-1.6	75.3 +/-9.5	44.7 +/-7.2	35.8	22.4
D5S818	39.2 +/-3.1	Homozygous	10.4 +/-5.3	Homozygous	-28.8	N/A
FGA	7.1 +/-0.7	6.4 +/-1.6	20.7 +/-7.2	14.4 +/-5.5	13.6	8

*Peak height ratio was calculated as (RFU of allele) / (RFU of tallest peak in the profile) x 100.

Table 3: Peak height ratios and percentage change before and after DSN treatment

Table 3 depicts the average peak height of each allele relative to the tallest peak in the profile. This methodology demonstrates the effect of UV irradiation on pre-DSN treated samples and the improvement and recovery of alleles on post-DSN treated samples.

In order to determine the effects of locus characteristics on normalization potential, locus class, repeat difference, the presence of a mobility modifier, and molecular weight were analyzed and compared to the percent improvement value of the first allele. Table 4 provides the results.

Table 4:

Locus	1st allele % Improvement	Class	Repeat Difference	Mobility Modifier	Molecular Weight
D8	42.1	Compound Nonconsensus	1	No	139, 143
TH01	30.5	Simple Nonconsensus	3.75	No	172, 187
CSF	19	Simple	Homozygous	Yes	325
D13	12.8	Simple	2	No	217, 225
D16	14.7	Simple	1	Yes	276, 280
D19	14.7	Compound	3	No	126, 138
FGA	13.6	Compound Nonconsensus	2	No	230, 238
D21	4.6	Complex	3.75	Yes	206, 221
vWA	10.5	Compound Nonconsensus	2	No	176, 184
TPOX	3.3	Simple	3	No	225, 237
D5	-29	Simple	Homozygous	No	150

Table 4: Percent improvement of the first allele, class, repeat difference, presence of a mobility modifier, and molecular weight of each allele

By looking at repeat difference and molecular weight, and comparing those factors to the percentage values of the first allele, it becomes evident that several factors contribute to the normalization capabilities of a particular locus. Those with fewer repeat differences and a higher molecular weight demonstrated the most improvement, while those with larger repeat differences had a decreased probability of normalization.

Table 4 gives indication that several factors affect the ability of a particular locus. The results above were analyzed to determine if any patterns were present among the locus characteristics that could allow for the prediction of normalization in future profiles. It is recommended that further studies with different donors be done in order to determine if the findings remain consistent.

Discussion

Modeling Degraded DNA Evidence

It is possible to artificially simulate degraded and LCN DNA samples through exposure to UV irradiation. When DNA is exposed to UV light, dimers between adjacent bases are created, which causes morphological changes to the overall structure of the DNA. Theoretically, those loci with a high molecular weight are more likely to be damaged because there is a larger area available for the insult to act upon. This research found that in addition to molecular weight, the repeat sequence was a factor in the severity of degradation at a particular locus. UV irradiation acts to create dimers between adjacent thymine bases or between adjacent thymine and cytosine bases, meaning repeats that had TT or CT in the sequence were more likely to be affected by degradation. For example, D7 dropped completely out of the first profile after only 300 seconds of UV exposure and CSF had significantly lower peak heights. D7 contains a simple repeat sequence of GATA, while CSF is comprised of simple AGAT repeats. These repeats mean their complementary strands contain high numbers of CT pairings where dimers are likely to be created. D8 and D21 also have many CT pairings, but due to the smaller molecular weight, were not as affected by degradation. During PCR amplification, *Taq* DNA polymerase is unable to replicate past the point where the dimers are located on the DNA strand, which results in truncated products. Analysis of the PCR products shows a reduction of peak heights, inter- and intra-locus peak imbalance, allelic dropout, and the preferential amplification of low-molecular weight loci, which is consistent with DNA profiles of degraded and low-copy number samples.

The quantification of DNA in extracted samples is performed in order to determine how much usable genetic material is available for analysis. This reserach utilized two methodologies

to determine if consistent results were achieved. It was hypothesized that UV irradiation would not alter the overall concentration of the sample DNA. UV spectroscopy measures the 260 nm absorbance of the bases in the DNA sample. While irradiation alters the backbone of the DNA molecule, nothing is lost from the sample, so absorbance levels remain the same before and after treatment. The Quantifiler® kit measures the amount of hTERT copies in a given sample and uses that data to determine the entire sample concentration. This sequence is only 94 base pairs long, and as a result, is rarely affected by degradation. This demonstrates the ineffectiveness of quantification to predict the outcome of a forensic DNA profile. Despite the indications of intact and pure DNA provided by the two different methodologies, analysis following PCR exhibited degradation artifacts and allelic dropout. Since the degradation approach of this study had little to no effect on the elements measured in both quantification methodologies, the overall DNA concentrations remained consistent and the values were utilized for subsequent DNA amplification.

Normalization of DNA Evidence

DSN normalization was performed to investigate the enzyme's ability to alleviate preferential amplification by selectively cleaving the double-stranded low-molecular weight loci. Following treatment, the samples were analyzed to determine the effects DSN had on the amplified product. Capillary electrophoresis produced results showing the dropout of the majority of the alleles, while those that remained had peak heights that were well below the validated threshold. The dramatic decrease in RFUs can be attributed to the fact that the samples were very dilute at this stage in the study due to the addition of DSN, buffer, and stop solution. As a result, no calculations or comparisons were performed.

Trial experimentation indicated that DSN would reactivate and continue to cleave DNA at temperatures similar to those in the annealing and extension phases of PCR. In order to prevent complete degradation of product, the enzyme was removed before the second round of amplification was preformed. The DSN treatment protocol called for the addition of EDTA, a common PCR inhibitor, to prevent further enzyme activity. Both enzyme and EDTA were removed using the MinElute® Reaction Cleanup Kit, a product specifically designed to purify enzymatic reactions. This methodology takes advantage of DNA's tendency to bind to silica membranes under high-salt conditions as contaminants are washed through. The DSN enzyme and EDTA are removed, preventing additional cleavage and inhibition, and elution of the purified product occurs through the addition of low-salt buffer.

Analysis of purified samples through capillary electrophoresis showed recovery of lost alleles, increased RFUs at all loci, and improved inter- and intra-locus peak balance. It was hypothesized that the purified product could be amplified for another 28 cycles to achieve consistent peak height ratios across all channels and alleviate dropout of the target loci. The purified samples were amplified a second time using the Identifiler® Plus PCR Amplification system, and the results were analyzed by capillary electrophoresis. The electropherograms showed an increase in off-ladder alleles and stutter peaks, but the entirety of the known DNA profile had either dropped out or was unable to be detected using the GeneMapper® ID software. Since this protocol did not produce viable results, interpretation and measurement of normalization was done using the results from the purified samples.

Standard protocol for the analysis of forensic DNA profiles calls for a single 28 or 29 cycle amplification (Applied Biosystems, 2012a). In order to remain consistent with standard forensic procedure, a split-amplification was performed with normalization occurring after 20

cycles of PCR. As before, capillary electrophoresis was done at each stage of analysis, but it was found that after only 20 cycles of PCR there was not enough DNA present to reach the analysis threshold, and no baseline calculations could be done. Analysis after the final 8 cycles showed numerous artifacts, off-ladder alleles, and complete dropout of the majority of the loci. As a result, no mathematical calculations were performed and this methodology was no longer investigated.

Normalization of alleles was determined through the examination of peak height ratios. The height of each allele was divided by the peak height of the tallest allele, D3, in the profile. This allowed for the measurement of normalization across the entire profile. The largest peak in this experiment was homozygous, which acted to alleviate some of the other variables during normalization analysis. As described later, homozygous loci normalized based solely on molecular weight, which provided a baseline for what other factors could affect the normalization of different loci. Table 3 shows that 21 of 30 alleles exhibited some improvement with their relative peak heights, but overall normalization was inconsistent among the loci.

Following mathematical interpretation, locus characteristics were analyzed to determine why some loci responded more to the normalization process than others. Each locus can be classified as either simple or compound based on its repeat motif. In this profile, D8 and TH01 demonstrated the most improvement relative to the tallest allele in the profile, but have opposite repeat motifs, indicating other elements determined locus normalization potential. In order to resolve which factors affected a particular locus' response to normalization, further analysis was performed examining the number of repeat differences between the alleles at each locus, the molecular weight of each allele, and whether or not a mobility modifier was added during

amplification. These factors combined with normalization percentages were ultimately used to determine what caused each locus' response to DSN treatment.

For example, the loci FGA and vWA can both be classified as compound, nonconsensus loci, meaning their sequences are made of two or more simple repeats. During amplification, each locus lacked a mobility modifier, and, in the profile studied, both contained two repeat differences between the sister alleles. Of the two, FGA had a significantly higher molecular weight and its alleles showed a much larger increase in peak height relative to the greatest allele in the profile compared to vWA. The purpose of this study was to achieve peak height balance across the channels and reduce the effects of preferential amplification on high-molecular weight loci. In this example, the locus with the higher molecular weight showed the most improvement, demonstrating the effectiveness of this procedure. However, D8, another compound, nonconsensus locus, demonstrated the most improvement of all loci analyzed, despite its relatively low molecular weight. Unlike FGA and vWA, D8 had only one repeat difference between the alleles, which indicated that repeat difference was the most important factor for normalization, followed by molecular weight.

Loci with simple repeat motifs were analyzed to determine if their normalization patterns were consistent with those loci that had compound motifs. Locus D16 was the only simple locus with one repeat difference between the sister alleles and showed the most peak height improvement relative to the tallest allele in the profile. Both D13 and TPOX are also classified as simple but have higher molecular weights. In this profile, D13 had a two repeat difference between the alleles while TPOX had three, and much like the compound loci, D13, the locus with the fewer repeat differences, demonstrated the most normalization. Both CSF and D5 were homozygous for the donor, but CSF had a significantly higher molecular weight. As expected,

CSF demonstrated the most normalization. The only simple repeat to not follow this pattern was TH01, which responded well to the normalization process despite having a low molecular weight and a large repeat difference. This could be attributed to the fact that, unlike the other simple loci in the profile, it contains a nonconsensus sequence which may reanneal more exactly than other simple loci. This locus cannot be compared to the other simple, nonconsensus loci because D18 and D7 completely dropped out from the profile. Other profiles may not exhibit this behavior at these loci and should be examined to determine if this anomaly remains the same.

The results indicated that multiple factors contribute to the ability of a particular locus to become normalized. Repeat difference was the major factor for allelic normalization, with those alleles with fewer differences demonstrating the most improvement. Molecular weight of the loci was determined to be the second most important factor in normalization. When few repeat differences were observed, those loci with a higher molecular weight achieved normalization better than low molecular weight loci with the same number of repeat differences. The presence of a mobility modifier did not seem to be a direct component in the ability of a particular locus to normalize, except to contribute to the molecular weight of some of the loci.

One hypothesis for the effect of allelic repeat difference on normalization potential could be the mechanism in which DSN cleaves sample DNA. This enzyme is capable of discriminating between DNA sequences that are perfectly matched and those with one base-pair mismatch, meaning those that are perfectly matched will be preferentially degraded. This study differs from previous research experiments utilizing DSN's function in that STRs are being analyzed, so sister alleles are not always the same length. Hypothetically, during the reannealing phase, the sister alleles join at the primers because the sequences are identical. If the alleles are of different repeat lengths, a hairpin loop is created, which DSN sees as a mismatch. DSN's

strong discriminatory capabilities mean that it will preferentially cleave those strands that are most like one another, or, for the purposes of this study, those with the fewest repeat difference due to the small hairpin loops. Evidence of this is seen in Table 4, because the loci that show the most improvement are those with the fewest repeat differences.

Although purification removed excess DSN and EDTA, an increased number of artifacts were present in the electropherograms of the purified product. Stutter peaks were seen infrequently across all channels, but off-ladder peaks were commonly observed. These off-ladder peaks were thought to be the result of inexact reannealing of the primers during DSN treatment. After denaturation, complementary strands began to bind to one another in the sample. If the strands did not reanneal exactly at the primers and were instead off by one or two repeats, there would be a section of single stranded DNA with the fluorescent label attached. DSN selectively degrades double-stranded DNA, so this piece would remain intact and have enough molecular weight to be analyzed during capillary electrophoresis. These off-ladder artifacts can be seen in Figure 15, and made the interpretation of the purified samples difficult. Further experimentation should amplify the loci without the fluorescent labels first, treat with DSN, and then amplify a second time with the fluorescent tag to see if there is a reduction in off-ladder artifacts.

Until this study, no prior research had been done utilizing DSN in a forensic setting. It is evident that the differences in repeat number between sister alleles affects the ability of STRs to normalize, which is an issue not seen in standard genomic research libraries. Further research needs to be performed to determine if the rules and patterns outlined above hold true for samples from other donors.

Conclusions

The purpose of this study was to determine if the use of DSN enzyme in conjunction with standard forensic DNA analysis protocols would alleviate the PCR artifacts due to stochastic effects of PCR on low quality and quantity DNA samples. In order to effectively model degraded and low-copy number DNA samples commonly encountered at crime scenes, sample DNA was exposed to UV irradiation. The use of UV irradiation on the extracted DNA was determined to be a suitable and reproducible methodology to model degraded DNA. Following exposure, shorter peak heights and inter- and intra- locus dropout was observed across all channels, with those of a higher molecular weight and TT or CT sequences being the most affected. UV irradiation affected the overall quality of the sample, but not the concentration. All UV treated samples were exposed to the DSN protocol and were purified to remove enzyme and EDTA in order prevent excess DNA cleavage and PCR inhibition.

Following normalization and purification, electropherograms of the samples demonstrated an increase in off-ladder peaks and artifacts. The majority of the off-ladder peaks are thought to be the due to the inexact reannealing of sister alleles during DSN treatment, which ultimately resulted in the incomplete degradation of DNA, and hypothetically, a release of a fluorescently labeled DNA fragment. In 21 of the 30 alleles analyzed, there was a recovery of lost alleles and an increase in peak height ratios relative to the tallest allele in the profile. Only D3 and D5 had statistically lower peak heights following treatment, and D7 and D18 dropped completely out of the profile after UV irradiation and were not recovered. Each locus responded to treatment with DSN differently, based primarily on the repeat difference of the sister alleles. Those loci that had identical repeat differences normalized based on the molecular weight, with those loci of a higher molecular weight showing the most improvement. Loci that were

homozygous demonstrated improvement based solely on molecular weight. Treatment of the DNA samples with DSN enzyme was shown to normalize samples that had been degraded through exposure to UV irradiation. Analysis of homozygous verses heterozygous loci provide strong indications that DSN treatment is not recommended for use in any case where a sample could have multiple donors. Mixed profiles are typically deconvoluted based on the peak heights of alleles from various contributors. Normalization will make all the peak heights relatively the same, making deconvolution an unlikely option. The ability to recover lost alleles and reduce preferential amplification may allow for the use of poor quality and quantity DNA in forensic casework, making it possible to obtain profiles from minute amounts of degraded biological samples.

Despite the prevalence of poor quality and quantity DNA in forensic casework, this methodology is not ready for implementation in a laboratory. The large number of artifacts present in the electropherograms makes interpretation of the DNA profiles difficult. The fluorescent label is thought to cause the majority of the off-ladder alleles, so further studies are needed without those labels to see if those peaks are alleviated. DSN normalization will have the potential to be an effective tool for the analysis of low-copy number and degraded DNA.

References

Adams, D. E., Presley, L. A., Baumstark, A. L., Hensley, K. W., Hill, A. L., Anoe, K. S., ... & Giusti, A. M. (1991). Deoxyribonucleic acid (DNA) analysis by restriction fragment length polymorphisms of blood and other body fluid stains subjected to contamination and environmental insults. *Journal of Forensic Sciences*, *36*(5), 1284.

Anisimova, V. E., Rebrikov, D. V., Shagin, D. A., Kozhemyako, V. B., Menzorova, N. I., Staroverov, D. B., … & Shcheglov, A. S. (2008). Isolation, characterization and molecular cloning of Duplex-Specific Nuclease from the hepatopancreas of the Kamchatka crab. *BMC Biochemistry*, *9*(1), 14.

Applied Biosystems. (2004). *GeneMapper® Software Version 3.7 User Guide.* Retrieved October 2012, from Applied Biosystems: Life Technology: http://mvz.berkeley.edu/egl/resources/manuals/GeneMapper%20Software%20v3.7%20User%20Guide.pdf

Applied Biosystems. (2010). *Quantifiler® Kits: Quantifiler® Human DNA Quantification Kit and Quantifiler® Y Human Male DNA Quantificaion Kit User Manual.* Retrieved October 2012, from Applied Biosystems: Life Technology: https://www3.appliedbiosystems.com/cms/groups/applied_markets_support/documents/generaldocuments/cms_041395.pdf

Applied Biosystems. (2012a). *AmpFISTR® Identifiler® PCR Amplification Kit user's guide.* Retrieved October 2012, from Applied Biosystems: Life Technology: http://www3.appliedbiosystems.com/cms/groups/applied_markets_support/documents/generaldocuments/cms_041201.pdf

Applied Biosystems. (2012b). *AmpFISTR® MiniFiler™ PCR Amplification Kit* Retrieved April, 2014 from Applied Biosystems: Life Technology: http://tools.lifetechnologies.com/content/sfs/manuals/cms_042748.pdf

Board, D.A. (1998). DNA advisory board quality assurance standards for forensic DNA testing laboratories. *Forensic Science Communications, 2(3).*

Boom, R. C. J. A., Sol, C. J., Salimans, M. M., Jansen, C. L., Wertheim-van Dillen, P. M., & Van der Noordaa, J. P. M. E. (1990). Rapid and simple method for purification of nucleic acids. *Journal of Clinical Microbiology*, *28*(3), 495-503.

Budowle, B., Moretti, T. R., Baumstark, A. L., Defenbaugh, D. A., & Keys, K. M. (1999). Population Data on the Thirteen CODIS Core Short Tandem Repeat Loci in African-Americans, US Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. *Journal of Forensic Sciences*, *44*, 1277-1286.

Budowle, B., Smith, J., Moretti, T., & DiZinno, J. (2000). *DNA typing protocols: molecular biology and forensic analysis*. Eaton Pub.

Buckingham, L. (2011). *Molecular Diagnostics: Fundamentals, Methods and Clinical Applications*. FA Davis.

Butler, J. M. (2011). *Advanced Topics in Forensic DNA Typing: Methodology*. Academic Press.

Butler, J. M., Buel, E., Crivellente, F., & McCord, B. R. (2004). Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis*, *25*(10-11), 1397-1412.

Butler, J. M., McCord, B. R., Jung, J. M., Lee, J. A., Budowle, B., & Allen, R. O. (1995). Application of dual internal standards for precise sizing of polymerase chain reaction products using capillary electrophoresis. *Electrophoresis*, *16*(1), 974-980.

Butler, J. M., & Reeder, D. J. (2012). *Short tandem repeat DNA internet database*. Retrieved October 10, 2012, from NIST Standard Reference Database: http://www.cstl.nist.gov/biotech/strbase/

Cohen, J. E., Lynch, M., & Tayplor, C. E. (1991). Forensic DNA tests and Hardy-Weinberg equilibrium. *Science*, *253*(5023), 1037-1038.

Clayton, T. M., Whitaker, J. P., & Maguire, C. N. (1995). Identification of bodies from the scene of a mass disaster using DNA amplification of short tandem repeat (STR) loci. *Forensic Science International*, *76*(1), 7-15.

Edwards, A. L., Hammond, H. A., Jin, L., Caskey, C. T., & Chakraborty, R. (1992). Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics*, *12*(2), 241-253.

Edwards, M. C., & Gibbs, R. A. (1994). Multiplex PCR: advantages, development, and applications. *Genome Research*, *3*(4), S65-S75

Evrogen. *Duplex-specific nuclease product insert.* Retrieved July 2013, from Evrogen: http://www.evrogen.com/protein-descriptions/UM-DSN.pdf

Federal Bureau of Investigation. (2011). *Quality Assurance Standards for Forensic DNA Testing Laboratories.* Retrieved April, 2014, from SWGDAM: http://swgdam.org/FBI%20Director%20Forensic%20Standards%20%20Revisions%20APPROVED%20and%20Final%20effective%209-01-2011.pdf

Garvan, F. L. (1964). Metal chelates of ethylenediaminetetraacetic acid and related substances. *Chelating Agents and Metal Chelates*, *Academic Press, New York*, 283-329.

Gijavanekar, C., Strych, U., Fofanov, Y., Fox, G. E., & Willson, R. C. (2012). Rare target enrichment for ultrasensitive PCR detection using cot–rehybridization and duplex-specific nuclease. *Analytical Biochemistry*, *421*(1), 81-85.

Gill, P. (2001). Application of low copy number DNA profiling. *Croatian Medical Journal*, *42*(3), 229-232.

Gill, P., Whitaker, J., Flaxman, C., Brown, N., & Buckleton, J. (2000). An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Science International*, *112*(1), 17-40.

Grossman, P. D., & Colburn, J. C. (Eds.). (1992). *Capillary electrophoresis: Theory and practice*. Academic Press.

Hares, D. R. (2012). Expanding the CODIS core loci in the United States. *Forensic Science International: Genetics*, *6*(1), e52-e54.

Hall, A., & Ballantyne, J. (2004). Characterization of UVC-induced DNA damage in bloodstains: forensic implications. *Analytical and Bioanalytical Chemistry*, *380*(1), 72-83.

Hartl, D. L., & Clark, A. G. (1997). *Principles of Population Genetics* (Vol. 116). Sunderland: Sinauer associates.

Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H., & Vogt, P. H. (1997). Multiplex PCR: critical parameters and step-by-step protocol. *Biotechniques*, *23*(3), 504-511.

Higuchi, R., Dollinger, G., Walsh, P. S., & Griffith, R. (1992). Simultaneous amplification and detection of specific DNA sequences. *Biotechnology*, *10*(4), 413-417.

Higuchi, R., Fockler, C., Dollinger, G., & Watson, R. (1993). Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology*, *11*, 1026-1030.

Holland, M.M. & Parsons, T.J. (1999).Mitochondrial DNA Sequence Analysis –Validation And Use for Forensic Casework. Forensic Science Review, 11: 21-50.

Innis, M. A., Myambo, K. B., Gelfand, D. H., & Brow, M. A. (1988). DNA sequencing with Thermus aquaticus DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proceedings of the National Academy of Sciences*, *85*(24), 9436-9440.

Iyer, R. R., Pluciennik, A., Burdett, V., & Modrich, P. L. (2006). DNA mismatch repair: functions and mechanisms. *Chem Rev*, *106*(2), 302-23.

Jeffreys, A. J., Wilson, V., & Thein, S. L. (1985). Hypervariable'minisatellite' regions in human DNA. *Nature*, *314*(6006), 67-73.

Jorgenson, J. W., & Lukacs, K. D. (1981). Zone electrophoresis in open-tubular glass capillaries. *Analytical Chemistry*, *53*(8), 1298-1302.

Kayser, M., & Schneider, P. M. (2009). DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. *Forensic Science International: Genetics, 3*(3), 154-161.

Kloosterman, A. D., & Kersbergen, P. (2003). Efficacy and limits of genotyping low copy number DNA samples by multiplex PCR of STR loci. In *International Congress Series* 1239, 795-798. Elsevier.

Lakowicz, J. R. (Ed.). (2007). *Principles of Fluorescence Spectroscopy*. Springer.

Lawyer, F. C., Stoffel, S., Saiki, R. K., Chang, S. Y., Landre, P. A., Abramson, R. D., & Gelfand, D. H. (1993). High-level expression, purification, and enzymatic characterization of full-length Thermus aquaticus DNA polymerase and a truncated form deficient in 5'to 3'exonuclease activity. *Genome Research, 2*(4), 275-287.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., ... & Grafham, D. (2001). Initial sequencing and analysis of the human genome. *Nature, 409*(6822), 860-921.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature, 362*(6422), 709-715.

McNally, L., Shaler, R. C., Baird, M., Balazs, I., DeForest, P., & Kobilinsky, L. (1989). Evaluation of deoxyribonucleic acid (DNA) isolated from human bloodstains exposed to ultraviolet light, heat, humidity, and soil contamination. *Journal of Forensic Sciences, 34*(5), 1059-1069.

Mullis, K. B., & Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology, 155*, 335.

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.

Nathans, D., & Smith, H. O. (1975). Restriction endonucleases in the analysis and restructuring of DNA molecules. *Annual Review of Biochemistry, 44*(1), 273-293.

Nicklas, J. A., & Buel, E. (2003). Quantification of DNA in forensic samples. *Analytical and Bioanalytical Chemistry, 376*(8), 1160-1167.

Peterson, D. G., Schulze, S. R., Sciara, E. B., Lee, S. A., Bowers, J. E., Nagel, A., Jiang, N., Tibbitts, D. C., Wessler, S. R., & Paterson, A. H. (2002). Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery. *Genome Research, 12*(5), 795-807.

Pang, B. C. M., & Cheung, B. K. K. (2007). One-step generation of degraded DNA by UV irradiation. *Analytical Biochemistry*, *360*(1), 163-165.

Puers, C., Hammond, H. A., Jin, L., Caskey, C. T., & Schumm, J. W. (1993). Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01 [AATG] n and reassignment of alleles in population analysis by using a locus-specific allelic ladder. *American Journal of Human Genetics*, *53*(4), 953.

QIAGEN. (2006). *DNeasy® Blood and Tissue Kit Handbook.* Retrieved July 2013, from QIAGEN: http://mvz.berkeley.edu/egl/resources/product%20inserts/DNeasy_Blood_&_Tissue_Handbook.pdf

Ravanat, J. L., Douki, T., & Cadet, J. (2001). Direct and indirect effects of UV radiation on DNA and its components. *Journal of Photochemistry and Photobiology B: Biology*, *63*(1), 88-102.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. T., Higuchi, R., Horn, G. T., Mullis, K.B., & Ehrlich, H. A. (1988). Primer-directed enzymatic amplification of DNA. *Science*, *239*, 487-491.

Schmutz, J., Martin, J., Terry, A., Couronne, O., Grimwood, J., Lowry, S., ... & Pollard, M. (2004). The DNA sequence and comparative analysis of human chromosome 5. *Nature*, *431*(7006), 268-274.

Shagina, I., Bogdanova, E., Mamedov, I. Z., Lebedev, Y., Lukyanov, S., & Shagin, D. (2010). Normalization of genomic DNA using duplex-specific nuclease. *Biotechniques*, *48*(6), 455.

Shagin, D. A., Rebrikov, D. V., Kozhemyako, V. B., Altshuler, I. M., Shcheglov, A. S., Zhulidov, P. A., ... & Lukyanov, S. (2002). A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. *Genome Research*, *12*(12), 1935-1942.

Sparkes, R., Kimpton, C., Gilbard, S., Carne, P., Andersen, J., Oldroyd, N., ... & Gill, P. (1996). The validation of a 7-locus multiplex STR test for use in forensic casework.(II), Artefacts, casework studies and success rates. *International Journal of Legal Medicine*, *109*(4), 195.

SWGDAM. (2010). SWGDAM interpretation guidelines for autosomal STR typing by forensic DNA testing laboratories. Available at http://www.fbi.gov/about-us/lab/biometric-analysis/codis/swgdam-interpretation-guidelines

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P., & Bouvet, J. (1996). Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, *24*(16), 3189-3194.

Urquhart, A., Kimpton, C. P., Downes, T. J., & Gill, P. (1994). Variation in short tandem repeat sequences—a survey of twelve microsatellite loci for use as forensic identification markers. *International Journal of Legal Medicine*, *107*(1), 13-20.

van Oorschot, R., & Jones, M. (1997). DNA fingerprints from fingerprints. *Nature*, *387*(6635), 767.

Vogelstein, B., & Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proceedings of the National Academy of Sciences*, *76*(2), 615-619.

Wang, D. Y., Chang, C. W., Lagacé, R. E., Calandro, L. M., & Hennessy, L. K. (2012). Developmental validation of the AmpFℓSTR® Identifiler® Plus PCR Amplification Kit: an established multiplex assay with improved performance. *Journal of Forensic Sciences*, *57*(2), 453-465.

Watson, J. D., & Crick, F. H. C. (1953). The structure of DNA. *Cold Spring Harbor*, 123-131.

Wellinger, R. E., & Thoma, F. (1996). Taq DNA polymerase blockage at pyrimidine dimers. *Nucleic Acids Research*, *24*(8), 1578-1579.

Yin, B. C., Liu, Y. Q., & Ye, B. C. (2012). One-Step, Multiplexed Fluorescence Detection of microRNAs Based on Duplex-Specific Nuclease Signal Amplification. *Journal of the American Chemical Society*, *134*(11), 5064-5067.

Zhulidov, P. A., Bogdanova, E. A., Shcheglov, A. S., Vagner, L. L., Khaspekov, G. L., Kozhemyako, V. B., ... & Shagin, D. A. (2004). Simple cDNA normalization using kamchatka crab duplex- specific nuclease. *Nucleic Acids Research*, *32*(3), e37-e37.

Appendix

Figure A: Sample 1 Blue Channel; 300 seconds UV, DSN treatment, Purification

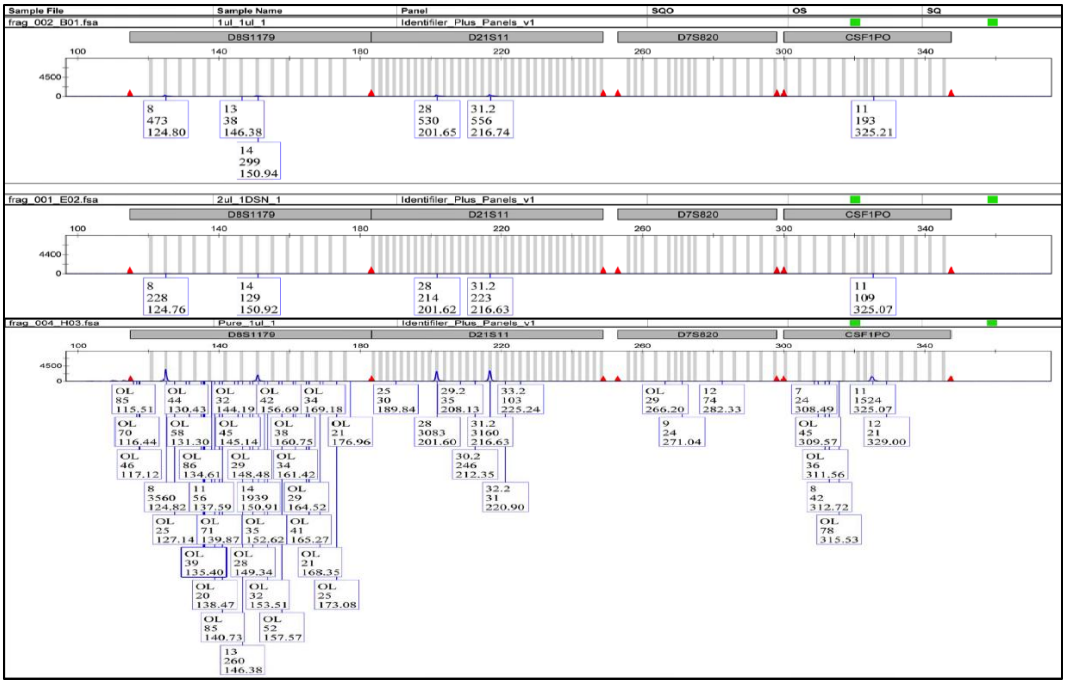

Figure B: Sample 1 Green Channel; 300 seconds UV, DSN treatment, Purification

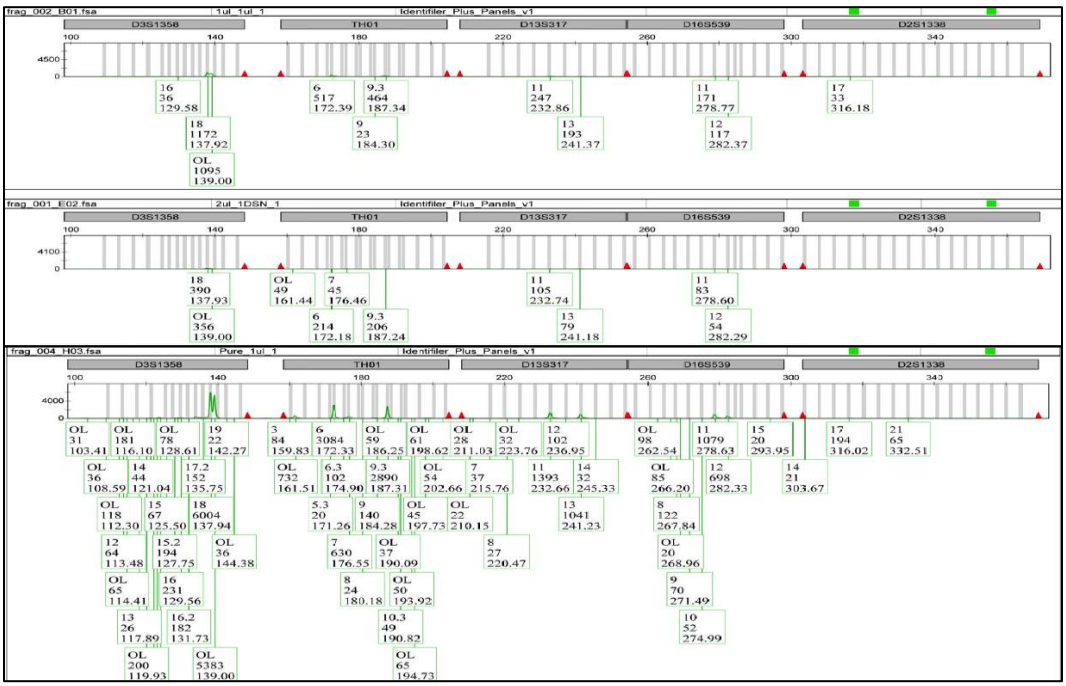

Figure C: Sample 1 Yellow Channel; 300 seconds UV, DSN treatment, Purification

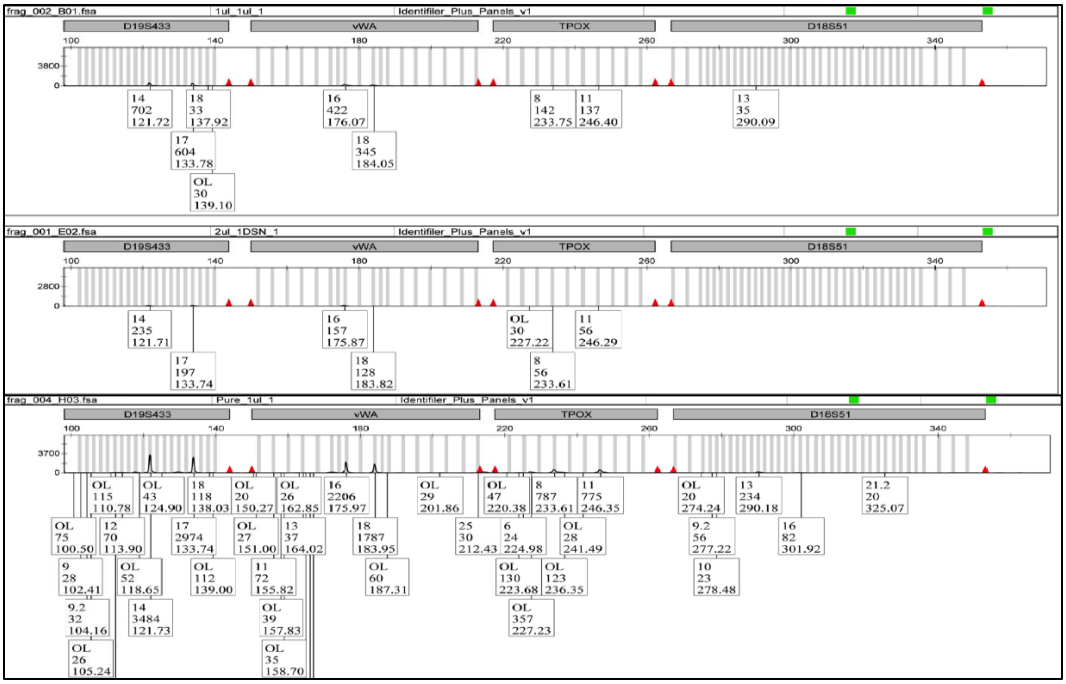

Figure D: Sample 1 Red Channel; 300 seconds UV, DSN treatment, Purification

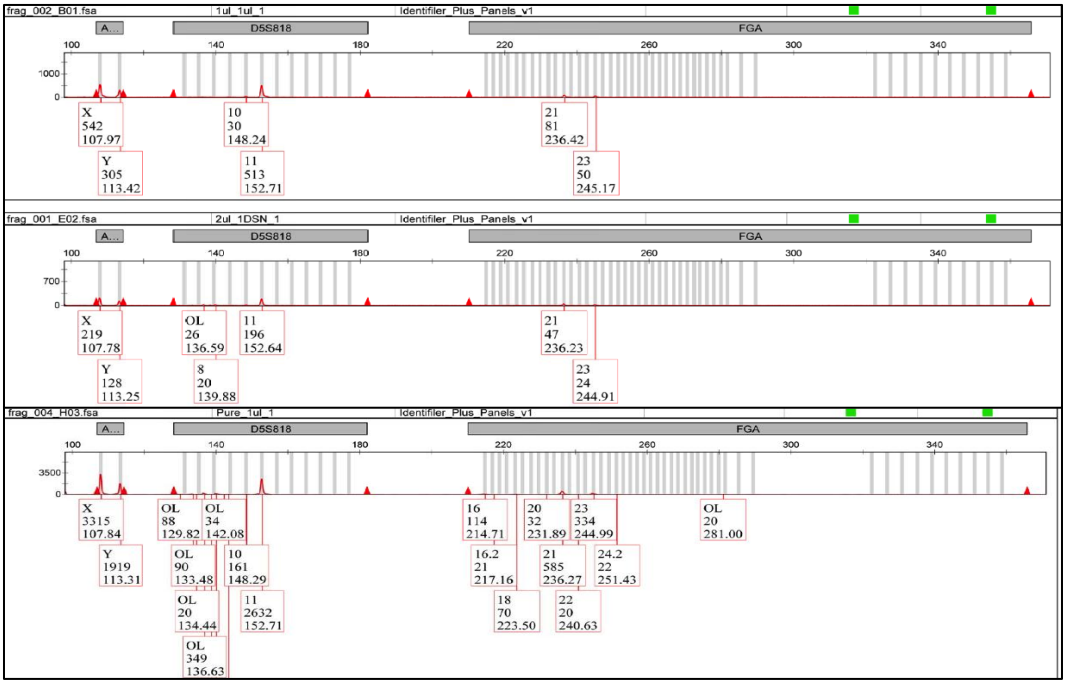

Figure E: Sample 2 Blue Channel; 300 seconds UV, DSN treatment, Purification

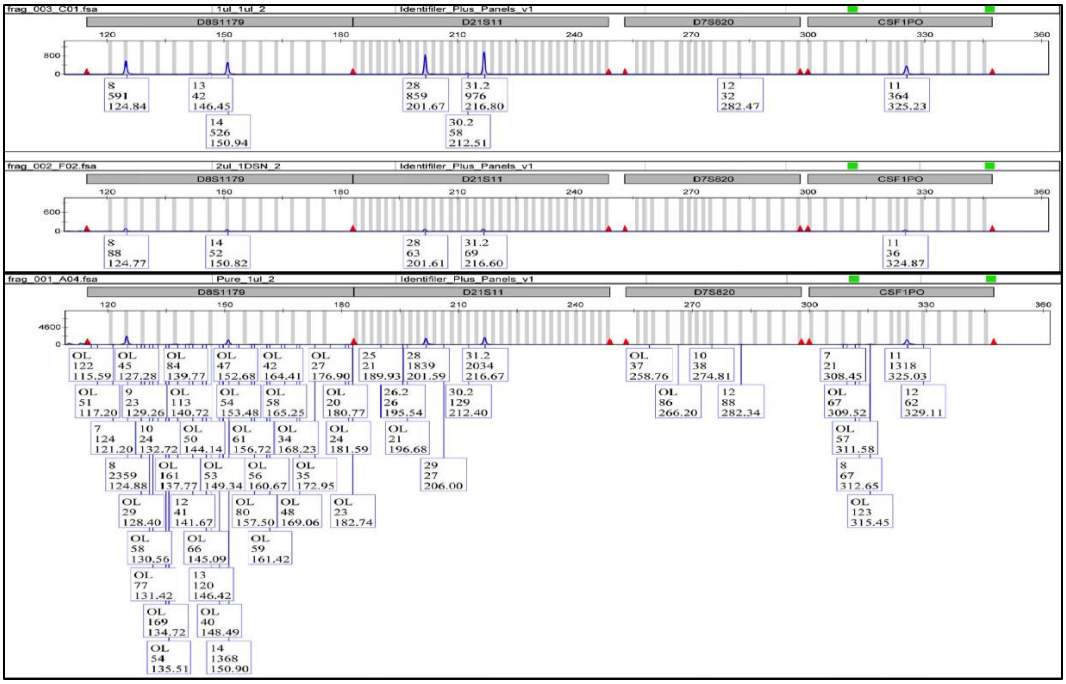

Figure F: Sample 2 Green Channel; 300 seconds UV, DSN treatment, Purification

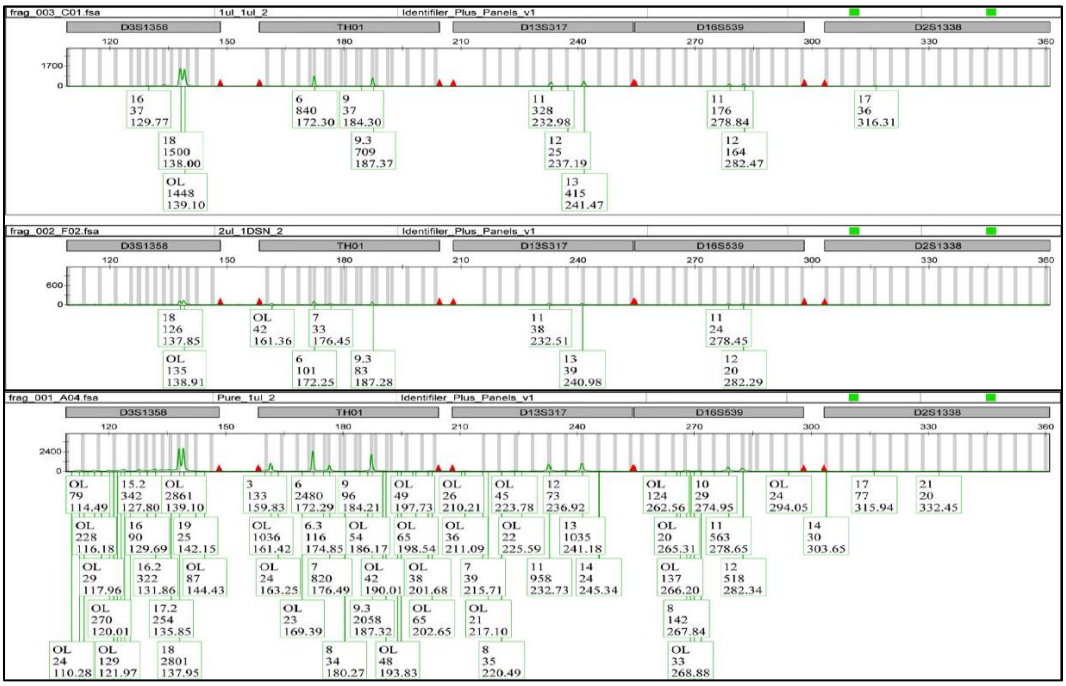

Figure G: Sample 2 Yellow Channel; 300 seconds UV, DSN treatment, Purification

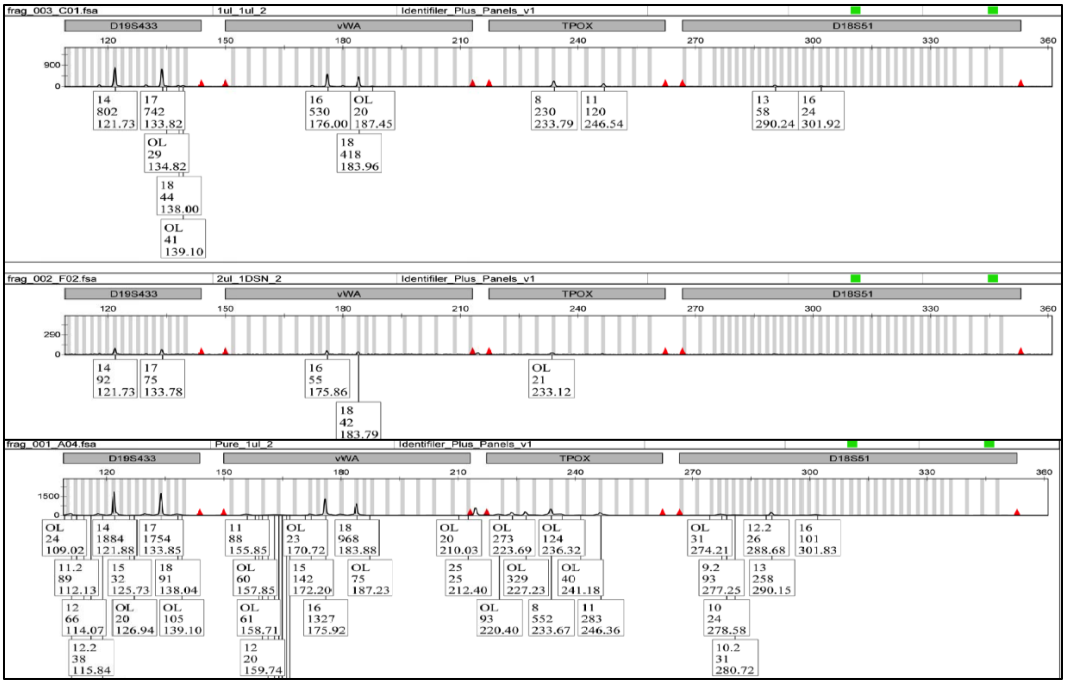

Figure H: Sample 2 Red Channel; 300 seconds UV, DSN treatment, Purification

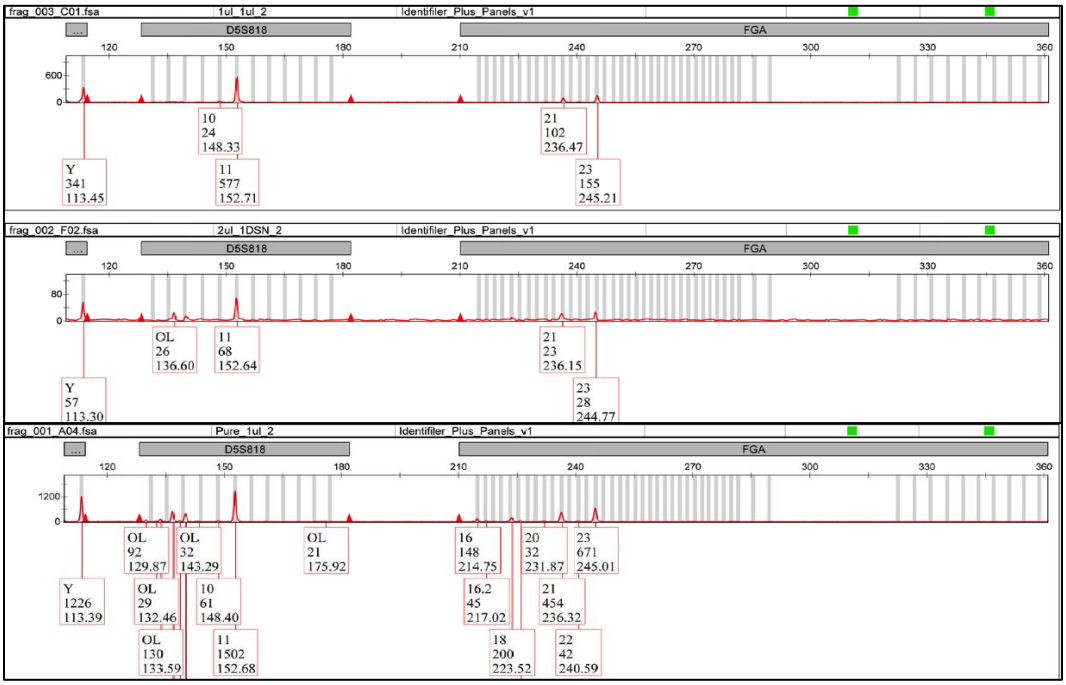

Figure I: Sample 3 Blue Channel; 300 seconds UV, DSN treatment, Purification

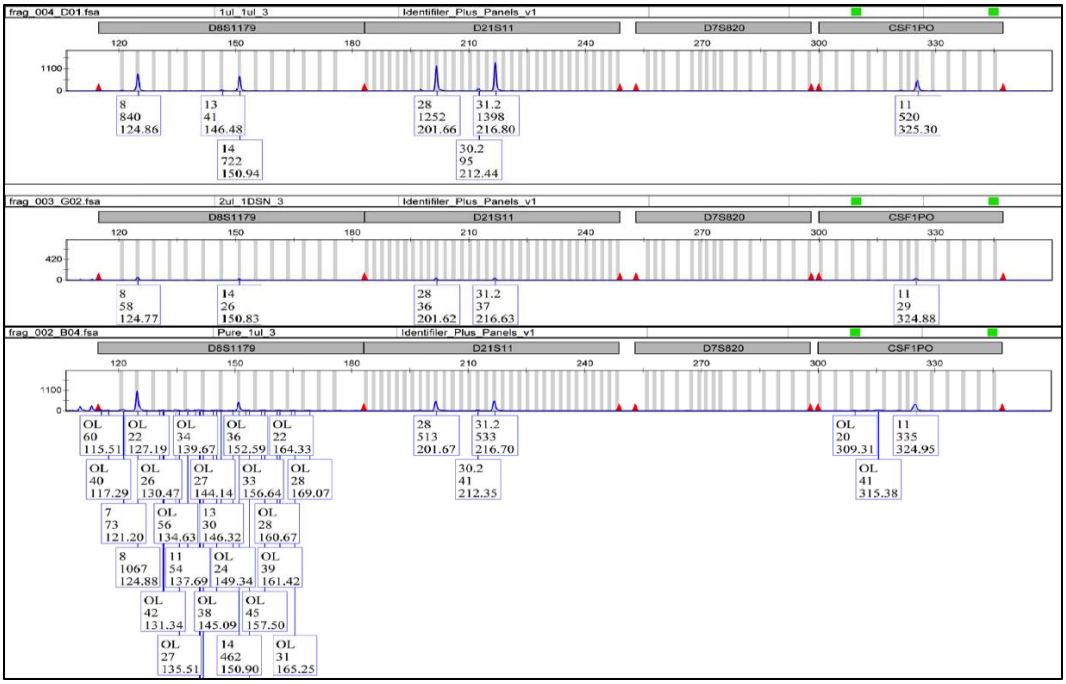

Figure J: Sample 3 Green Channel; 300 seconds UV, DSN treatment, Purification

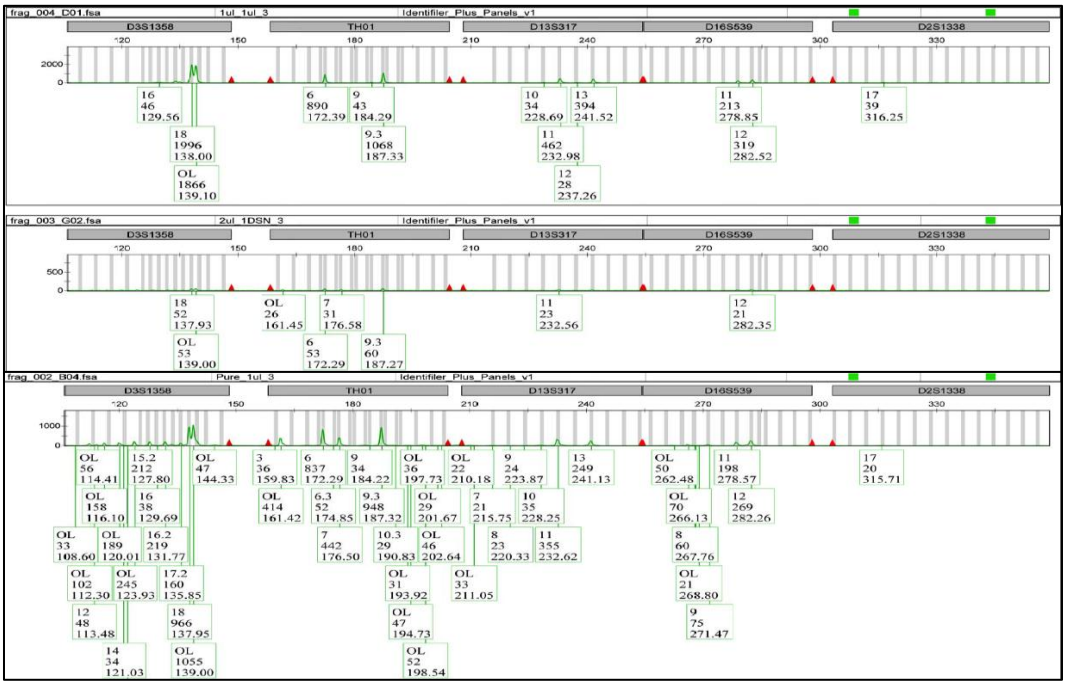

Figure K: Sample 3 Yellow Channel; 300 seconds UV, DSN treatment, Purification

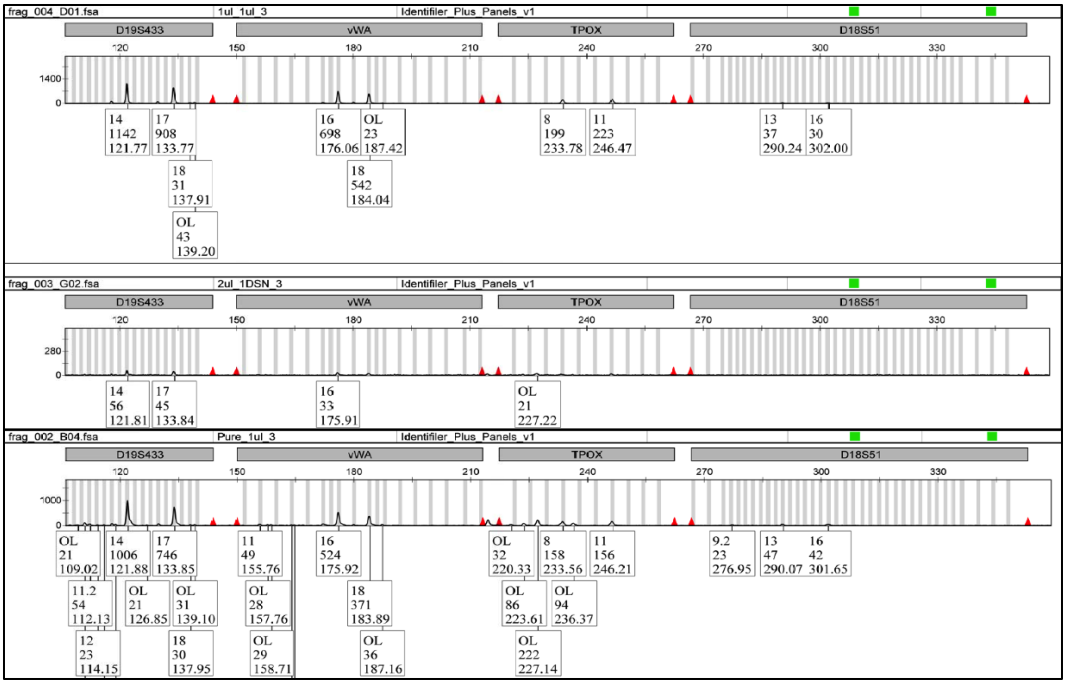

Figure L: Sample 3 Red Channel; 300 seconds UV, DSN treatment, Purification

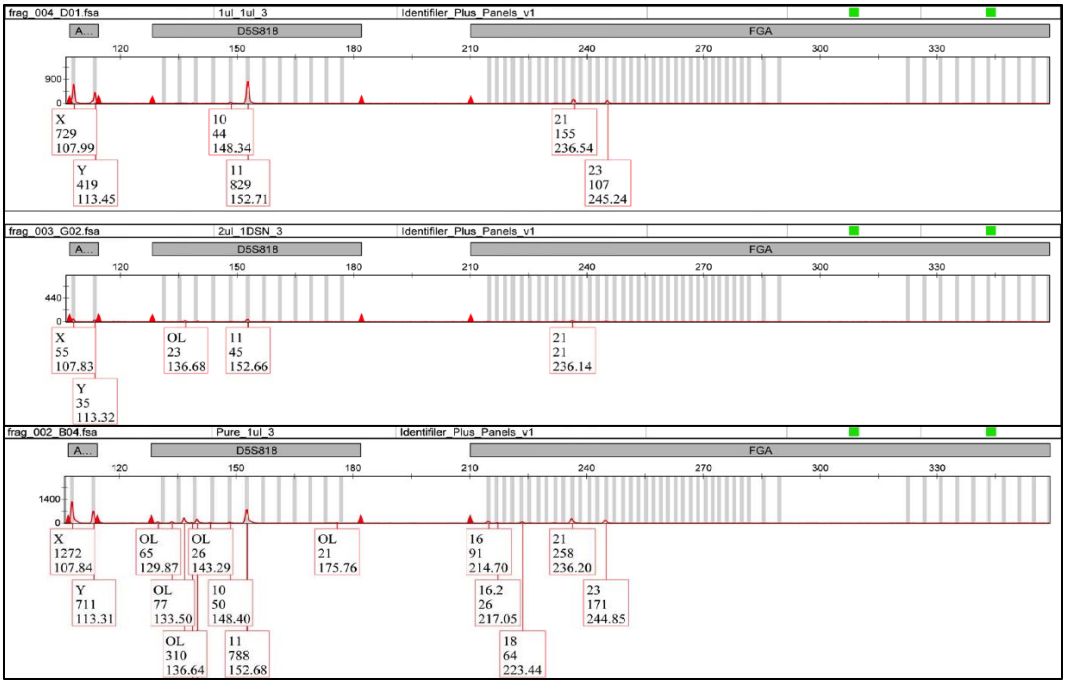

Figure M: Sample 4 Blue Channel, 300 seconds UV, DSN treatment, Purification

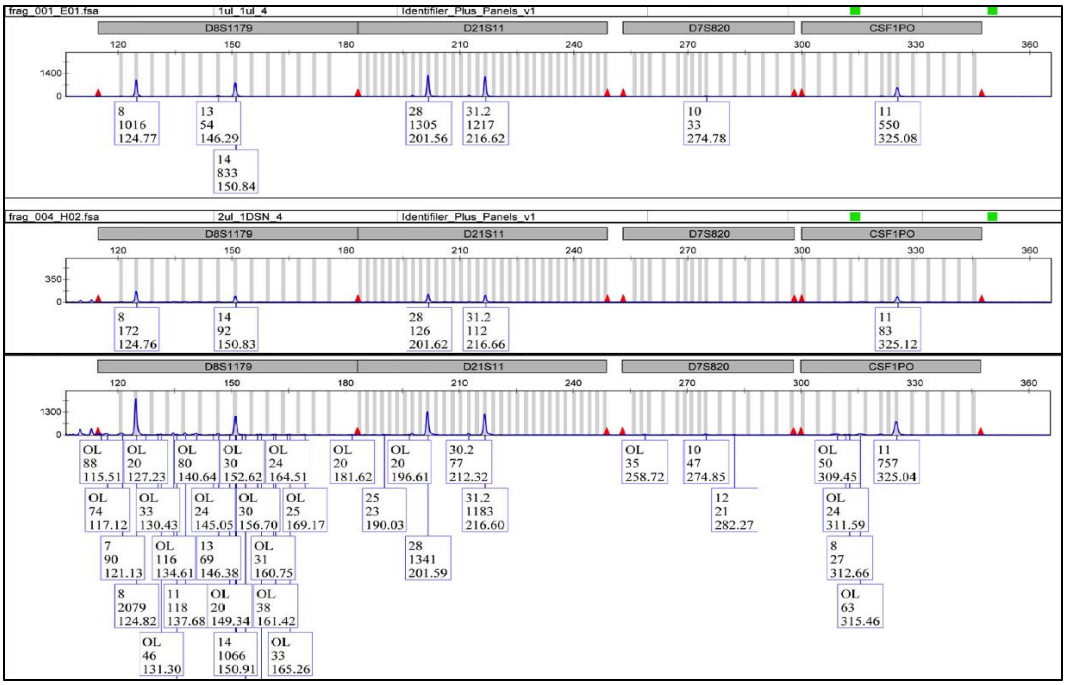

Figure N: Sample 4 Green Channel, 300 seconds UV, DSN treatment, Purification

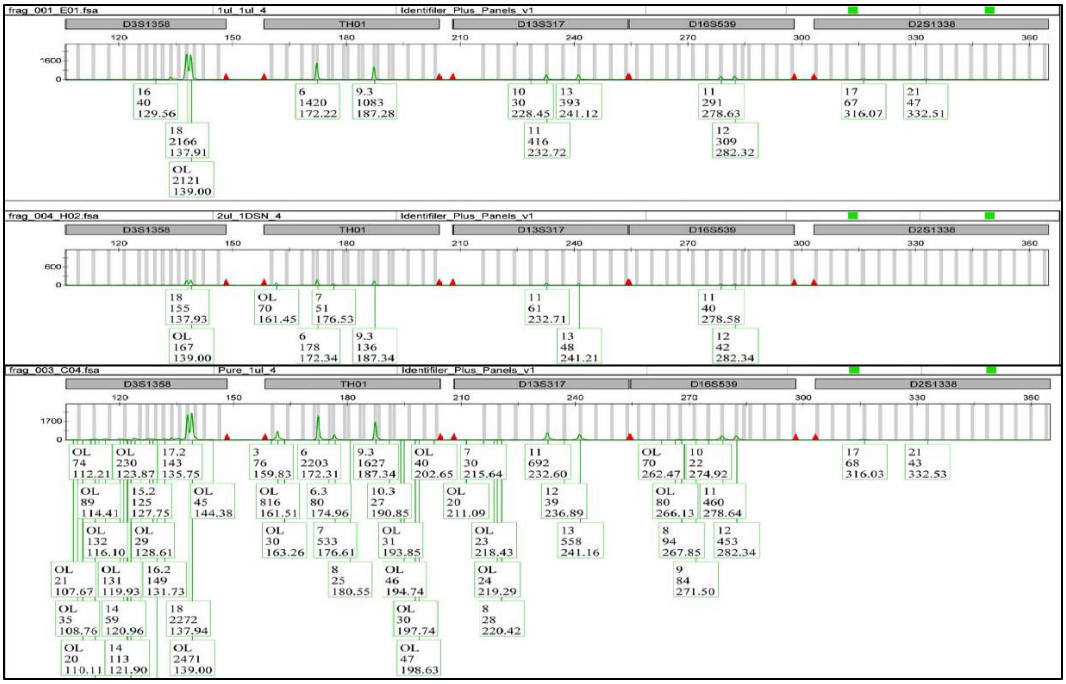

Figure O: Sample 4 Yellow Channel; 300 seconds UV, DSN Treatment, Purification

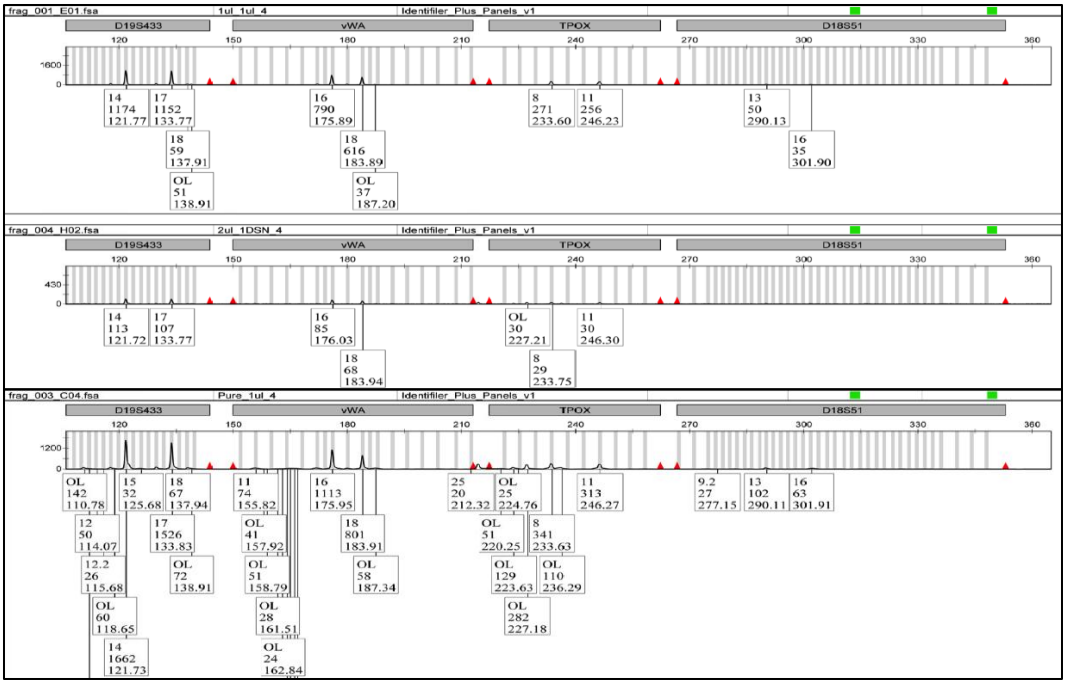

Figure P: Sample 4 Red Channel; 300 seconds UV, DSN treatment, Purification

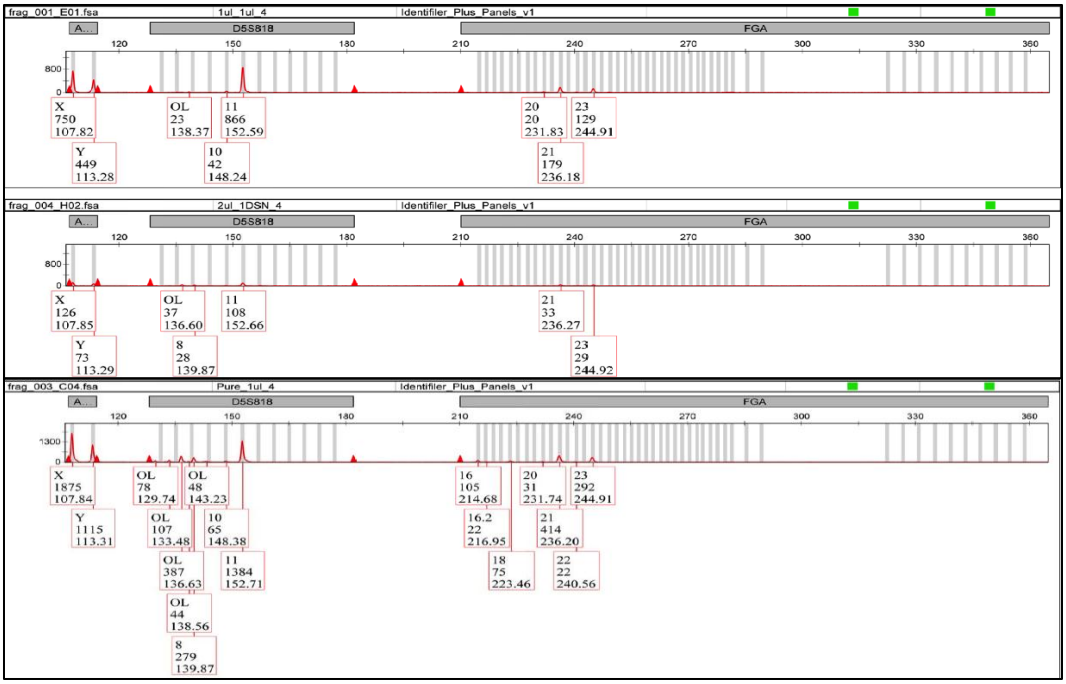

Figure Q: Sample 5 Blue Channel; 300 seconds UV, DSN treatment, Purification

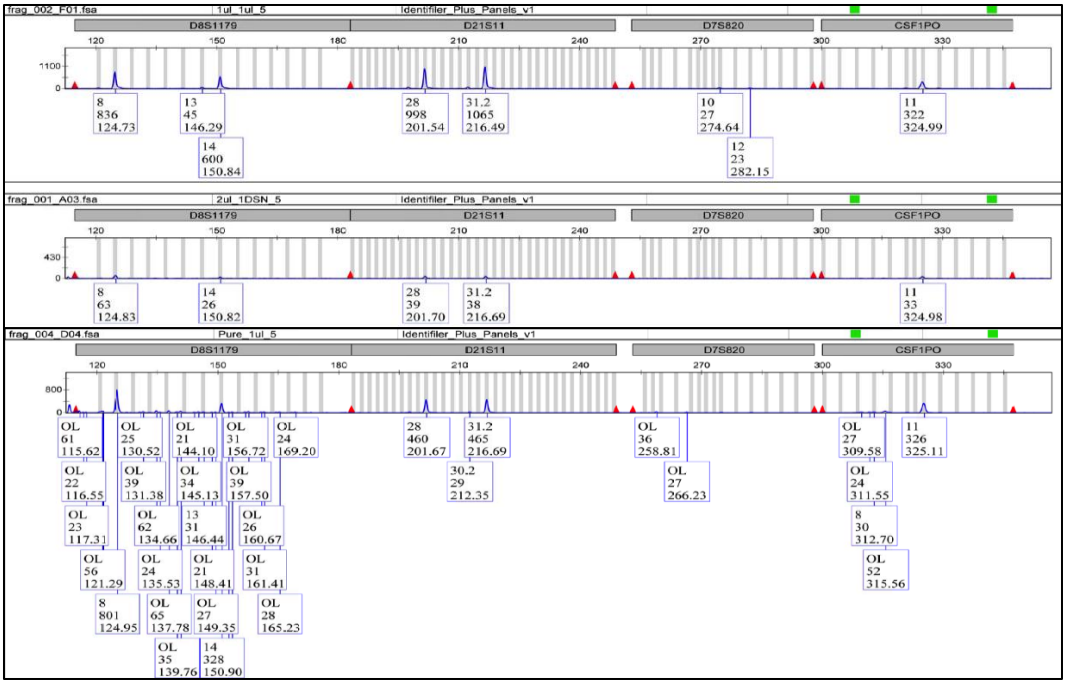

Figure R: Sample 5 Green Channel; 300 seconds UV, DSN treatment, Purification

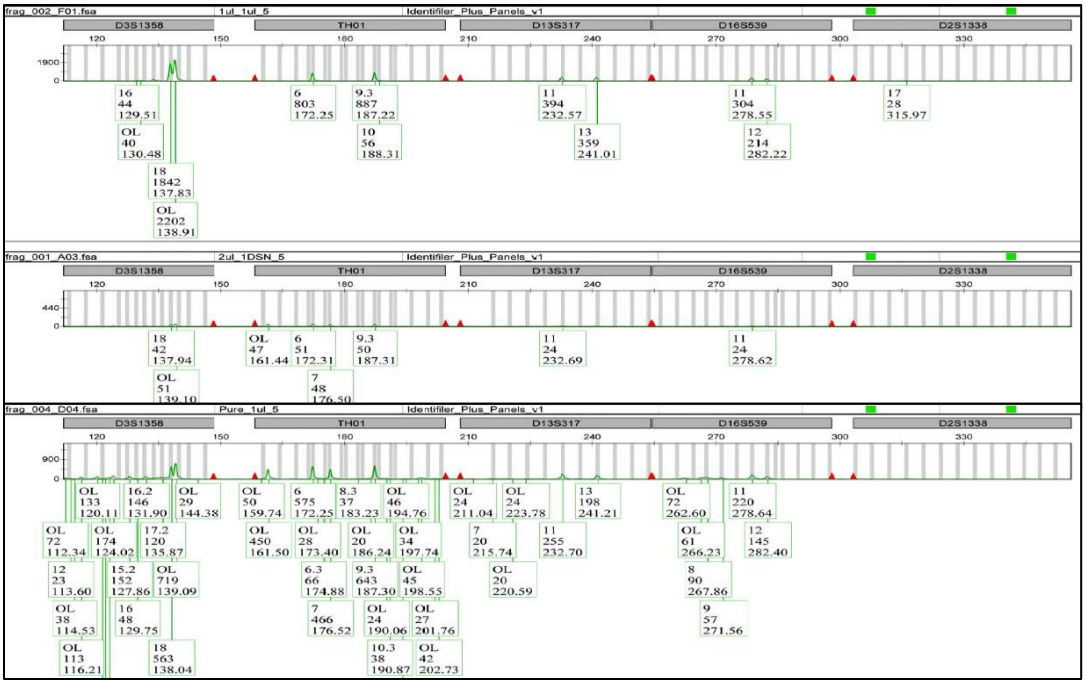

Figure S: Sample 5 Yellow Channel; 300 seconds UV, DSN treatment, Purification

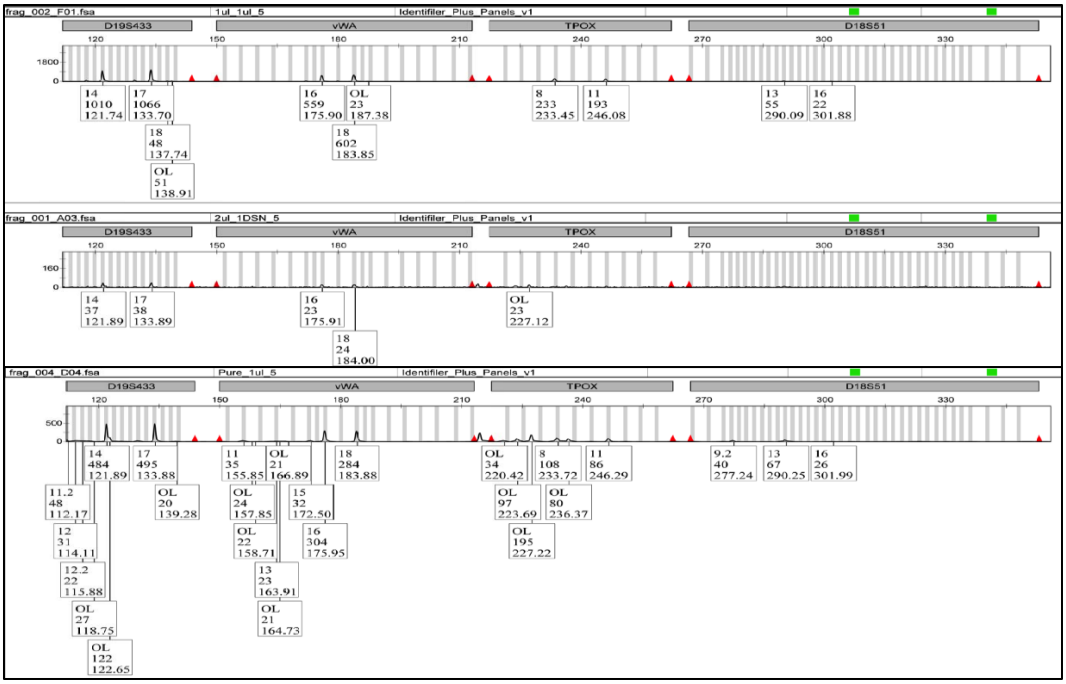

Figure T: Sample 5 Red Channel; 300 seconds UV, DSN treatment, Purification