THE OPTIMIZATION OF WHOLE GENOME AMPLIFICATION AND MOLECULAR CROWDING FOR USE IN FORENSIC LOW COPY NUMBER SAMPLES

A Paper
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Megan Frances Palmer

In Partial Fulfillment of the Requirements
For the Degree of
MASTER OF SCIENCE

Major Department:
Chemistry and Biochemistry
Degree: Biochemistry

March 2011

Fargo, North Dakota

North Dakota State University
Graduate School

Title

The Optimization of Whole Genome Amplification and Molecular Crowding for Use in Forensic Low Copy Number Samples

By

Megan Frances Palmer

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

ABSTRACT

Palmer, Megan Frances, M.S., Department of Chemistry and Biochemistry, College of Science and Mathematics, North Dakota State University, March 2011. The Optimization of Whole Genome Amplification and Molecular Crowding for use in Forensic Low Copy Number Samples. Major Professor: Dr. Stuart Haring.

Whole genome amplification has been used as a powerful tool to increase the amount of DNA template used for microarrays, STR and SNP assays in clinical and forensic settings. Our laboratory observed that multiple displacement whole genome amplification demonstrated a higher reliability for increasing DNA template than PCR primer extension pre-amplification of human genomic DNA. We also demonstrated a truncated reaction time for whole genome amplification was necessary to decrease artifact, while still increasing authentic DNA species. We determined that molecular crowding using polyethylene glycol and non-human DNA also increased sensitivity of our assays. Combining the modified whole genome amplification protocol with macromolecular crowding increased STR signals with as low as one haploid cell DNA equivalent.

ACKNOWLEDGMENTS

This thesis would not have been possible without the guidance and feedback of my adviser, Dr. Stuart Haring, along with my graduate committee members Dr. Berch Henry, Dr. DK Srivastava, and Dr. Lisa Montplaisir.

I am indebted to my many colleagues that supported me, namely: Thomas A. Wahl for his expertise, constructive criticism, and the time and effort invested in providing priceless training in the field of forensic DNA analysis. His dedication is irreplaceable and truly valued. Also, a thank you to Travis Van der Steen and Naomi Light for augmenting their experiments to help me with mine.

I owe my deepest gratitude to Dr. Berch Henry, as his support and direction have been unconditional during my time of study. Through Dr. Henry's mentorship I have learned lifelong skills to help me succeed working in a lab as an analyst, in the courtroom as a teacher and expert witness, and in my everyday life as an independent, young woman. I am truly grateful for his patience and enthusiasm.

I would also like to thank my family for their unrestricted encouragement in the achievement of my many goals. And finally, my friends, who are my family away from home; the people always willing to lend an ear to listen or offer a joke in exchange for a smile. It's my family and friends who really make it all worth it at the end of the day.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AB.....Applied Biosystems

Ab.....antibody

bp.....base pairs

CCD.....charged coupled device

CE.....capillary electrophoresis

CODIS..... Combined DNA Index System

DOP-PCR..... degenerate oligonucleotide primed PCR

DNA.....deoxyribonucleic acid

DTT.....dithiothreitol

ERC.....extraction reagent blank control

HLA..... human leukocyte antigen

HRP.....horse radish peroxidase

hTERT.....human telomerase reverse transcriptase gene

HUT.....human T-lymphocyte

HVI and HVII..... hypervariable regions I and II

I-PEP..... improved PCR primer extension

LCM.....laser catapult microdissection

LCN.....low copy number

mDNA.....mouse DNA

mtDNA.....mitochondrial DNA

MDA.....multiple displacement amplification

NP.....nonyl phenoxypolyethoxylethanol

PBS.....phosphate buffered saline

PCR.....polymerase chain reaction

PEG.....polyethylene glycol

PEP..... primer extension pre-amplification

RFLP.....restriction fragment length polymorphism

RFU.....relative fluorescent unit

RMP..... random match probability

RNA.....ribonucleic acid

SDS v1.2.....sequence detection system

SDS.....sodium dodecyl sulfate

SRY.....sex-determining region Y gene

SSO..... .sequence specific oligonucleotides

STR.....short tandem repeat

TMB..... tetramethylbenzidine

VNTR.....variable number tandem repeats

WGA.....whole genome amplification

CHAPTER 1.

THESIS OVERVIEW

Over the past hundred years human identification has evolved from the use of fingerprinting, to protein analysis, and finally to deoxyribonucleic acids (DNA). Each technology possesses its own advantages and disadvantages as described in the literature review in Chapter 2. Forensic DNA analysis has become the method of choice for the identification of individuals from biological evidence. Increasingly sensitive over the years from a minimum of over 2 μg DNA needed for identification 20 years ago to now the present requirement of only 1 ng template DNA. Due to these technological advances in DNA analysis throughout the decades, the limit of detection of current methods has finally been reached with the limited DNA samples often seen today in routine crime lab casework. The scope of this research project was to explore avenues that would decrease the limit of detection for implementation in low copy number DNA analysis.

The first approach to increase the sensitivity of our DNA analysis human identification assays was to employ whole genome amplification (WGA). Chapter 3 covers in detail how WGA was used and evaluated for forensic DNA short tandem repeat (STR) analysis.

With the success of WGA, the limits of detection were achieved using different sample types (e.g., diploid cells and pre-extracted DNA) as described in Chapter 4. Along with determining the detection limit, the difficulties of employing adhesive cap tubes by Zeiss to collect intact cells is also included in this chapter. The troubleshooting

experiments eventually exhausted the available time and budget of the project. Therefore, to move forward with the project, the remaining experiments were performed only with pre-extracted DNA.

Chapter 5 presents experiments exploring samples of only extracted DNA. It is in this chapter that the implementation of macromolecular crowding is used to further increase the detection level of DNA. The excluded volume strategy proved to be a promising outlet to increase reaction efficacy and enhance sample signals. By incorporation macromolecular crowding and WGA into the DNA STR analysis, we were able to generate a partial DNA STR profile in the single cell range using more than one human identification assay.

Chapter 6 is a general discussion relating the successes of each attempt to a real casework setting. A guideline of when to use each technique, if necessary is included to streamline analyses, as WGA and macromolecular crowding may not be required for all samples brought into the lab. Though only a partial profile was generated from a single diploid AND haploid cell equivalent, these data were forensically significant as the information gained can be used to exonerate or incarcerate an individual. Furthermore, it can be argued that since there is no perfect method for DNA quantitation and the troubleshooting described in Chapter 4 involving the difficulty in retrieval of sample from the adhesive tube cap that the limit of detection could actually be in the single cell range. Additionally, the use of this low copy number (LCN) research can be applied to authentic case samples which are shelved due to the low template available at the time of collection. Such samples are increasingly adding to the back-log of crime labs and the time an

innocent person may spend behind bars waiting for their case evidence to be examined. Until now, LCN samples could not routinely be analyzed as the confidence in information recovered was low. By incorporating the methods and techniques optimized in this thesis, LCN samples can now demand the attention needed for analysis.

CHAPTER 2.

GENERAL INTRODUCTION

Historically, fingerprinting has been treated like a signature to bind contracts and establish identity among the human population (Block, 1979). It wasn't until the early 1900s that human fingerprinting was accepted as the main method of identification to aid in criminal investigations (Frauds, 1905). Though this mode of identification is traditionally the hub of human identification in law enforcement, the science of latent fingerprints is subjective, potentially leading to different interpretations of the same skin furrow (Edwards and Gotsonis, 2009). It wasn't until the 1930s that a more reliable and less subjective method (ABO blood typing) was developed for identification purposes. This method was developed and implemented in Italy, and was the primary method of identification used until the identification and development of restriction length polymorphism (RFLP) examination in the 1980s, and finally short tandem repeat (STR) analysis in the 1990s. The benefit of RFLP and STR analysis is that results are less subjective than analyses such as fingerprinting or handwriting analysis; and thus, are more desirable and carry more weight for identification purposes.

Blood typing technique was developed in the early 1900s by Karl Landsteiner, when he discovered three blood factors that could potentially discriminate one donor from the next (Inman and Rudin, 2000). Though Landsteiner originally discovered agglutinogens of alleles M, N and P, the breakthrough led to many more serological isoantibodies, which ultimately led to Franz Josef Holzer's discovery of the ABO typing

system in the 1930s (Inman and Rudin, 2000). There are four primary blood typing alleles, A, B, AB, and O. Each contains unique phenotypes on the protein level to discern one type from another. The ABO blood typing system is based on a protein's secondary structure that is recognized by a specific antibody (Ab) which is determined by a physical coagulation reaction. It was later discovered that the secretion of the same group-specific antigens were found in other bodily fluids such as semen, which also assisted in solving sexual assault cases that remained open (Inman and Rudin, 2000). In the Unites States, over 70% of the population possesses a blood type of either Type A or Type O (Bloodbook), suggesting that ABO typing is very helpful for exclusion. In other words, if a stain left at a crime scene exhibited an AB blood type, those individuals possessing Type A, Type B or Type O blood types would all be exempt from further investigation. However, when inclusion types are exhibited, the science of blood typing is less distinguishing as a means of identification. For example, if the stain left at a crime scene is Type A, which represents approximately 42% of Americans, then that particular type can only exclude 58% of the population (all individuals that exhibit a blood type other than Type A; Butler, 2009). There are additional problems to the inclusion issue that lie in sample collection and determination. Since the ABO typing assay relies on the secondary structure of the proteins on the red blood cells, it is important that the proteins remain intact and not denatured; when the denatured proteins unfold, the reliability of the test decreases due to the lack of recognition of the antibody for the antigen (Gaensslen, 1983).

In almost all cases, those blood samples could have offered more information in years to come than the low statistical inclusion or exclusion obtained by protein

characterization. It should be noted that mature red blood cells do not contain any genetic information (*i.e.*, DNA) due to anucleation. It is the other cells (*i.e.*, white blood cells) that contain DNA. Although the structure and significance of DNA would not be completely realized for another 20 years, the blood used to assign alleles also consumed the DNA-carrying cells in the blood. The shift from protein to DNA would offer a more versatile range of items that could be tested, even after exposure to the elements.

The human population contains DNA, that is over 99.9% identical, to develop and maintain the functions of our bodies (Casey, 2008). The DNA blueprint in humans remains relatively constant; these regions are responsible for encoding DNA sequences that are transcribed and translated into protein. It is the remaining 0.3% that lends significant change from one individual to the next; these regions are denoted as hypervariable regions. In 1980, Alec Jeffreys was inspired by the discovery of the hypervariable regions used in the restriction length polymorphism (RFLP) technique developed by Arlene Wayman and Ray White (Jeffreys *et al*, 1985). RFLPs are DNA sequences within the conserved regions of DNA that are recognized by certain endonuclease enzymes. When the enzymes cleave the DNA sequence, DNA fragments of different lengths are produced and detected via gel electrophoresis (Southern, 1975). Jeffreys' innovative consideration that DNA was holding the key to millions of polymorphic hypervariable regions led to the advancement of applying DNA fingerprinting to solve the problem of human identification. The first case to utilize RFLPs was an immigration dispute in the United Kingdom (Jeffreys, 2005). The blood group determination combined with serological and allozyme testing were performed, but even with the combination of the assays, the information gleaned from the

experiments was not discriminating enough, requiring the need for development of the then novel discipline of RFLP examination.

Restriction fragment length polymorphisms, formerly known as variable number tandem repeats (VNTRs) were used extensively in the 1980s and 1990s for human identification. The numerous single locus probes used for RFLP analysis offer a high degree of discrimination between individuals, but also require more time, labor, and training in the many techniques involved to develop an RFLP DNA pattern over the ABO typing method. For RFLP technology to be successful, it requires relatively large amounts of DNA (500ng-2μg), attention to time and detail and the use of radioactive ^{32}P for development of a signal (Budowle *et al*, 2000). Extracted DNA was quantitated using a standard curve on a slot blot, which is a variation of the traditional southern blot technique. The method of quantity evaluation was subjective to the analyst as the probe was conjugated to horseradish peroxidase which produced a color changing reaction when the 40bp probe specific for humans and higher primates (D17Z1) hybridized to the target DNA (Walsh *et al*, 1992). It was the intensity of the color which assigned a quantitative value to the sample by comparing the unknown samples to those of the standard curve. The 48-well manifold that was used for quantitation was then photographed to document the results of the quantitative analysis performed because the signals would fade over time. The probe used to identify the quantity of DNA was specific for higher primates and humans, while containing a high copy number of the sequence for higher sensitivity. After quantitation, the DNA was then subjected to a series of restriction enzymes for digestion of the DNA at certain sites in the genome. Before proceeding to the southern blot analysis, a mini-gel was

performed and stained with ethidium bromide to ensure complete digestion of the DNA. If complete digestion had not taken place, resulting data would mimic that of a mixture or be rendered uninterpretable. Once complete digestion was confirmed, the analyst could move onto the southern blot analysis to estimate the sizes of the RLFPs produced by digestion. The DNA fragments were electrophoresed on a 14-30cm 1% agarose gel, often running overnight (Budowle *et al*, 2000). The most critical step of the procedure was the transfer of DNA from the agarose gel to the nitrocellulose membrane. If the gel were to break, all of the time, effort, and evidence dedicated to the analysis would be lost. A minimum of 4 probes were required to evaluate the DNA pattern, but often 6 or more probes were assessed during hybridization for identity. It was standard practice to *sandwich* the hybridized membrane between two pieces of x-ray film, which was then placed at -70°C for approximately two days (Budowle *et al*, 2000). This allowed development of one autoradiogram to evaluate the strength of the signal and determine an optimal exposure time for the remaining autoradiogram. This process then had to be repeated for each of the sequential probes and required a great deal of time to complete the analysis of one single sample. The skills and performance of the forensic analyst were often tested during the RFLP DNA process due to the balancing of many cases being analyzed at one time, as it could easily take 3 months to process a single case. Time and the amount of available DNA had an inverse relationship with respect to RFLPs; if only 500ng of DNA were available, the exposure time during the reverse hybridization process could reach up to 4 weeks *per probe,* extrapolating to 4-6 months of exposure time alone when testing a minimum of 6 probes. In many instances, insufficient DNA was available to allow RFLP

analysis, therefore, it was critical that new methods be developed which could be performed faster, using less DNA, while still providing a powerful statistical analysis (Budowle *et al*, 1995[a]).

In 1983, after 6 months of failed trials, Kary Mullis performed his first successful experiment using polymerase chain reaction (PCR; Mullis, 1998). His ability to foresee the importance of reading the blueprint of DNA, or as he puts it "like reading a particular license plate out on Interstate 5 at night from the moon", has changed biochemistry and molecular biology from small advancements to a tsunami of endless possibilities. The invention of PCR has revolutionized the world of DNA; impacting, fields of medicine, genetics, forensic science and many other disciplines. Polymerase chain reaction can be thought of as analogous to a molecular Xerox in that exact copies are reproduced. Using primers to guide the DNA polymerase to the site of replication enables scientists to amplify a specific location of the genome, which grows exponentially with each round of replication. Figure 2-1 illustrates the amplification of a single location on the genome.

The first STR segment to be applied to forensic DNA analysis was the human leukocyte antigen DQα (HLA-DQα; Blake *et al*, 1992). The HLA-DQα gene is found on chromosome 6 and exhibits 6 distinguishable alleles; however, due to limitations inherent in the testing procedure, in practice only 28 different genotypes could be discerned (Comey *et al*, 1993). It was clear that the need for a higher level of discrimination was required to further aid in the identification process as the number of possible genotypes of the former example can display up to 64 combinations (Butler, 2009). In 1995, Perkin-Elmer released the first multiplex kit for DNA typing containing five loci in addition to the DQα marker

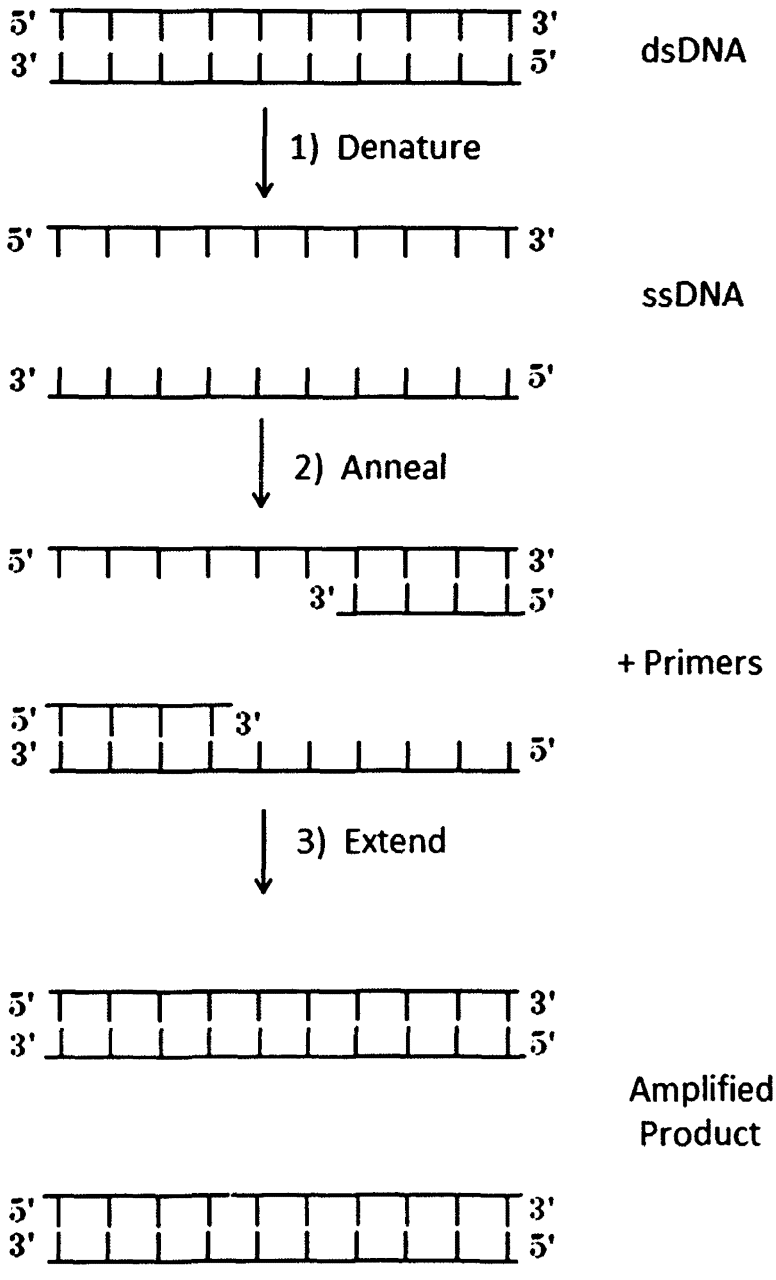

Figure 2-1: DNA amplification using polymerase chain reaction. There are three major steps in PCR: 1) Denaturing the double stranded DNA product into single stranded DNA species, 2) annealing primers that are homologous to the consensus sequence of the target DNA, and 3) extension of the primers to synthesize new DNA replicates.

(Fildes and Reynolds, 1995). The multiplex kit was called AmpliType PM + DQA1 or Poly-Marker, and was designed as reverse dot blot, or reverse hybridization technology. This multiplex kit was able to statistically yield probabilities up to the one in ten-thousand ranges (Budowle *et al*, 1995[b]).

The reverse hybridization method is a variation of a traditional Southern Blot in which instead of the target DNA being cross-linked to the nitrocellulose membrane, each specific oligonucleotide probe is bound to the membrane. The membranes are then exposed to the PCR product generated using a biotinilated forward primer. After hybridization, the membrane is washed with a streptavidin-horseradish peroxidase complex to further react with tetramethylbenzidine (TMB) to produce a colored product (Butler, 2009). Though this process obtained a higher order of discrimination, the technique required a skillful hand to be successful; many assays were ineffective due to renaturation of the substrate or cross contamination (Figure 2-2).

The statistical significance of any profile is determined by how rare or how common each of the specific alleles is observed in a population. This was established empirically by profiling 100-200 unrelated individuals and concluding how often a single allele was present in the given population (Budowle *et al*, 1995[b]). An evidentiary profile could then be compared to the population database, and a number can be assigned to each corresponding allele. Based on Mendelian genetics, the frequencies of all alleles can be multiplied to give a comprehensive frequency of the DNA profile.

The first multiplex STR amplification system compatible with silver staining techniques was introduced in 1994 (Budowle *et al*, 1997). The CTT kit was designed with

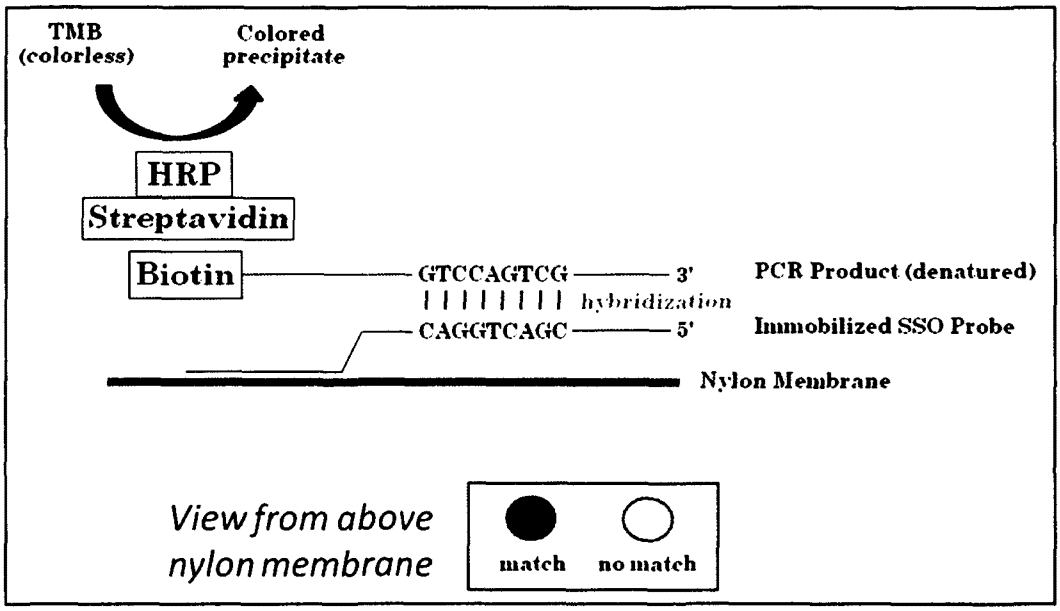

Figure 2-2: Reverse hybridization dot blot. Sequence specific oligonucleotides (SSO) are immobilized onto the nylon membrane. After hybridization of the biotin-labeled PCR product, the complex is exposed to TMB for a color producing reaction.

primers that were developed to exam the regions of three loci: CSF1PO, TH01 and TPOX (hence the name CTT). Silver staining is more sensitive than the routinely used ethidium bromide. Furthermore, it can detect down to 10ng of DNA on a polyacrylamide gel, while ethidium bromide at its most sensitive can detect levels of approximately 50ng. However, there are many cases that contain DNA samples with less than 10ng. Therefore, a more sensitive mode of detection was needed in the field of forensic DNA.

In the mid-1990s, fluorescence was implemented into STR technology by labeling the primers used in the PCR reaction. With this innovative technology, the ability to multiplex DNA fragments of similar sizes and lower input DNA amounts had finally been developed (Buse *et al*, 2003). Though initially the discriminating power was less than that of RFLPs due to a lower number of alleles per locus, the relatively rapid process enabled

the identification of additional discriminating loci in a timely manner. This allowed an analyst to test enough loci to provide for statistical analysis that is both incriminating and discriminating with respect to identification (Chakraborty *et al*, 1999).

Thousands of polymorphic regions of the human genome have been profiled to determine the flanking regions and the degree of hypervariability among modern humans (Urquhart *et al*, 1994). Since over 99.7% of the human genome is the same from one individual to the next, it is important to distinguish which regions in the genome will better suit the requirements of the identification process over others. Dinucleotide and trinucleotide STRs were developed in the beginning stages, however, tetra-, penta-, and hexa-nucleotides are considered to be a better choice when using PCR technology (Gill *et al*, 2006). The shorter repeats are more apt to produce mistakes called stutter. Stutter occurs when the enzyme slips off the DNA strand and produces a fragment that is one repeat short of the authentic DNA pattern. It may also be harder to distinguish a size difference on a polyacrylamide gel if only one repeat difference exists between a di- and tri-nucleotide.

In the late 1990s, it was apparent that the forensic community in the United States was using many different DNA techniques for human identification. The decision was made to conduct a study to evaluate 22 different loci and establish which candidates were polymorphic enough to yield statistical discrimination (*i.e.*, have a frequency of occurrence in the range of one in a trillion (10^{12}) individuals when all loci are combined). Following exhaustive studies involving many laboratories, the 13 combined DNA index system

(CODIS) core STR loci that were chosen and still are used today are included in Table 2-1 (Butler, 2009).

Upon the unified decision of which loci should be analyzed, a number of commercially available kits for STR analysis arose. The early kits were limited by the ability to only examine 3-4 loci at a time; however, when used in combination, provided statistically significant discriminatory data. AmpFℓSTR® Blue™ (Applied Biosystems) and AmpFℓSTR® Green I™ (Applied Biosystems) were the first kits available amplifying loci D3S1358, vWA, FGA, Amelogenin, TH01, TPOX, and CSF1PO, respectively. A random match probability (RMP) is the estimated frequency at which a specific STR profile would be expected to occur in a given population; that is, if a person is selected at random from a population, it is the likelihood that he or she will have the DNA profile in question by random chance. AmpFℓSTR® Blue™ (Applied Biosystems) had the ability to generate a RMP of approximately 1 in 1000, while AmpFℓSTR® Green I™ (Applied Biosystems) could manage one more order of magnitude. Combining the two calculated probabilities using the product rule allowed the power of discrimination to grow exponentially.

In order to save time and sample, the industry produced multiplex kits that contained seven and ten primer sets named AmpFℓSTR® COfiler™ (Applied Biosystems) and AmpFℓSTR® Profiler™ (Applied Biosystems) respectively; this time with more than one fluorophore per reaction. Currently, AmpFℓSTR® Identifiler™ (Applied Biosystems), which simultaneously amplifies fifteen loci in addition to the sex determining locus, amelogenin, is most widely used in the United States. In recent years, AmpFℓSTR®

Table 2-1: Combined DNA Index System STR Loci.

Locus Designation	PCR Product Size (bp)	Chromosome Location	Category and Repeat Motif
TPOX	222-249	2p23-2per	* GAAT
D3S1358	112-140	3p	‡ TCTG/TCTA
FGA	215-355	4q28	‡ CTTT/TTCC
D5S818	134-172	5q21-31	* AGAT
CSF1PO	305-342	5q33.3-34	* TAGA
D7S820	255-292	7q11.21-22	† GATA
D8S1179	123-170	8	‡ TCTA/TCTG
TH01	163-202	11p15.5	† TCAT
vWA	155-207	12p12-pter	‡ TCTG/TCTA
D13S317	217-245	13q22-31	* TATC
D16S539	252-293	16q24-qter	* GATA
D18S51	262-345	18q21.3	† AGAA
D21S11	185-240	21q11.2-q21	§ TCTA/TCTG

* Simple repeat: consists of one repeating sequence *i.e.,* GAATGAATGAAT (n=3, three full repeats)
† Simple repeat with nonconsensus alleles: consists of one repeating sequence often missing at least one base pair in one repeat *i.e.,* GATAGATAGAT (n=2.3, two repeats plus three bases)

‡ Compound repeat with nonconsensus alleles: consists of two or more different repeat sequences *i.e.,* TCTGTCTATCTGTCTA (n=4, four full repeats alternating sequences)

§ Complex repeat: consists of repeats of variable length and/or sequences *i.e.,* TCTATCTCTG (n=2.2, two full repeats plus two base pairs)

Yfiler™ (Applied Biosystems) was developed, which tests sixteen Y-chromosome sites, along with AmpFℓSTR® MiniFiler™ (Applied Biosystems) which was designed for nine smaller, possibly degraded, microsatellites.

In addition to nuclear genomic DNA, human cells contain mitochondria with a separate genome; referred to as mitochondrial DNA (mtDNA). The information in this

mtDNA genome can also be utilized for human identification. However, mtDNA lacks the discriminating power of nuclear DNA due to a number of features (Budowle *et al*, 2003; Melton, 2004). First, the circular genome contains just over 16,000 base pairs and does not contain the discriminating power that nuclear DNA holds (Anderson *et al*, 1981). Second, the mitochondria are inherited through the maternal lineage, leading to the fact that all maternal relatives will possess the same mitochondrial DNA sequence (Chen and Butow, 2005). Third, in addition to the lack of discrimination between maternal relatives, mtDNA also is characterized by heteroplasmy. In other words, mitochondria from the same individual may contain more than one mtDNA pattern within a single cell, may contain one type of mtDNA in one tissue that is different from another tissue, or display a homoplasmic mtDNA pattern in one tissue and heteroplasmic in another (Ingman *et al*, 2000). Finally, the mutation rate of mitochondria is up to 5-10 times greater than the mutation rate of the nuclear DNA genome (Scheffler, 1999) making mtDNA testing difficult and often inconclusive.

Despite the disadvantages listed above with respect to mtDNA, one of the biggest advantages of mtDNA testing is the fact that mitochondria have the ability to remain intact in extreme conditions. This allows for analysis of samples that are aged or exposed to a harsh environment (*i.e.*, archeological). The entire circular genome can be sequenced using the Sanger sequencing method to compare point mutations in the two hypervariable regions (HVI and HVII) in the non-coding region of the mitochondrial DNA (Bendall *et al*, 1996). The sequences are then compared to a database for genotype assignment and statistical determination.

In addition to sample stability, a major obstacle for forensic DNA analysis is limited sample acquisition. One technique to potentially overcome this limitation is whole genome amplification (WGA). The basis of this technique is to use one or more enzymes to replicate (*i.e.*, amplify) the entire genome of the sample for further downstream testing. There are many different approaches to whole genome amplification, including multiple displacement amplification (MDA) which relies on the φ29 polymerase to displace the double stranded DNA and synthesize two complementary strands with the guidance of random hexamer primers. Figure 2-3 illustrates the MDA technique in a detailed manner.

Other techniques (*e.g.*, PCR primer extension pre-amplification (PCR-PEP) and degenerate oligonucleotide primed PCR (DOP-PCR)) involve two enzymes, *Taq* DNA polymerase and another polymerase containing proofreading ability. The problem with these thermal cycling methods lies in the fact that they both produce amplification bias caused by the uneven distribution of the GC –rich regions in the genome (Sun *et al*, 2005). It has also been observed that these techniques produce non-specific artifacts in addition to incomplete coverage of the genome (Cheung *et al*, 1996; Sun *et al*, 1995). Methods like *improved* PCR primer extension (I-PEP) have been developed and have alleviated some of the mentioned concerns, but MDA is the leader in WGA due to the wide array of downstream applications it can adapt to (Ng *et al*, 2005).

Another potential course to increase the efficacy of STR analysis is not by increasing the amount of sample available for use, but the utilization of macromolecular crowding. Molecular crowding has been used to increase efficiency of DNA ligation reactions and cloning into yeast by the addition of polyethylene glycol (PEG; Pheiffer *et al*,

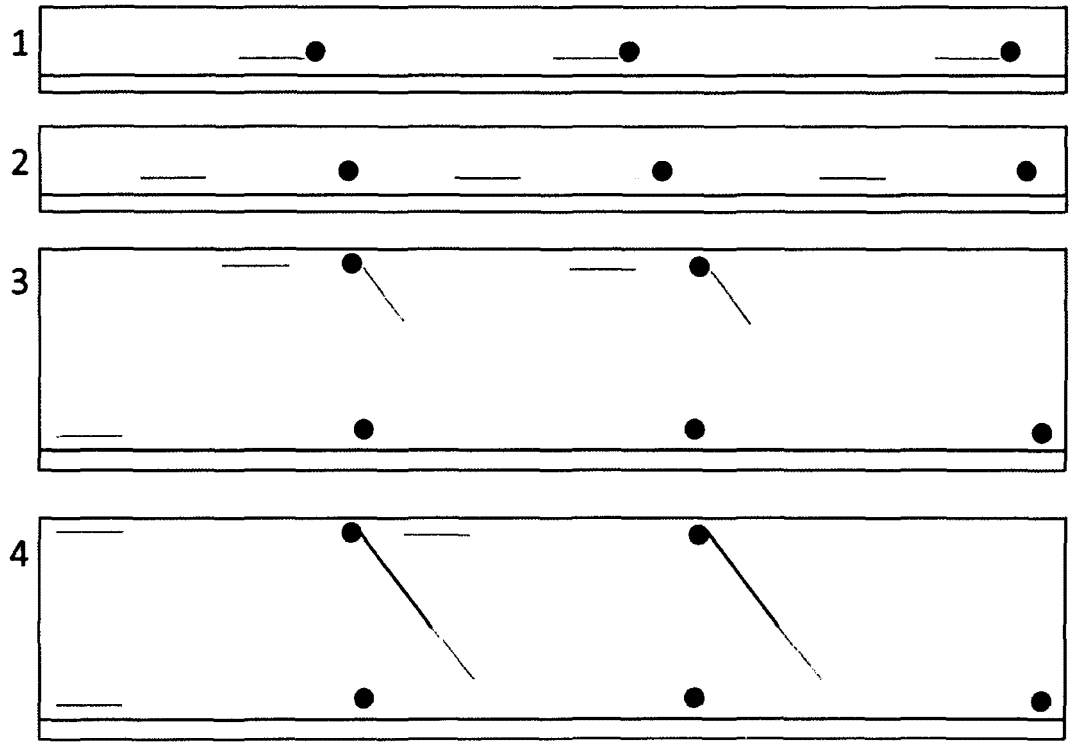

Figure 2-3: Whole genome amplification (WGA) by multiple displacement amplification (MDA). (1) Random hexamer primers in light blue anneal to the alkaline denatured DNA, represented by the green line; (2) The dark blue circle, representing the φ29 DNA polymerase synthesizes new double stranded DNA (orange line) by extending the random primers (3 and 4). As synthesis of new DNA continues, the enzyme continues to displace the double stranded characteristic of the DNA, enabling the random primers to anneal to the newly synthesized DNA for another round of replication.

1983). Another approach to molecular crowding is instead of adding a synthetic crowding agent like PEG, a natural nucleic acid can also be added to act as the molecular crowding agent. The basis of molecular crowding is to reduce the excess volume of the reaction, by driving the intermolecular forces into proximity of one another for a more efficient reaction than without the crowding agent. The excluded volume alters the T_m property of the DNA, encouraging increased polymerase binding affinity (Jarvis *et al*, 1990; Karimata *et al*, 2004).

This thesis explores the two techniques of MDA and PEP, along with synthetic and natural macromolecular crowding, in an effort to determine methodology best suited to conserve and potentially amplify limited amounts of DNA. The desired outcome was to maximize and improve the ability to obtain identity-related information from such limited samples that are often unsuitable for testing, with the ultimate goal of developing a full STR DNA pattern from a single cell.

REFERENCES

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, et al. Sequence and organization of the human mitochondrial genome. Nature. 1981 (290): 457-465.

Applied Biosystems. AmpFℓSTR® Identifiler™ PCR Amplification kit user's manual. Foster City, CA: Applied Biosystems, 2001.

Applied Biosystems. AmpFℓSTR® YFiler™ PCR Amplification kit user's manual. Foster City, CA: Applied Biosystems, 2006.

Applied Biosystems. Quantifiler® Human DNA Quantificaiton Kit and Quantifiler® y Human Male DNA Quantification Kit user's manual. Foster City, CA: Applied Biosystems, 2006.

Bendall KE, Macaulay VA, Baker JR, and Sykes BC. Heteroplasmic point mutations in the human mtDNA control region. Am J Hum Genet. 1996 Dec (59):1276-1287.

Blake E, Mihalovich J, Higuchi R, Walsh PS, and Erlich H. Polymerase chain reaction (PCR) amplification and human leukocyte antigen (HLA)-DQα oligonucleotide typing on biological evidence samples: Casework experience. J Forensic Sci. 1992 May (37):700-726.

Block EB. Science vs. Crime: The Evolution of the Police Lab. Cragmont Publications. 1979.

Bloodbook. http://www.bloodbook.com/type-facts.html January 31, 2011.

Budowle B, Baechtel FS, Comey CT, Guisti AM, and Klevan L. DNA protocols for typing forensic biological evidence: Chemiluminescent detection for human DNA quantitation and restriction fragment length polymorphism (RFLP) analyses and

manual typing of polymerase chain reaction (PCR) amplified polymorphisms. Electrophoresis. 1995 (16): 1559-1567.

Budowle B, Lindsey JA, Decou JA, Koons BW, Giusti AM and Comey CT. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci), and HLA-DQ alpha using a multiplex amplification and typing procedure. J Forensic Sci. 1995 Jan (40):45-54.

Budowle B, Moretti TR, Keys KM, Koons BW, and Smerick JB. Validation studies of the CTT multiplex system. J Forensic Sci. 1997 (42): 701-707.

Budowle B, Moretti T, Smith J. DNA typing protocols: Molecular biology and forensic analysis. Natick, MA: Eaton Publishing. 2000.

Budowle B, Smith JA, Moretti T, and DiZinno J. Forensics and Mitochondrial DNA: Applications, debates, and foundations. Annu Rev of Gen Hum Genet. 2003(4):119-141.

Buse JM, Putinier JC, Hong MM, Yap AE, and Hartmann JM. Performance evaluation of two multiplexes used in fluorescent short tandem repeat DNA analysis. J Forensic Sci. 2003 (48): 348-357.

Butler JM. Fundamentals of Forensic DNA Typing. New York: Elsevier, 2009.

Casey B, Taylor GR, and Linacre A. Review of the science of low template DNA Analysis. 2008. http://police.homeoffice.gov.uk/publications/operational-policing/Review_of_Low_Template_DNA_1.pdf January 24th, 2011.

Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B. The utility of short tandem repeat loci beyond human identification: Implications for development of new DNA typing systems. Electrophoresis. 1999 (20): 1682-1696.

Chen XJ and Butow RA. The organization and inheritance of the mitochondrial genome. Nature Rev Gen. 2005 (6): 815-825.

Cheung VG and Nelson SF. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. Proc Natl Acad Sci USA. 1996 (93): 14676-14679.

Comey CT, Budowle B, Adams DE, Baumstark AL, Lindsey JA, Presley LA. PCR amplification and typing in the HLA DQ alpha gene in forensic samples. J Forensic Sci. 1993 Mar (38):239-249.

Edwards H and Gotsonis C. Strengthening Forensic Science in the United States: A Path Forward. Washington DC: National Academies Press. 2009 Feb. url: http://nationalacademies.org/morenews/20090218.html.

Fildes N and Reynolds R. Consistency and reproducibility of AmpliType PM results between seven laboratories: Field trial results. J Forensic Sci. 1995 Mar (40):279-286.

Frauds H. Fingerprint Identification. Nature. 1905 Oct (72):iv-v.

Gaensslen RE. Sourcebook in Forensic Serology, Unit IX: Translations of Selected Contributions to the Original Literature of Medicolegal Examination of Blood and Body Fluids. National Institute of Justice. 1983.

Gill P, Fereday L, Morling N, Schneider PM. The evolution of DNA databases—Recommendations for new European STR loci. Forensic Sci Int. 2006 (163): 155-157.

Ingman M, Kaessmann H, Pääbo S, Gyllensten U. Mitochondrial genome variation and the origin of modern humans. Nature. 2000 (408): 708-713.

Inman K and Rudin N. Principles and Practice of Criminalistics: The Profession of Forensic Science. CRC Press. 2000 Sept.

Jarvis, Ring DM, Daube SS, von Hippel PH. "Macromolecular crowding": thermodynamic consequences for protein-protein interactions within the T4 DNA replication complex, J boil Chem. 1990 (265): 15160-15167.

Jeffreys AJ, Wilson V, Thein SL. Hypervariable "minisatellite" regions in human DNA. Nature. 1985 (314):67-73.

Jeffreys AJ. Genetic Fingerprinting. Nature Medicine. 2005 (11):1035-1039.

Karimata H, Nakano S, Ohmichi T, Kawakami J and Sugimoto N. Stabilization of a DNA duplex under molecular crowding conditions of PEG. Nucleic Acids Symp Ser. 2004 (48): 1871-1875.

Melton T. Mitochondrial DNA heteroplasmy. Forensic Sci Rev. 2004(16):1-20.

Mullis, K. Dancing Naked in the Mind Field. New York, NY: Vintage Books. 1998.

Ng G, Roberts I, Coleman N. Evaluation of 3 methods of whole-genome amplification for subsequence metaphase comparative genomic hybridization. Diagn Mol Pathol. 2005 (14): 203-212.

Pheiffer BH and Zimmerman SB. Polymer-stimulated ligation: enhanced blunt or cohesive-end ligation of DNA or deoxyribooligonucleotides by T4 DNA ligase in polymer solutions. Nucleic Acid Res. 1983 (11): 7853-7871.

Scheffler IE. Mitochondria. New York: Wiley-Liss. 1999.

Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol. 1975 (98): 503-517.

Sun FZ, Arnheim N, and Waterman MS. Whole genome amplification of single cells: mathematical analysis of PEP and tagged PCR. Nucleic Acids Res. 1995 (23): 3034-3040.

Sun G, Kaushal R, Pal P, Wolujewicz M, Smelser D, Cheng H, Lu M, Chakraborty R, et al. Whole-genome amplification: Relative efficiencies of the current methods. Leg Med. 2005 (7): 279-286.

Urquhart A, Kimpton CP, Downes TJ, and Gill P. Variation in short tandem repeat sequences—A survey of twelve microsatellite loci for use as forensic identification markers. Int J Legal Med. 1994 (107): 13-20.

Walsh PS, Varlaro J, and Reynolds R. A rapid chemiluminescent method for quantitation of human DNA. Nucleic Acids Res. 1992 (20): 5061-5065.

CHAPTER 3.

COMPARISON OF PCR PRIMER EXTENSION PRE-AMPLIFICATION VERSUS MULTIPLE DISPLACEMENT AMPLIFICATION

INTRODUCTION

A common problem encountered in forensic science is the limited sample size available for DNA analysis. Such cases, both new and old, present challenges to the analyst whether to consume the sample in its entirety and hope for a usable result or save the sample, in anticipation, that future technologies will develop assays sensitive enough to allow successful analysis. Short tandem repeat (STR) analysis itself has become increasingly sensitive over the years, but the limitations of the human identification kits currently in use have been reached, thus requiring the scientific and forensic communities to explore new avenues to expand the types of samples that can be successfully analyzed. One of the most promising is whole genome amplification (WGA); we have shown it to be a reliable technique that increases authentic DNA template thereby alleviating the problem of limited sample.

There are many techniques available to accomplish WGA; however, each technique varies in efficiency and ease of use. Our studies examined the differences between PCR primer extension pre-amplification (PEP) and multiple displacement amplification (MDA). PCR primer extension pre-amplification's initial step is a cell lysis combined with a lengthy fragmentation step prior to library preparation and amplification. Traditionally, PEP uses *Taq* DNA polymerase to produce multiple copies of DNA template by applying

the traditional PCR steps of denaturation, annealing, and extension. However, unlike PCR that is targeted to amplify a specific DNA fragment, this PCR amplification employs random primers to encourage an even amplification of the entire initial DNA template (Zhang *et al.* 1992). In contrast, MDA, utilizes an alkaline denaturation buffer with a short incubation time immediately before the isothermal amplification takes place. Unlike traditional PCR, continuous amplification takes places using the φ29 DNA polymerase. In addition to producing remarkably long DNA products (up to 70kbp) the φ29 DNA polymerase also possesses 3' → 5' exonuclease activity (Ballantyne *et al.* 2007). Proofreading activity is an important attribute some enzymes have to ensure the fidelity of the enzyme to produce authentic copies of the original DNA template.

Although WGA is a powerful tool for increasing the amount of template DNA in samples with limited starting material, WGA also increases the background noise (Applied Biosystems, 2001). This background noise consists of peak imbalances often associated with heterozygote genotypes, allele drop-in of spurious peaks not observed in traditional genotyping methods, allele drop-out with low levels of template DNA, and amplification-related artifacts such as stutter peaks and incomplete 3' nucleotide addition as *Taq* polymerase arbitrarily adds an adenine to the end of each DNA copy produced (Butler, 2009).

The goal of this set of experiments was to determine which technique was superior at producing a reliable DNA template as measured by the downstream applications of real-time PCR for quantitation and STR analysis for identification.

MATERIALS AND METHODS

Laser Catapult Microscopy

Laser catapult microscopy is a powerful tool which provides absolute precision during sample collection. Duplicate sets of 50 and 25 HUT-78 (TIB-161) cells were collected using micro dissection via a photoactivated localization microscopy (P.A.L.M.™) Microbeam Laser Catapult Microscope (LCM; Zeiss). All quantities of input template DNA were converted to cellular equivalents based on 6.6 picograms (pg) per diploid cell and 3.3 pg per haploid cell (Butler, 2009). The cells were then lysed in the adhesive tube cap (Zeiss) using Qiagen's REPLI-g® Ultrafast Mini Kit denaturation buffer for Genomic DNA from Blood or Cells or by following the cell lysis and fragmentation step using Sigma-Aldrich's GenomePlex kit.

Duplicate sets of 50 and 25 HUT-78 and DLD-1 cells and 50 Ramos cells were also collected and genomic DNA was isolated using Qiagen's QIAamp® DNA Micro Kit. These samples were not subjected to WGA, but were processed with the downstream application of STR analysis only.

Genomic DNA Sample Preparation

Genomic DNA was isolated by following Qiagen's QIAamp® DNA Micro Kit protocol for Isolation of Genomic DNA from Laser-Microdissected tissues. Cells collected via LCM were lysed by adding 15µl Buffer ATL and 10µl Proteinase K and incubating at 56°C for 3 hours. The volume was doubled at the end of the incubation by adding an additional 25µl Buffer ATL. After mixing the sample by vortexing, 50µl Buffer AL was added to ensure complete lysis. Once the solution is homogenous, 50µl Absolute Ethanol

(Sigma) was added and the samples were incubated at room temperature for 5 minutes. The entire lysates was transferred to a QIAamp MinElute silica based column and centrifuged at 6000 x *g* for 1 minute. Five hundred microliters of Buffer AW1 was used to wash salts and proteins from the samples by centrifuging for another minute at 6000 x *g*. A final wash was performed using 500µl of Buffer AW2, again centrifuging for 1 minute at 6000 x *g*. The membrane was dried by spinning at maximum revolutions for 3 minutes. Prewarmed at 50°C, 20µl of water was added to elute the DNA from the column by incubating at room temperature for 5 minutes before a last centrifugation of 1 minute at maximum revolutions.

Qiagen REPLI-g® Ultrafast Mini Kit: Amplification of Genomic DNA from Blood or Cells (Qiagen)

Sufficient denaturation buffer (D2 Buffer) was prepared for the total number of whole genome amplification reactions (DLB Buffer with a final DTT concentration of 83mM). 1.5µl of PBS was added to the cellular material on the adhesive tube cap followed by 1.5µl D2 Buffer to cellular material and PBS on adhesive tube cap. The entire volume was manually aspirated around the tube cap to encapsulate all cellular material into the limited volume reaction. The tubes were incubated upside-down for 10 minutes on ice and 1.5µl Stop Solution was added after incubation. Vortexing was used to mix the samples before they were briefly and centrifuged. The WGA master mix was prepared as follows per reaction: 15µl REPLI-g® UltraFast Reaction Buffer and 1µl REPLI-g® UltraFast DNA Polymerase. 16µl of the master mix was added to the 4.5µl of prepared denatured DNA

and incubated at 30°C for 1.5 hours. The REPLI-g® UltraFast DNA Polymerase was inactivated by heating the samples for 5 minutes at 65°C.

Time titration experiments of the Qiagen REPLI-g® Ultrafast Mini Kit were as stated above, except that after addition of 16 µl of the master mix, each duplicate set was incubated at 30°C for 15, 30, 45, 50, 55, 60 or 90 minutes. The positive and negative controls remained in the 90 minute incubation time as recommended by the manufacturer. The REPLI-g® UltraFast DNA Polymerase was inactivated by heating the samples for 5 minutes at 65°C.

Sigma-Aldrich GenomePlex Single Cell Whole Genome Amplification Kit (Sigma-Aldrich)

Single Cell Lysis and Fragmentation: Cells were isolated into a PCR-ready vessel using laser capture micro-dissection. A sufficient volume of water was added to the cell sample for a final volume of 9µl. A working solution of Lysis and Fragmentation Buffer was prepared by adding 2µl of Proteinase K Solution into 32µl of the 10X Single Cell Lysis & Fragmentation Buffer. After vortexing thoroughly 1µl of the freshly prepared Proteinase K Solution—10X Single Cell Lysis & Fragmentation Buffer was added to the cell sample. The samples were mixed thoroughly and incubated at 50°C for 1 hour, then heated to 99°C for exactly four minutes. Note that the incubation is very time sensitive; any deviation may alter results according to the manufacturer. Samples were cooled on ice and collected at the bottom of the tube via centrifugation prior to proceeding to Library Preparation.

Library Preparation: 2µl of 1X Single Cell Library Preparation Buffer was added to each sample. 1µl of Library Stabilization Solution was then added and mixed thoroughly before placing in heat block at 95°C for 2 minutes. The samples were cooled on ice, and consolidated by centrifugation before returning to ice. 1µl Library Preparation Enzyme was added before the samples were placed in a thermal cycler and incubated as follows: 16°C for 20 minutes, 24°C for 20 minutes, 37°C for 20 minutes, 75°C for 5 minutes, 4°C hold. Samples were removed from thermal cycler and centrifuged briefly. The samples were amplified immediately or stored at -20°C for three days.

Amplification: The following reagents were added to the entire 14µl reaction: 7.5µl of 10X Amplification Master Mix, 48.5µl of Nuclease-Free Water, 5.0µl of WGA DNA Polymerase. Components were mix thoroughly, centrifuged briefly and subjected to thermal cycling as follows: 95°C for 3 minutes, 25 cycles of 94°C for 30 seconds, 65°C for 5 minutes, hold at 4°C.

Sample clean up using post amplification purification

Microconcentration clean up of each sample was performed using the Amicon® Ultra-0.5 Centrifugal Filter Device (100,000 MW exclusion) after WGA (Millipore). Microcon filters were placed into Micron 1.5ml microcentrifuge tubes and 100µl of TE^{-4} (0.1mM EDTA, 10mM Tris, pH 8.0) buffer was added to pre-wet the membrane. The deactivated WGA reaction was transferred into the Microcon 100 and centrifuged for 5 minutes at 14,000 RPM. The flow through was discarded before another 200µl TE^{-4} buffer was added to the Microcon 100 for a second wash step. Centrifugation for 20 minutes at 14,000 RPM was performed and again the flow through was discarded. The Microcon

filter was inverted into a new Micron 1.5ml microcentrifuge tube and centrifuged for 5 minutes at 2500 RPM. The recovered volume was measured and adjusted to 12µl if needed using TE^{-4} buffer.

Applied Biosystems Quantifiler™ Human Real Time Analysis

This kit was designed to quantify the total amount of amplifiable human DNA in a sample to determine if sufficient DNA is present to proceed with STR analysis. The DNA quantification assay combined two 5′ nuclease assays: A target-specific human DNA assay (Human- hTERT) and an internal PCR control assay (a synthetic sequence not found in nature). For each assay, a specific fluorescent probe was utilized. The amplified product was detected using fluorescence emission data with tungsten-halogen lamp excitation and CCD camera detection.

Sample yields were estimated using the Real Time PCR Quantifiler™ Human DNA Quantification Kit (Applied Biosystems) as measured by AB Prism® 7500 Sequence Detection System (SDS v1.2). Thermocycler conditions were set at the manufacturer's recommendations of 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds, 60°C for 1 minute.

DNA Quantification Standards were prepared using a TE-4 buffer containing 0.1% Glycogen as the diluent. See Table 3-1 for more information on the standard dilution series.

Absolute quatitation assays were performed following the manufacturer's guidelines at half volume reactions as follows per reaction: 5.25µl Quantifiler™ Human Primer Mix and 6.25µl Quantifiler™ PCR Reaction Mix. One microliter of standard

Table 3-1: Standard curve dilution series used for absolute quantitation via Real Time PCR.

Standard	[DNA] (ng/µl)	Example Amounts
Std. 1	50.000	50µl [200ng/µl] stock + 150µl TE^{-4}/glycogen buffer
Std. 2	16.700	50µl [Std. 1] + 100µl TE^{-4}/glycogen buffer
Std. 3	5.560	50µl [Std. 2] + 100µl TE^{-4}/glycogen buffer
Std. 4	1.850	50µl [Std. 3] + 100µl TE^{-4}/glycogen buffer
Std. 5	0.620	50µl [Std. 4] + 100µl TE^{-4}/glycogen buffer
Std. 6	0.210	50µl [Std. 5] + 100µl TE^{-4}/glycogen buffer
Std. 7	0.068	50µl [Std. 6] + 100µl TE^{-4}/glycogen buffer
Std. 8	0.023	50µl [Std. 7] + 100µl TE^{-4}/glycogen buffer

and/or sample was added to each well for a total reaction volume of 12.5µl. The reaction plate was sealed with an Optical Adhesive Cover and spun at 3000 RPM for 5 minutes to remove any bubbles before loading plate into the Real Time 7500 instrument, as described above.

Applied Biosystems AmpFℓSTR® Identifiler™ and Yfiler™ and Promega's PowerPlex®16 amplification

Multiplex PCR for STR genotyping was performed with the AmpFℓSTR® Identifiler™ and Yfiler™ multiplex kits using an Applied Biosystems GeneAmp® 9700 thermal cycler instrument (Applied Biosystems), amplifying the gender identifying amelogenin locus and fifteen autosomal STR loci and sixteen Y-chromosome specific loci, respectively (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, FGA, Amelogenin, DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385 a/b, DYS393,

DYS391, DYS439, DYS635, DYS392, Y GATA H4, DYS437, DYS438, and DYS448). Approximately 1.0ng of each WGA sample was amplified using the AmpFℓSTR® Identifiler™ and Yfiler™ kits. The sample DNA was added to 7.5µl Identifiler™ master mix composed with the following components: 5.25µl AmpFℓSTR® PCR Reaction Mix, 2.75µl AmpFℓSTR® Identifiler™ Primer Set, and 0.25µl AmpliTaq® Gold DNA Polymerase. For the Yfiler™ reaction, 1.0ng of DNA was added to 7.5µl of the following master mix components: 4.6µl AmpFℓSTR® PCR Reaction Mix, 2.5µl AmpFℓSTR® Yfiler™ Primer Set, and 0.4µl AmpliTaq Gold® DNA Polymerase. Thermal cycling for Identifiler™ was performed in a AB9700 (Applied Biosystems) for 11 minutes at 95°C, followed by 28 cycles of 94°C for 1 minute, 59°C for 1 minute, and 72°C for 1 minute, as recommended by the manufacturer. A final extension step at 60°C for 60 minutes completed the amplification. Thermal cycling conditions for Yfiler™ were 11 minutes at 95°C, followed by 30 cycles of 94°C for 1 minute, 61°C for 1 minute, and 72°C for 1 minute, and a final extension step at 60°C for 80 minutes.

PCR amplifications using Promega's PowerPlex® 16 System and AmpliTaq Gold® DNA Polymerase (Applied Biosystems) were performed using an AB GeneAmp® 9700 Thermal Cycler to examine sixteen loci, including amelogenin (D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, CSF1PO, FGA, TH01, TPOX, vWA, amelogenin, Penta D, and Penta E). The amplification cycling parameters as described in the PowerPlex® 16 System technical manual were as follows: 95°C for 11 minutes, 96°C for 1 minute, ramp 100% to 94°C for 30 seconds, ramp 29% to 60°C for 30 seconds, ramp 23% to 70°C for 45 seconds for 10 cycles. The second set of amplifications

are as follows: ramp 100% to 90°C for 30 seconds, ramp 29% to 60°C for 30 seconds, ramp 23% to 70°C for 45 seconds for 22 cycles. A final extension set was performed at 60°C for 30 minutes.

Capillary electrophoresis and genotyping

The Capillary Electrophoresis (CE) instrument separates the amplified DNA fragments by size. In combination with the internal size/lane standard, it allows accurate sizing of the fragments, which are subsequently compared to the allelic ladder to determine the fragment allele.

An AB Prism 3130 Genetic Analyzer fitted with a 50μm by 36cm capillary loaded with POP-4 polymer (Applied Biosystems) was used for electrophoresis and product detection. PCR product (1μl) was loaded with 9.0μl Hi-Di formamide (Applied Biosystems) and 0.5μl of LIZ 600 size standard which incorporates fragments ranging from 40-600bp (Applied Biosystems).

The PowerPlex®16 HS System employs four fluorescent dyes. Spectral resolution was established using the Promega PowerPlex® Matrix Standards to allow evaluation of each fluorescent dye contained in the kit. All analyses used the ILS 600 size standard and allelic ladder mix provided with the PowerPlex®16 HS System. Separation of amplification products was performed on the Applied Biosystems 3130 Genetic Analyzer. Generally, 1 μL of amplified sample or ladder was prepared in 9μL deionized formamide containing 1μl ILS-600 (fragments ranging from 60-600bp).

A positive control sample was analyzed with every analysis batch. The documented human DNA control cell line 9947a, of a known phenotype/genotype produces

a known allele pattern for each STR locus. This control serves as a system's check for the following functions: electrophoretic resolution, sizing precision, locus identity and completion of the amplification process. A negative amplification control and DNA extraction reagent blank control (DNA ERC) were also included in each assay. These controls should produce no callable peaks above the allele interpretation thresholds within the allele calling range with the exception of any reproducible anomalies. These controls function as system checks for reagent and cross-contamination.

One microliter of allele ladder was added as a reference sample for the human identification system used. A septum strip was placed across the tray making sure it was flush with the edges of the 96-well plate. If air bubbles were present, the 96-well plate was centrifuged briefly. All samples were heated at 95°C for 3 minutes in the thermal cycler or on a heat block and then snap cooled for at least 3 minutes in a thermal cycler or cooling block. Samples were injected for 3, 5 and 10s at 3kV at 60°C. Genemapper® 3.2 software (Applied Biosystems) was used for genotype analysis employing a detection threshold of 50 and 100 RFU. No stochastic threshold was set for designation of heterozygotes.

RESULTS

The achievement of generating a DNA profile from 50 and 25 HUT-78 and DLD-1 cells and 50 Ramos cells using Identifiler™, Yfiler™ and PowerPlex®16 (where applicable) was successful without utilizing WGA (Figures 3-1, 3-2 and 3-3). However, the samples exhibited allele drop-out, moderate-to-high peak imbalance, and perhaps most importantly, total consumption of the sample template. To remedy the existing low level templates, WGA was then applied to new duplicate samples to produce more DNA for

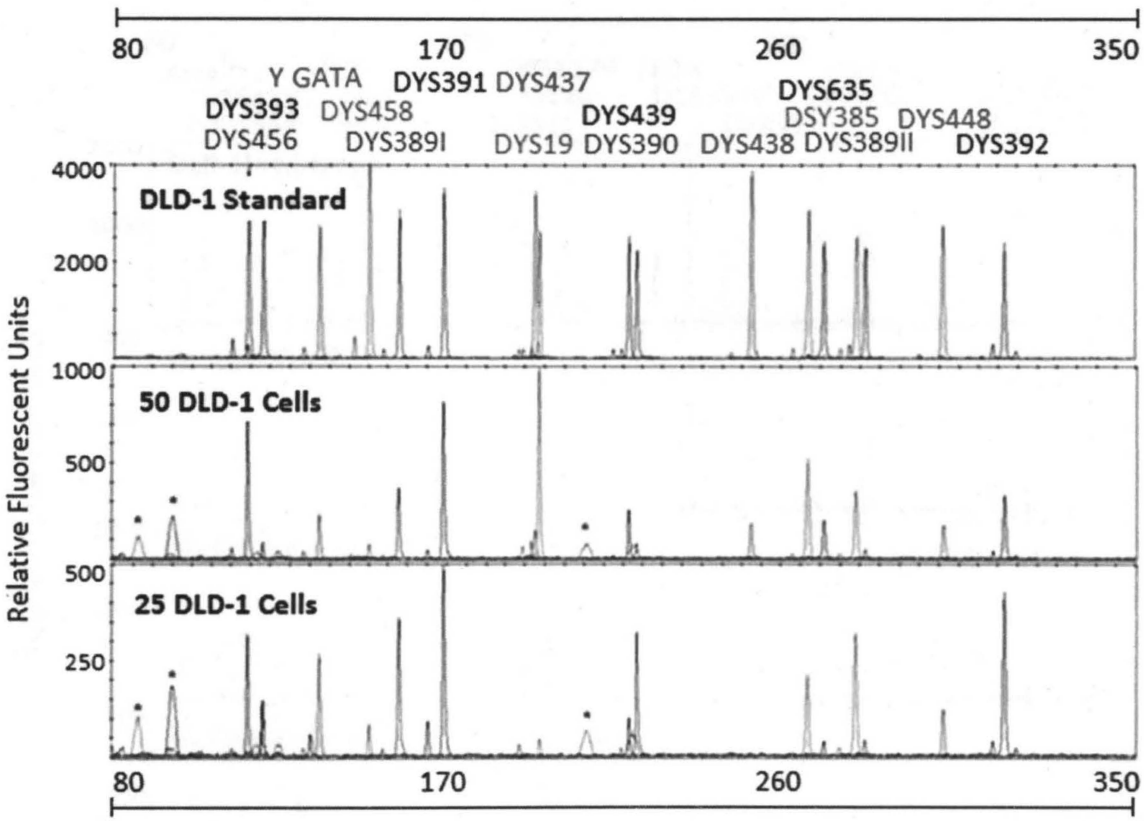

Figure 3-2: Genetic profiles produced by Yfiler™ from cells collected via laser catapult micro dissection. DNA extraction was performed using QiaAMP DNA Micro Kit on 50 or 25 DLD-1 cells. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values. * indicates a known reported artifact, most likely due to the fluorescent dye used.

quantification of the DNA from the WGA reactions, we found the MDA technique to consistently produce more reliable results than PEP by steadily producing orders of magnitudes more DNA, as shown in Table 3-2. The quantitation by real time PCR revealed MDA to be a much more robust and efficient technique for WGA than PEP by two orders of magnitude in all samples; not only did MDA produce more DNA than PEP, genotyping both products also revealed less allele drop-in (artificial allele calls of PCR products produced during amplification or developed due to slippage of the enzyme) in the

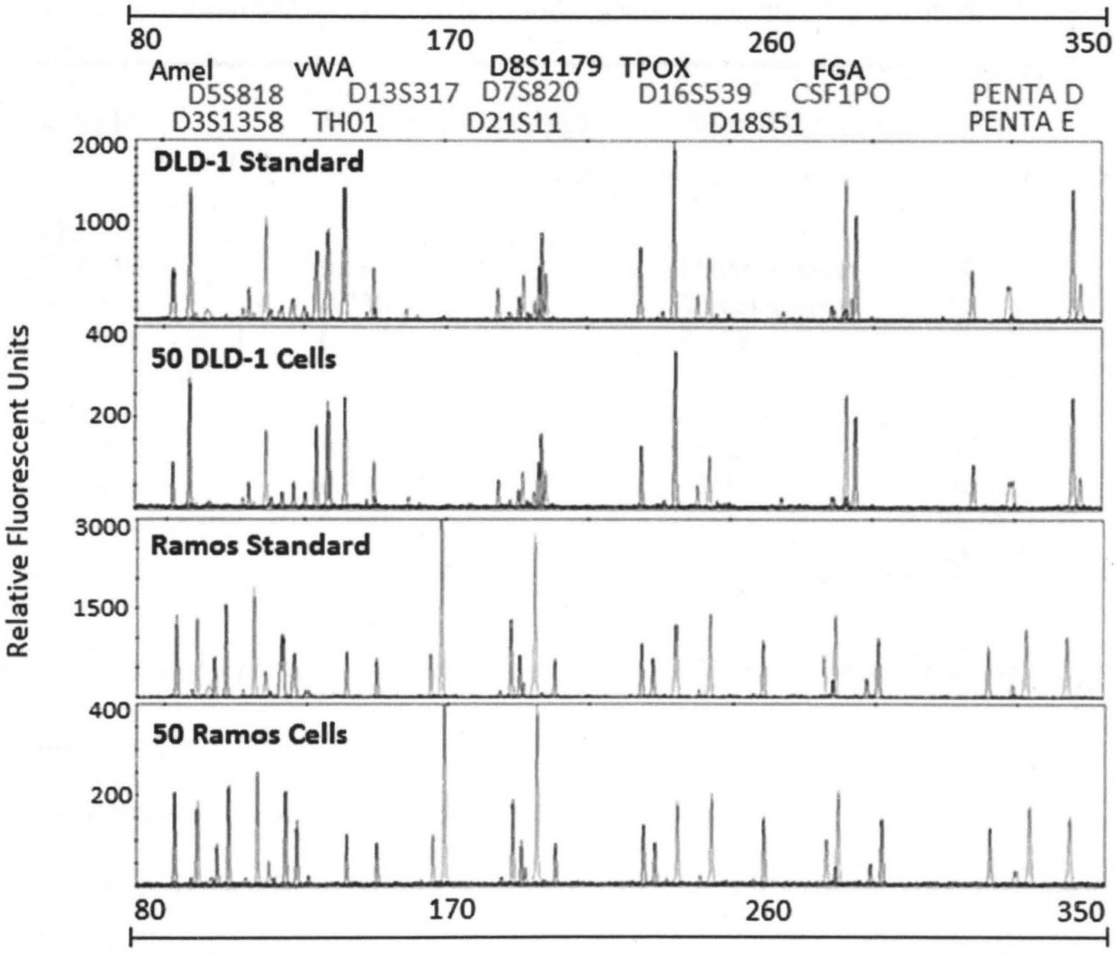

Figure 3-3: Genetic profiles produced by PowerPlex® 16 from cells collected via laser catapult micro dissection. DNA extraction was performed using QiaAMP DNA Micro Kit. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values.

MDA samples. Genotyping using Identifier™ confirmed these data by showing full DNA profiles for the MDA samples, while PEP samples exhibited high allele dropout (low or no signals in the electropherogram) and in many cases no detectable pattern at all (Figure 3-4). This observation is best explained by the fact that an insufficient amount of DNA template was added to the PEP amplification samples simply because the PEP WGA reaction did not

Table 3-2: Quantification of DNA following whole genome amplification.

Sample	DNA equivalent (ng)	After WGA (ng)	STD
PEP			
25 Cells	0.165	Undetectable	0
50 Cells	0.330	Undetectable	0
Positive Control	0.100	0.019	0.027
MDA			
25 Cells	0.165	4596.54	1258.62
50 Cells	0.330	3739.92	691.21
Positive Control	0.100	7.32	9.14

produce any additional template. However, raised baselines were observed in the MDA samples along with allele drop-in in some samples.

Multiple Displacement Amplification: Time Titration

The manufacturer's recommendation of a 90 minute isothermal incubation was effective for enhancing the signal of our DNA template, yet in addition to reproducing our authentic DNA template, the reaction also increased artificial signals. Artificial signals are often detected when the enzyme is compromised or the instruments used to detect samples are in need maintenance. Our artificial signals can be attributed to the WGA reaction was ruled out as a scheduled preventative maintenance regimen was strictly followed. In efforts to decrease the background signals produced by the MDA technique, we decided to modify the protocol by changing the length of the incubation time during which WGA takes place. Like the reduced cycles performed in the human identification kits, the truncated amplification step in the WGA reaction produced less artificial signals and lower baseline

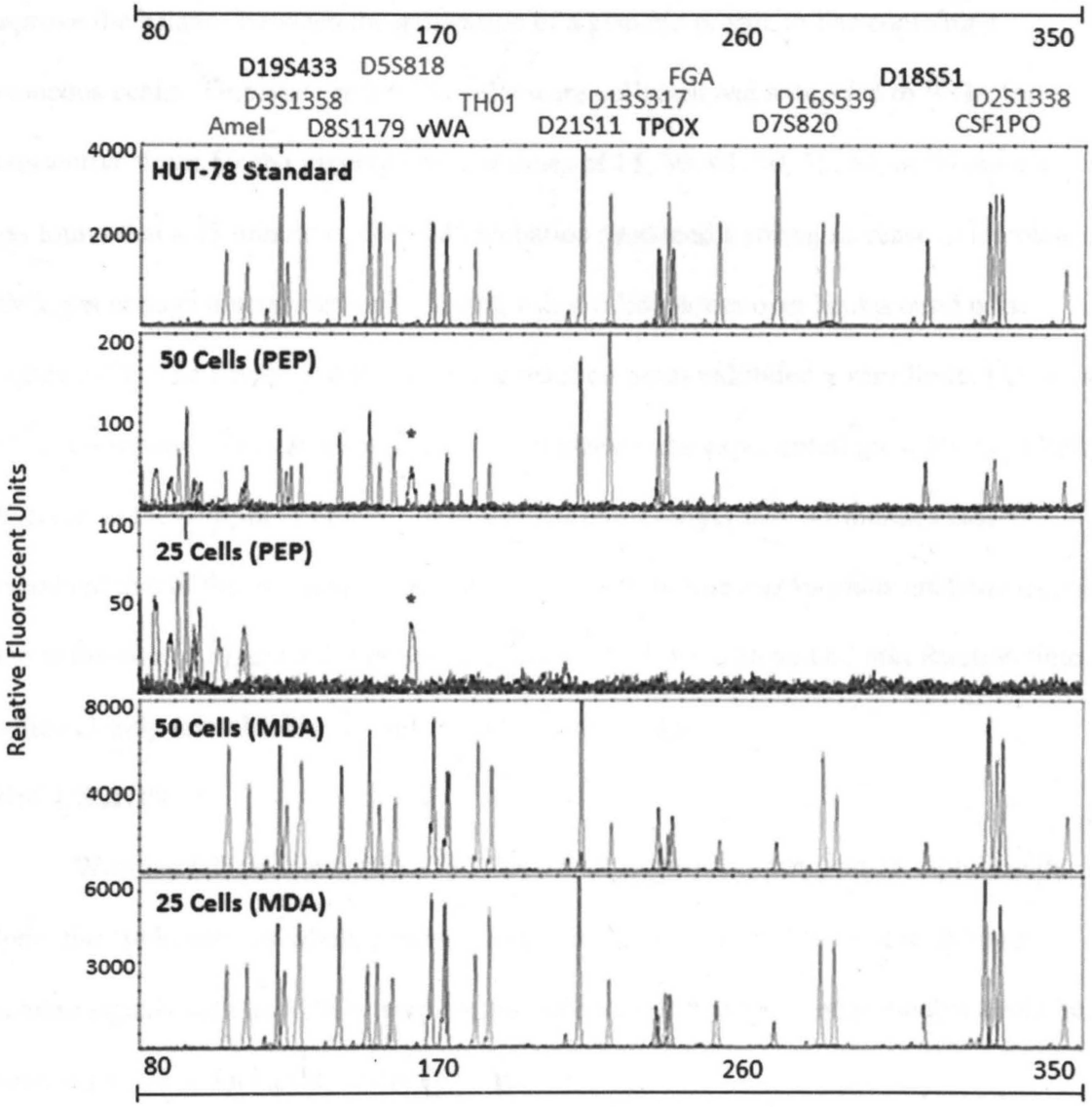

Figure 3-4: Comparison of whole genome amplification strategies. PCR Primer Preamplification (PEP) compared to Multiple Displacement Amplification (MDA) shows increased STR amplification using MDA while PEP shows minimal progression in whole genome amplification. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values. * indicates a known reported artifact, most likely due to the fluorescent dye used.

while still enriching the sample with authentic DNA template. Therefore, a time titration experiment was performed to determine whether a shorter incubation period would

improve the balance between the generation of a genuine profile to one containing erroneous peaks. Duplicate sets of six cells were collected and subjected to WGA three consecutive times for the varying reaction times of 15, 30, 45, 50, 55, 60, or 90 minutes. It was found that a 45 minute isothermal incubation produced a strong increase in template DNA, yet enough discrimination to distinguish a valid pattern over background noise (Figure 3-5). The fifteen and thirty minute reaction times exhibited a very limited increase in DNA template. Though the signals were increased, the exponential growth was initially observed in the 45 minute sample. The reaction times longer than 45 minutes closely resembled that of the 90 minute reaction with raised baseline and spurious artificial signals. Due to the size of Figure 3-5, the data was not shown for the 55 and 60 min reaction times as they closely resemble the 50 and 90 minute panels' data.

DISCUSSION

With the failure to develop a full DNA STR pattern from 50 and 25 diploid cells alone, the application of whole genome amplification was utilized to test the ability to increase signals with low DNA template and also to determine if enough sample could be produced to allow for further testing or future analysis.

We found that using an initial starting material of 50 and 25 HUT-78 cells (which are approximately equal to 0.165ng and 0.330ng, respectively) generated DNA equivalents of hundreds of thousands of cells (3,700-4,500ng of DNA) using MDA. One might hypothesize that twice as much DNA should have been generated during whole genome amplification for the sample with a higher initial DNA concentration (0.330ng *vs.* 0.165ng); however, we propose that the lack of DNA concentration correlation between the

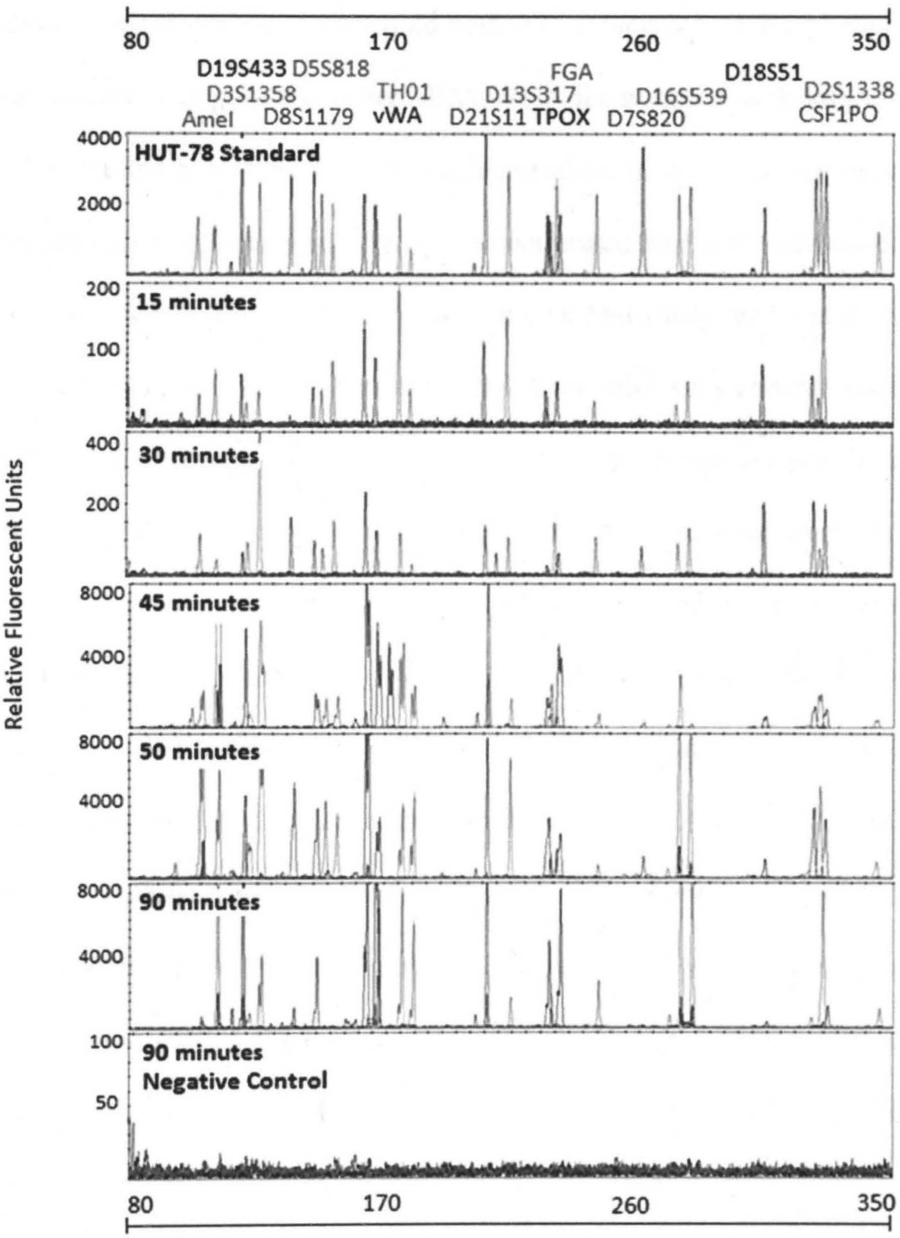

Figure 3-5: Comparison of varying incubation times for whole genome amplification. Multiple Displacement Amplification (MDA) shows increased STR amplification of 6 HUT-78 cells even at limited incubation times. Molecular grade water was added to the reaction instead of DNA template for the negative control. The positive control contained 100 picograms (pg) pre-extracted DNA, while the HUT-78 standard was profiled using 1.0ng DNA without WGA. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values.

50 and 25 diploid cell starting material could be due to exhaustion of the whole genome amplification reaction components. After MDA WGA the samples were subjected to STR analysis which resulted in high levels of amplified product, in fact, in some instances the CCD fluorescence data collection camera was oversaturated due to the amount of PCR product produced. The excessive signals produced could be attributed to an inaccurate quantification measurement, as the absolute quantitation utilizes a standard curve with a range of 0.023ng/µl-50.0ng/µl, therefore, any value that is estimated outside of this range is an extrapolation of the standard curve. In short, the PEP was not as successful when compared to MDA. Furthermore, even after samples were forwarded to STR analysis by amplifying the entire DNA sample, they yielded very little information due to allele drop out caused by lower than optimal input DNA. Since the quantitation value was undetectable, we can confidently assume the sample contained a concentration that was less than minimum DNA standard used of 0.023ng/µl. Therefore, it can be concluded that less than 115 pg were subjected to STR analysis, as the samples were concentrated to a volume of 5µl. Since starting DNA template was approximately 0.330ng and 0.165ng, one would expect at least that amount of DNA to be present during quantitation and STR analysis. It is possible the DNA was extensively nicked causing degradation during the WGA reaction of the PEP procedure.

The failure to generate an increase in template (or the ability to detect even the template by real-time PCR DNA quantification) using PEP WGA could be explained by the prolonged cell lysis and fragmentation step. Though the kits used contain premade buffers with proprietary reagents, in theory, the Lysis and Fragmentation buffer included in

the kit would induce denaturing conditions which lead to fragmentation most likely by DNase I, thus nicking the DNA extensively. If the DNA is fragmented significantly, the primers used to prepare the genomic library will not be able to anneal properly, thus inhibiting sufficient amplification of the genome. This extensive unraveling of the DNA template into short fragments could explain the inadequate results obtained from PEP's protocol. One might also speculate that the GenomePlex kit contains an inhibitor to the downstream PCR applications, thus resulting in decreased quantification of DNA. Furthermore, even though each sample was subjected to a column dialysis cleanup, there may still be enough residual inhibitor present in the sample to prevent subsequent successful PCR amplification necessary for Identifiler to generate complete and accurate profiles. It was demonstrated that MDA was superior in our hands when compared to PEP whole genome amplification.

When applied to a forensic laboratory setting, the analyst must know how much material is available and how much material is required for an assay to produce readily interpretable and admissible results. Since forensic DNA testing is destructive to the evidence, decisions about how to utilize and analyze evidence are not to be taken lightly. We have demonstrated that MDA is the most reliable and consistent method for amplifying DNA, based on the fact that a quantifiable signal was produced in the downstream applications of real-time and traditional PCR. Although the DNA is amplified significantly, the signals produced in the real-time PCR quantitation were not always consistent, even in the duplicate samples. Additionally, the traditional PCR multiplex kit (Identifiler™) produced correct allele calls, but also displayed an abundance of background

noise. Background signals go hand in hand with any sensitive test, but the goal is the keep the false signals at a minimum to avoid misinterpretation of the data. The level of background initially observed in the MDA was at a level too great for this assay to be applied to forensic casework. In the field of forensic DNA, all procedures and protocols must be performed multiple times and evaluated to set certain guidelines and thresholds as each instrument from laboratory to laboratory varies in sensitivity. One of the validation criterion practiced is to set a minimum threshold value. This value is determined empirically by examining the relative fluorescent unit (RFU) values of the baseline of the output data of many samples (30-40 samples), taking an average of the highest signal seen that cannot be attributed to the authentic genotype expected of the sample and multiplying that number by three to achieve the desired confidence level. The threshold serves as a demarcation that anything above the RFU threshold can be confidently reported as an authentic allele, while anything below the threshold is examined by the analyst and considered to be authentic or not based on the size and morphology of the signal.

Since forensic DNA analysis is often scrutinized and examined extensively, it is necessary to maximize protocols for efficiency, and perhaps most importantly, reproducibility. We observed some inconsistency in whole genome amplification; however, this is not too surprising, since we are starting with extremely low levels of cells/DNA. The more important part is to ensure that profiles generated using this amplified material be as reproducible and interpretable as possible to minimize the arguments that can be made against it. To make this protocol suitable for use in a forensic setting it was necessary to decrease the background. The background signals seen are

similar to those observed when using increased PCR cycles, therefore, it was decided that the next series of experiments was to perform time titrations for the WGA protocol. This process would determine if the 90 minute incubation was necessary, or if a shorter isothermal amplification would generate a high level of true signals, while minimizing the production of artificial signals.

The manufacturer's recommendations specify a 90 minute reaction time, however in our hands; this resulted in a high level of background even with the low starting template amount of approximately 0.040 nanograms. In order to decrease the production of artifacts, a time titration was performed using a reliable amount of starting template (6 diploid cells). A dramatic decrease in background noise and artifact was observed, exhibiting an inverse relationship with the isothermal incubation time. However, in addition to the decrease in artificial signal, the authentic DNA signals were also lower. Comparison of the experimental data showed that an incubation time of 45 minutes provided the best overall result. This length of time enabled the reaction to produce a sufficient amount of product, while limiting artifacts in most cases. This incubation time was adopted and applied to most of the remaining experiments in this thesis where WGA was used.

The goals of this set of experiments was to: 1) determine whether or not WGA could indeed amplify the starting template for accurate and reliable profile analysis and 2) to optimize a protocol for WGA that would result in maximization of amplification and minimization of background noise. The ultimate goal of this thesis work was to be able to generate a full DNA profile using material from a single cell. The next chapter focuses on

the detection limits of the human identification kits after samples are subjected to optimized WGA.

REFERENCES

Applied Biosystems. AmpFℓSTR® Identifiler™ PCR Amplification kit user's manual. Foster City, CA: Applied Biosystems, 2001.

Applied Biosystems. AmpFℓSTR® YFiler™ PCR Amplification kit user's manual. Foster City, CA: Applied Biosystems, 2006.

Applied Biosystems. Quantifiler® Human DNA Quantificaiton Kit and Quantifiler® y Human Male DNA Quantification Kit user's manual. Foster City, CA: Applied Biosystems, 2006.

Ballantyne KN, van Oorschot RAH, Iman Muharam, Angela van Daal and R. John Mitchell. Decreasing amplification bias associated with multiple displacement amplification and short tandem repeat genotyping. Anal Biochem. 2007 March (368):222-229.

Butler JM. Fundamentals of Forensic DNA Typing. New York: Elsevier, 2009.

Promega Corporation. PowerPlex® 16 System Technical Manual: Instruction for Use of Products DC6530 and DC6531. Madison, WI: Promega Corporation. 2008 May.

Promega Corporation Technical Bulletin #TBD022-PowerPlex® Matrix Standards 3100/3130. Madison, WI: Promega Corporation. 2009.

Qiagen. QIAamp® DNA Micro Handbook 2nd Edition. 2009 May: 28-30.

Qiagen. REPLI-g® UltraFast Mini Handbook. 2006 October: 13-15.

Sigma. GenomePlex Single Cell Whole Genome Amplification Kit Technical Bulletin.

Zhang L, Cui X, Schmitt K, Hubert R, Navidi W and Arhheim N. Whole genome amplification from a single cell: Implications for genetic analysis. Proc Natl Acad Sci USA. 1992 July (89):5847-5851.

CHAPTER 4.

DETECTION LIMITS OF CELLS AND GENOMIC DNA FOLLOWING WHOLE GENOME AMPLIFICATION

INTRODUCTION

Whole genome amplification (WGA) is a powerful technique used to increase minute quantities of DNA to levels sufficient to allow testing using many methodologies. Our objective was to use WGA to generate a DNA pattern from only one diploid cell. We've previously demonstrated the ability to generate sufficient sample for thousands of reactions from starting material significantly less than the recommended amount. Currently, one nanogram (ng; approximately 150 diploid cells) is optimal for most STR analyses and any amount under 100 picograms (pg; approximately 15 diploid cells) is considered low copy number (LCN). Low copy number samples are to be handled with care and analyzed with caution, as sporadic contamination can be seen (allele drop-in) along with the loss of signal due to the failure to amplify (allele drop-out, preferential amplification; Caddy, 2008). This work is important in the field of forensic DNA as many cold cases and cases that were previously shelved due to insufficient amounts of DNA for analysis are being revisited with the new PCR technology. However, there are still those cases with DNA insufficient to allow analysis using currently available PCR methods. Unfortunately, forensic labs around the country are not equipped with extra funds and scientists to dedicate their expertise and knowledge toward a scientific remedy to address LCN samples. With the achievement of increasing DNA template from 50 and 25 cells by

three orders of magnitude, we tailored our experiments to reach our objective of utilizing the DNA equivalent of one diploid cell to generate a full DNA STR profile using WGA.

The employment of the LCM was useful to collect a precise number of cells for each reaction, however, the inability to remove the cellular material from the adhesive cap in a consistent dependable manner proved extremely difficult. Therefore, a number of experiments were performed to test the degree of adhesiveness on the cap to determine an effective method to remove the cellular substrate from the cap in a way that would not hinder downstream applications.

The lack of a method to consistently remove material from the adhesive cap eventually exhausted the resources available for further investigation of this problem. This realization led to the decision to apply the same techniques used to increase the amount of starting template obtained from cells to purified genomic DNA. The experiments in this chapter were performed to determine the limitations of MDA WGA using diploid cells and genomic DNA.

MATERIALS AND METHODS

Laser Catapult Microscopy

Laser catapult microscopy is a powerful tool to control precision during sample collection. For one set of experiments, duplicate samples of 12, 6, 3 and 1 HUT-78 cell(s) were collected using micro dissection via a photoactivated localization microscopy (P.A.L.M.™) Microbeam Laser Catapult Microscope (LCM; Zeiss). The cells were then lysed in the adhesive tube cap (Zeiss) using Qiagen's REPLI-g® Ultrafast Mini Kit denaturation buffer for Genomic DNA from Blood or Cells. For another set of

experiments, duplicate samples of 8, 6, 4 and 2 HUT-78 cells were collected and lysed as described above.

Removing substrate from the adhesive tube cap was often challenging, yielded no sample in the reaction tube. Many approaches were performed in efforts to overcome this problem, while still abiding by the 1.5µl volume limit for the downstream application of WGA by Qiagen REPLI-g® Ultrafast Mini Kit. Though the same protocol (using 1.5µl) was used to perform the WGA reaction as in the 50 and 25 cells described in Chapter 3, it is hypothesized that, indeed, the full cellular sample was not removed from the adhesive cap, yet enough was used in the reaction to generate successful data. This hypothesis could also explain why the amount of DNA generated via WGA yielded inconsistent amounts when quantitated as seen in Table 3-2.

Recovery of Cells from Adhesive Tube Cap

One duplicate set of 20 HUT-78 cells were collected via LCM. The samples were centrifuged at 10,000 x g for cellular collection at the base of the tube for further DNA analysis.

Four duplicate sets of 12 HUT-78 cells were collected via LCM. Sufficient water was added to the adhesive tube cap to cover the surface (20µl) to two duplicate sample sets. Two sample sets were carried on without water added to the cap. Four samples (Two wet and two dry) were then frozen at -80°C or -20°C for 2 hours. The tubes were then centrifuged at 10,000 x g for cell collection at the base of the tube.

Another attempt for DNA collection was from 20 HUT-78 cells and was approached using SDS as a substitution (2% final concentration) for the PBS to remove

samples from the adhesive tube cap and Proteinase K was added for a final concentration of 1mg/mL to augment cell lysis.

The failure to disrupt the bonds of the adhesive cap to the HUT-78 cells was further investigated by layering other adhesives onto the adhesive cap itself. Prior to collecting 20 HUT-78 cells with the LCM, the silicon cap was blanketed by a number of contestants: 0.1% Poly-L-Lysine, Elmer's glue, and Elmer's glue sticks. Twenty microliters of DNA grade water was then added to the adhesive cap and incubated upside down at 56°C for 15 minutes before centrifugation at 10,000 x g. Mechanical dislodging was also tried by aspirating DNA grade water around the surface area of the adhesive cap.

Many different combinations of detergents with and without DTT were used with 100pg of Applied Biosystems Quantifiler™ Standard male DNA in attempt to ascertain which reagents would not denature the φ29 DNA Polymerase enzyme. The following are final concentrations of reagents used: [DTT]:40mmol, [SDS]: 2% and 1%, [TritonX100]:1%, [NP40]:1%, [Deoxycholate]: 0.5%, [Tween20]:1%, [Dodecyl-β-maltoside]:4mg/mL, [Proteinase K]:1mg/mL.

Each sample was examined and photographed by the LCM before and after the chemical and mechanical attempts for cellular removal took place. This confirmed that the observed inconsistencies in template material were due to the inability to successfully remove the cells from the adhesive cap.

Genomic DNA Sample Preparation

Since the enormous efforts of dislodging the cellular substrate from the adhesive tube caps failed, we shifted our focus to the ultimate outcome of genomic DNA. Duplicate

samples of Quantifiler™ Standard male DNA were prepared accordingly: 1ng, 0.5ng, 0.25ng, 0.125ng, 0.0625ng, 0.0312ng, 0.0156ng, 0.0078ng and 0.0035ng (Applied Biosystems). All quantities of input template DNA were converted to cellular equivalents based on 6.6 picograms (pg) per diploid cell and 3.3 pg per haploid cell (Butler, 2009).

Multiple Displacement Amplification

Whole genome amplification by MDA was performed (as described in the Materials and Methods of Chapter 2) when working with assays containing sample duplicates of 12, 6, 3 and 1 HUT-78 cell(s). These samples utilized an incubation time of 90 min at 30°C. All other experiments containing HUT-78 cells and genomic DNA were performed with the adaptation of the truncated incubation time of 45 min at 30°C.

DNA Quantification, STR Analysis, Capillary Electrophoresis, and Genotyping

All methods were performed as described in the Materials and Methods of Chapter 2. Only Identifiler was used for STR analysis and genotyping.

RESULTS

Previous amplification of 50 and 25 HUT-78 cells by MDA demonstrated the high degree of sensitivity for downstream assays such as Identifiler™ and Yfiler™. To develop a better understanding of the limitations of MDA, we collected 3 sets of duplicate samples from 12 diploid cells down to a single diploid cell and subject the samples to the full 90 minutes of WGA. Detection of HUT-78 samples, confirmed by real time PCR quantitation and DNA genotyping, was achieved for samples containing as little as 6 cells, while the samples containing 1 and 3 cells exhibited high levels of background noise and artifact along with minimal correct allele calls when subjected to STR analysis (Figure 4-1).

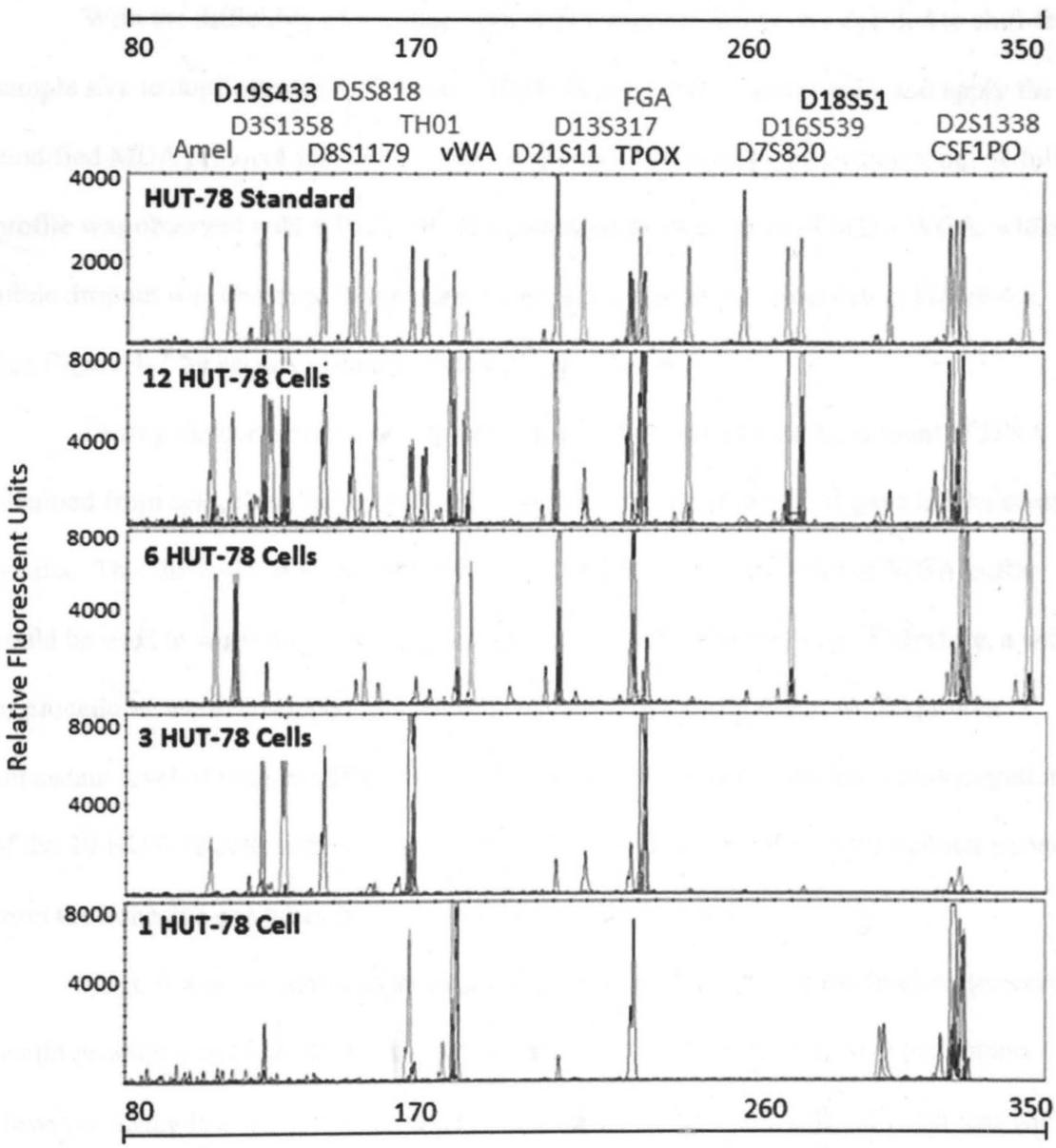

Figure 4-1: Genetic profiles produced from decreasing cell numbers with 90 min incubation. DNA extraction and a 90 minute WGA was performed using MDA. DNA profiling was performed using AB's AmpFℓSTR® Identifiler™ multiplex human identification kits at half volume reactions. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values.

With the difficultly of working with such a minute sample, we decided to shift the sample size to duplicates of 8, 6, 4 and 2 HUT-78 cells and 12 sperm cells and apply the modified MDA protocol for WGA using the truncated 45 minute incubation time. A full profile was observed with 6 HUT-78 cells subjected to 45 minutes of MDA WGA, while allele dropout was observed in the other samples, similar to that observed in Figure 4-1. See Figure 4-2 for results from the truncated reaction time.

During the course of these experiments it was observed that the amount of DNA obtained from cells collected on adhesive cap tubes by use of the LCM gave inconsistent results. This difficulty was thought to be due to the fact that only 1.5μl of WGA buffer could be used to wash the cells before collection from the adhesive cap. Therefore, a series of procedures were developed to attempt to standardize a method that would provide a consistent level of template DNA from a known number of cells. Initially, centrifugation of the 20 HUT-78 cells without the presence of buffer failed to remove the cellular material from the adhesive cap tubes in a dependable manner (Figure 4-3).

Next, it was thought with the expansion of 20μl water during the freezing process would provide a mechanical method to separate the cells from the adhesive membrane. However, using low temperatures to freeze the substrate with and without water was shown to be unsuccessful at removing the cellular material from the adhesive tube cap (Figure 4-4).

The addition of 6% SDS to the buffer used to wash the cells from the adhesive cap into the WGA reaction protocol without a dialysis clean-up step demonstrated another unsuccessful approach. Though, the photographs in Figure 4-5 suggest some adhesion

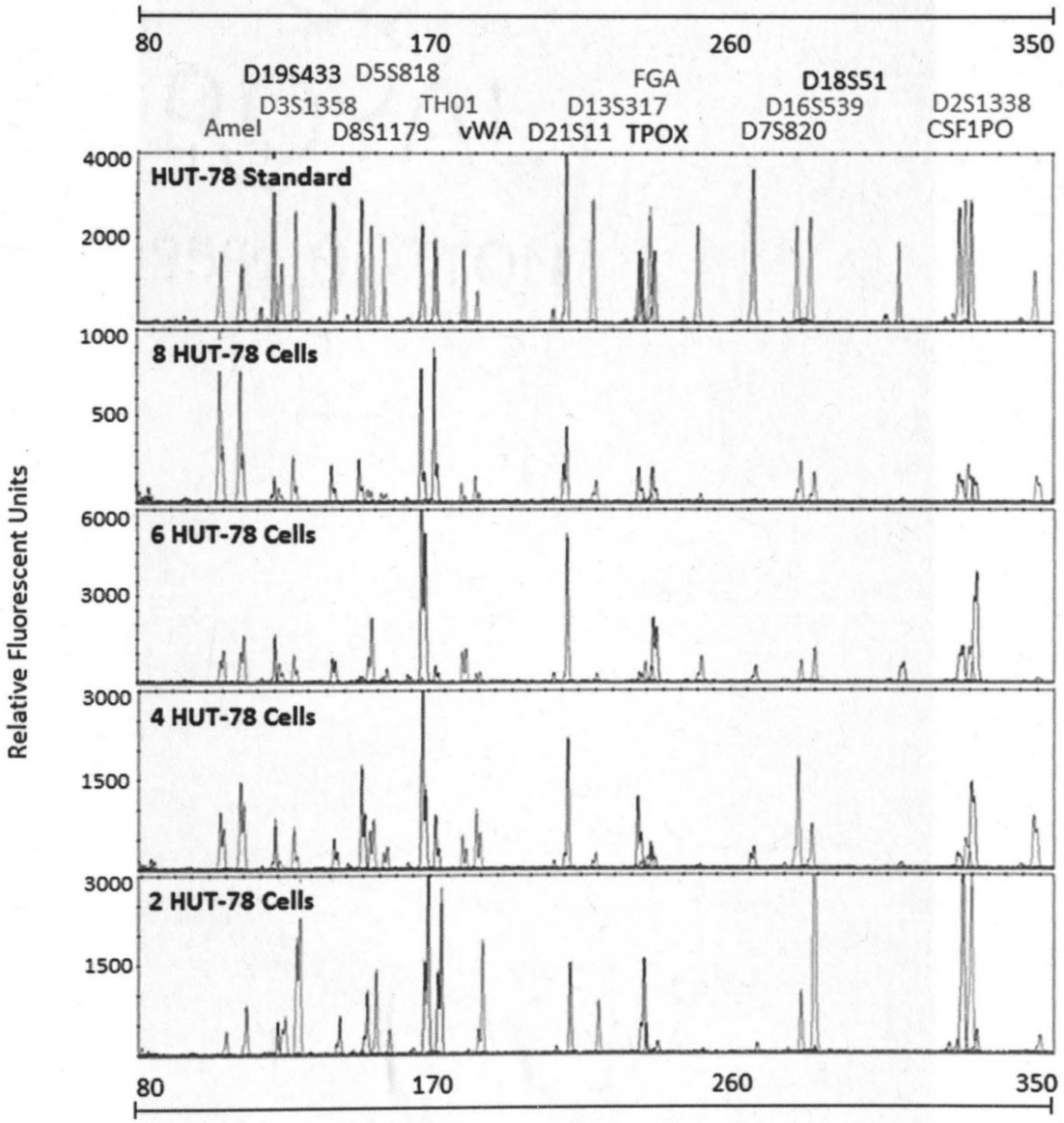

Figure 4-2: Genetic profiles produced from decreasing cell numbers with 45 min incubation. DNA extraction and a 45 minute WGA was performed using MDA. DNA profiling was performed using AB's AmpFℓSTR Identifiler™ multiplex human identification system. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values.

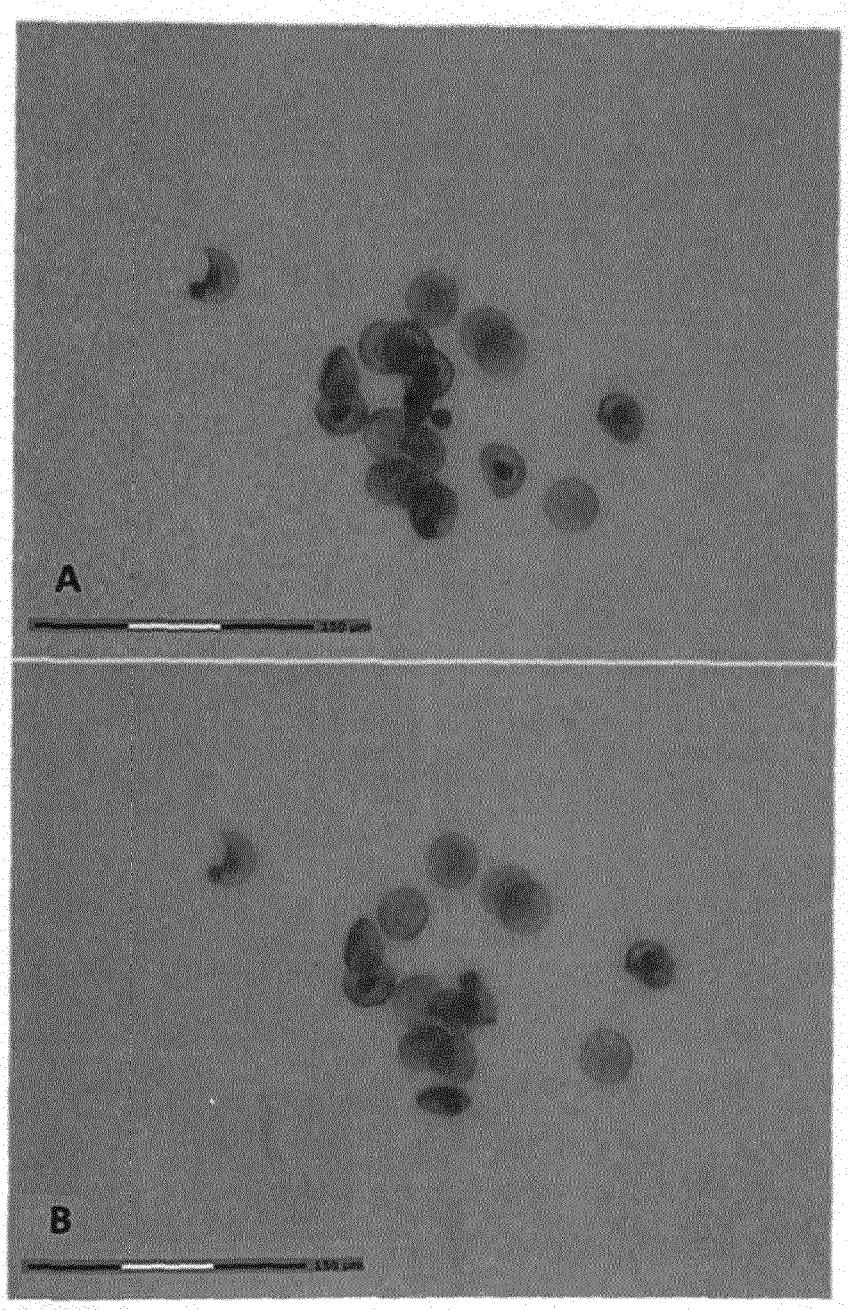

Figure 4-3: Dry HUT-78 cells subjected to maximum revolution RPM centrifugation in attempt to collect cellular material inside the tube. Panel A shows the HUT-78 cells after collection by LCM, while panel B is a photograph taken after the centrifugation process. Photographs were taken using 40X magnification.

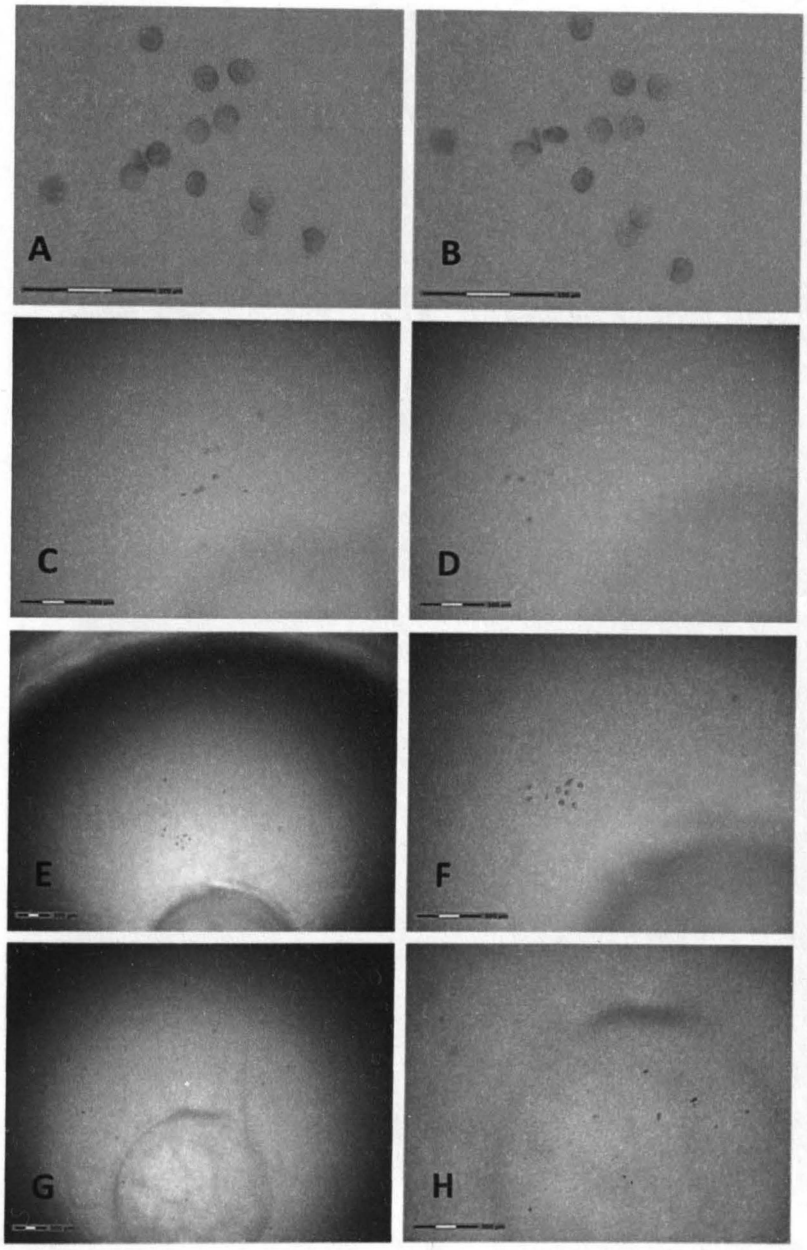

Figure 4-4: 12 HUT-78 cells collected via LCM and frozen to aid removal from the cap. Two duplicate sample sets were collected and frozen at -80°C (A-D), while the remaining two duplicate sample sets were frozen at -20°C after collection (E-H). Samples A, B, E and F were frozen dry, while samples C, D, G and H were frozen with 20µl DNA Grade water (Sigma-Aldrich). The panels on the left (A, C, E, G) are photographs before freezing and centrifugation was performed while the panels on the right (B, D, F and H) are photographs taken after 10 minutes centrifugation at maximum revolutions. Panels A and B were taken using 40X magnification, while C-H were taken at 20X magnification due to the lack of focusing ability.

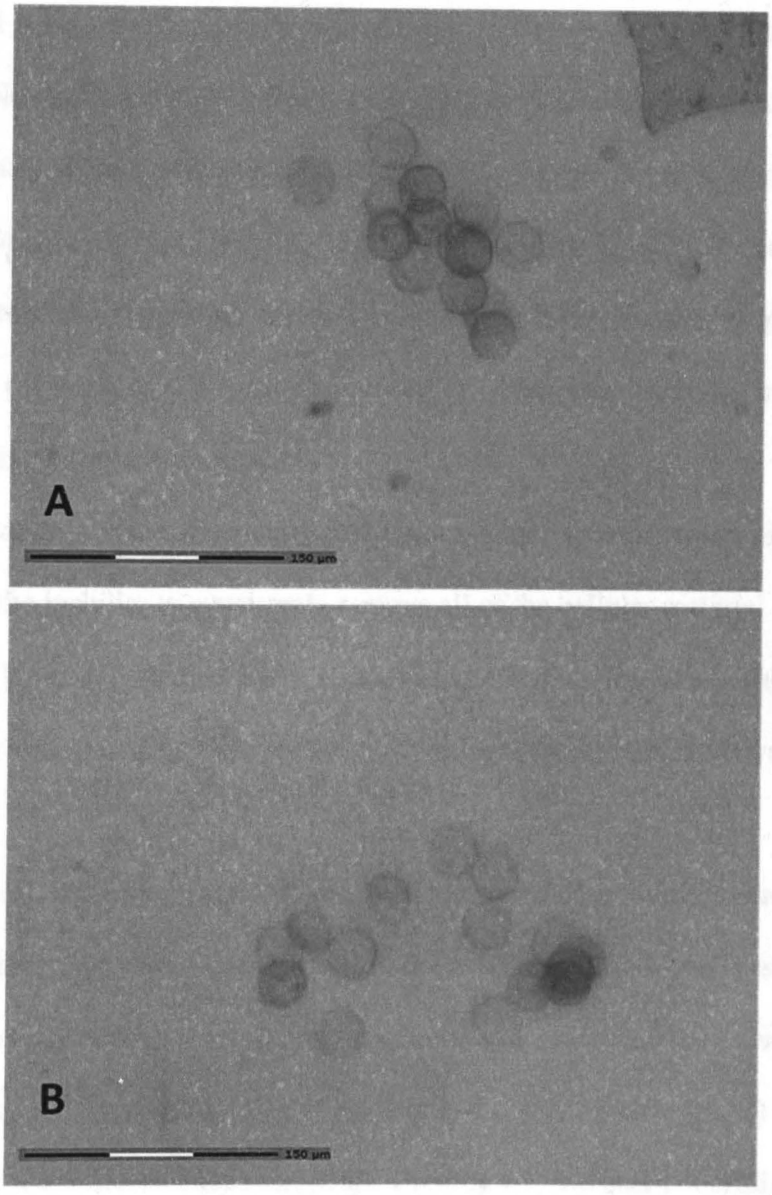

Figure 4-5: Treatment of HUT-78 cells with SDS to aid removal from the cap. Shown in each panel are photographs of the cap prior to (A) and post (B) 6% SDS treatment and centrifugation at 10,000 x g. These photographs were taken at 40X magnification.

disruption possibly due to the addition of SDS to the reaction, or mechanical pipetting of the analyst, full collection of the sample was not achieved. The photograph shows the positioning of the cells had changed; however, the cells were still present on the cap.

Additionally, panel A shows only 14 cells of the 20 cells collected due to the limits of the field of vision, while panel B exhibits 15 cells in its field of vision. This is due to position the cellular substrate took on after being subjected to SDS and Proteinase K treatment. DNA produced from this reaction following WGA using the Real Time PCR quantitation method was not detectable, suggesting the SDS inhibited the WGA reaction by possible denaturation of the polymerase as the positive control of 100pg was detectable using real time PCR, confirming a successful assay.

While a number of WGA procedures provided successful results when using cells obtained by LCM, the inability to consistenly remove all of the cellular material from adhesive cap tubes led to experiments using chemicals in efforts to mask the adhesive nature of the silcon cap. Results of WGA after chemical treatments can be viewed in Table 4-1.

Alternative cell collections methods involved the use of other adhesives applied directly to the adhesive cap tubes. The idea was to decrease the HUT-78 cell's affinity for the silicon adhesive cap containing an alternative adhesive. However, these attempts were unsuccessful. Figure 4-6 shows duplicates of 0.1% Poly-L-Lysine before (A, C) and after (B, D) subjected to incubation and centrifugation. Figures 4-7, 4-8 and 4-9 exhibit the cells layered with Elmer's glue (4-7), Elmer's glue stick, (4-8) and mechanical dislodging (4-9) before and after incubation and centrifugation. Though it appears that some methods worked to a greater extent than that seen when using adhesive cap tubes alone to retrieve the cells in the bottom of the tube, the downstream application of WGA was unsuccessful

Table 4-1: Quantification of DNA following whole genome amplification from cells displaced by various detergents using 100 pg of pre-extracted-pre-quantified DNA as a substrate.

Sample	Chemical Concentrations	After WGA (ng)	STD
DTT	40mmol (all)	2.5	3.22
SDS	2%	0	0
DTT + SDS	1%	0	0
DTT + TritonX100	1%	0.04	0.05
DTT + NP40	1%	0.02	0.02
DTT + Deoxycholate	0.5%	0.035	0.05
DTT + Tween20	1%	3.14	4.39
DTT + dodecyl-β-maltoside	4mg/mL	0	0
DTT + Proteinase K	1mg/mL	0	0

as determined by quantitation. Furthermore, it should be noted all of the methods used a 20µl volume to ensure complete coverage of the tube cap.

Additional experimental protocols involved the use of many different combinations of detergents with and without DTT. All of these attempts were used with 100pg of Applied Biosystems Quantifiler™ Standard male DNA in attempt to identify the reagents that would not denature the φ29 DNA Polymerase enzyme. The following are final concentrations of reagents used: [DTT]:40mmol, [SDS]: 2% and 1%, [TritonX100]:1%, [NP40]:1%, [Deoxycholate]: 0.5%, [Tween20]:1%, [Dodecyl-β-maltoside]:4mg/mL, [Proteinase K]:1mg/mL. (see Table 4-1).

Detection Limits of Genomic DNA using MDA WGA

To remove the complication of complete cell dislodgement from the tubes and since

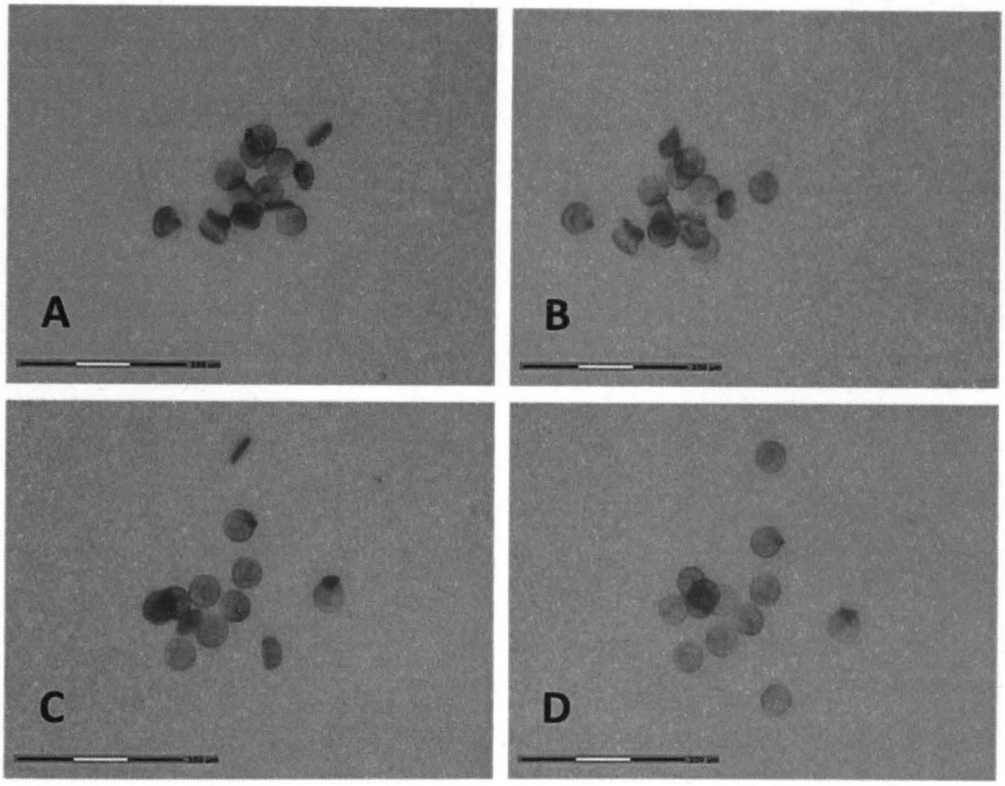

Figure 4-6: Poly-L-Lysine was added to the silicon adhesive cap to decrease the HUT-78's affinity for the silicon prior to cellular collection. 20µl of DNA grade water was added prior to incubation and centrifugation at maximum revolutions. Panels A and C are duplicate samples of 20 HUT-78 cells before manipulation of the collected sample. Panels B and D are photographs after incubation and centrifugation. All photographs in this figure were taken using 40X magnification.

most forensic laboratories do not have access to a laser capture microscope anyway, the same experimental protocol was applied to pre-extracted, purified DNA to imitate the origin of sample routinely used in a forensic facility. Quantifiler™ Standard male DNA was titrated from 1ng down to near the theoretical equivalent of a single cell (0.0066ng) via a 1:2 serial dilution series as can be seen in Table 4-2. The presence of male DNA subjected to WGA was demonstrated in Figure 4-10 down to the single cell range of 0.0078ng while a full profile was achieved with 0.0312ng of DNA (the equivalent to less

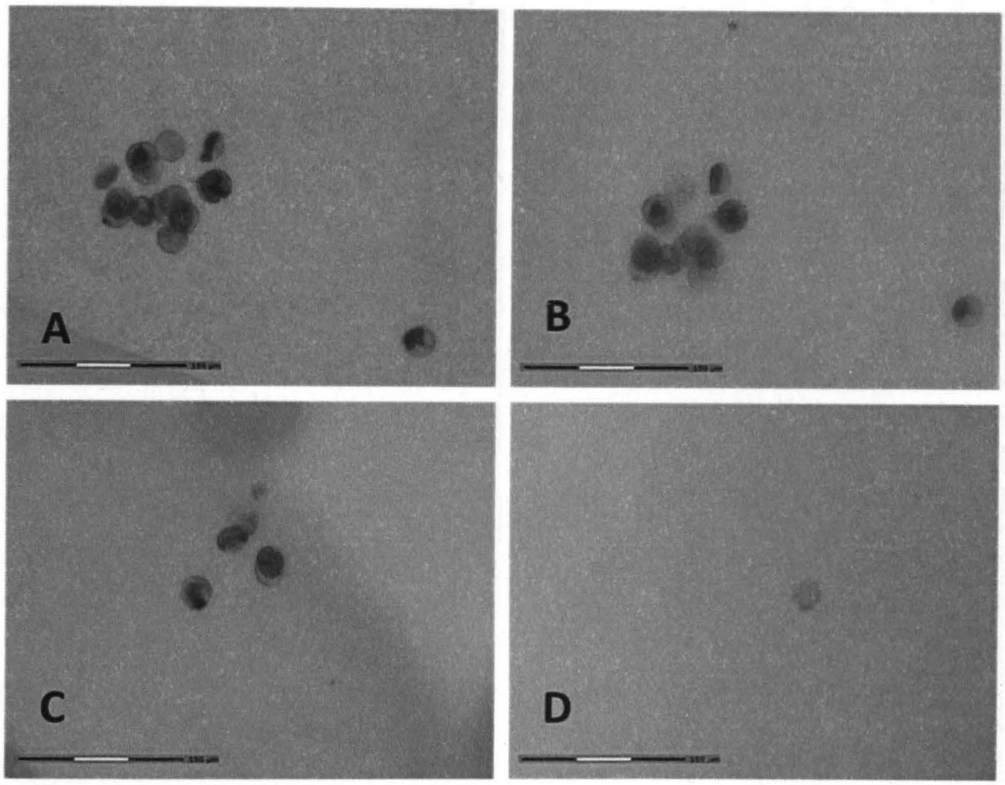

Figure 4-7: Examination of Elmer's Glue used to decrease the HUT-78's affinity for the silicon prior to cellular collection. 20µl of DNA grade water was added prior to incubation and centrifugation at maximum revolutions. Panels A and C are duplicate samples of 20 HUT-78 cells before manipulation of the collected sample. Panels B and D are photographs after incubation and centrifugation. These photographs were taken using 40X magnification.

than 5 diploid cells) using the Applied Biosystems AmpFSTR® Identifiler™ kit. Allele drop-out was observed in samples with decreasing amounts of DNA, however, 75% of the expected alleles were detected in the 0.0156ng sample (the theoretical equivalent of two cells). Two alleles were detected in the 0.0078ng sample and one different correct allele in the duplicate of the 0.0078ng sample. Incorrect allele calls were not observed after alteration of the WGA reaction time, even at very low DNA concentrations. It should be noted that even though quantitation methods reported undetectable, it has been observed

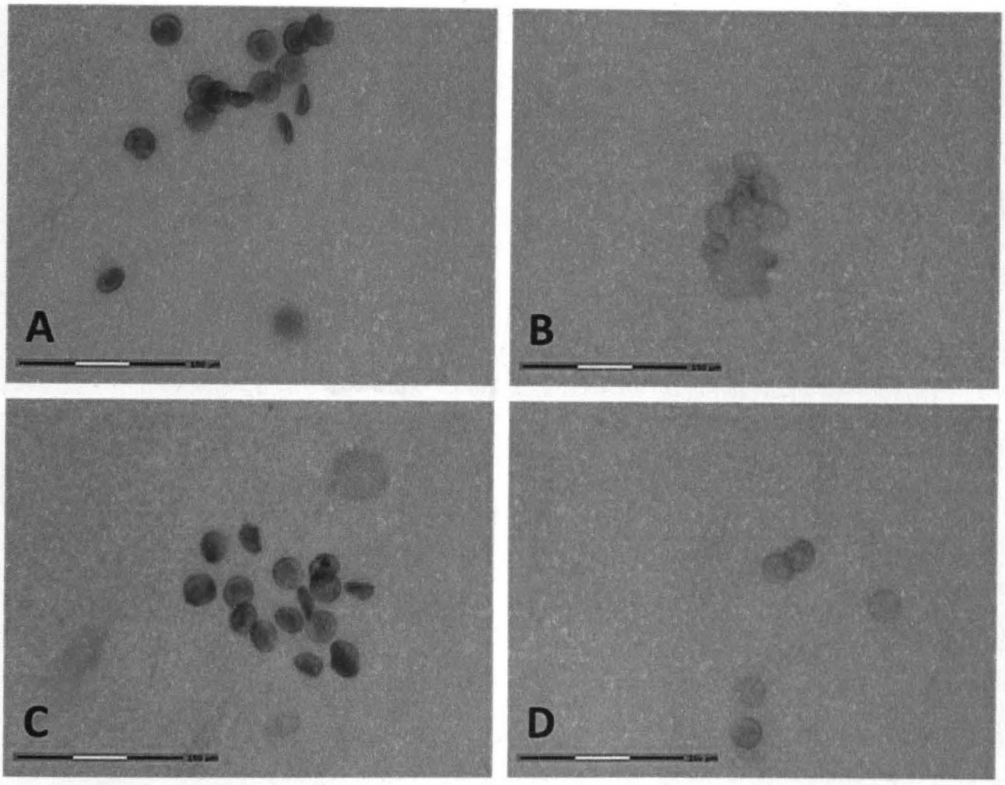

Figure 4-8: Elmer's GlueStick was layered onto the silicon adhesive cap to decrease the HUT-78's affinity for the silicon prior to cellular collection. 20µl of DNA grade water was added prior to incubation and centrifugation at maximum revolutions. Panels A and C are duplicate samples of 20 HUT-78 cells before manipulation of the collected sample. Panels B and D are photographs after incubation and centrifugation. Photographs were taken using 40X magnification.

the AB STR multiplex kits are more sensitive than the quantitation methods used. (see Figure 4-8).

In many forensic DNA cases, the availability of sample is limited due to its use in previous assays (*i.e.*, cold cases) or the sample collected was minimal to begin with (*i.e.*, touch DNA, guns). Our preliminary experiments using templates of 50 and 25 HUT-78 cells resulted in the generation of high levels of target product. Therefore, we decided to extend our assays into the lower range of template amount to determine the limits of

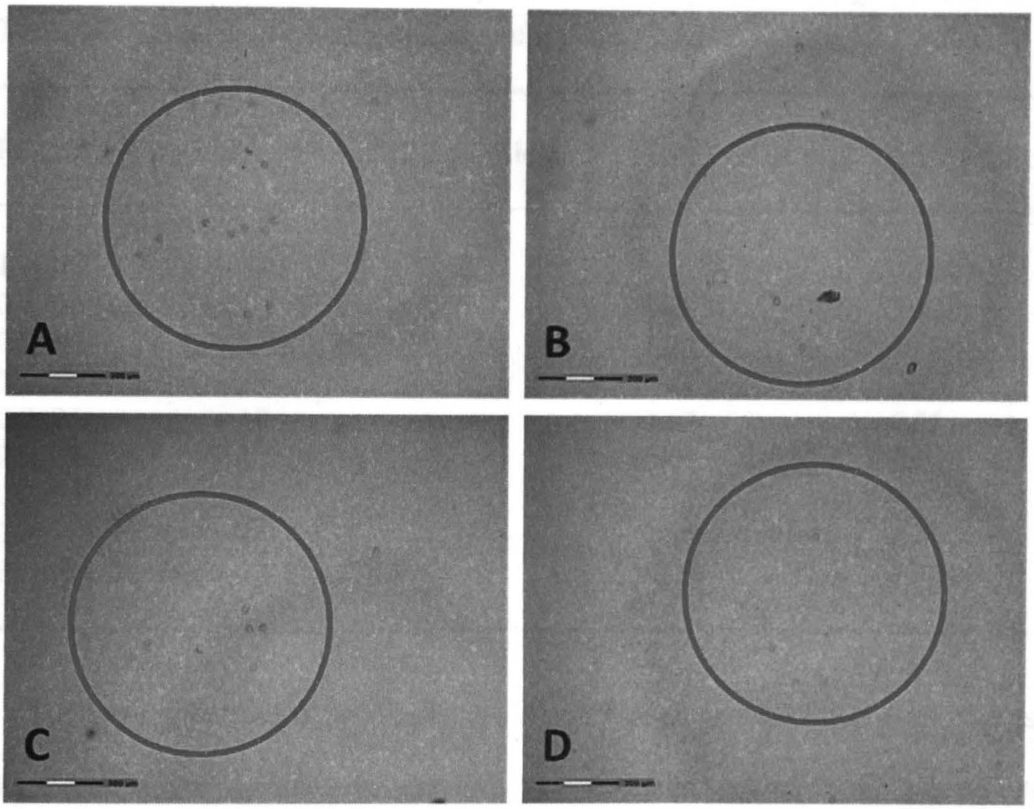

Figure 4-9: No additional adhesive was layered onto the silicon adhesive cap to decrease the HUT-78's affinity for the silicon prior to cellular collection. 20µl of DNA grade water was added and vigorously aspirated prior to incubation and centrifugation at maximum revolutions. Panels A and C are duplicate samples of 20 HUT-78 cells before manipulation of the collected sample. Panels B and D are photographs after incubation and centrifugation. Photographs were taken using 20X magnification due to lack of focusing ability at 40X. The red circles indicate where the cellular material can be visualized.

detection of the MDA protocol. The limitations of the kit were approached in those samples containing less than 6 diploid cells prior to the WGA reaction. However, samples that contained 6 diploid cells or higher remained consistent and exhibited reliable results when carried through STR analysis. Though samples with less starting material than 6 cells did produce signals and partial profiles, our analyses showed that WGA samples containing 6 cells were established to be the limit of reliable detection for this system.

Table 4-2: Quantification of serial-diluted starting template followed by MDA.

Input DNA (ng)	Theoretical Cell Equivalent	After WGA (ng)	STD (ng)
1.0000	~ 150	419.83	154.64
0.5000	~ 75	209.37	59.85
0.2500	~ 38	53.46	11.95
0.1250	~ 19	4.41	2.33
0.0625	~ 10	1.67	0.70
0.0312	~ 5	0.64	0.20
0.0156	~ 2	0.07	0.12
0.0078	~ 1	Undetectable	0.00
Positive Control	~ 15	41.36	47.54

The nomination of 6 diploid cells may actually be an upper limit of detection, as it was demonstrated that signals were achieved at the three and one cell level. We suggest that a certain amount of cellular material remains on the adhesive cap of the collection tube, rendering any sample with less material than thought to be present. Many approaches were evaluated to shift the affinity for the silicon cap to a water soluble agent that would carry the biological matter from the tube cap to the bottom of the reaction tube for further processing.

DISCUSSION

The dry centrifugation approach was ineffective. We predict that the adhesion of the cells to the surface is much greater than that which we can generate to remove the cells without damaging the tube. The freezing of water on the tube cap could potentially be used to dislodge the cells was also shown to be unable to remove the cells from the tube cap. A potential problem could be that the small amount of water did not affect the cells, and since

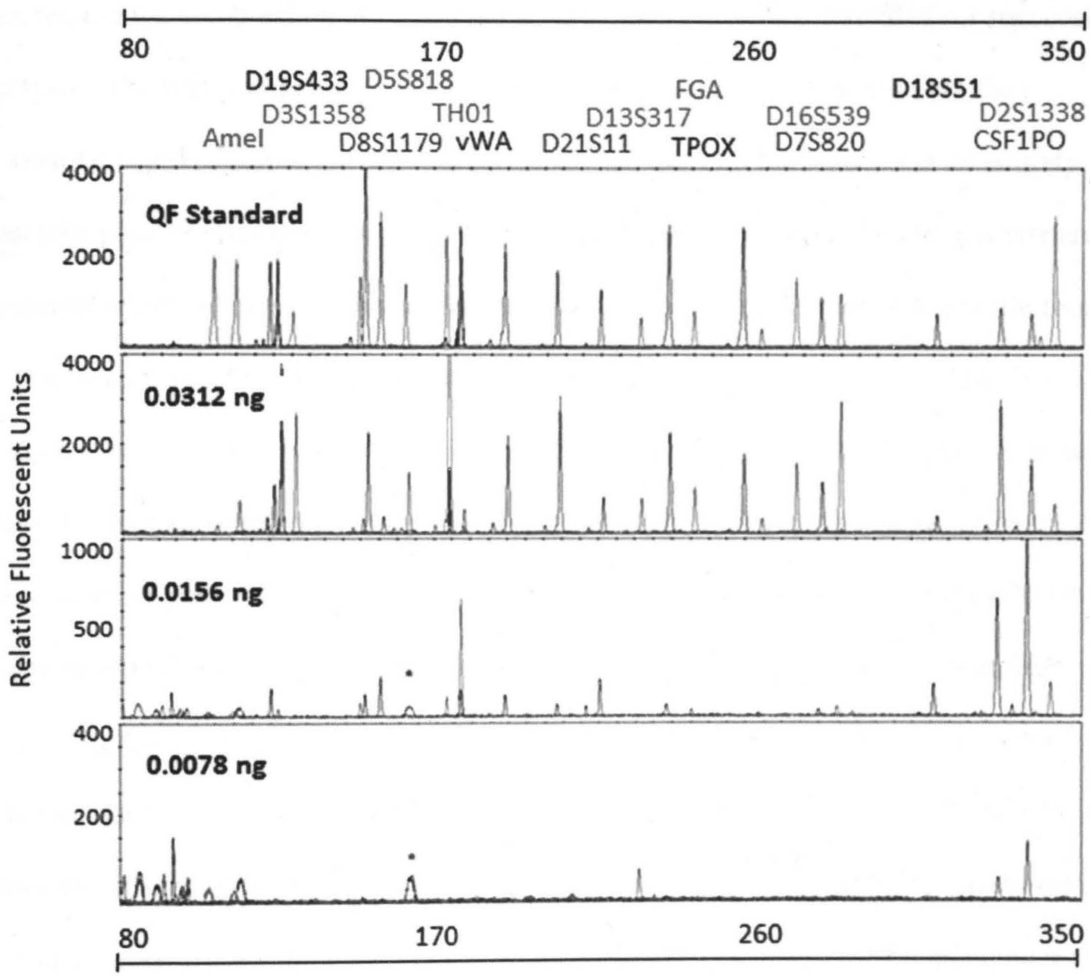

Figure 4-10: Examination of profiles produced following MDA of serial-diluted starting template DNA. MDA WGA and Applied Biosystem's AmpFSTR® Identifiler™ kit. The 0.0078ng sample is an overlay of duplicate samples to show all the information collected. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values. * indicates a known reported artifact, most likely due to the fluorescent dye used.

the frozen water quickly thawed when removed from the freezer, the result was similar to the previous attempt at using DNA grade water for removal.

It was demonstrated that the D2 buffer provided in MDA kit was not sufficient in lysing the acrozomal cap of the sperm cell. Traditional extraction protocols for sperm cell

lysis require the combination of Proteinase K, DTT and a detergent like SDS for successful cell lysis. The first problem with these component requirements when utilizing the downstream application of the MDA procedure is the fact that the reaction in its entirety must take place in the same tube to prevent unacceptable sample loss. While reagents are constantly added to the reaction vessel containing minimal sample, there leaves little room for removing reagents in a clean-up step. The Proteinase K and SDS, even at low final concentrations are still a source of inhibition of DNA synthesis by φ29 DNA polymerase and in downstream PCR applications. The data in Table 4-1 indicates that the presence of one or more of these reagents denatured or somehow inhibited the activity of the φ29 DNA polymerase before enough WGA had taken place to produce high amounts of template.

The final attempt at removing the cellular material from the adhesive cap of the collection tube was to add another adhesive material to the domed silicone collection lid to reduce the affinity of the HUT-78 cells to the silicone. The two Elmer's glue variations were unsuccessful and both inhibited the WGA reaction. The mechanical dislodging by aspiration appeared to disrupt the adhesion of the cells to the cap; however, the agitation was too little for whole sample collection. With more failure than success at attempting sample collection from the LCM, a new approach was taken to estimate the detection levels using pre-extracted purified DNA in place of whole cells.

As stated above, the use of a restricted 45 minute incubation time prevented false signals while also decreasing relative fluorescent unit (RFU) values of valid alleles. Since the reduced WGA reaction time significantly lowered the detection limit when using intact cells it was anticipated that a similar decrease would be observed with purified DNA.

Therefore, instead of cells, pre-extracted, pre-quantitated DNA from Applied Biosystems was tested in a series of serial dilutions at the new reduced incubation time of 45 minutes. It was demonstrated that 0.0625ng (equivalent to approximately 9-10 diploid cells) of input template DNA consistently produced sufficient amounts of DNA after WGA to generate a full profile. Also, it should be noted that the samples containing approximately 0.0312ng (equivalent to approximately 4-5 diploid cells) of DNA produced a full profile, however, there was considerable peak imbalance observed in those samples. The limit of detection was approached in those samples containing approximately 0.0156ng (approximately 2-3 diploid cells) as the number of expected alleles was no longer detectable. Though full profiles were not produced, it is noteworthy that usable partial profiles were developed down to the single diploid cell range of 0.0078 ng.

It has been shown that macromolecular crowding is a methodology that can enhance the efficiency of many assays. Traditionally, polyethylene glycol (PEG) is used to augment the reaction efficacy. The next chapter explores the concept of macromolecular crowding using a synthetic crowding agent (PEG) and also a natural biomolecule (DNA) to see if a further increase in the amount of template DNA produced with the WGA reaction could be achieved.

REFERENCES

Butler JM. Fundamentals of Forensic DNA Typing. New York: Elsevier, 2009.

Casey B, Taylor GR, and Linacre A. Review of the science of low template DNA Analysis. 2008. http://police.homeoffice.gov.uk/publications/operational-policing/Review_of_Low_Template_DNA_1.pdf January 24th, 2011.

CHAPTER 5.

INCREASING SENSITIVITY BY USING MACROMOLECULAR CROWDING

INTRODUCTION

Macromolecular crowding *in vivo* uses a high concentration of cellular components including protein and RNA, which decrease nonspecific steric repulsion, thus enabling the cell to efficiently carry on mandatory biological functions (Al-Habori, 2001). Molecular crowding *in vivo* can be imitated *in vitro* using agents with a high molecular weight such as dextran, polyethylene glycol, protein and nucleic acids; this premise has been demonstrated to influence the thermodynamics and kinetic effects of reactions upon addition of macromolecules (Ellis, 2001). These approaches, in theory, decrease the stochastic effects that a reaction vessel can exhibit with a low copy number template or scarce sample availability. By using an effective natural or synthetic macromolecule the efficiency of the reaction can be enhanced by driving the reagents into proximity with one another.

Our experiments explored MDA WGA in tandem with macromolecular crowding in an effort to further enhance the sensitivity of the assays. Our findings led us to the optimization of the WGA MDA protocol to reduce artifacts and baseline, along with increasing the DNA template. Macromolecular crowding further increased the ability to detect DNA template in the single cell range, while providing an option for sample retention as well. This chapter will focus on the effects of natural and synthetic crowding

agents during the WGA and STR amplification processes to determine which type of agent, if any, will help further increase the signals generated from low template DNA.

MATERIALS AND METHODS

Sample Preparation

Duplicate samples of Quantifiler™ Standard male DNA were prepared containing 0.0625ng (Applied Biosystems). Samples were amplified as described in Chapter 2 with the truncated 45 minute reaction time. Polyethylene glycol (PEG) was added to these reactions prior to WGA at a final concentration of 0.5%. Molecular weight explored were: 2000, 3350, 600 and 10,000 KDa.

Duplicate sample of Quantifiler™ Standard male DNA were prepared containing 0.0625 ng, 0.0312 ng, 0.0156 ng, 0.0078 ng and 0.0035 ng (Applied Biosystems). Samples were amplified by MDA with one nanogram mouse DNA (C57B1/6) included in each reaction. Samples were amplified as described in Chapter 3 with the truncated 45 minute reaction time.

Duplicate samples of Quantifiler™ Standard male DNA were prepared accordingly: 1 ng, 0.5 ng, 0.25 ng, 0.125 ng, 0.0625 ng, 0.0312 ng, 0.0156 ng, 0.0078 ng and 0.0035 ng (Applied Biosystems). Samples were amplified with 0.5% PEG (3350KDa) added to each reaction. Samples were not subjected to WGA, but instead forwarded straight to STR analysis with Identifiler™ (Applied Biosystems).

Two sets of duplicate samples of Quantifiler™ Standard male DNA were prepared accordingly: 1 ng, 0.5 ng, 0.25 ng, 0.125 ng, 0.0625 ng, 0.0312 ng, 0.0156 ng, 0.0078 ng and 0.0035 ng (Applied Biosystems). Samples were amplified with one nanogram mouse

DNA (C57B1/6) added to each reaction. Samples were not subjected to WGA, but instead forwarded straight to STR analysis with Identifiler™ and Yfiler™ (Applied Biosystems).

One nanogram of 9947a female DNA (Promega) was added to each sample containing the following amounts of Quantifiler™ Standard male DNA: 1 ng, 0,5 ng, 0.25 ng, 0.125 ng, 0.0625 ng, 0.0312 ng, 0.0156 ng, 0.0078 ng and 0.0035 ng (Applied Biosystems). Samples were amplified by MDA as described in Chapter 3 with the truncated 45 minute reaction time.

With the realization that, when used separately, both WGA and molecular crowding via mouse DNA were both successful techniques to increase DNA template and signal, we wanted to determine whether the use of two techniques were additive. Duplicate samples of Quantifiler™ Standard male DNA were prepared as follows and subjected to WGA as described in Chapter 2: 1 ng, 0,5 ng, 0.25 ng, 0.125 ng, 0.0625 ng, 0.0312 ng, 0.0156 ng, 0.0078 ng and 0.0035 ng. After clean-up and quantitation, 1 ng of mouse DNA (C57B1/6) was added to each sample prior to STR analysis. This experiment was performed one time in duplicate.

Multiple Displacement Amplification, DNA Quantification, STR Analysis, Capillary Electrophoresis, and Genotyping

The above were performed as described in the Materials and Methods of Chapter 3.

RESULTS

Polyethylene glycol (PEG) is a reagent commonly used for applications involving molecular crowding. In order to determine the effect of this reagent on whole genome amplification, we tested a series of PEGs with varying molecular weights using 0.0625 ng

input DNA template prior to the WGA reaction. The use of these different polyethylene glycols did not show a consistent increase in sensitivity for these assays (Figure 5-1). The highest molecular weight tested of 10,000 kDa seemed to show the highest signals approaching 4000 RFU values, while the subsequent molecular weight tested of 6,000 kDa exhibited the lowest signal strength of under 2,000 RFUs. The two smallest molecular weights tested contained signals under 3,000 RFU. The 3,350 kDa sample showed higher signals in the larger DNA fragments while the smaller fragments show higher expression in the 2,000 kDa sample. Polyethylene glycol was also tested as a molecular crowding agent in the STR analysis protocol alone (without WGA). It was found that all of the PEG species inhibited the PCR reaction most likely due to the omission of a clean-up step.

In addition to using polyethylene glycol, we experimented using a non-human (*i.e.*, mouse) nucleic acid as a macromolecular crowding agent in the WGA reaction. With the addition of the mouse DNA (mDNA) followed by WGA, it can be seen in Figure 5-2 the lowest DNA concentration that resulted a complete profile was 0.0312 ng starting DNA human template, while allele dropout was observed in the 0.0156 ng sample. While these results are similar to the limit of detection experiment previously conducted, it should be noted number of the detected alleles more than doubled in the 0.0078 ng sample. Therefore, sensitivity was slightly increased by using mouse DNA as a molecular crowding agent during WGA.

The success of increased sensitivity using mDNA with Identifiler™ encouraged us to examine other human identification systems. Our next experiment tested Applied Biosystem's AmpFℓSTR® Yfiler™ using 1 ng mDNA as a molecular crowding agent in

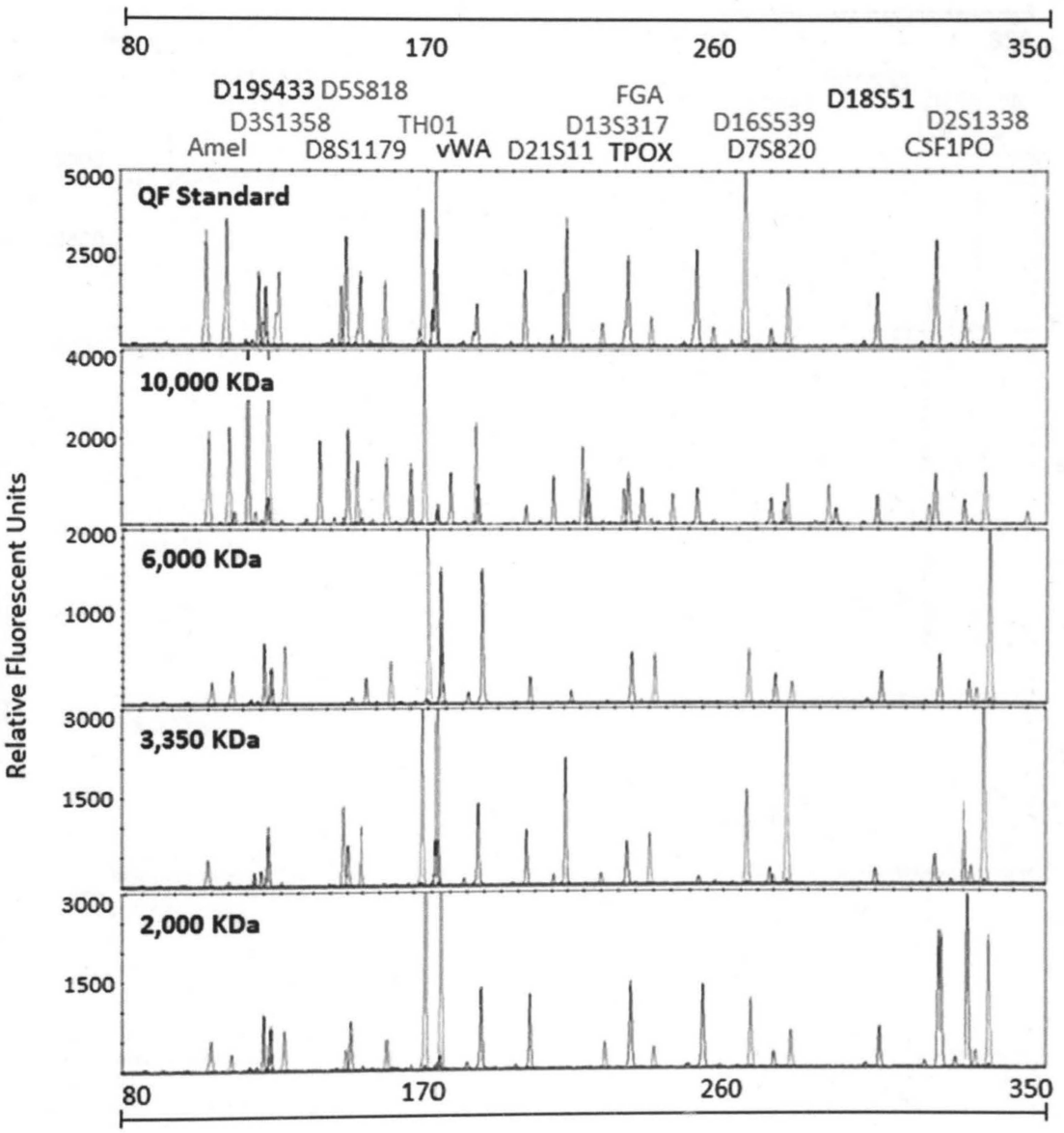

Figure 5-1: Comparison of Identifiler™ profiles generated in the presence of varying molecular weight polyethylene glycol during WGA. A final concentration of 0.5% PEG with varying molecular weights was added to the MDA WGA reaction and profiled with AB AmpFℓSTR® Identifiler™ human identification system. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values.

the STR analysis reaction. The results were similar to that of the Identifiler™, in that there indeed was increased sensitivity. Figure 5-3 shows genotype detection in the haploid DNA

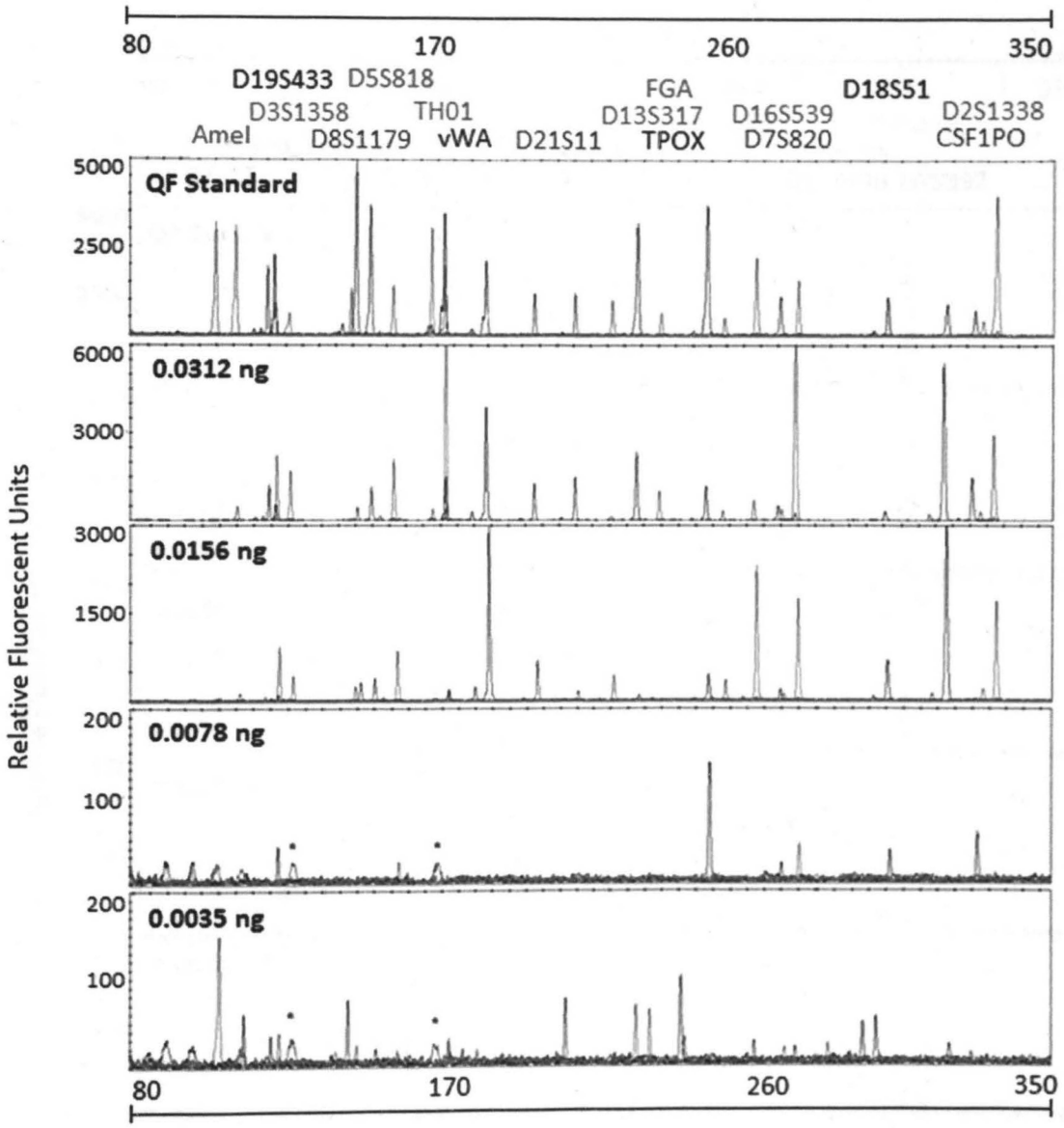

Figure 5-2: Limit of detection using AmpFℓSTR® Identifiler™ for starting templates with mouse DNA that has been amplified by MDA. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values. * indicates a known reported artifact, most likely due to the fluorescent dye used.

range using only mDNA as a molecular crowding agent without WGA to augment the original DNA template. This experiment shows that mDNA does not need to be added to the WGA reaction, as WGA increases the time it takes to analyze the DNA. Rather,

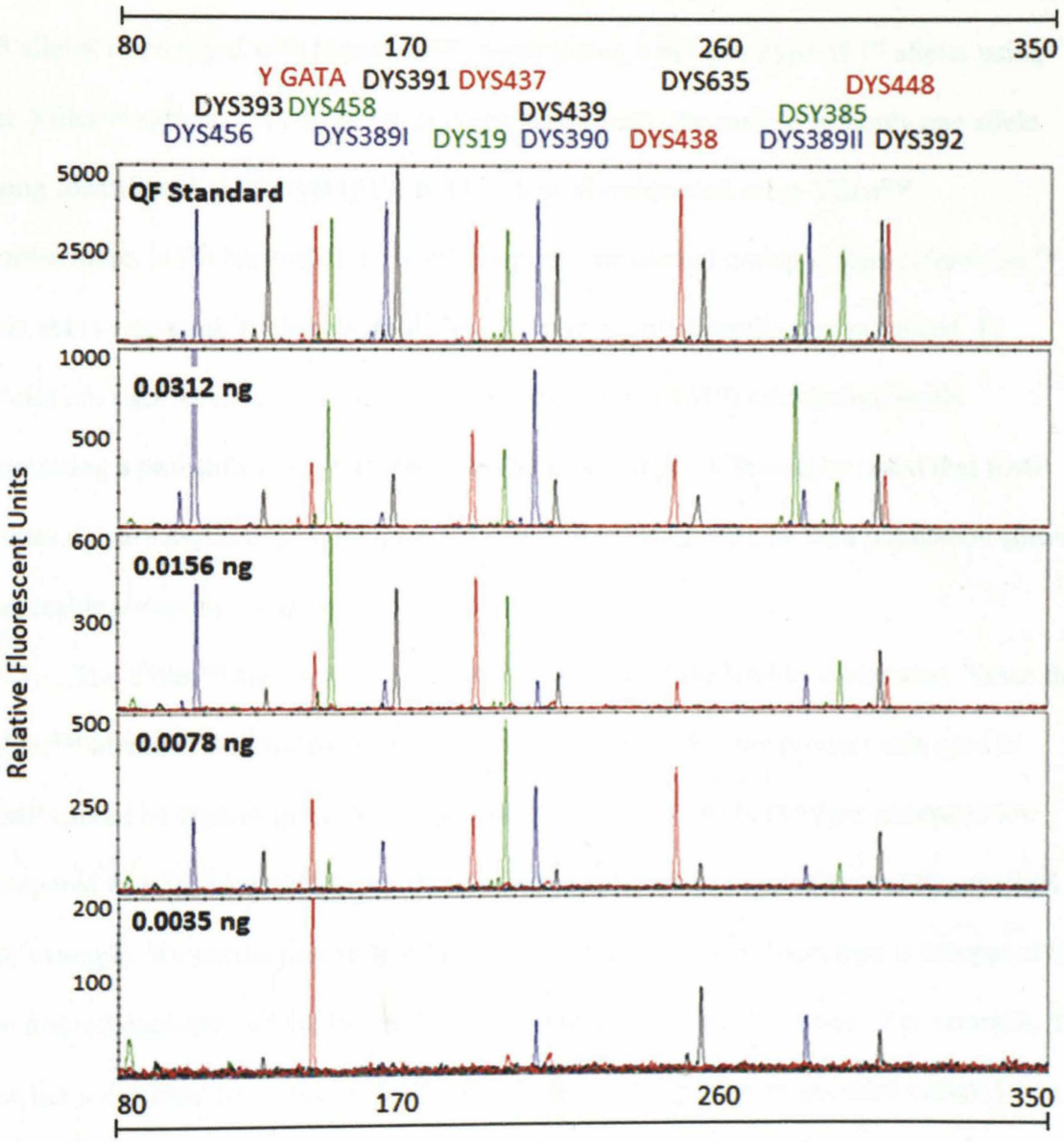

Figure 5-3: Short tandem repeat analysis (AmpFℓSTR® Identifiler™) without WGA, but in the presence of mouse DNA The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values.

mDNA can be added straight into the genotyping reaction to achieve similar results. Table 5-1 illustrates the achieved genotypes in the Identifiler™ and Yfiler™ systems using only molecular crowding.

The male DNA that was used for these experiments has a full genotype exhibiting 28 alleles when typed with Identifiler™, while typing a full genotype of 17 alleles using the Yfiler™ system. At 0.0312 ng, or 5 cell equivalents, the sample lost only one allele using Identifiler™ while typing the full DNA profile expected using Yfiler™. Furthermore, in the haploid range (0.0035 ng), the autosomal typing system (Identifiler™) was able to generate 10 detectable alleles. Though a partial profile was produced, 10 alleles are significant using a random match probability (RMP) calculation, easily generating a probability into the one in ten-thousand range. It should be noted that RMP values do vary depending on the population database being used, as each population shows noticeable variations in allele frequencies over time.

The Yfiler™ typing system was able to generate 6 detectable alleles also. Since the Yfiler™ identification system uses only one sex chromosome, the product rule used in RMP cannot be applied to the Y-chromosome haplotype. The haplotypes generated are compared to established databases which use a calculation known as the counting method. For example, if a partial profile detects 6 alleles, then that partial haplotype is compared to the desired database and the frequency of that combination is determined. For example, if one has a database consisting of 1000 individuals and the pattern in question occurs 10 times in that database, then the probability is 1 in 100. In effect, even this partial pattern allows the forensic analyst to exclude 99% of the general population as contributors of this DNA. While these statistics are not as powerful as those obtained using an autosomal profiling method, Yfiler™ is often the only option in sexual assault cases.

Table 5-1: STR profiling using 1ng mDNA instead of WGA.

Identifiler™ Amount DNA (ng)	Theoretical Cell Equivalent	Observed Alleles out of 28 Expected
0.0312	~5	27
0.0156	~2	N/A
0.0078	~1 Diploid Cell	18.5
0.0035	~1 Haploid Cell	10
0.0018	< ~ 1 Haploid Cell	0
Yfiler™ Amount DNA (ng)	**Theoretical Cell Equivalent**	**Observed Alleles out of 17 Expected**
0.0312	~5	17
0.0156	~2	15
0.0078	~1 Diploid Cell	8
0.0035	~1 Haploid Cell	6
0.0018	< ~ 1 Haploid Cell	2.5

Based on our macromolecular crowding using mouse DNA, we tested the efficacy of human DNA as the molecular crowding agent. Although mixtures of DNA are often undesirable in forensic DNA analysis, it is possible to distinguish male DNA from female DNA using identification of alleles specific for the Y chromosome. Our previous mouse study indicated it is possible that a DNA mixture could actually be beneficial when examining low amounts of male DNA in a sample. By adding 1 ng female human DNA to the same titration of human male DNA (1 ng-0.0035 ng) without WGA, we were able to produce a partial profile with as little as 0.0035 ng of male target DNA, the approximate equivalence to a single haploid cell (*i.e.*, one spermatozoa; Figure 5-4 and Table 5-2).

Table 5-2: Qiagen WGA Female: Male Mixture DNA Titration.

Input DNA (ng)	Input DNA (ng)	Total DNA Post WGA (ng)	STD	Male DNA Post WGA (ng)	STD
Female	Male	Average		Average	
1.0	1.0000	232.95	100.18	283.5	104.73
1.0	0.5000	187.26	51.08	208.2	75.52
1.0	0.2500	114.03	55.24	100.8	35.42
1.0	0.1250	48.51	24.43	71.48	82.54
1.0	0.0625	33.68	26.06	11.95	11.31
1.0	0.0312	19.79	13.4	10.94	13.37
1.0	0.0156	17.96	13.69	0.89	1.1
1.0	0.0078	29.14	26.77	1.28	1.83
1.0	0.0035	8.33	6.33	0.83	1.66

The single trial designed to evaluate if using WGA and macromolecular crowding were synergistic concluded that the two methods were not additive with one another. This was determined based on the total number detected alleles, as well as, the relative strength of each peak based on RFU values. (Data not shown).

DISCUSSION

Macromolecular crowding has improved the efficiency of many techniques ranging from molecular cloning techniques and intracellular signaling to volume regulation in protein phosphorylation experiments. An inconsistency was observed in PEG molecular weight, leading to the conclusion that increasing molecular weight has little influence on reaction efficiency at least with respect to PEG. It would be interesting to see if lower molecular weights at higher concentrations exhibited similar peak heights and signals to

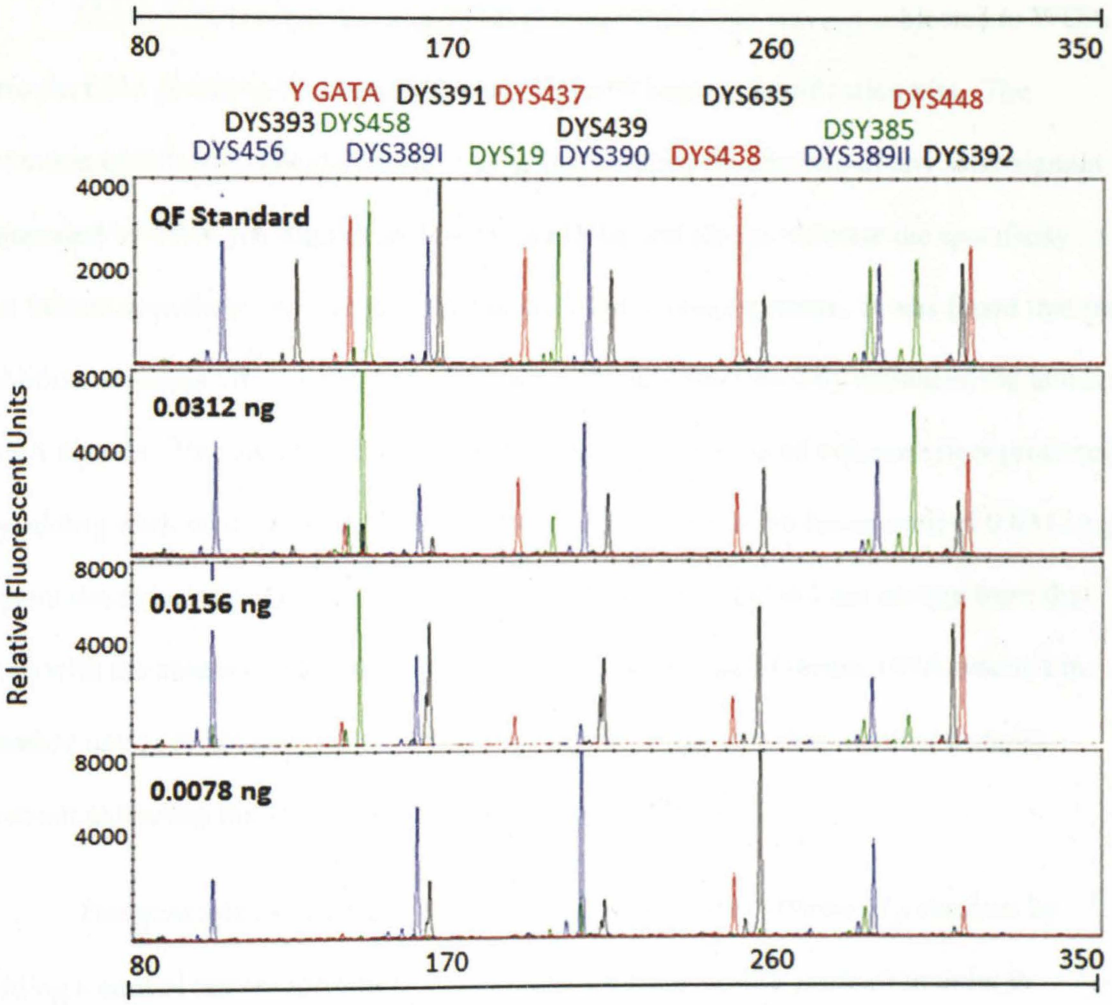

Figure 5-4: Samples of Quantifiler™ Standard male DNA using 9947a female DNA as a molecular crowder after STR analysis by Applied Biosystem's AmpFℓSTR® Yfiler™ was used here. The X-axis is a measurement of the base pairs in ascending order from 80-350bp. The Y-axis is a measurement of the relative fluorescent units (RFU) values. * indicates a known reported artifact, most likely due to the fluorescent dye used.

that of the higher molecular weight macromolecules used at a lower concentration to see if the molecular weight of a PEG molecule to further confirm or refute if the molecular weight of the PEG makes a difference in assays that utilize molecular crowding.

Mouse DNA was added to purified genomic DNA that was not subjected to WGA prior to DNA profiling by Identifiler™ and Yfiler™ human identification kits. The intention of this experimental avenue was to demonstrate the absence of any false signals generated by DNA profiling caused by mouse DNA and also to reiterate the specificity of the human specific primers used in the human identification systems. It was found that the addition of mouse DNA alone to the STR amplification reaction only enhanced the human DNA signals. This augmentation can be explained by the reduced volume effect produced by adding additional DNA from a non-human species. While the lower limit of 0.0312 ng (approximately 5 diploid cells) needed to generate a full profile did not change from that seen with the absence of a molecular crowding agent, the use of mouse DNA resulted in another option to obtain useful information when handling low copy number samples without extending the analysis process by employing WGA.

The principle of molecular crowding is to reduce dead volume of a reaction by adding a neutral macromolecule (*i.e.*, PEG, nucleic acids, and/or protein) in order to promote a more efficient reaction. Mouse DNA was investigated to determine if mimicking the intact cell with a natural macromolecule would provide a better condensing environment for reactions to take place rather than a synthetic polymer. Indeed, it was shown that the addition of mouse DNA at a high concentration to the WGA reaction containing low levels of human DNA did noticeably increase the signals from the human DNA during STR analysis. These data are interpretable because in theory, the WGA reaction containing human and mouse DNA species are both being subjected to WGA, thus consequently resulting in a similar DNA species ratio to the beginning of the assay.

In many sexual assault cases, the evidence samples taken contain an overwhelming amount of female DNA relative to male DNA. Additionally, when such disparate mixtures occur, traditional chemical separation using differential extraction is often unsuccessful due to preferential amplification; frequently resulting in loss of the secondary (*i.e.*, male) profile. By employing WGA one can overcome the preferential amplification by increasing the template DNA of the sample. It was shown that adding human female DNA as a molecular agent to the WGA reaction increased the signals of the human male DNA template. It was also observed after the WGA; the male pattern with starting DNA equal to one haploid cell (*i.e.*, 0.0035 ng) was detected and yielded a partial profile using Yfiler™. Therefore, the application of WGA in sexual assault cases mimics the use of a macromolecular crowding agent and WGA may offer a solution to those sexual assault samples containing prohibitively low levels of male DNA.

Whole genome amplification as the names implies makes multiple copies representative of the entire genome using random hexamers as primers. However, only those fragments which contain the primer binding site and the STR region together will be detected by the use of forensic multiplex kits. Therefore, many primers will bind to shorter fragments which will not be detected by the fluorescent detection system. Since the limiting factor is the number of primers, decreasing the reaction space via molecular crowding would have a small effect, at best. This is one reason that could account for the two systems not exhibiting an additive characteristic when used together. To test this, one could increase the relative concentration of the primers while all other components remain

the same. If primers are the limiting factor, then one would expect an increase in overall signals.

The advantage of using WGA extends the usefulness of a limited sample by increasing the amount of template to a high enough degree to gain a genetic profile rather than consuming the entire sample without generating any information. For example, if only enough sample is available for one identification reaction after traditional extraction, that limited sample can be subjected to WGA to increase the available DNA template to provide the opportunity for genetic testing using more than one identification system while preserving the option for re-analysis. On the other hand, WGA does increase the time spent processing any given DNA sample, which could act as a deterrent to laboratories with heavy case loads and minimal time. Keeping time management in mind, molecular crowding using mouse DNA can be employed after the extraction step in tandem with the STR analysis to increase signals of low level DNA samples. For instance, if again only enough sample is available for one reaction, but profiling with two systems would be ideal, instead of consuming the entire sample in one assay, the sample could be rationed over two or more aliquots with the addition of a non-human nucleic molecular crowding agent to allow for additional testing on different assays or re-analysis.

It seems WGA and molecular crowding play significant roles in the augmentation of DNA template and STR signals. The previous identification systems routinely used have approached their limitations in sensitivity. We have shown here that the ability to generate a full DNA STR profile from a single cell sample is within reach. Molecular

crowding and WGA may be important in attaining consistent and reliable identity profiles from the smallest sample possible, the single cell.

REFERENCES

Al-Habori M. Macromolecular crowding and its role as intracellular signaling of cell volume regulation. Int J Biochem Cell Biol. 2001 (33): 844-864.

Ellis JR. Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci. 2001 (26): 597-604.

CHAPTER 6.

GENERAL DISCUSSION

DNA typing is used in many disciplines including, but not limited to: medicine, microbiology, animal typing, mass disasters, terrorism, genetics, and forensic science (Watson *et al*, 2007). Applications for DNA typing can be helpful to control animal poaching, a useful tool for genetic reconstruction in mass disasters, in addition to playing a pivotal role for identification in the prevention of a terrorist attack, while typing biopsies and other samples is useful in the practice of medicine and microbiology. The field of genetics studies gene frequencies which can be applied to population statistics used in forensic science.

DNA typing is specifically used in forensic science to identify the source of biological evidence left at a crime scene. There are three different analyses used in forensic DNA, each utilizing a different DNA target. Autosomal DNA and Y-chromosome analysis are practiced routinely using STR technology, while mitochondrial DNA (mtDNA) analysis is achieved via Sanger sequencing methods. In order to appropriately consume the DNA, as DNA analysis is destructive to the evidence, it is important to review the questions asked of the evidence. Autosomal DNA analysis yields the most discriminating statistical evaluation, but this target is easily compromised by environmental elements along with bacteria and chemical degradation. Y-chromosome analysis also is a form of chromosome analysis; however, the target is the Y-chromosome. Y-STR analysis is useful when a high female to male mixture is present, commonly observed in sexual

assaults or to achieve paternal lineage. Chromosomal analysis is very informative, but as mentioned previously, DNA in suboptimal conditions can be easily degraded, resulting in the loss of useful DNA template. Mitochondrial DNA, on the other hand, is housed in the heavily protected mitochondria. Analysis of mtDNA is often utilized as a last resort for DNA analysis in instances where autosomal DNA is unavailable (*i.e.,* mass disasters such as 9-11, missing person cases, and ancient DNA events as those found at archeological sites). Autosomal DNA and Y-chromosome assays are optimal when using approximately 0.5-2ng starting template DNA. Mitochondrial DNA analysis requires several hundred mitochondria. Too much input DNA template can lead to signals that are off scale, making data difficult to interpret, while too little DNA often exhibits allele dropout and limited information.

Unfortunately, samples acquired often are small and exhibit low levels of DNA, so that even if the entire sample is consumed, the amount of available DNA template that can be isolated is often lower than the optimum level required to achieve a successful DNA profile. Sample consumption is a major issue encountered in the practice of forensic DNA typing. Rules and guidelines have been set in place to discourage the consumption of an entire sample; half is allocated for the prosecution, while the other half is reserved for the defense. Experience comes into play when a stain is examined; the analyst must use discretion while deciding how much of the stain is necessary to remove for analysis. If the biological stain is large enough to use a small portion (less than half), then the analyst can proceed with extraction, one the other hand, if the stain is a size where half may not yield the optimal level of DNA, the analyst must obtain permission from the opposing side

requesting for testing more than half the sample. In the instance that a request is submitted, the opposing team may reject the request and instead demand a third party to analyze and consume the sample. These regulations have been put in place to protect the rights of individuals, but they do require additional time and effort, by the analyst, along with the legal support as well, to appropriately process limited evidence. These struggles paint a clear picture that it is imperative to develop new protocols that are more sensitive than current procedures.

Our exploration of WGA to increase low level DNA template for further downstream applications has turned out to be a successful approach to widen the possibilities of analysis. Our first series of experiments demonstrated the MDA technique was much more reliable than the PEP protocol. The failure to generate an increase in template using PEP WGA could be explained by the prolonged cell lysis and fragmentation steps required with the method. In theory, the Lysis and Fragmentation buffer included in the kit would induce denaturing conditions which lead to fragmentation, thus nicking the DNA extensively. If the DNA is fragmented significantly, the primers used to prepare the genomic library will not be able to anneal properly, thus inhibiting sufficient amplification of the genome. This extensive unraveling of the DNA template into short fragments could explain the inadequate results obtained from the PEP protocol. It was demonstrated that MDA was superior in our hands in all cases when compared to the PEP whole genome amplification system. Many techniques of WGA have been used as comparative methods for downstream applications in genomic hybridization assays, however, gene coverage and

allele bias are regarded with relaxed requirements as combining WGA with each current technology is relatively new.

A major change which leads to a decrease in the level of artifact and peak imbalance was to limit WGA reaction time, though; at first it seemed counter intuitive as our goal was to increase the quantity of template sample. For admissible evidence, there is a fine line between quantity and quality. The ultimate goal was to increase the amount of authentic DNA template to increase the viability of downstream DNA analyses, but to also increase the quality of the profile results. By limiting the φ29 DNA polymerase reaction time, we were able to confidently make allele calls due to the lowered baseline heights observed in the data. This adjustment of the protocol led to stringent, yet reliable way to obtain usable results from template DNA of 4-6 diploid cells. The additional use of macromolecular crowding further enhanced the efficiency of the assay by increasing the sensitivity of allele identification from a haploid cell equivalent.

The application of clinical laser catapult micro dissection led to a limiting factor of our protocols. Though there were sporadic assays that were successful in the development of a partial genetic profile from a single cell, the limits of detection were attained in all procedures examined. The LCM delivered precision in sample selection, however, lacked in ease for sample retrieval. Our laboratory used membrane slides that are catapulted into the tube cap of an adhesive tube. It was demonstrated the adhesiveness of the sample tubes was inconsistent from lot-to-lot as the level of sample recovery varied from one lot of tubes to the next. It was also confirmed by Zeiss that quality assurance of these tube styles is minimal as it is not "cost effective to monitor because of the small amount of that tube

style distributed." Our efforts to mask the adhesive silicone of the cap only heightened the stickiness of the adhesive cap while the sample remained on the tube cap or in other cases, alternatively managed to inhibit downstream reactions. The fact that we were getting signals from 6 diploid cells exploits the possibility that those signals were not from 6 cells, they were more likely generated from a fraction of that amount, as the remaining sample very likely adhered to the tube cap. It was the realization of this fact after months of troubleshooting that switching to pre-extracted DNA in fact the proper way to proceed. Forensic DNA facilities work with purified DNA on a daily basis, so experiments based on this template more closely reflected the "real world" of forensic analysis.

Adding WGA to the DNA processing workflow does add additional hours to the analysis protocol; however, the benefits of generating maximum information with the alternative of re-analysis significantly favor this approach. Conventionally, it takes a minimum of 1 ng (approximately 150 diploid cells) of template DNA to generate a full acceptable DNA pattern. By initiating WGA, we have developed a way to confidently generate a full pattern for samples containing as little as 0.0312 ng, the equivalent of approximately 5 diploid cells. Traditional analysis does not have the power to provide such extensive information from such a limited sample, which is why the use of WGA and/or macromolecular crowding is useful, if for nothing else than to extend the number of assays that can be done with a limited original sample.

In practice, many cases can only provide a fraction of the amount needed using current methods to produce full profiles with reliability. Rather than using the less than desirable amount of DNA to produce minimal information, the sample can be subjected to

a different route of analysis using whole genome amplification and/or macromolecular crowding. The flow chart in Figure 6-1 illustrates alternate approaches with different samples to provide the best overall results. The first consideration is to determine how much sample is available for use. If there is at least 2 ng, one can proceed with confidence using conventional techniques accepted in the forensic community and be assured that there is sufficient sample reserved for one additional testing or re-analysis, if necessary. With the inclusion of WGA and/or macromolecular crowding, the option for alternative processing increases by a thousand-fold. WGA can be applied to the sample to generate multiple copies of the genome for downstream analysis and sample retention. If the sample is very minimal, mouse DNA can be added to the whole genome amplification reaction to ensure efficient replication of the DNA template. On the other hand, if the sample needs to be processed immediately, molecular crowding during the STR amplification has also shown promising results. Those situations that are time sensitive may be more apt to utilize molecular crowding to prevent another victim in a serial rapist/murder case or a potential terrorist attack.

The combination of WGA and macromolecular crowding could help solve the dilemma of samples containing limited DNA. Further examination of these alternatives holds a promising outcome for the scientific and forensic communities. It would be meaningful to study the effects of extracting low levels of human DNA with the addition of a carrier DNA during the extraction process to promote a higher human DNA yield. It is evident that during every tube transfer and assay, sample may be lost, but with the addition of less critical nucleic acids that do not cross react in downstream applications may lead to

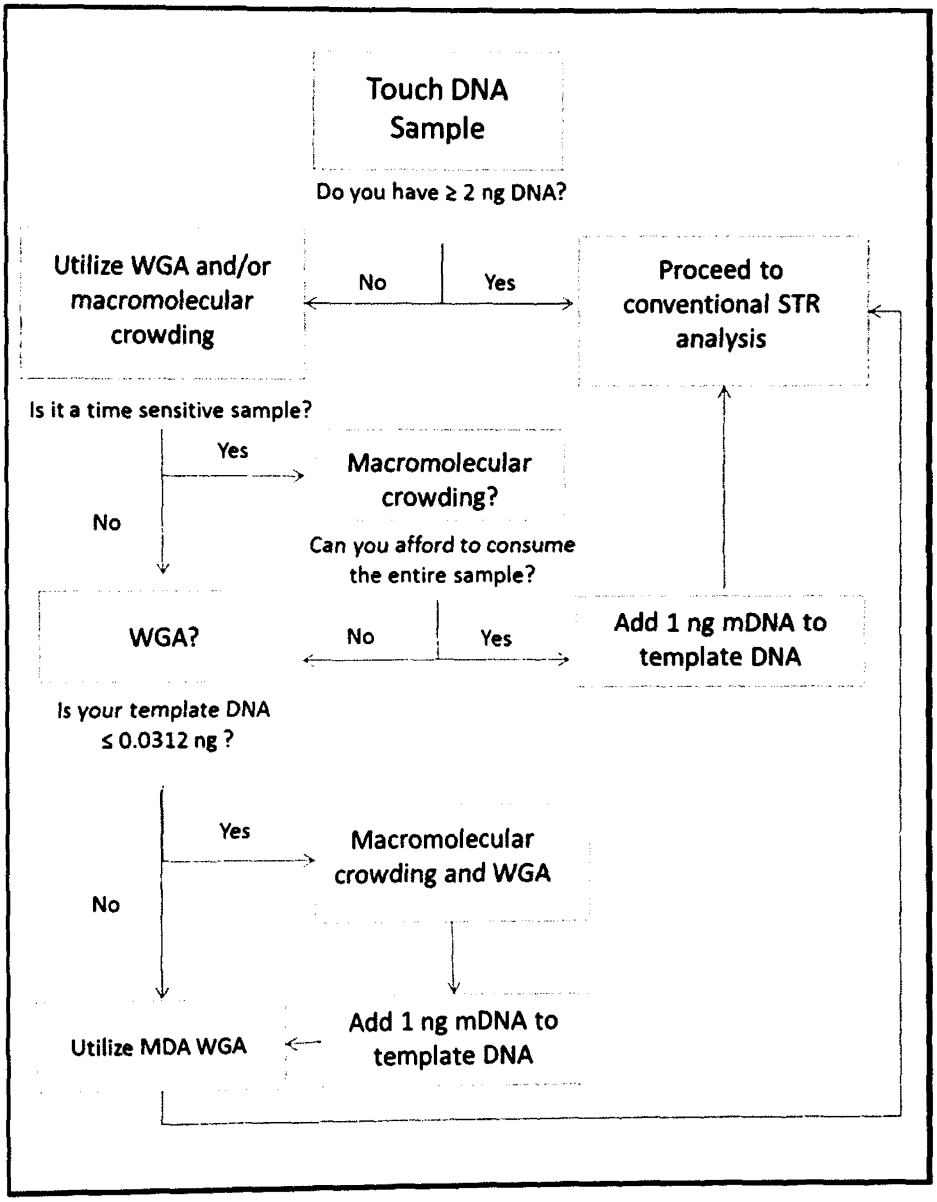

Figure 6-1: Outline of potential workflow incorporating WGA and molecular crowding using low copy number samples.

a protocol with an even higher rate of success.

The application of these techniques in a modern forensic laboratory would generate viable options for those cases that don't have enough DNA initially to be processed. Providing alternatives to just processing the sample or not would offer law enforcement

agencies the opportunity to remove more perpetrators from the public arena for a reduction in crime and also the option to process cold cases. Another important application in recent years is ability to allow analysis of marginal evidence, much like those included in the Innocence Project (The Innocence Project). Unfortunately, many requests for analysis are on items that contain one nanogram or less. Though, it is possible to conduct one assay with the single nanogram, there is not any sample retention for future testing, if needed. Too many cases are shelved for characteristics relating to low levels of DNA template and/or the sample would require much of the analyst's time, in turn, taking them away from other *routine* casework. The number of unprocessed cases nationwide is astounding and continues to grow. These techniques can also be applied to a clinical lab setting when often times the amount of sample taken is not enough for more than one analyses (*e.g.,* biopsy).

The availability of a protocol that has the ability to achieve an STR DNA profile from a single cell range is the "Holy Grail" of forensic DNA analysis. Though there are statistical limitations to the pattern obtained from a haploid cell (*i.e.,* the calculation would alternatively, continued inclusion as a potential perpetrator. In the instance of generating a DNA profile from a diploid cell, the random match probability calculation can be used, be a combined paternity index or using the counting method rather than a random match probability),the obtained information can still be used to exclude a portion of the population; therefore, single cell analysis would provide for immediate exclusion, or resulting in a very discriminating genotype frequency, making low copy number samples manageable and alleviating this as a roadblock for DNA analysis. I have demonstrated that

five cells is sufficient for reliable analysis, and that 'tweaking' these protocols may soon allow for single cell analysis to become a reality.

REFERENCES

The Innocence Project. http://innocenceproject.org

Watson JD, et al. DNA fingerprinting and forensics. Recombinant DNA: Genes and genomes—A short course 3rd Edition. Cold Spring Harbor, NY: Cold Spring Harbor Press. 2007.