2020

Single Cell Forensic Genomics - DNA Profiling of Micromanipulated Single Spermatozoa

Haley Hardin
University of Central Florida

SINGLE CELL FORENSIC GENOMICS – DNA PROFILNG OF MICROMANIPULATED SINGLE SPERMATOZOA

by

HALEY HARDIN
B.S. UNIVERSITY OF CENTRAL FLORIDA, 2018

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Sciences in Biotechnology
in the Burnett School of Biomedical Sciences
in the College of Medicine
at the University of Central Florida
Orlando, Florida

Summer Term
2020

ABSTRACT

Current U.S. National guidelines allow for the collection of sexual assault evidence up to 5 days after the incidence occurs. In these cases, the ability to obtain an autosomal STR (aSTR) profile of the male donor in these cases diminishes as the time interval increases. This inability to recover an aSTR profile from the semen donor is not due to a complete lack of sperm cells, as studies have frequently shown that sperm persists in the vaginal canal or cervix up to 10 days post coitus. Thus, the inability to recover an aSTR profile of the sperm donor is likely due to a low quantity of sperm cells and/or degradation of sperm cells, which pose significant problems to existing DNA extraction and typing methods. A typical DNA workflow for this type of evidence in a forensic casework laboratory includes the use of a differential extraction to separate sperm cells from non-sperm cells. These often harsh extraction methods can cause degraded and fragile sperm cells to be prematurely lysed into the non-sperm cell fraction. The significant amounts of vaginal epithelial cells in the sample can overwhelm the minute number of sperm cells present in this fraction, resulting in a complete masking of the male profile. For most sexual assault samples collected more than 48-72 hours after an incident, Y-STR analysis might be used instead of aSTR analysis, as it allows for an increased time frame of DNA recovery by detecting only the male donor Y-haplotype, circumventing the need for a differential extraction and avoiding potential competition during amplification. However, Y-STR loci are part of the non-recombining region of the Y-chromosome, and thus do not have the same discrimination power of aSTR loci. Therefore, enhanced methods for the recovery of sperm cells that allow researchers to bypass the limitations of a typical DNA workflow and obtain an aSTR profile need to be developed. This study aimed to do so by utilizing enhanced, non-standard methods to collect individual sperm cells via direct physical recovery from semen containing

samples. Optimized direct lysis and micro-volume aSTR amplification were also used in order to obtain DNA profiles of the sperm donor. The quality of the obtained DNA profiles was evaluated using metrics such as allele recovery, stutter occurrence and percentage, and drop-in allele levels. Using these developed methods, the ability to analyze single sperm cells was demonstrated and the minimum number of sperm cells required in order to obtain probative and reliable DNA profiles was determined.

This thesis is dedicated to all victims of the violent crimes for which we search for answers.

ACKNOWLEDGMENTS

I would like to thank my parents, family, and friends for their unending support during my time in school. To my committee chair and PI, Dr. Jack Ballantyne, for welcoming me into his lab and giving me guidance through the last two and a half years. To Dr. Erin Hanson for being a mentor and a friend and having endless patience with me. To my committee members Dr. Annette Khaled and Dr. Saleh Naser for their support and encouragement. To my fellow lab members, past and present: Paris Volk, Nikhita Shankar, Anna Kimball, and Kaitlin Huffman. Finally, I would like to thank all of the body fluid donors, without whom none of this research would be possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

μL	Microliter
μM	Micromolar
ADE	Acoustic Differential Extraction
ADO	Allele Drop Out
aSTR	Autosomal Short Tandem Repeats
bp	Base pair
CODIS	Combined DNA Index System
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
LCN	Low Copy Number
LDO	Locus Drop Out
LMD	Laser Microdissection
Low-TE	Low Tris-EDTA
LTDNA	Low Template Deoxyribonucleic Acid
min	Minutes
mL	Milliliter
ng	Nanograms
NFW	Nuclease Free Water
PCR	Polymerase Chain Reaction
pg	Picograms
RFU	Relative Fluorescent Unit

RMP	Random Match Probability
RPM	Revolutions Per Minute
SE	Semen
sec	Seconds
SD	Standard Deviation
SR	Stutter Ratio
STR	Short Tandem Repeat
SWGDAM	Scientific Working Group on DNA Analysis Methods
Tris	Tris (Hydroxymethyl) Aminomethane
U.S.	United States
UV	Ultraviolet
Y-STR	Y-chromosome Short Tandem Repeat

CHAPTER 1: INTRODUCTION

DNA profiling is considered the gold standard of forensics [1], as it allows for the identification of individuals involved in a crime. In sexual assault cases, identification of the male perpetrator is critical for the prosecution of such criminal cases [2]. Therefore, it is imperative to isolate sperm cells from other epithelial cells in the sample to obtain a single-source profile of the male donor [3-6]. If the cell types are not successfully separated, the major contributor (typically the victim) can cause titration of PCR reagents during amplification, which effectively masks the male DNA in the sample [7]. In order to perform such a separation, forensic laboratories will typically use a differential extraction (Figure 1) to separate non-sperm cells (consisting primarily vaginal epithelial cells obtained from the victim, as well as male epithelial cells) from sperm cells [4-6], allowing the analysts to amplify the two groups separately. This is achieved by utilizing the difference in the composition and structure of the nuclear membranes of either cell type, as epithelial cells will preferentially lyse while sperm cells require the addition of a reducing agent, such as dithiothreitol (DTT) for complete lysis [4-6].

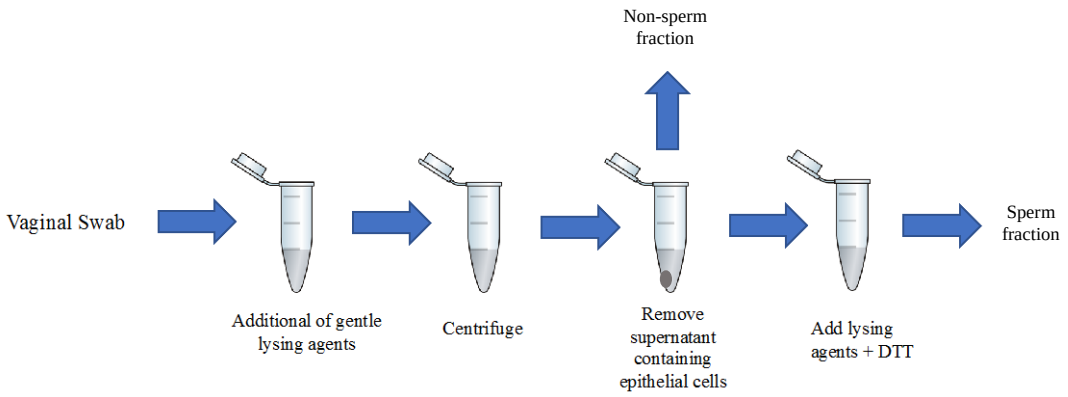

Figure 1: Differential extraction workflow

Sexual assault victims might not report an assault immediately (i.e. within the 5 day evidence collection period), due to a variety of reasons such as incapacitation via drugs or alcohol, fear to come forward, or the inability to recognize and report the assault, which often occurs in cases with young children [8, 9]. In these instances, the sperm cells might have been subjected to vaginal lavage and drainage, menstruation, and/or general degradation from the cervicovaginal environment [10-14]. Performing a differential extraction in these instances is not always possible or successful, as the sperm cells are fragile and have the potential to prematurely lyse, resulting in a loss of sample into the non-sperm fraction [7, 10, 15]. In cases where performing a differential extraction is not possible, an alternative option for analysis of these samples is the use of Y-chromosome STR (Y-STR) analysis [10, 16-23]. Y-STR analysis amplifies short tandem repeats found on the Y-chromosome. Short tandem repeats (STR) are DNA regions with repeat units that are 2-6 base pairs long [24]. The number of repeat units in an STR is highly polymorphic, which makes them extremely effective for human identification. Y-STR analysis eliminates the need for separation of the male cells from the sample, as the amplification process only targets male DNA [10, 16-22].

Y-STR typing is a valuable tool in forensics and has allowed for the extension of the time interval in which we are able to obtain probative profiles from sperm donors in extended post-coital time intervals [10, 20-22]. While the use of Y-STR analysis is valuable to forensic analysts, the loci in Y-STR analysis are not independent of each other and therefore have a low discriminatory power, as the males in one lineage are likely to have the same Y-STR profile [16, 19, 23, 24]. Since Y-STR loci are not independent of each other, Y-STR genotype frequencies in a population cannot be generated using the product rule. Instead, the counting method is used which observes the number of times a haplotype is seen in a database size of N [23]. While there

are several Y-STR databases in the U.S., these databases are limited in the number of samples and loci provided and thus are limited in their overall usefulness. For this reason, the gold standard of DNA profiling is autosomal STR (aSTR) analysis, which consists of amplified STRs found on autosomal chromosomes [24]. Autosomal STR loci are inherited independently, which allows for greater variation in DNA profiles [25-27]. When aSTR profiles are used, the profiles obtained can be evaluated for discriminatory power by calculating the random match probability (RMP) [6, 25-27]. The RMP is calculated using the product rule to combine the frequency of a genotype at each locus analyzed. The frequency of a heterozygote genotype (f) is calculated for alleles "p" and "q" using the equation f=2pq [6, 25]. The frequency of a homozygous genotype (f) is calculated for allele "p" using the equation $f=p^2 + p(1-p)\theta$ where $\theta=0.01$ for the general population [6, 25]. Autosomal STR analysis is advantageous in this regard as it allows for the analysts to determine how common or rare a profile is in the general population. [6, 24-27]. Additionally, autosomal loci are used in the Combined DNA Index System (CODIS), a database created by the Federal Bureau of Investigation in which data from core aSTR loci are uploaded to create the ability to record and search for suspects in a crime [6, 24, 27, 28]. When the CODIS loci are tested, the average RMP, or the chance that two unrelated individuals in a population share the same profile, is typically rarer than one in a trillion among unrelated individuals [6, 24, 26, 27, 29]. The ability to produce aSTR profiles allows for a higher discrimination power and the use of a national database which is contributed to by federal, state, and local forensic laboratories, making it highly advantageous in casework analysis.

Due to the limitations of current forensic analysis methods in regard to late-reported (i.e. more than 5-day) sexual assault samples, there is the need to develop an alternative method for sperm cell recovery that allows for aSTR analysis. Several methods have been proposed, such as

acoustic differential extraction and laser microdissection [15, 30-33]. Acoustic differential extraction (ADE) uses a microfluidic device that manipulates acoustic forces to sort cells into different groups and has been used to process sexual assault samples [30, 32]. However, it has been demonstrated that use of this method can cause cells in the sample to adhere to the trapping component of the microfluidic device and hinder sorting or cause a loss of the sample. Another method used to separate sperm cells from samples is laser microdissection (LMD) [15, 31, 33], which uses UV laser ablation to physically remove the cell from a slide and capture it into a reaction mix using gravitational forces for downstream analysis. For this method to work, accurately, however, requires a large sample size which is not always available in these cases. While these methods have been able to successfully isolate sperm cells, they have not yet been implemented into U.S. forensic laboratories and remain used mainly in research settings. This could be for a number of reasons – including the fact that both alternative approaches would require the use of equipment not found regularly in U.S. crime laboratories and would require additional training for analysts.

Another alternative method to those described previously is the use of simplified micromanipulation. Micromanipulation is a technique that was developed by Dr. Robert Chambers in 1921 to isolate bacterium from a mixture of cells for dissection of the microbe [34]. This technique has since been used in a variety of applications since its inception, such as *in vitro* fertilization [35-37]. In forensic research, simplified micromanipulation has been used to successfully recover trace DNA from physical assault samples via stereomicroscope, by capturing "touch DNA" on a Gel-Film® (Gel-Pak, Hayward, CA) slide and successfully isolating the "touch" DNA for downstream lysis and amplification, resulting in probative aSTR profiles [38, 39]. In this context, micromanipulation is performed using a tungsten needle coated with

water-soluble adhesive to "pick up" cells from a slide and deposit them into a solution. In this work, simplified micromanipulation was evaluated for use in the recovery of sperm cells from semen samples to produce an aSTR profile.

After retrieval of the sperm cells, the genetic material must be extracted for downstream amplification. This can provide unique challenges, as most purification techniques can cause a loss of genetic material when removing contaminants, which is not suitable for use in small cell samples. In these instances direct lysis can be used, which allows for the extraction of genetic material by using a proteinase that will lyse cells and degrade proteins and nucleases at higher temperatures while leaving nucleic acids intact [40, 41]. By performing a direct lysis, a closed environment is provided without multiple washing steps which would lead to critical loss of nucleic acids [40, 41]. In cases where the sample is only several cells, any loss of genetic material would be extremely detrimental. Direct lysis also cuts down on potential contaminations that might occur with repeated opening and closing of the tube that takes place during a typical extraction. Any contamination to a sample with such little input can drastically affect the results.

After lysis, amplification of the sample is required for the production of a DNA profile. In single- or few-cell samples, the small amount of genetic material available for amplification can prove troublesome to the production of a full DNA profile [42-44]. Most aSTR amplification kits have an optimal DNA input of 0.5 to 1 nanogram (ng) of DNA. Single cells contain approximately 6 picograms (pg) of DNA, while haplotype cells (e.g. sperm cells) contain approximately 3 pg of DNA. Any samples containing less than >100 pg of DNA are considered low copy number samples [42-45], and require enhanced analysis methods to recover a probative profile [43, 44, 46-49]. Several enhanced methods were used in this work to obtain

optimal DNA recovery. The PCR cycle number was increased from 28 to 34, which adds more cycles of PCR in order to increase the amount of amplified product [44, 45, 47, 50]. Micro-volume amplification was also used, where the reaction is scaled down from 25 μL to 5 μL to concentrate the reagents of the mixture allowing for a more efficient amplification [46, 48, 51]. Additional DNA polymerase was also added to the amplification mix to increase the overall activity of the polymerase in the reaction [48].

While these methods may increase allele recovery for the samples provided, they do not eliminate, and often cause an increase in stochastic artifacts observed with the amplification of low template DNA (LTDNA) [43, 44, 47, 49, 52, 53]. These artifacts include allele drop-in, allele drop-out, heterozygous peak imbalance and increased stutter [44, 45, 50, 52-55]. Allele drop-in is the presence of alleles in the profile that are not attributable to the donor, while allele drop-out is the loss one allele at a locus. This is especially troublesome if only one heterozygous allele is amplified, which would then appear as a homozygous genotype for an individual locus. Stutter is a result of polymerase slippage during replication that results in an amplification product containing one repeat unit less (n-1) than the true allele [56-58]. While n-1 stutter products are the most common, additional stutter products can be formed that are more than one repeat unit less (e.g. n-2, two repeat units less) or one repeat unit more (e.g. n+1) [56-58]. Stutter is a standard occurrence in STR profiling and can complicate profile analysis typically with mixture samples. Thresholds based on average or typical stutter percentages are employed in order to filter out these artifacts [56-58]. However, these are based on the standard analysis, with standard input DNA amounts and standard amplification conditions. When LTDNA samples are used, especially in combination with increased cycle number, increased stutter will be observed and will likely exceed these thresholds and cannot be distinguished from genuine alleles in most

cases [46, 55, 59, 60]. While stochastic effects cannot be avoided in low template samples, the profiles obtained in this work were evaluated and analysis thresholds were developed to help combat their effects on the quality of the profiles.

Another method for the evaluation of LTDNA is through consensus profiling [55, 59]. To produce consensus profiles, multiple replicates of the same LTDNA sample is amplified and the separate profiles are combined together to create one profile. This helps ensure accuracy of the profiles, as an allele must be present in more than one profile to be included in the final profile. Since LTDNA artifacts are stochastic in nature, they would not be expected to be present in each replicate amplification. The use of consensus profiling was evaluated in this work to determine if it would be an effective tool to aid in the profile interpretation of the single- or few sperm cell sample profiles.

Additionally, the probative value of the profiles obtained in this work were evaluated by calculating the RMP. Typically, the Scientific Working Group on DNA Analysis Methods (SWGDAM) recommends that LCN samples are analyzed with a stochastic threshold to distinguish between a homozygous determination at a locus and a heterozygous genotype with allele drop-out [61]. However, due to the nature of the cells used in this work (i.e. haplotype cells), heterozygote peak imbalance would be increased thus the creation and use of a stochastic threshold to determine homozygosity would not likely be accurate. For this reason, when analyzing the discriminatory power of the aSTR profiles obtained in this work, all alleles were analyzed as heterozygote genotypes, and the frequency of homozygote genotypes (f) were

calculated as a heterozygote genotype with only one allele detected, using the equation f=2p, as to not conflate the statistical value of the profiles obtained [62].

The results of this work demonstrate that sperm cells can be collected using simplified micromanipulation and probative aSTR profiles can be obtained from them using enhanced analysis techniques. Using this work, it is the hope that further research can be done to fully develop and incorporate this new method for isolation of sperm cells from late-reported sexual assault samples in U.S. forensic laboratories to produce probative aSTR profiles that can be used to identify and prosecute the perpetrator(s) involved.

CHAPTER 2: MATERIALS AND METHODS

2.1. Semen Samples

Semen was collected from anonymous donors using procedures approved by the University of Central Florida's (UCF) Institutional Review Board. Informed consent was obtained from each donor. Freshly ejaculated semen was collected in 15 μL conical tubes and stored at -20°C. Additional semen samples were commercially purchased (donors from BIOIVT, Nassau, NY and Lee Biosolutions, Maryland Heights, MO) and stored at -20°C. Semen swabs were prepared by placing the swab directly in thawed liquid semen to saturate, dried overnight at room temperature (~22°C), and stored at -20°C.

2.2. Micromanipulation Sample Preparation

Gel-Film® slides were prepared by securing a piece of Gel-Film® on to a clean microscope slide. To transfer the semen sample to the Gel-Film® slide, 1/4th of a saturated semen swab was suspended in 500 μL of Nuclease Free Water (NFW) in a sterile 1.5 mL microcentrifuge tube, agitated for 30 seconds then incubated at room temperature for 30 minutes. The swab piece was then placed in a spin basked inside the microcentrifuge tube and centrifuged for 5 minutes at 13,000 rpm. The spin basket containing the swab was then discarded, along with the supernatant, and the pellet was resolubilized in 400 μL NFW. The suspension was the transferred to the Gel-Film® slide and dried overnight at room temperature (~22°C). The slide was then stained using Christmas tree stain (Serological Research Institute (SERI), Richmond, CA). The nuclear fast red stain was placed on the Gel-Film® surface and incubated at room temperature (~22°C) for 10 minutes and rinsed with NFW by gentle flooding (e.g. slide tipped on an angle and water gently runs down from the top of the slide in order to prevent sample loss).

The picroindingocarmine counter stain was placed on the Gel-Film® surface for 10 seconds and rinsed with 100% ethanol by gentle flooding. The slide was then air dried at room temperature before micromanipulation. Slides were stored at room temperature (~22°C) in a sterile slide box.

2.3. Micromanipulation

Sperm cells were collected using a Caltex LX-100 Digital Microscope Video Inspection System (Caltex Scientific, Irvine, CA) at 400x magnification. Each collection consisted of five 1-cell samples, and two 2-, 5-, 10- and 20- cell samples. For all collections, two "0" cell samples were collected from a blank area of the Gel-Film® slide to test for the presence of 'cell-free' DNA. Samples were collected using 3M™ (3M™, St. Paul, MN) water-soluble wave solder tape on the end of a tungsten needle to adhere to the sperm cells. First, 3M™ adhesive tape was attached to a clean slide using sterilized tweezers, then the tungsten needle was scraped along the adhesive to create a small ball (Figure 2).

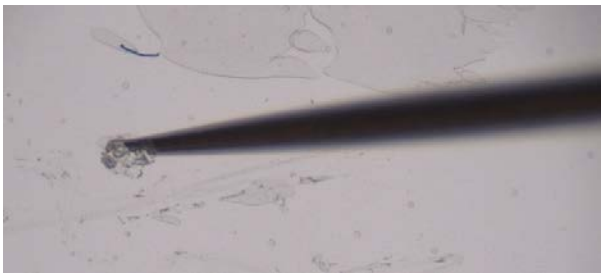

Figure 2: Microscope image of a tungsten needle coated in 3M™ adhesive for micromanipulation.

To collect the sperm cells, the needle with adhesive was placed over the sperm cell of interest, and the pressure was applied (Figure 3). The needle was then lifted up from the slide, and a visual confirmation was used to ensure the cell(s) of interest were successfully collected from the slide and all other cells in the visual field remain (Figure 3).

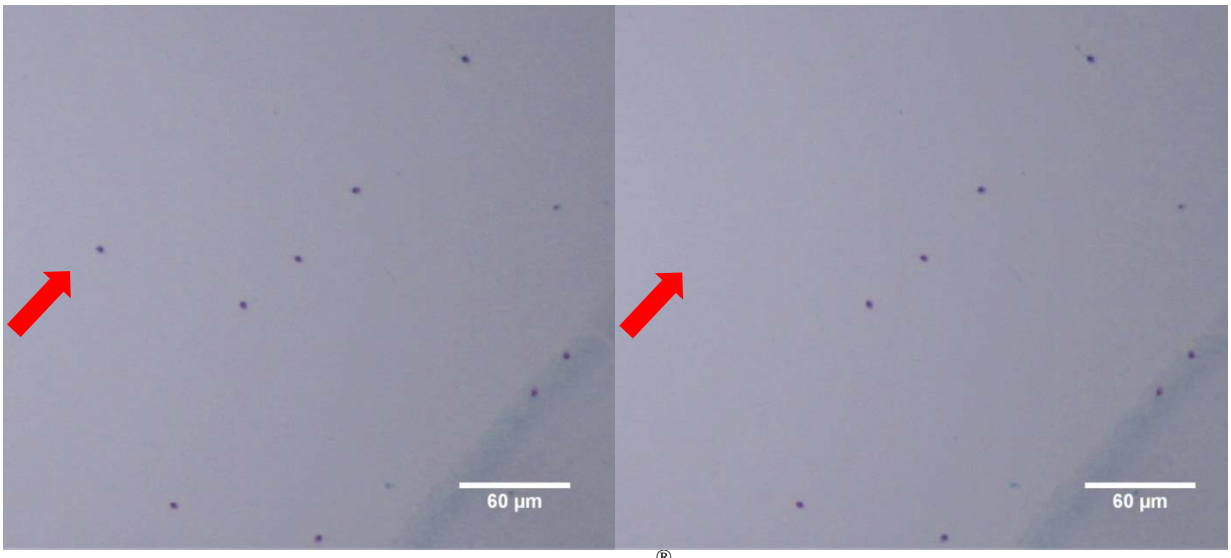

Figure 3: Microscope image of a sperm cell on a Gel-Film® slide using 400X magnification, before (left) and after (right) micromanipulation

Next, the needle containing the cell(s) of interest and water-soluble adhesive were dissolved into the lysis mix (either 1.5 µL or 2 µL depending on the kit) in a sterile 0.2 mL PCR flat-cap tube (Figure 4).

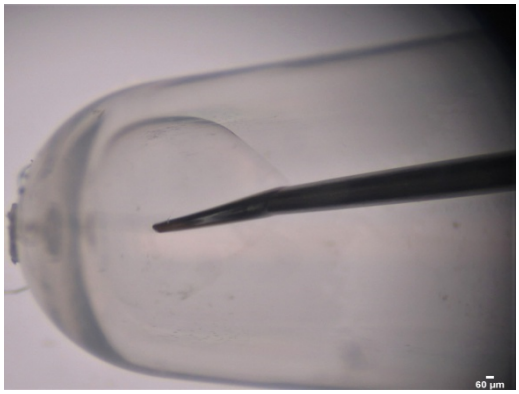

Figure 4: Microscope image of a 3M™ adhesive coated needle attached to a sperm cell in a 0.2 mL PCR tube.

2.4. DNA Extraction

Reference profiles were created for all semen donors used for micromanipulation. A manual extraction was performed on semen donors using QIAmp® DNA Investigator kit (QIAGEN, Germantown, MD) according to the manufacturer's protocol. Samples were incubated for 1 hour at 56°C in 400µL Buffer ATL, 20 µL Proteinase K (Fisher Scientific, Grand Island, NY) and 20

μL DTT (Fisher Scientific, Grand Island, NY) in a sterile microcentrifuge tube. Samples were agitated and spun down every 10 minutes to improve lysis. Samples were then incubated for 10 min at 70°C, agitated and spun down every 3 minutes. After lysis, swabs were transferred to a DNA IQ spin basket (Promega, Madison, WI) in the microcentrifuge tube and centrifuged at 14,000 rpm for 5 minutes remove all liquid from the swab. After centrifugation, the spin basket and swab were discarded, and the lysate was transferred to a QIAmp® MinElute® spin column. Next, the samples were washed with 500μL Buffer AW1 (QIAGEN), 700μL Buffer AW2 and 700μL 100% ethanol. The column was then dried via centrifugation at 14,000 rpm for 3 min, followed by an incubation at room temperature for 10 min. Samples were eluted into 60 μL of Buffer EB. Each extraction included an extraction blank as a negative control. Extracts were quantified using Quantfiler™ Trio (Applied Biosystems™ by ThermoFisher, Foster City, CA) according to the manufacturer's protocol on an ABI Prism® 7000 real time PCR instrument (Applied Biosystems™ by ThermoFisher, Foster City, CA).

2.5. Direct Lysis Methods

2.5.1 forensicGEM Sperm

Sperm cells collected via micromanipulation were deposited into a direct lysis mix. The first method tested was the forensicGEM Sperm kit (ZyGEM™, Solona Beach, CA). Sperm cells were collected directly into a 1.5 μL lysis mix. This lysis mix was prepared using 1.1 μL nuclease free water, 0.35 μL 10X Buffer Orange, 0.15 μL ACROSOLV, and 0.1 μL forensicGEM enzyme. Cells were then lysed in a thermocycler using the following program:

52°C for 5 min, 75°C for 3 min, 95°C for 3 min, and a 4°C hold. Lysates were used immediately in STR amplification.

2.5.2 LysePrep

Sperm cells collected via micromanipulation were deposited into a direct lysis mix. The second method tested was the LysePrep kit (DEPArray™ by Menari Silicon Biosystems, Bologna, Italy). Sperm cells were collected directly into a 2.0 µL lysis mix. This lysis mix was prepared using 1.18 µL NFW, 0.13 µL Reagent 1, 0.13 µL Reagent 2, 0.2 µL Buffer, 0.3 µL Enzyme and 0.1 µL 100 µM DTT. Cells where then lysed in a thermocycler using the following program: 42°C for 45 min, 65°C for 30 min, 80°C for 15 min, and a 4°C hold. Lysates were used immediately in STR amplification.

2.6. Autosomal STR Amplification

2.6.1 AmpFLSTR™ IdentiFiler™ Plus

For micromanipulated samples, a 5 µL (forensicGEM sperm) or a 5.5 µL (LysePrep) AmpFLSTR™ IdentiFiler™ Plus amplification (Applied Biosystems™ by ThermoFisher, Foster City, CA) was performed by adding 3.5 µL of the STR amplification mix to each tube, which consisted of 2.2 µL master mix, 1.2 µL primer mix and 0.2 µL AmpliTaq Gold® enzyme (Applied Biosystems™ by ThermoFisher, Foster City, CA). The samples were then amplified using the following PCR program: 95°C for 11 min 34 cycles of 94°C for 20 sec and 59°C for 3 min, 60°C for 10 min and a hold at 4°C. Amplification products were stored at 4°C until needed.

For amplification of donor extracts, a 25 µL AmpFLSTR™ IdentiFiler™ Plus amplification (Applied Biosystems™ by ThermoFisher, Foster City, CA) was performed by

adding 25 µL of the STR amplification mix to each tube, which consisted of 10 µL master mix, 5 µL primer mix. Extracts were diluted with Low Tris-EDTA (TE) buffer (Applied Biosystems™ by ThermoFisher, Foster City, CA) to achieve an input of 1 ng and a final volume of 10 µL. and The samples were then amplified using the following PCR program: 95°C for 11 min, 28 cycles of 94°C for 20 sec and 59°C for 3 min, 60°C for 10 min and a hold at 4°C. Amplification products were stored at 4°C until needed.

2.6.2 GlobalFiler™

For micromanipulated samples, a 5 µL (forensicGEM sperm) or a 5.5 µL (LysePrep) GlobalFiler™ amplification (Applied Biosystems™ by ThermoFisher, Foster City, CA) was performed by adding 3.5 µL of the STR amplification mix to each tube, which consisted of 1.5 µL Reagent Mix, 1.3 µL Low-TE Buffer, 0.5 µL Primer Mix and 0.2 µL AmpliTaq Gold. The samples were then amplified using the following PCR program: 95°C for 1 min, 34 cycles of 94°C for 10 sec and 59°C for 90 sec, 60°C for 10 min and a hold at 4°C. Amplification products were stored at 4°C until needed.

For amplification of donor extracts, a 25 µL GlobalFiler™ amplification (Applied Biosystems™ by ThermoFisher, Foster City, CA) was performed by adding 25 µL of the STR amplification mix to each tube, which consisted of 7.5 µL Reagent Mix, and 2.5 µL Primer Mix. Reference samples were diluted with Low-TE buffer to achieve an input concentration of 1.0 ng, with a final volume of 15 µL. The samples were then amplified using the following PCR program: 95°C for 1 min, 29 cycles of 94°C for 10 sec and 59°C for 90 sec, 60°C for 10 min and a hold at 4°C. Amplification products were stored at 4°C until needed.

2.6.3. PowerPlex® Fusion 6C

For micromanipulated samples, a 5 µL (forensicGEM sperm) or a 5.5 µL (LysePrep) PowerPlex® Fusion 6C amplification (Promega Corporation, Madison, WI) was performed by adding 3.5 µL of the STR amplification mix to each tube, which consisted of 1 µL Reagent Mix, 1.3 µL NFW (Invitrogen™ by ThermoFisher Scientific, Carlsbad, CA), 0.5 µL Primer Mix and 0.2 µL AmpliTaq Gold. The samples were then amplified using the following PCR program: 96°C for 1 min, 34 cycles of 96°C for 10 sec and 60°C for 1 min, 60°C for 10 min and a hold at 4°C. Amplification products were stored at 4°C until needed.

For donor extracts, a 25 µL PowerPlex® Fusion 6C amplification (Promega Corporation, Madison, WI) was performed by adding 25 µL of the STR amplification mix to each tube, which consisted of 5 µL Reagent Mix, and 5 µL Primer Mix. Reference samples were diluted with NFW to achieve an input concentration of 1.0 ng, with a final volume of 15 µL. The samples were then amplified using the following PCR program: 96°C for 1 min, 29 cycles of 94°C for 10 sec and 60°C for 1 min, 60°C for 10 min and a hold at 4°C. Amplification products were stored at 4°C until needed.

2.6.4. VeriFiler™ Plus

For micromanipulated samples, a 5 µL (forensicGEM sperm) or a 5.5 µL (LysePrep) VeriFiler™ Plus amplification (Applied Biosystems™ by ThermoFisher, Foster City, CA) was performed by adding 3.5 µL of the STR amplification mix to each tube, which consisted of 1 µL Reagent Mix, 1.8 µL Low-TE Buffer, 0.5 µL Primer Mix and 0.2 µL AmpliTaq Gold. The samples were then amplified using the following PCR program: 95°C for 1 min, 2 cycles of 96°C

for 10 sec and 62°C for 90 sec, 32 cycles of 96°C for 10 sec and 59°C for 90 sec, 60°C for 5 min and a hold at 4°C. Amplification products were stored at 4°C until needed.

For donor extracts, a 25 μL VeriFiler™ Plus amplification (Applied Biosystems™ by ThermoFisher, Foster City, CA) was performed by adding 25 μL of the STR amplification mix to each tube, which consisted of 5 μL Reagent Mix, and 2.5 μL Primer Mix. Reference samples were diluted with Low-TE Buffer to achieve an input concentration of 0.5 ng, with a final volume of 17.5 μL. The samples were then amplified using the following PCR program: 95°C for 1 min, 2 cycles of 96°C for 10 sec and 62°C for 90 sec, 29 cycles of 96°C for 10 sec and 59°C for 90 sec, 60°C for 5 min and a hold at 4°C. Amplification products were stored at 4°C until needed.

2.7. Capillary Electrophoresis and Data Analysis

2.7.1 AmpFLSTR™ IdentiFiler™ Plus

Electrophoresis was performed on all samples amplified using the AmpFLSTR™ IdentiFiler™ Plus Amplification Kit (dye set G5: 6-FAM™, blue; VIC™, green; NED™, yellow; PET®, red; and LIZ™, orange). A MicroAmp™ optical 96-well reaction plate (Applied Biosystems™ by ThermoFisher, Foster City, CA) was prepared using 9.5 μL of Hi-Di Formamide (Applied Biosystems™ by ThermoFisher, Foster City, CA) and 0.5 μL of GeneScan™ 500 LIZ™ dye Size Standard v2.0 (Applied Biosystems™ by ThermoFisher, Foster City, CA) per well. One microliter of the amplified samples was added to the corresponding wells, along with one well containing 1 μL of AmpFLSTR™ IdentiFiler™ Plus Amplification Kit allelic ladder (Applied Biosystems™ by ThermoFisher, Foster City, CA). The plate was then covered with a plate septa and spun down on a Labnet MPS 1000 vertical mini plate spinner (Stellar Scientific, Baltimore,

MD) at 2,500 rpm. Samples were injected through a 4-capillary 36 cm array under the run module HIDFragementAnalysis36_POP4_1 (15 sec injection, 3kV, 1,500 sec runtime) on a 3130 Genetic Analyzer (Applied Biosystems™ by ThermoFisher, Foster City, CA; Data Collection v3.0) or injected through a 8-capillary 36 cm array using POP-4 polymer under run module HID36_POP4 (15 sec injection, 1.2 kV, 1,500 sec runtime) 3500 Genetic Analyzer (Applied Biosystems™ by ThermoFisher, Foster City, CA; Data Collection v3.1). The data was then analyzed using the GeneMapper™ ID-X Software v1.6.

2.7.2 GlobalFiler™

Electrophoresis was performed on all samples amplified using the a GlobalFiler™ PCR Amplification Kit (dye set J6: 6-FAM™, blue; VIC™, green; NED™, yellow; TAZ™, red; SID™, purple; and LIZ™, orange). A MicroAmp™ optical 96-well reaction plate was prepared using 9.6 μL of Hi-Di Formamide and 0.4 μL of GeneScan™ 600 LIZ™ dye Size Standard v2.0 (Applied Biosystems™ by ThermoFisher, Foster City, CA) per well. Then 1 μL of the amplified samples was added to the corresponding wells, along as a well with a GlobalFiler™ PCR Amplification Kit allelic ladder. The plate was then covered with a plate septa and spun down on a Labnet MPS 1000 vertical mini plate spinner at 2,500 rpm. Samples were injected through a 4-capillary 36 cm array under the run module HIDFragementAnalysis36_POP4_1 (15 sec injection, 3kV, 1,500 sec runtime) on a 3130 Genetic Analyzer or injected through a 8-capillary 36 cm array using POP-4 polymer under run module HID36_POP4 (15 sec injection, 1.2 kV, 1,500 sec runtime) 3500 Genetic Analyzer. The data was then analyzed using the GeneMapper™ ID-X Software v1.6.

2.7.3. PowerPlex® Fusion 6C

Electrophoresis was performed on all samples amplified using the a PowerPlex® Fusion 6C Amplification Kit (Dye Set Promega J6: FL-6C, blue; JOE-6C, green; TMR-6C, yellow; CXR-6C, red; TOM-6C, purple; WEN-6C, orange). A MicroAmp™ optical 96-well reaction plate was prepared using 9.5 μL of Hi-Di Formamide and 0.5 μL of WEN ILS 500 Internal Lane Standard (Promega Corporation, Madison, WI) per well. Then 1 μL of the amplified samples was added to the corresponding wells, along with a well containing 1 μL PowerPlex® Fusion 6C Amplification Kit allelic ladder. The plate was then covered with a plate septa and spun down on a Labnet MPS 1000 vertical mini plate spinner at 2,500 rpm. Samples were injected through an 8-capillary 36 cm array using POP-4® polymer under the run module HID36_POP4 (15 sec injection, 1.2kV, 1,500 sec runtime) using a 3500 Genetic Analyzer. The signals were then analyzed using the GeneMapper™ ID-X Software v1.6.

2.7.4 VeriFiler™ Plus

Electrophoresis was performed on all samples amplified using the a VeriFiler™ Plus Amplification (Dye Set J6-T: 6-FAM™, blue; VIC™, green; TED™, yellow; TAZ™, red; SID™, purple; and LIZ™, orange). A MicroAmp™ optical 96-well reaction plate was prepared using 9.6 μL of Hi-Di Formamide and 0.4 μL of GeneScan™ 600 LIZ™ dye Size Standard v2.0 per well. Then 1 μL of the amplified samples was added to the corresponding wells, along with a well containing 1 μL of VeriFiler™ Plus Amplification Kit allelic ladder. The plate was then covered with a plate septa and spun down on a Labnet MPS 1000 vertical mini plate spinner at 2,500 rpm. Samples were injected through a 8-capillary 36 cm array using POP-4 polymer under

the run module HID36_POP4 (15 sec injection, 1.2KV, 1,500 sec runtime) using a 3500 Genetic Analyzer. The signals were then analyzed using the GeneMapper™ ID-X software v1.6.

CHAPTER 3: MICROMANIPULATION RESULTS

3.1. Adhesive Size Testing

Before testing sperm recovery using micromanipulation, evaluation of the adhesive used to recover the cells on our downstream lysis and amplification reaction was performed. The proposed workflow of this project separated the lysis and amplification steps, with collection occurring directly into the reaction mix for either the forensicGEM Sperm or LysePrep lysis method. Both methods utilize small lysis volumes (1.5 and 2 μL, respectively) therefore testing was done to ensure that the adhesive used to pick up cells during micromanipulation would not affect the overall volume and potentially cause downstream biochemical effects to the reaction when dissolved. Two different sizes of adhesive were dissolved into water of corresponding volumes to that of the lysis volumes used, then the volume of the solution was measured to determine if there was a significant difference. The two adhesive sizes used were “small” (Figure 5), in which the tip of the tungsten needle was lightly coated in adhesive, and “large” (Figure 5) in which a ball of adhesive was attached to the end of the needle.

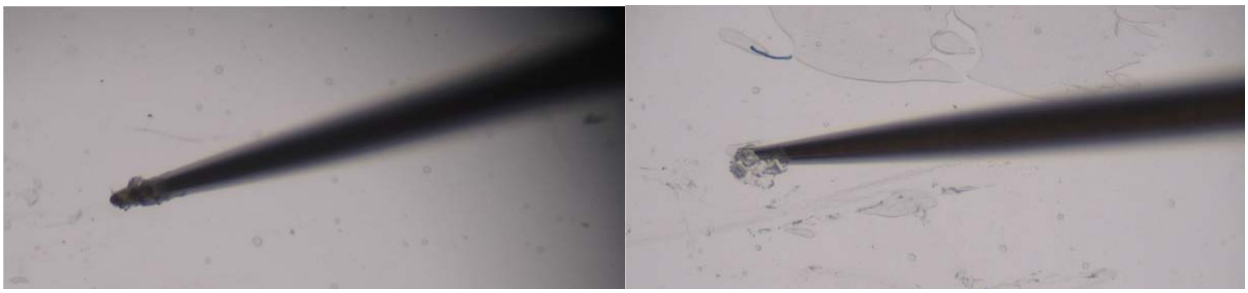

Figure 5: Microscope image of “Small Adhesive” (left) and “Large Adhesive” (right).
A tungsten needle lightly coated in 3M™ adhesive to produce the “small adhesive” tested, and a tungsten needle attached to a ball of 3M™ adhesive to produce the “large adhesive” tested.

The small adhesive size was used to pick up 1- and 2-cells while large adhesive size was used to collect larger groups of cells. After the experiment was completed, the percent difference in the volumes before and after addition of the adhesive was calculated (Table 1). The results showed

that there was no significant increase in the volume of the lysis mixture after dissolution of adhesive, regardless of lysis mixture volume or adhesive volume, as the largest change in volume was only 1.03%. These results ensured confidence moving forward that the amount of adhesive used in the collection of cells would not significantly impact the volume of the lysis mix and thus potentially affect the lysis reaction.

Table 1: Adhesive Size Testing.
Small and large sizes of adhesive were dissolved in 1.5 and 2 µL volumes, and the final volumes were measured. The volume difference was calculated as well as the percent difference. For each replicate n=10.

Lysis Volume	Adhesive Size	Volume Difference	Percent Difference
1.5 µL	Small	0.016 µL	1.03 %
1.5 µL	Large	0.004 µL	0.25 %
2.0 µL	Small	0.004 µL	0.19 %
2.0 µL	Large	0.008 µL	0.39 %

3.2. Method Evaluation

To evaluate the capacity to recover aSTR profiles from single and few sperm cells, two lysis methods, forensicGEM Sperm and LysePrep, along with four autosomal amplification kits (AmpFLSTR™ IdentiFiler™ Plus, Globalfiler™ Plus, PowerPlex® Fusion 6C, and VeriFiler™ Plus) were evaluated. Each lysis method uses different proprietary enzymes as well as different cycling methods to extract the genetic material. The aSTR amplification kits each contain all core CODIS loci, but then also may contain additional different aSTR loci for increased discrimination. Each lysis method and amplification kit were tested using four donors (one collection per donor). Each collection consisted of five 1-cell samples, and two 2-, 5-,10-, and 20-cell samples. Additionally, two 0-cell samples were included with each collection as a negative control to ensure that aSTR results were originating from collected cells and not due to

possible “cell-free” DNA or other contamination. Each amplification consisted of an amplification positive control with 0.0125 ng of control DNA supplied with each kit and an amplification negative control consisting of water or Low-TE buffer, per the kit’s instructions. After successful lysis, amplification, and capillary electrophoresis, the resulting DNA profiles were evaluated. The accuracy of any obtained DNA profile was verified using a donor reference profile, obtained using normal extraction methods and the suggested DNA input per kit. Profiles were evaluated in terms of number of alleles recovered per sample, only including alleles that were attributable to the donor. Homozygous alleles were counted twice to represent both copies of the detected allele if above 500 RFUs. If the RFU value was below 500 allele drop-out could be possible and therefore only one copy of the allele was counted. Percent recovery of each profile was calculated by dividing the number of autosomal alleles observed per profile by the total number of autosomal alleles available for recovery. For one cell samples, only one allele is available for recovery at each locus as sperm cells are haplotype cells. Therefore, when calculating the percent recovery of 1-cell samples, the total number of alleles available for recovery is halved.

3.2.1. AmpFLSTR™ IdentiFiler™ Plus Results

The first kit used to evaluate the micromanipulation technique with was AmpFLSTR™ IdentiFiler™ Plus kit, which amplifies 15 autosomal loci and Amelogenin, the sex determination loci, allowing for a maximum of 30 alleles to be recovered. This was the first kit used to evaluate sperm recovery as it had been used in previous experiments to successfully amplify single epithelial cells, therefore acting as a control to ensure successful collection and lysis of the sperm cells (Table 2, Figure 6).

Table 2: Allele Recovery Data for AmpFLSTR™ IdentiFiler™ Plus
The average number of alleles recovered and the average percent recovery per cell number and lysis kit using AmpFLSTR™ IdentiFiler™ Plus.

Cell Number	n	Average Alleles Recovered		Average Percent Recovery	
		Forensic Gem Sperm	DEPArray LysePrep	Forensic Gem Sperm	DEPArray LysePrep
1-Cell	5	1	5	7	33
2-Cell	2	2	8	7	27
5-Cell	2	9	14	30	47
10-Cell	2	8	17	27	57
20-Cell	2	12	25	40	83

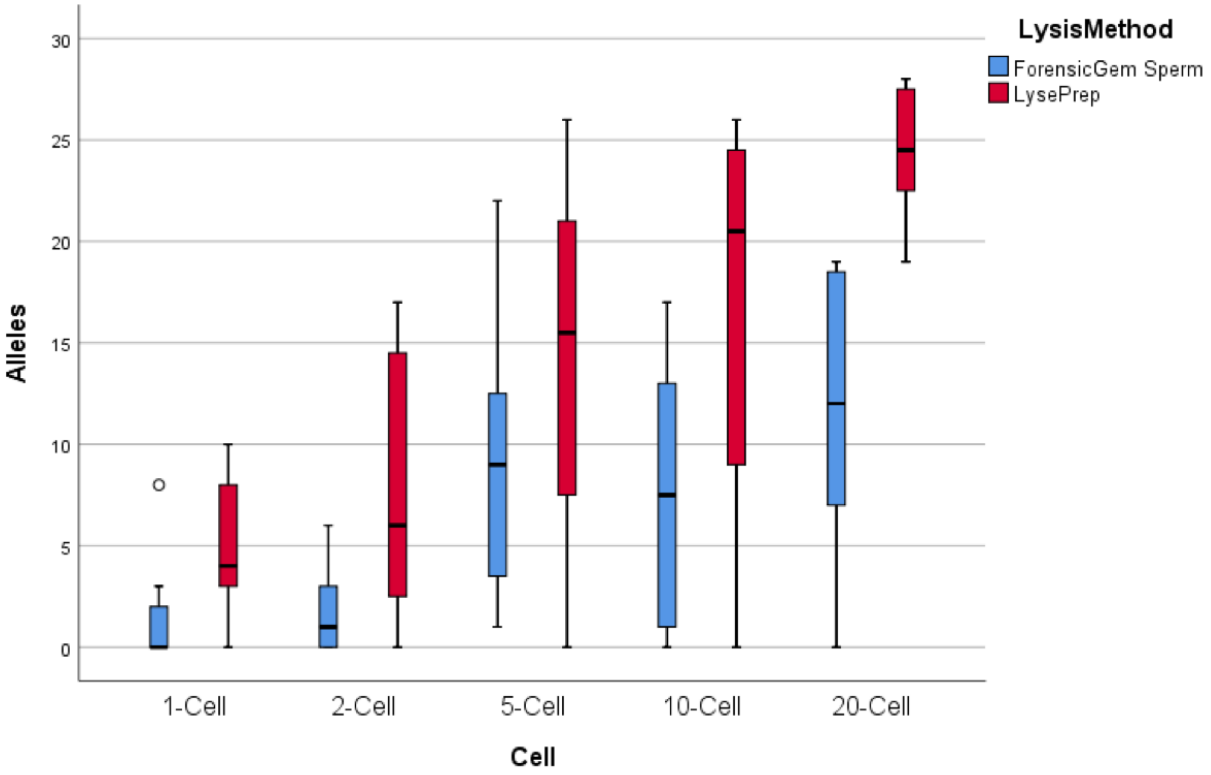

Figure 6: Boxplot of AmpFLSTR™ IdentiFiler™ Plus Allele Recovery Data.
The boxplot shows the range of alleles recovered for each sperm cell using AmpFLSTR™ IdentiFiler™ Plus, lysed with either forensicGEM Sperm (blue) or LysePrep (red). Y-axis: number of alleles observed; x-axis: number of sperm cells collected. Error bars are shown in black, outliers indicated by a black circle, extremes indicated by a star. For 1-cell collections, n=5. For 2-, 5-, 10-, and 20-cell collections n=2.

With 1-cell samples, 10 alleles were recovered (out of 15 possible; 67%) using LysePrep (Figure 7). This profile resulted in several instances of locus drop-out, where no alleles were

recovered, however this is to be expected when amplifying samples of such low quantity DNA. When comparing the allele recovery of 1-cell samples using the two lysis methods, a significant amount of variability was observed, with an average allele recovery of 5 (out of 15 possible) for LysePrep as compared to 1 (out of 15 possible) with forensicGEM Sperm. This difference is quite significant, suggesting LysePrep to be superior in recovering genetic information from LCN samples. When considering the 2-cell samples, there was an increase in the maximum number of alleles recovered with both lysis methods, obtaining 17 alleles (out of 30 possible; 57%) using LysePrep and 6 alleles (out of 30 possible; 20%) using forensicGEM Sperm. The average allele recovery for both methods also increased with the 2-cell samples, with an average recovery of 8 alleles (out of 30 possible; 27%) using LysePrep and 2 alleles (out of 30 possible; 7%) using forensicGEM Sperm. This is to be expected, as when two cells are collected there is a possibility of increasing the amount of genetic information with haplotype cells, as both the maternal and paternal chromosomal locus of DNA can be recovered. Even if both cells have the same parental chromosomal locus of DNA, the starting material is also increased prior to PCR which should result in an increased allele recovery. With the 2-cell samples, the LysePrep method was also superior to forensicGEM Sperm. This trend continues when more cells are collected, such as with 5-cell samples, which were able to obtain an allele recovery up to 26 alleles (out of 30 possible; 87%) using LysePrep, almost a full profile. The average allele recovery also increased with the 5-cell samples, as 14 alleles (out of 30 possible; 47%) were recovered using LysePrep and 9 alleles (out of 30 possible; 30%) were recovered using forensicGEM Sperm. When increasing the cell collection size from 5 to 10 cells, the maximum recovery of alleles did not change, with 26 alleles (out of 30 possible; 87%) recovered using LysePrep. However, the average recovery per profile did increase in the 10-cell samples when

using LysePrep, with an average allele recovery of 17 alleles (out of 30 possible: 57%) but decreased with forensicGEM Sperm, which recovered an average of 8 alleles (out of 30 possible: 27%). When increasing the collection size to 20-cells, the maximum number of alleles recovered increased to 28 alleles (out of 30 possible; 93%) using LysePrep (Figure 7), recovering information at all loci, and observing only two instances of allele drop-out (ADO). With 20-cell samples, the average allele recovery also increased, recovering 25 alleles (out of 30 possible; 83%) using LysePrep and 12 alleles (out of 30 possible; 40%) using forensicGEM Sperm. For all cell samples, LysePrep was observed to significantly recover more alleles per profile than forensicGEM Sperm, with LysePrep recovering the same number of alleles per profile as forensicGEM Sperm using 5-cells opposed to the 20-cells necessary for forensicGEM Sperm. The results obtained using AmpFLSTR™ IdentiFiler™ Plus indicated that the use of micromanipulation to isolate sperm cells for downstream autosomal STR analysis was indeed possible as there was significant recovery seen in the larger cell numbers. With these successful results, the next step was to evaluate additional aSTR kits as the AmpFLSTR™ IdentiFiler™ Plus kit does not contain all current CODIS loci and therefore would no longer be used by operational laboratories.

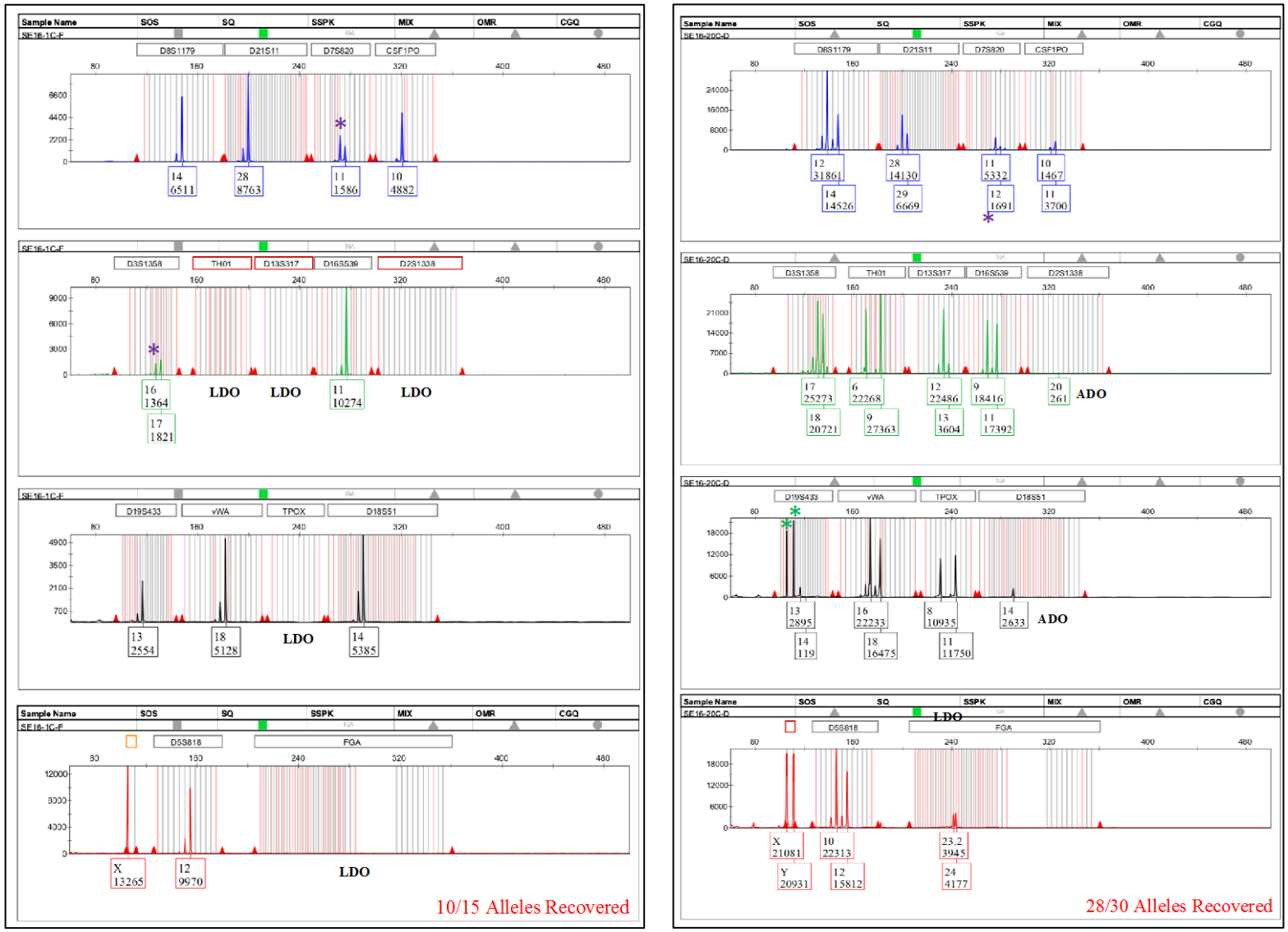

Figure 7: DNA Profile of 1-Cell sample (left) and 20-Cell sample (right) using AmpFLSTR™ IdentiFiler™ Plus and LysePrep
Profiles generated using a 3500 genetic analyzer. LDO indicates locus drop out. ADO indicates allelic drop out. Green star indicates pull up. Purple star indicates increased stutter. Allele numbers and relative fluorescent units are listed below each allele. X-axis represents size in base pair (bp); y-axis represents peak height (RFU).

3.2.2. GlobalFiler™ Results

The next amplification kit evaluated was GlobalFiler™ amplification kit (Table 3, Figure 8), which amplifies 21 autosomal loci, a Y-Indel and Amelogenin.

Table 3: Allele Recovery Data for GlobalFiler™
The average number of alleles recovered and the average percent recovery per cell number and lysis kit using Globalfiler™.

Cell Number	n	Average Alleles Recovered		Average Percent Recovery	
		Forensic Gem Sperm	DEPArray LysePrep	Forensic Gem Sperm	DEPArray LysePrep
1-Cell	5	2	7	10	33
2-Cell	2	1	13	2	31
5-Cell	2	5	17	12	40
10-Cell	2	7	19	17	45
20-Cell	2	12	20	29	48

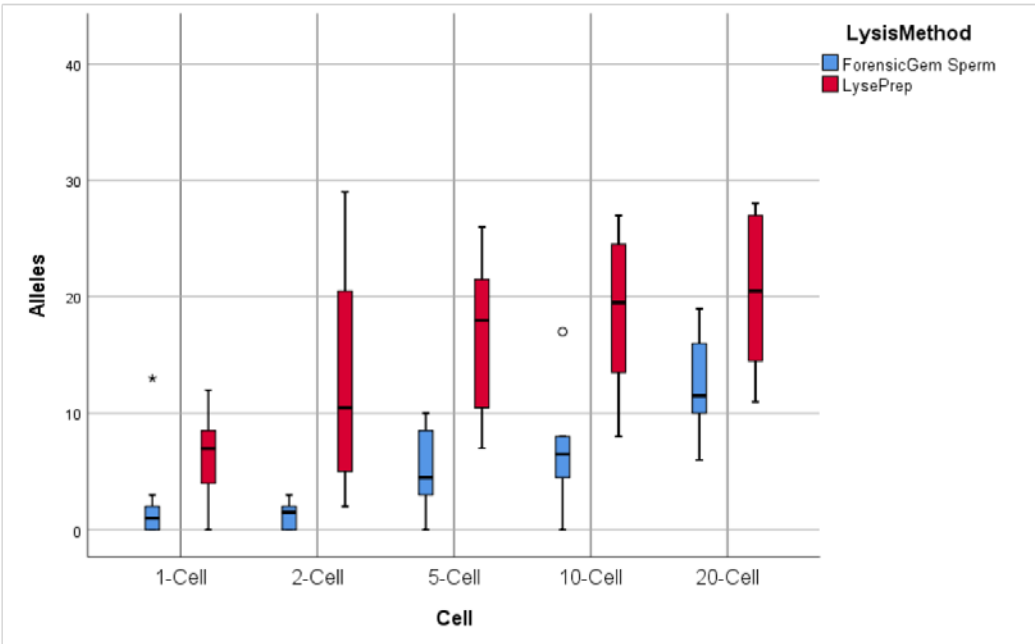

Figure 8: Boxplot of GlobalFiler™ Allele Recovery Data
The boxplot shows range of alleles recovered for each sperm cell sample using Globalfiler™ amplification kit, lysed with forensicGEM Sperm (blue) or LysePrep (red). Y-axis: number of alleles observed; x-axis: number of sperm cells collected. Error bars are shown in black, outliers indicated by a black circle, extremes indicated by a star. For 1-cell collections, n=5. For 2-, 5-, 10-, and 20-cell collections n=2.

With GlobalFiler™, the number of alleles recovered using 1-cell increased, obtaining 11 alleles (out of 21 possible; 52%) using LysePrep (Figure 9) and 13 alleles (out of 21 possible; 61%) using forensicGEM Sperm. While forensicGEM Sperm did recover more alleles in one profile than LysePrep, the latter method still showed an improvement over the former in average allele recovery, obtaining 7 alleles (out of 21 possible; 33%) using LysePrep and 2 (out of 21 possible; 9%) using forensicGEM Sperm, a significant difference. Therefore, LysePrep was still considered the superior lysis method for 1-cell samples with GlobalFiler™. With 2-cells, a higher allele recovery was obtained with the GlobalFiler™ kit, recovering 29 alleles (out of 42 possible; 69%) using LysePrep compared to a maximum recovery of only 17 alleles (out of 30 possible) using AmpFLSTR™ IdentiFiler™ Plus. The average number of alleles recovered using 2-cells also increased with GlobalFiler™ when using LysePrep, obtaining an average of 13 alleles (out of 42 possible; 31%) per sample compared to that of AmpFLSTR™ IdentiFiler™ Plus, which obtained an average of 8 alleles (out of 30 possible). However, the average allele recovery decreased for 2-cell samples when using the forensicGEM Sperm method, recovering an average of 1 allele (out of 42 possible; 2%) with GlobalFiler™, compared to an average of 2 alleles (out of 30 possible) using AmpFLSTR™ IdentiFiler™ Plus. The maximum recovery of alleles using 5-cells did not increase with GlobalFiler™, recovering 26 alleles (out of 42 possible; 62%) when using LysePrep, the same number of alleles achieved when using the same lysis method and AmpFLSTR™ IdentiFiler™ Plus. However, the average number of alleles recovered per 5-cell sample was increased when using the LysePrep method and GlobalFiler™, obtaining 17 (out of 42 possible; 40%) using LysePrep compared to only 14 alleles (out of 30 possible) obtained using the same lysis method and AmpFLSTR™ IdentiFiler™ Plus. The average number of alleles recovered from 5-cell samples decreased when using the forensicGEM

Sperm method and GlobalFiler™, recovering an average of 5 alleles (out of 42 possible; 12%), opposed an average of 9 alleles (out of 30 possible) recovered with AmpFLSTR™ IdentiFiler™ Plus. When using LysePrep and GlobalFiler™, the maximum allele recovery increased with 10-cells, obtaining 27 alleles (out of 42 possible; 64%), an increase in the number of alleles recovered from our 5-cell samples using the same lysis method and amplification kit, as well as an increase in the number of alleles recovered using 10-cells with AmpFLSTR™ IdentiFiler™ Plus. The average number of alleles obtained for 10-cells increased using LysePrep and GlobalFiler™, with an average allele recovery of 19 alleles (out of 42 possible; 48%) compared to an average allele recovery of 17 alleles (out of 30 possible) using AmpFLSTR™ IdentiFiler™ Plus. As seen with the other cell numbers, the average allele recovery for the 10-cell samples decreased in comparison to AmpFLSTR™ IdentiFiler™ Plus, recovering an average of 7 alleles (out of 42 possible; 16%) using GlobalFiler™ and 8 alleles (out of 30 possible) using AmpFLSTR™ IdentiFiler™ Plus. For 20 cells, the maximum number of alleles recovered did not increase when using Globalfiler, recovering 28 alleles (out of 42 possible; 93%) using LysePrep (Figure 9), which was achieved using AmpFLSTR™ IdentiFiler™ Plus. Additionally, the average of number of alleles recovered from 20-cells decreased when using LysePrep and GlobalFiler™, recovering an average of 20 alleles (out of 42 possible; 48%) compared to an average of 25 alleles (out of 30 possible) using AmpFLSTR™ IdentiFiler™ Plus. With forensicGEM Sperm, the average allele recovery from 20-cells did not change, recovering an average of 12 alleles (out of 42 possible; 29%), the same as AmpFLSTR™ IdentiFiler™ Plus. Some improvement in allele recovery was observed for GlobalFiler™ in the lower cell numbers, increasing the number of alleles recovered overall for both 1- and 2-cell samples. However, overall allele recovery did not improve for the larger cell numbers. Additionally, fewer alleles

were observed with forensicGEM sperm using GlobalFiler™ than with AmpFLSTR™ IdentiFiler™ Plus. Therefore, other aSTR kits were evaluated.

Figure 9: DNA Profile of a 1-cell sample (left) and 20-cell sample (right) using GlobalFiler™ and LysePrep
Profiles generated using a 3130 genetic analyzer (left) and 3500 genetic analyzer (right). LDO indicates locus drop out. ADO indicates allelic drop out. Green star indicates pull up. Purple star indicates increased stutter. Allele numbers and relative fluorescent units are listed below each allele. X-axis represents size (bp); y-axis represents peak height (RFU).

3.2.3. PowerPlex® Fusion 6C Results

The next amplification kit tested was PowerPlex® Fusion 6C (Table 4, Figure 10), which amplifies 46 autosomal loci, 3 Y-chromosome loci, as well as a Y-Indel and Amelogenin.

Table 4: Allele Recovery Data for PowerPlex® Fusion 6C Results
The average number of alleles recovered and the average percent recovery per cell number and lysis kit using PowerPlex® Fusion 6C.

Cell Number	n	Average Alleles Recovered		Average Percent Recovery	
		Forensic Gem Sperm	DEPArray LysePrep	Forensic Gem Sperm	DEPArray LysePrep
1-Cell	5	9	8	39	35
2-Cell	2	4	12	9	26
5-Cell	2	11	23	24	50
10-Cell	2	15	20	32	43
20-Cell	2	25	30	54	65

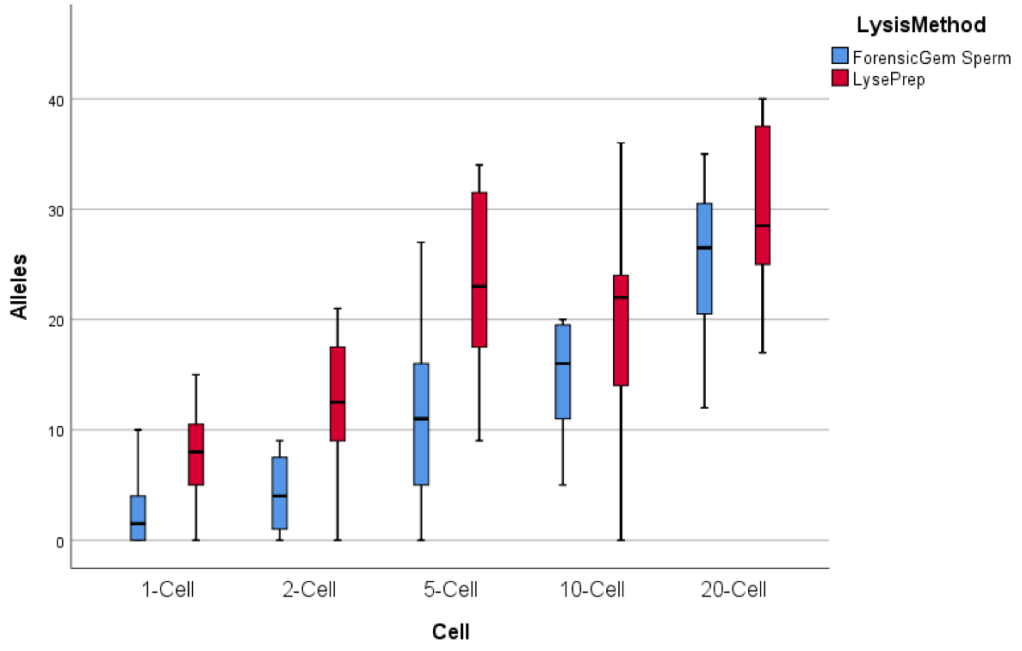

Figure 10: Boxplot of PowerPlex® Fusion 6C Allele Recovery Data
The boxplot shows the range of alleles recovered for each sperm sample using PowerPlex® Fusion 6C, lysed with forensicGEM Sperm (blue) or LysePrep (red). Y-axis: number of alleles observed; x-axis: number of sperm cells collected. Error bars are shown in black, outliers indicated by a black circle, extremes indicated by a star. For 1-cell collections, n=5. For 2-, 5-, 10-, and 20-cell collections n=2.

With PowerPlex® Fusion 6C, a higher number of alleles were recovered with 1-cell, recovering 15 alleles (out of 23 possible; 65%) using LysePrep (Figure 11), an improvement over the two other aSTR kits evaluated. Additionally, the average allele recovery for 1-cell improved over the other a STR kits with PowerPlex® Fusion 6C, recovering an average of 8 alleles (out of 23 possible; 35%) using LysePrep and 9 alleles (out of 23 possible; 39%) using forensicGEM Sperm. When increasing the sample size to 2-cells, the maximum number of alleles recovered did not increase, recovering 21 alleles (out of 46 possible; 46%) using LysePrep, compared to a maximum allele recovery of 29 alleles (out of 42 possible) using GlobalFiler™ and LysePrep. Additionally, the average number of alleles recovered for 2-cells did not increase with PowerPlex® Fusion 6C, recovering an average of 12 alleles (out of 46 possible; 26%) using LysePrep compared to GlobalFiler™, which recovered an average of 13 alleles (out of 42 possible) with the same lysis method. However, the average recovery for 2-cells was improved in comparison to the other aSTR kits when using forensicGEM Sperm, recovering an average of 4 alleles (out of 46 possible; 9%) using PowerPlex® Fusion 6C, compared to 1 allele (out of 42 possible) achieved with GF and 2 alleles (out of 30 possible) with AmpFLSTR™ IdentiFiler™ Plus. PowerPlex® Fusion 6C did show an increase in the maximum alleles recovered using 5-cells, obtaining 34 alleles (out of 46 possible; 74%) using LysePrep, an improvement over the previous two aSTR kits examined at any cell number. Additionally, the average allele recovery for 5-cells was increased when using both lysis methods and PowerPlex® Fusion 6C, recovering an average of 23 alleles (out of 46 possible; 50%) using LysePrep and 11 alleles (out of 46 possible; 24%) using forensicGEM Sperm. This trend continued for the 10-cell samples, recovering a total of 36 alleles (out of 46 possible; 78%) using LysePrep and PowerPlex® Fusion 6C, further improving the total allele recovery obtained using

micromanipulation. The average alleles recovered for 10-cells also increased for both lysis methods using PowerPlex® Fusion 6C, recovering an average of 20 alleles (out of 46 possible; 44%) using LysePrep and 15 alleles (out of 46 possible; 33%) using forensicGEM Sperm. For 20-cells, the maximum alleles recovered increased to 40 alleles (out of 46 possible; 87%) using LysePrep and PowerPlex® Fusion 6C (Figure 11) as well as the average allele recovery, obtaining 30 alleles (out of 46 possible; 65%) with LysePrep and 25 alleles (out of 46 possible; 54%) using forensicGEM Sperm. For this amplification kit, the LysePrep lysis method showed a significant improvement in the recovery of alleles achieved at each cell number when compared to forensicGEM sperm. PowerPlex® Fusion 6C showed a significant improvement over AmpFLSTR™ IdentiFiler™ Plus and Globalfiler™ in allele recovery at 5-, 10- and 20-cells, ultimately improving the total number of alleles recovered using micromanipulation. Additionally, this amplification kit showed an improvement in the total allele recovery for 1-cell samples as well. While these results were indeed promising, an additional aSTR kit was tested to determine if increased allele recovery could be achieved from our micromanipulated samples.

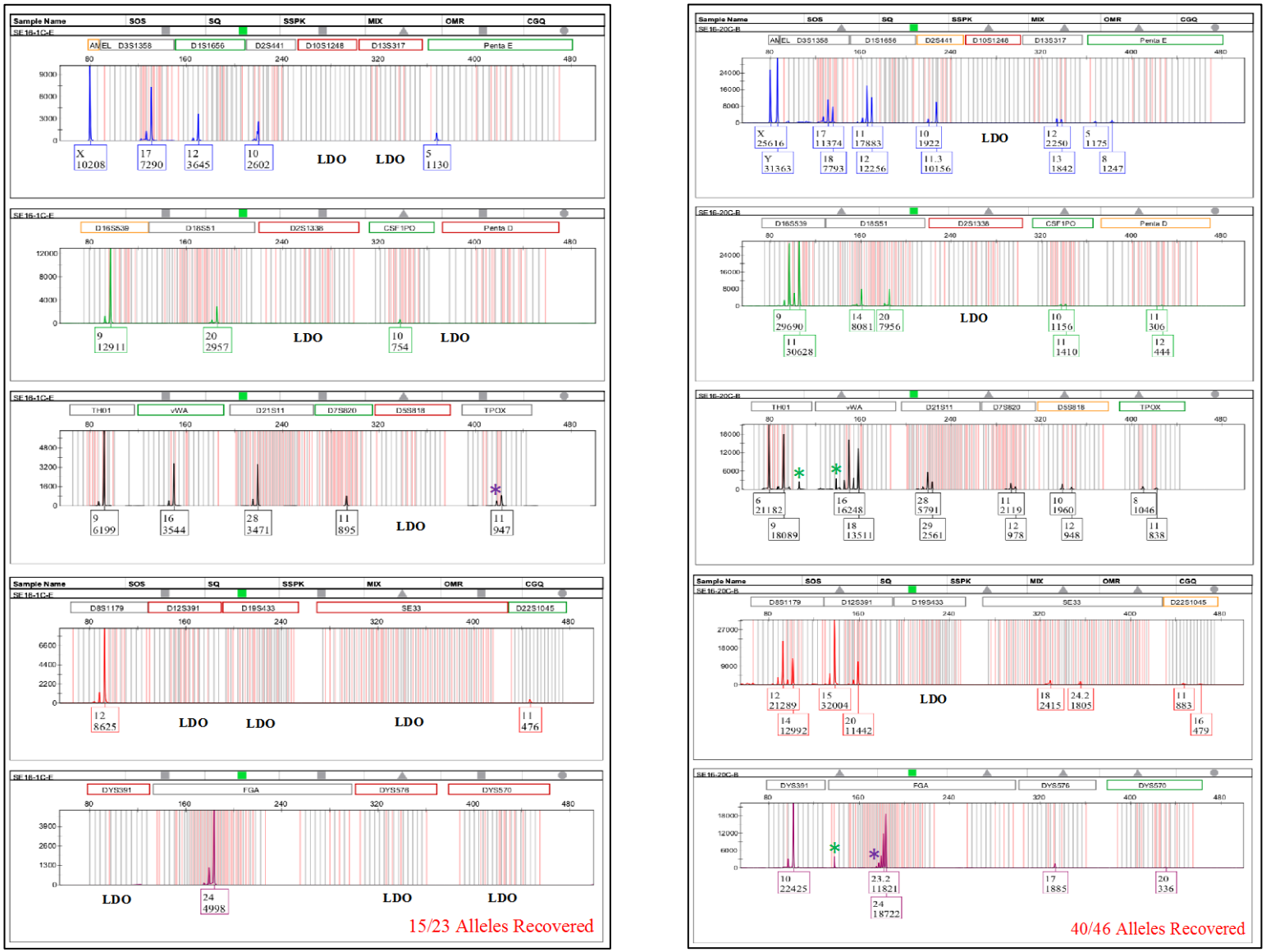

Figure 11: DNA Profile of a 1-cell sample (left) and 20-cell sample (right) using PowerPlex® Fusion 6C and LysePrep

Profiles generated using a 3500 genetic analyzer. LDO indicates locus drop out. ADO indicates allelic drop out. Green star indicates pull up. Purple star indicates increased stutter. Allele numbers and relative fluorescent units are listed below each allele. X-axis represents size (bp); y-axis represents peak height (RFU).

3.2.4. VeriFiler™ Plus Results

The last amplification kit tested was VeriFiler™ Plus (Table 5, Figure 12). This kit amplifies 46 autosomal loci, as well as a Y-Indel and Amelogenin. VeriFiler™ Plus also includes inhibition markers for each dye channel, which is a helpful tool for distinguishing between possible effects from a degraded DNA sample versus PCR inhibition, the latter of which can be resolved typically using various analysis strategies.

Table 5: Allele Recovery Data for VeriFiler™ Plus
The average number of alleles recovered and the average percent recovery per cell number and lysis kit using Verifiler™ Plus.

Cell Number	n	Average Alleles Recovered		Average Percent Recovery	
		Forensic Gem Sperm	DEPArray LysePrep	Forensic Gem Sperm	DEPArray LysePrep
1-Cell	5	4	9	17	39
2-Cell	2	5	13	11	28
5-Cell	2	14	25	30	54
10-Cell	2	16	27	35	59
20-Cell	2	21	29	46	63

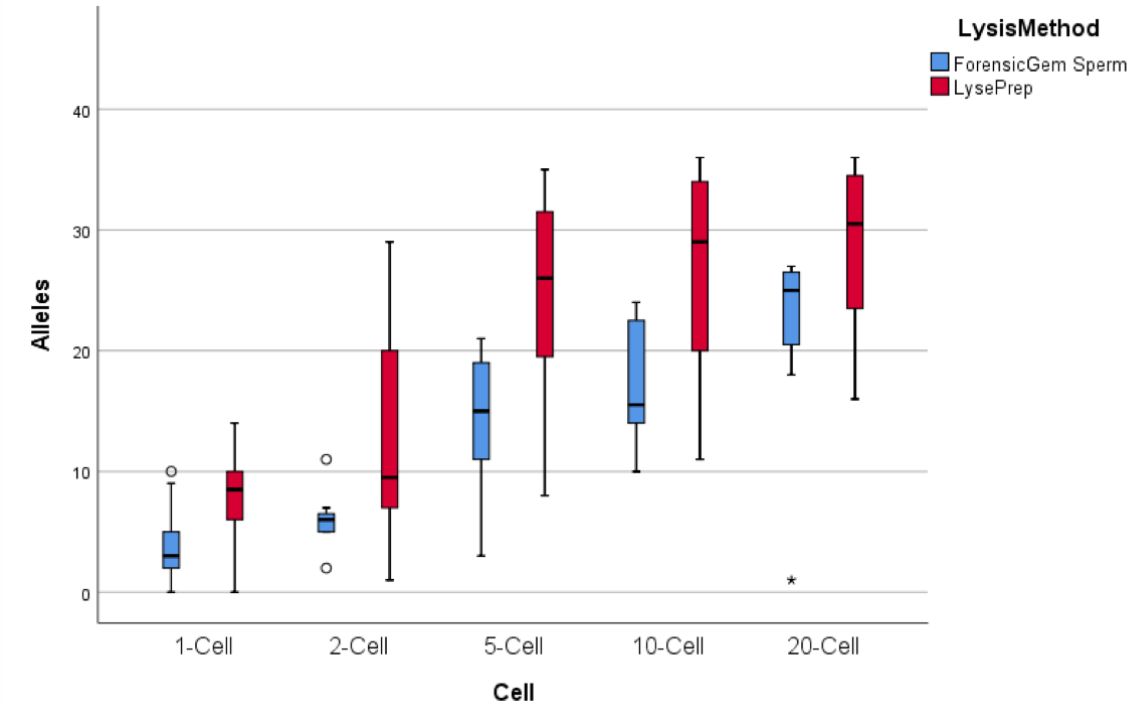

Figure 12: Boxplot of VeriFiler™ Plus Allele Recovery Data
The boxplot shows the range of alleles recovered for each sperm cell sample using VeriFiler™ Plus, lysed with forensicGEM Sperm (blue) or LysePrep (red). Y-axis: number of alleles observed; x-axis: number of sperm cells collected. Error bars are shown in black, outliers indicated by a black circle; extremes indicated by a star. For 1-cell collections, n=5. For 2-, 5-, 10-, and 20-cell collections n=2.

Using VeriFiler™ Plus, the maximum number of alleles recovered using a 1-cell sample did not increase, recovering 14 alleles (out of 23 possible; 61%) using the LysePrep (Figure 13), compared to the 15 alleles (out of 46 possible) recovered using PowerPlex® Fusion 6C and the same lysis method. The average number of alleles recovered for 1-cell did increase with VeriFiler™ Plus, recovering an average allele recovery of 9 alleles (out of 23 possible; 39%) using LysePrep and 4 alleles (out of 23 possible; 17%) using forensicGEM Sperm, an improvement over the other aSTR kits tested. The maximum allele recovery using 2-cells did not increase with VeriFiler™ Plus, which recovered 29 alleles (out of 46 possible; 63%) using LysePrep, which was achieved using GlobalFiler™ and LysePrep. However, again the average allele recovery did increase when using VeriFiler™ Plus, recovering an average of 13 alleles using LysePrep (out of 46 possible; 28%) and 5 alleles (out of 46 possible; 11%) using

forensicGEM Sperm for the 2-cell samples, an improvement over all other aSTR kits tested. This trend continued with 5-cells – while VeriFiler™ Plus did not show an improvement in the maximum number of alleles recovered for the cell numbers, an improvement in the average allele recovery was observed, with VeriFiler™ Plus recovering an average of 25 alleles (out of 46 possible; 54%) using LysePrep and 14 alleles (out of 46 possible; 30%) using forensicGEM Sperm was achieved. This was seen again in the 10-cell samples, with VeriFiler™ Plus recovering a maximum of 36 alleles (out of 46 possible; 78%) using LysePrep, the same that was achieved with PowerPlex® Fusion 6C, however an increased average allele recovery of 27 alleles (out of 46 possible; 59%) using LysePrep and 16 alleles (out of 46 possible; 35%) using forensicGEM Sperm was observed, an improvement over the other aSTR kits tested. Finally, this trend was seen as well in the 20-cell samples, with VeriFiler™ Plus recovering a maximum of 36 alleles (out of 46 possible; 78%) was achieved using LysePrep (Figure 13), which was less than that recovered using PowerPlex® Fusion 6C and LysePrep, but improved the average allele recovery for 20-cells, recovering an average of 29 alleles (out of 46 possible; 63%) with LysePrep and 21 alleles (out of 46 possible; 46%) using forensicGEM Sperm. When comparing the lysis methods used with this amplification kit, LysePrep was found to recover significantly more alleles per profile than forensicGEM Sperm at every cell number. When comparing amplification kits, VeriFiler™ Plus recovered more alleles than both AmpFLSTR™ IdentiFiler™ Plus and GlobalFiler™ and performed similarly to PowerPlex® Fusion 6C. These kits will be further compared in the next section.

Figure 13: DNA Profile of a 1-cell sample (left) and a 10-cell sample (right) using VeriFiler™ Plus and LysePrep
Profiles generated using a 3500 genetic analyzer. LDO indicates locus drop out. ADO indicates allelic drop out. Green star indicates pull up. Purple star indicates increased stutter. Allele numbers and relative fluorescent units are listed below each allele. X-axis represents size (bp) and y-axis represents peak height (RFU).

3.2.5. Method Determination

When comparing the two lysis methods, a higher allele recovery for each cell number at each amplification kit was obtained using LysePrep and therefore was determined to be the superior lysis method. When comparing the amplification kits (Figure 14), more allele recovery was observed with PowerPlex® Fusion 6C and VeriFiler™ Plus compared to Globalfiler™ or AmpFLSTR™ IdentiFiler™ Plus. For both 1-cell and 20-cell samples, PowerPlex® Fusion 6C had a higher maximum number of alleles recovered, but VeriFiler™ Plus had a higher number of alleles on average. After considering the four amplification kits and the different results achieved, the method determined to be most suitable over the others was the VeriFiler™ Plus amplification kit and LysePrep. While a greater number of alleles were observed using PowerPlex® Fusion 6C, VeriFiler™ Plus had a more consistently high recovery of alleles, and had better success using less cells than PowerPlex® Fusion 6C. Therefore, the optimized method for recovery of sperm cells using micromanipulation is direct lysis using LysePrep, followed by autosomal amplification using VeriFiler™ Plus.

.

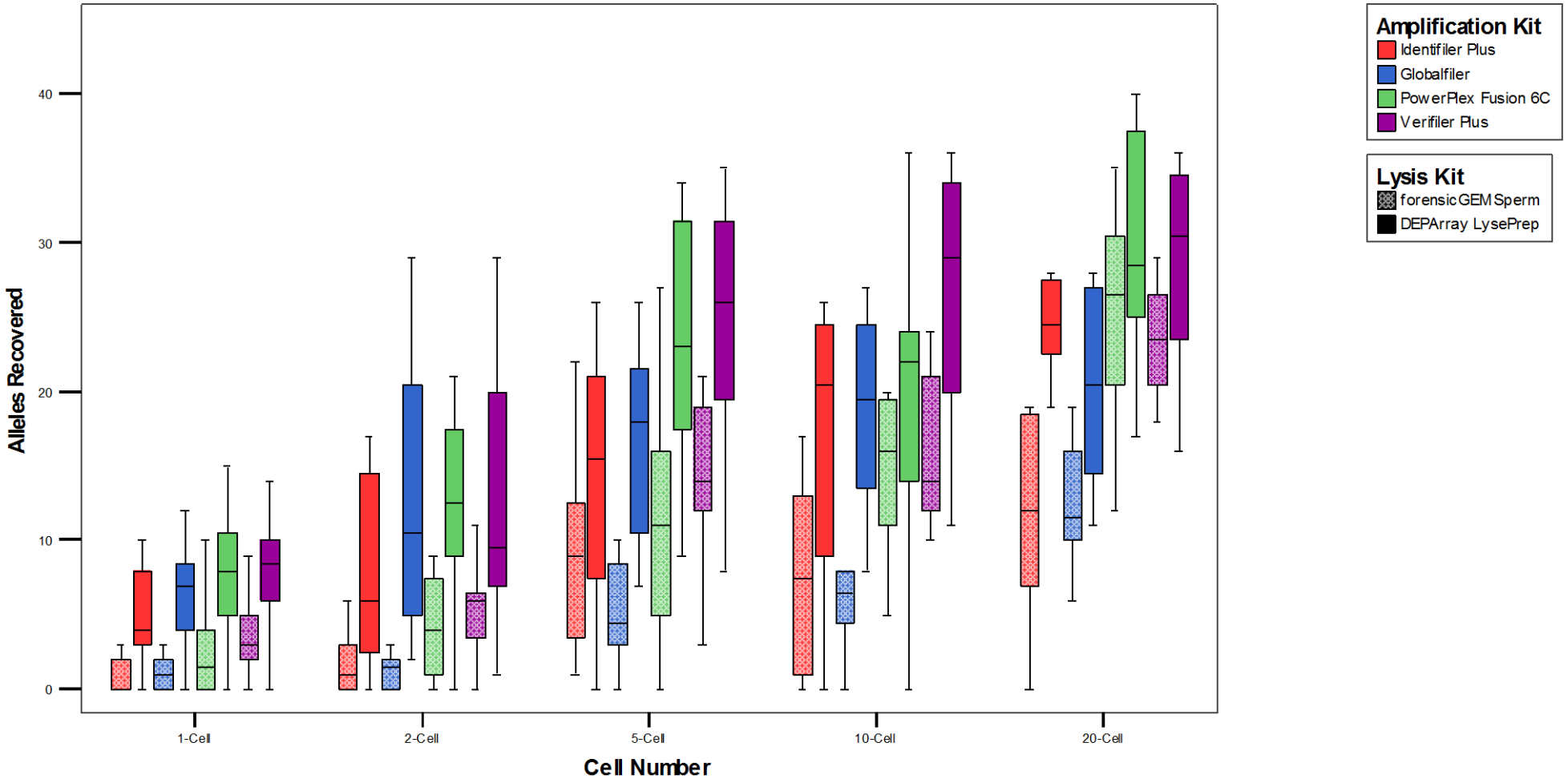

Figure 14: Boxplot of alleles recovered from each lysis and amplification kit.
Figure shows the number of alleles recovered in each profile obtained from using all four amplification kits with either lysis method. Y-axis shows the number of alleles recovered; x-axis lists shows the number of cells collected. Red boxes show results obtained using AmpFLSTR™ IdentiFiler™ Plus, blue boxes show results obtained Globalfiler™, green boxes show results obtained using PowerPlex® Fusion 6C, purple boxes show results obtained using Verifiler™ Plus. Boxes with hash marks indicated samples lysed using forensicGEM Sperm, solid boxes indicate samples lysed with LysePrep. Error bars are shown in black, outliers indicated by a black circle. For 1-cell collections, n=5. For 2-, 5-, 10-, and 20-cell collections n=2.

CHAPTER 4: DETERMINING LIMIT OF DETECTION

Once the optimized method for recovering, lysing and amplifying sperm cells was established, the next steps were twofold. First, the probative value of the DNA profiles obtained using LysePrep and VeriFiler™ Plus was determined, and this information was used to recommend the number of sperm cells needed to routinely recover probative information. Second, the quality of the profiles was analyzed by evaluating the amount of allele drop-in and stutter observed, which could potentially detract from the usefulness of the profiles in a casework situation and thresholds were developed to overcome these effects.

4.1. Determining Probative Value

The probative value of the profiles obtained were evaluated by calculating the RMP of each profile for each donor and comparing it to the RMP of the full donor profile. The RMP was calculated by multiplying the population frequency of each loci. Heterozygote genotype frequency (f) was calculated using f=2pq for each heterozygous allele, with “p” and “q” representing the population frequency of each allele obtained, respectively, and the frequency of heterozygote genotype with only one allele amplified due to allelic dropout (f) was calculated using f=2p. Due to the nature of the profiles obtained, a stochastic threshold to determine the homozygosity of an allele was not used; as such homozygous alleles were calculated as heterozygous loci with allelic dropout as to not conflate the statistical value of the profiles obtained. While there is no defined limit for what constitutes a probative profile, for this work any profile with an RMP of 10^{-6} or 1 in 1 million was considered adequately probative for potential use in casework.

Table 6: Average Random Match Probability Per Donor.
Table shows the average of all random match probabilities calculated for each profile by cell number and donor using the optimized method. Rightmost column shows the random match probability of the full donor profile with all alleles recovered. For 1-cell collections, n=5. For 2-, 5-, 10- and 20-cell collections n=2

Donor	1-Cell	2-Cell	5-Cell	10-Cell	20-Cell	Full Profile
SE5	E-05	E-04	E-08	E-09	E-14	E-30
SE15	E-05	E-06	E-12	E-14	E-15	E-34
SE16	E-07	E-19	E-20	E-25	E-25	E-32
SE25	E-04	E-03	E-16	E-15	E-20	E-29

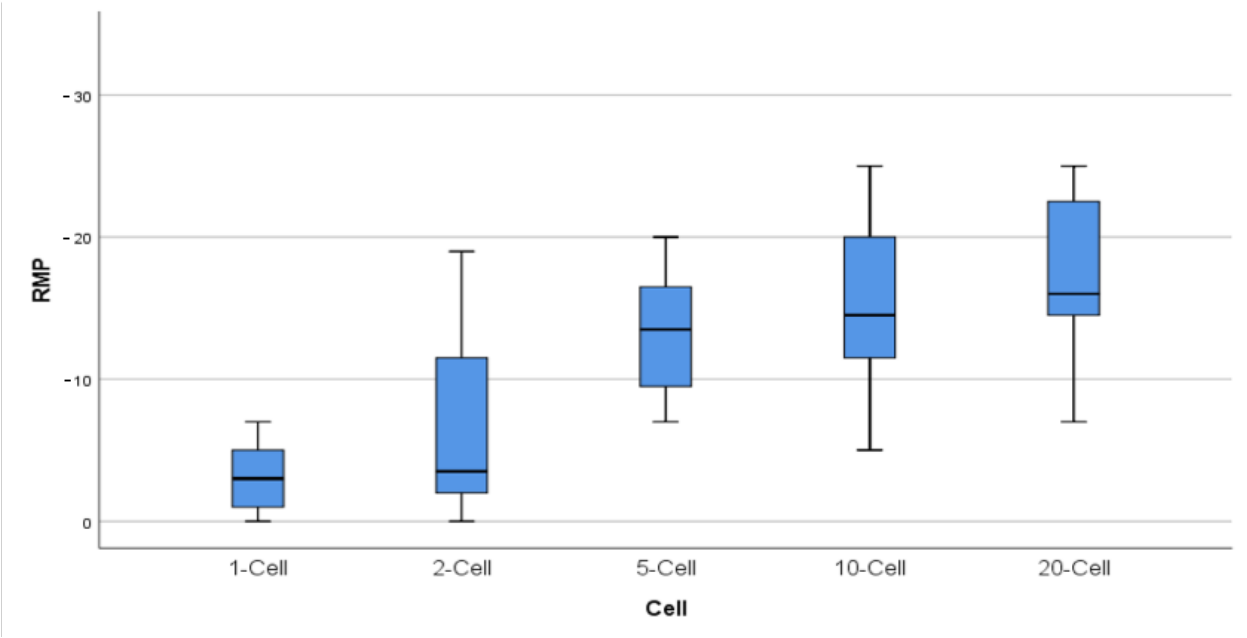

Figure 15: Boxplot of Average Random Match Probability Per Donor.
Figure shows the random match probability calculated for each profile obtained using the optimized method. Y-axis shows the random match probability; x-axis shows number of cells collected. Error bars are shown in black; outliers are indicated by a black circle. For 1-cell collections, n=5. For 2-, 5-, 10-, and 20-cell collections n=2

The results show that a probative profile was obtained for some of the 1-cell samples (Figure 15), with one donor recovering an average RMP over our threshold, 10^{-7} (Table 6). For 2-cell samples, 2 donors recovered an average RMP over our threshold, with averages of 10^{-6} and 10^{-19}. For 5-, 10-, and 20-cell samples, each donor achieved an average RMP over our probative threshold, with no profiles recovered with an RMP below the threshold. Therefore, when

suggesting a limit of detection with which to provide a probative DNA profile from the minimum number of micromanipulated sperm cells, at least 5-cell samples should be used. However, if the situation only allowed for 1- or 2-cell recoveries, there is still the possibility to recover a probative profile.

4.1.1 Consensus Profiling

Another method used to evaluate LCN samples is with consensus profiling, which involves amplification of multiple replicates of the same sample in order to create a single profile. This method is used help overcome stochastic effects and ensure accuracy of the alleles recovered. To produce a consensus profile, the different replicate profiles are compared and only alleles which occur two or more times are included in the consensus profile. To test the validity of using consensus profiling on micromanipulated sperm cells, consensus profiles were produced with the replicates of each donor using both lysis methods for VeriFiler™ Plus, as several replicates per donor were needed. For each 1-cell consensus profile, ten 1-cell samples were used. For each 2-, 5-, 10- and 20-cell profile, four samples were used. To determine if allelic recovery was impacted using this method, percent recovery (PR) of each consensus profile was calculated (Table 7) along with the RMP (Table 8) to determine the probative value of the consensus profiles created. These values were then compared to the PR (Figure 16) and RMP (Figure 17) of the individual profiles.

Table 7: Percent Recovery calculated for each consensus profile for each donor
Percent recovery was calculated per profile obtained for each donor and cell number using VeriFiler™ Plus with both lysis methods. Replicates used to create consensus profiles: for 1-cell consensus profiles, n=10. For consensus profiles created using 2-, 5-, 10-, and 20-cell samples n=4. Each cell number per donor contains data for one consensus profile (n=1).

Donor	1-Cell	2-Cell	5-Cell	10-Cell	20-Cell
SE5	32	24	34	41	44
SE15	41	10	51	73	66
SE16	64	64	64	78	77
SE25	43	17	74	76	88

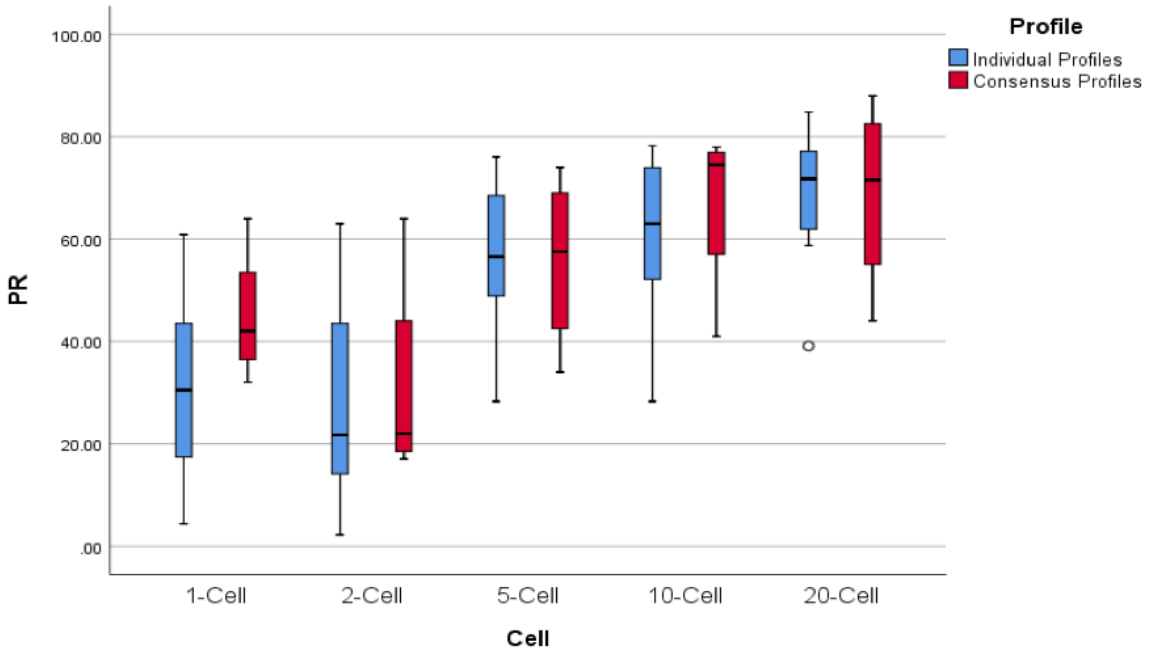

Figure 16: Boxplot of Percent Recovery for individual profiles compared to consensus profiles
A comparison of the percent recovery calculated for the consensus profile of each donor to the percent recovery calculated for each profile obtainied per sample. Y-axis shows the percent recovery; x-axis lists shows the number of cells collected. Red boxes show results obtained using consensus profiles, blue boxes show results obtained from each individual profile. Error bars are shown in black, outliers indicated by a black circle. For the individual profiles, 1-cell samples n=10 and 2-, 5-, 10- and 20-cell samples n=4. Replicates used to create consensus profiles for 1-cell consensus profiles, n=10. For consensus profiles created using 2-, 5-, 10-, and 20-cell samples n=4. Each cell number per donor contains data for one consensus profile (n=1).

Table 8: Random match probability calculated for each consensus profile for each donor
Random match probability was calculated per profile obtained for each donor and cell number using VeriFiler™ Plus with both lysis methods Replicates used to create consensus profiles: for 1-cell consensus profiles, n=10. For consensus profiles created using 2-, 5-, 10-, and 20-cell samples n=4. Each cell number per donor contains data for one consensus profile (n=1).

Donor	1-Cell	2-Cell	5-Cell	10-Cell	20-Cell
SE5	E-07	E-05	E-08	E-09	E-10
SE15	E-11	E-01	E-12	E-19	E-16
SE16	E-22	E-20	E-20	E-25	E-26
SE25	E-07	E-02	E-16	E-18	E-20

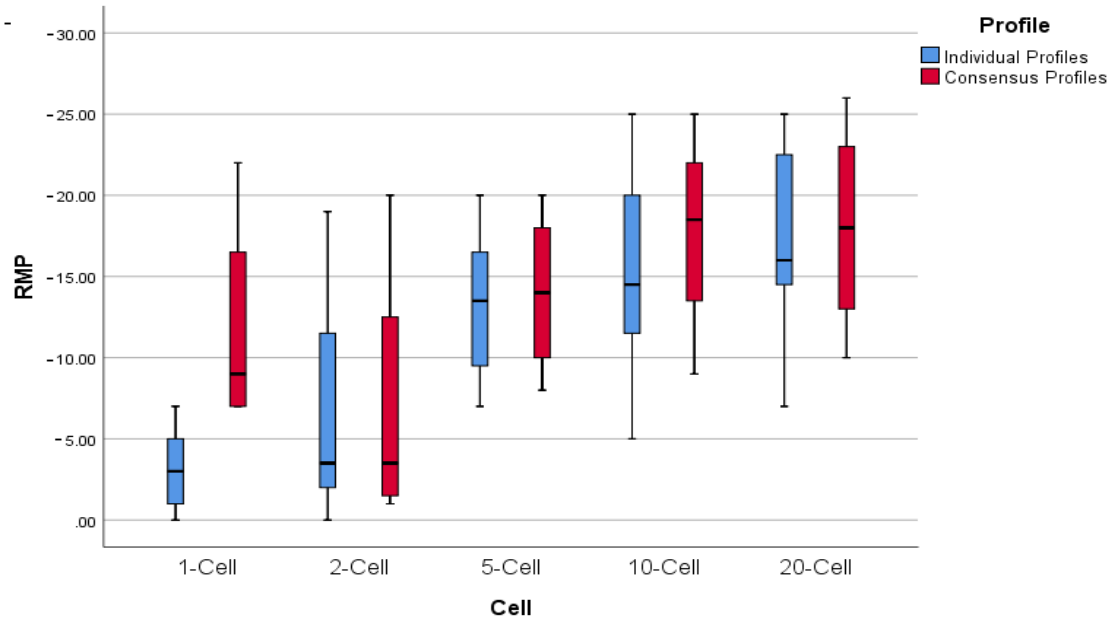

Figure 17: Boxplot of random match probability calculated for consensus profiles compared to individual profiles.
A comparison of the random match probability calculated for the consensus profile of each donor to the percent recovery calculated for each profile obtainied per sample. Y-axis shows the random match probability; x-axis shows the number of cells collected. Red boxes show results obtained using consensus profiles, blue boxes show results obtained from each individual profile. Error bars are shown in black, outliers indicated by a black circle. For the individual profiles, 1-cell samples n=10 and 2-, 5-, 10- and 20-cell samples n=4. Replicates used to create consensus profiles for 1-cell consensus profiles, n=10 For consensus profiles created using 2-, 5-, 10-, and 20-cell samples n=4. Each cell number per donor contains data for one consensus profile (n=1).

For 1-cell samples, both the average PR and RMP were increased when using consensus profiling. For 2-cell samples, the average PR and RMP did not change. For 5-cell samples, the average PR did not change, and the average RMP increased slightly using consensus profiling. For 10-cell samples, both the average PR and RMP increased using consensus profiling. For 20-cell samples, the PR of alleles did not change, but the average RMP increased when using consensus profiling. These results show that the 1-cell samples are more informative when consensus profiling is used, as the consensus profile for each donor was able to achieve an RMP over the probative threshold of 10^{-6}. Furthermore, the PR and RMP did not decrease for any cell numbers when using consensus profiling, which indicates that any future casework performed using this method could use consensus profiling to increase the reliability of the alleles recovered without detracting from the overall value of the profiles if enough replicates are available for collection.

4.2. Analyzing Quality of DNA Profiles

The quality of profiles obtained using micromanipulation were analyzed for stochastic effects that could potentially detract from the usefulness of a profile in a casework setting, such as drop-in alleles or increased stutter which might be falsely attributed as an allele.

4.2.1. Stutter Analysis

To analyze the stutter produced from the profiles obtained, the height of the allele and the height of stutter peak were recorded, and then the height of the stutter peak was divided by the allelic peak height to produce a stutter ratio (SR) [63, 64]. The SR was then multiplied by 100 to produce the percent height of the stutter peak compared to the allelic peak. Each locus already had a threshold determined by the manufacturer of VeriFiler™ Plus to filter out stutter produced

in a reaction using the manufacture's recommended conditions including DNA input and cycle number. These filters were not used when examining the profiles for instances of stutter at each locus. Once the percent height of each stutter peak was determined, each instance of stutter was plotted on a graph with the manufacturer's reported stutter thresholds indicated by a red line. Next, the average of the stutter peak heights was calculated as well as the standard deviation. The average stutter peak height plus a number of standard deviations higher were used to calculate the new stutter thresholds that would filter out the increased stutter seen in the profiles. The two thresholds that were considered used one standard deviation higher than the average, a more conservative threshold, as well as three standard deviations higher. By using more standard deviations, the likelihood of filtering out more stutter peaks increases as does the chance to filter out true alleles. The new threshold using one standard deviation is indicated by a green line on the graphs, while the new threshold using three standard deviations is indicated by a yellow line on the graphs. The first stutter type analyzed this way was the most common stutter peak observed, n-1, evaluated by loci and dye channel [56-58].

Table 9: Stutter peak data (n-1) for 6-FAM™ (blue channel) using VeriFiler™ Plus
The percent stutter average was calculated along with the standard deviation. The manufacturer's threshold is listed for reference, as well as the new threshold calculated using 1 standard deviation and 3 standard deviations. Replicates for each locus: D3S1358 n = 124, vWA n = 116, D16S539 n = 55, CSF1PO n = 20, D6S1043 n = 11.

Locus	Percent Stutter Average	Std. Dev	Kit Threshold	New Threshold (+1 SD)	New Threshold (+3 SD)
D3S1358	20.19	7.08	12.61	27.3	41.4
vWA	21.56	12.39	11.86	34.0	58.7
D16S539	17.07	12.73	11.19	29.8	55.3
CSF1PO	12.09	5.21	11.00	17.3	27.7
D6S1043	15.39	7.26	12.05	22.7	37,1

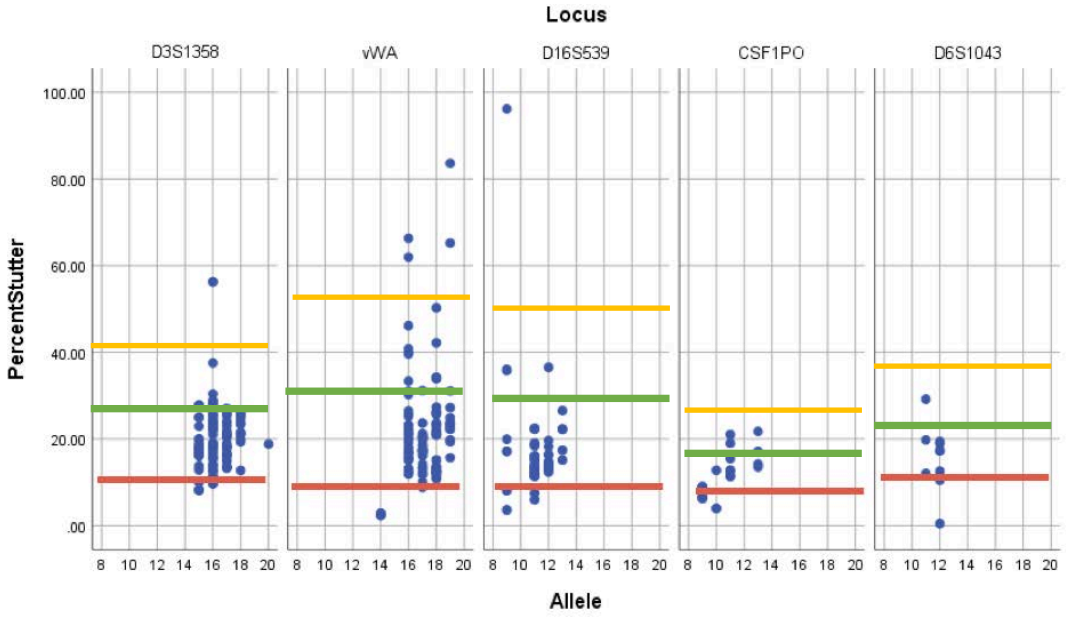

Figure 18: Stutter peaks (n-1) observed in 6-FAM™ (blue channel) using VeriFiler™ Plus
The percent stutter height is shown for each instance of stutter. Y-axis indicates percent stutter height; top x-axis shows locus; bottom x-axis indicates allele designation. Data points are indicated with blue circles. The manufacturer's threshold is represented by the red line, the new calculated threshold using 1 standard deviation is indicated by the green line, the new calculated threshold using 3 standard deviations is indicated by the yellow line. Replicates per locus: D3S1358 n = 124, vWA n = 116, D16S539 n = 55, CSF1PO n = 20, D6S1043 n = 11.

For the blue dye channel (Figure 18, Table 9), a significant amount of stutter is shown above the kit threshold for each locus, showcasing that increased stutter products are indeed present in our samples. The increase in stutter height frequently produces stutter peaks above the kit threshold, with an average stutter height ranging from 12-21%. Some stutter peaks were observed up to 95% of the height of the true allele (locus D16S539), which could prove to be challenging for those analyzing these profiles. While it would be difficult to filter out stutter of that height without losing alleles, the stutter thresholds developed using 1 standard deviation are able to filter out the 80-96% of the stutter seen, while the stutter thresholds using three standard deviations successfully filtered out 97-100% of the stutter seen. The results also showed an increase in instances of stutter peaks in the first two loci, which could be attributed to the smaller size of those loci, which could allow them to be more frequently amplified to completion by the DNA polymerase thus increasing the stochastic effects seen.

Table 10: Stutter peak data (n-1) for VIC ™ (green channel) using VeriFiler™ Plus
The percent stutter average was calculated along with the standard deviation. A new threshold was calculated using 1 standard deviation and 3 standard deviations. The kit threshold is listed for reference**.** Replicates used to calculate the averages: D8S1179 n = 142, D21S11 n = 61, D18S51 n = 51.

Locus	Percent Stutter Average	Std. Dev	Kit Threshold	New Threshold (+1 SD)	New Threshold (+3 SD)
D8S1179	18.20	11.03	11.54	29.2	51.3
D21S11	16.45	6.16	13.83	22.6	34.9
D18S51	25.77	17.07	13.73	42.8	77.0
D5S818	n.d.	n.d.	10.90	n.d.	n.d.

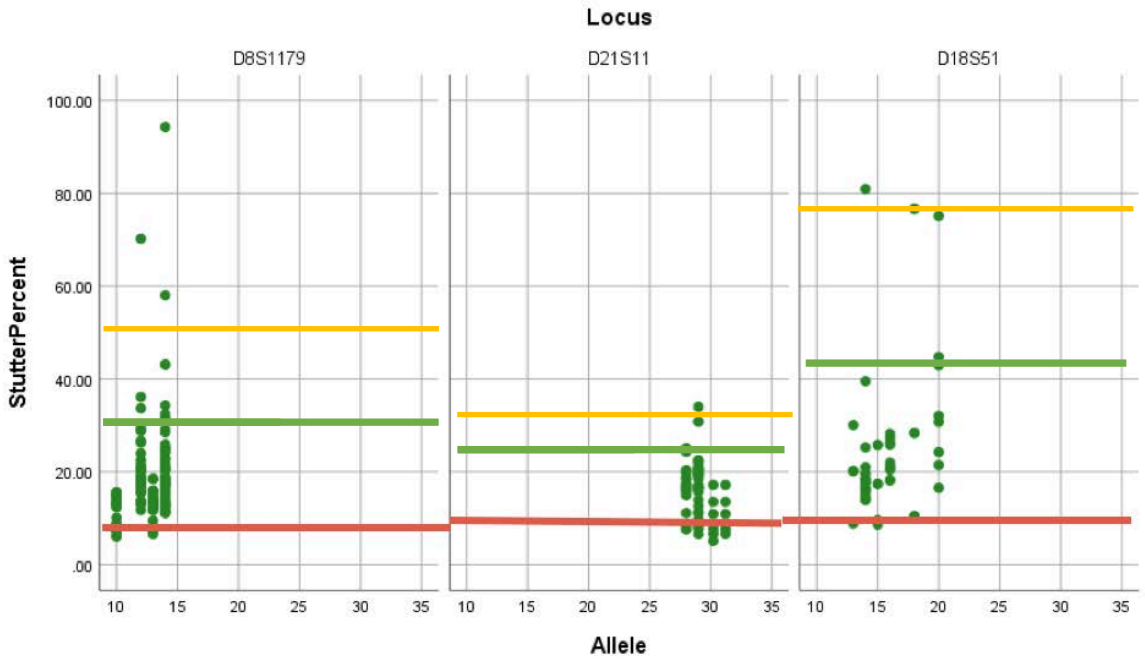

Figure 19: Stutter peaks (n-1) observed in VIC ™ (green channel) using VeriFiler™ Plus
The percent stutter height is shown for each instance of stutter. The percent stutter height is shown for each instance of stutter. Y-axis indicates percent stutter height; top x-axis shows locus; bottom x-axis indicates allele designation. Data points are indicated with green circles. The kit threshold is represented by the red line, the new calculated threshold using 1 standard deviation is indicated by the green line, the new calculated threshold using 3 standard deviations is indicated by the yellow line. Replicates used per locus: D8S1179 n = 142, D21S11 n = 61, D18S51 n = 51.

In the green channel, a high instance of increased stutter peaks was observed above the kit threshold, with an average stutter height ranging from 16-25% (Figure 19, Table 10). Additionally, stutter peaks were observed that were up to 90% of the allele height (seen in locus D8S1179). However, the majority of the stutter peaks observed in the samples were below 40% of the allele height. Using the new thresholds calculated for each locus, 91-95% of the stutter peaks observed were successfully filtered out using the threshold with one standard deviation, while 98-100% of the stutter was successfully filtered out using the threshold with three standard deviations. No data was obtained for the locus D5S818, as it was not frequently amplified in the profiles, and any alleles that were present did not produce stutter. In these instances, the kit threshold should be used since a new threshold could not be determined.

Table 11: Stutter peak data (n-1) for TED™ (yellow channel) using VeriFiler™ Plus
The percent stutter average was calculated along with the standard deviation. A new threshold was calculated using 1 standard deviation and 3 standard deviations. The kit threshold is listed for reference**.** Replicates used to calculate the averages: D2S441 n = 118, D19S43 n = 19, FGA n = 94.

Locus	Percent Stutter Average	Std. Dev	Kit Threshold	New Threshold (+1 SD)	New Threshold (+3 SD)
D2S441	14.96	10.82	9.71	25.8	47.4
D19S43	11.05	8.67	11.00	19.7	37.0
FGA	15.76	7.83	14.01	23.6	39.3
D10S1248	n.d.	n.d.	12.48	n.d.	n.d.

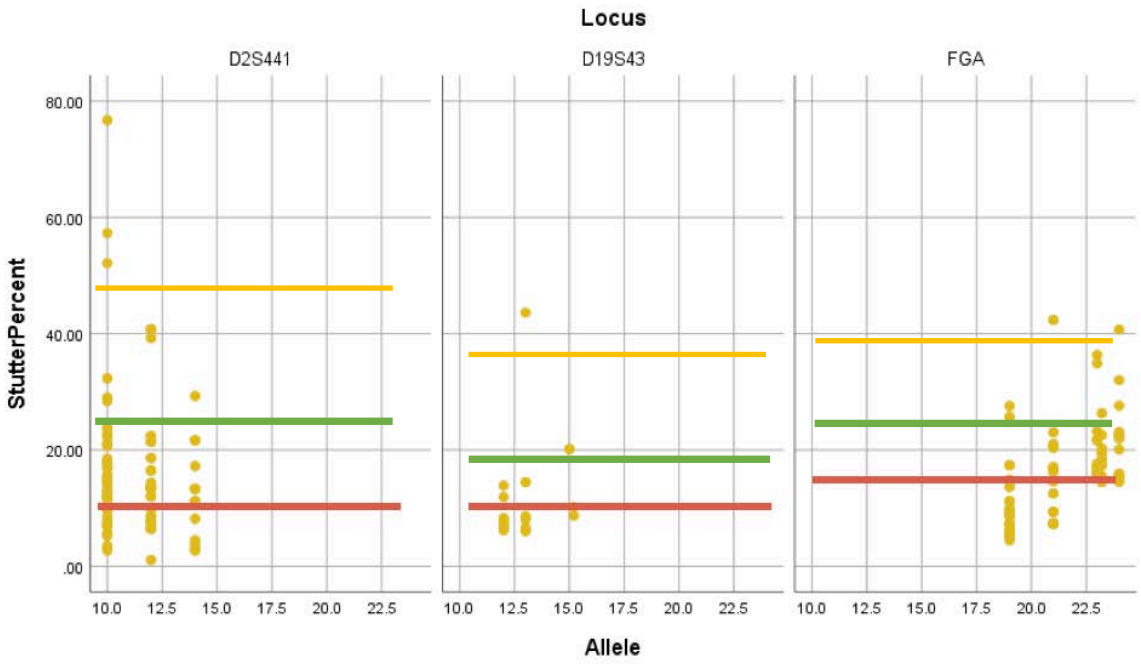

Figure 20: Stutter peaks (n-1) observed in TED™ (yellow channel) using VeriFiler™ Plus
The percent stutter height is shown for each instance of stutter. The percent stutter height is shown for each instance of stutter. Y-axis indicates percent stutter height; top x-axis shows locus; bottom x-axis indicates allele designation. Data points are indicated with yellow circles. The kit threshold is represented by the red line, the new calculated threshold using 1 standard deviation is indicated by the green line, the new calculated threshold using 3 standard deviations is indicated by the yellow line Replicates per locus: D2S441 n = 118, D19S43 n = 19, FGA n = 94.

For the yellow dye channel, an increased stutter peak height was again observed, with the average stutter height between 11-16% (Figure 28, Table13) with a peak height up to 80% of the allele height (seen at locus D2S441). With the implementation of new stutter thresholds, 89-94% the of the stutter seen at was successfully filtered out when using a threshold with one standard deviation, and 95-98% was successfully filtered out using a threshold with three standard deviations. Additionally, no stutter data was retrieved for locus D10S1248.

Table 12: Stutter peak data (n-1) for TAZ™ (red channel) using VeriFiler™ Plus
The percent stutter average was calculated along with the standard deviation. A new threshold was calculated using 1 standard deviation and 3 standard deviations. The kit threshold is listed for reference**.** Replicates used to calculate the averages: D22S1045 n = 73, D1S1656 n = 107, D13S317 n = 73, D7S820 n = 6, Penta E n = 15.

Locus	Percent Stutter Average	Std. Dev	Kit Threshold	New Threshold (+1 SD)	New Threshold (+3 SD)
D22S1045	20.64	13.39	19.04	34.0	60.8
D1S1656	20.60	7.22	15.20	27.8	42.3
D13S317	9.89	9.23	11.00	19.1	37.6
D7S820	5.61	2.40	9.83	9.8	12.8
Penta E	10.77	5.49	7.99	16.3	27.2

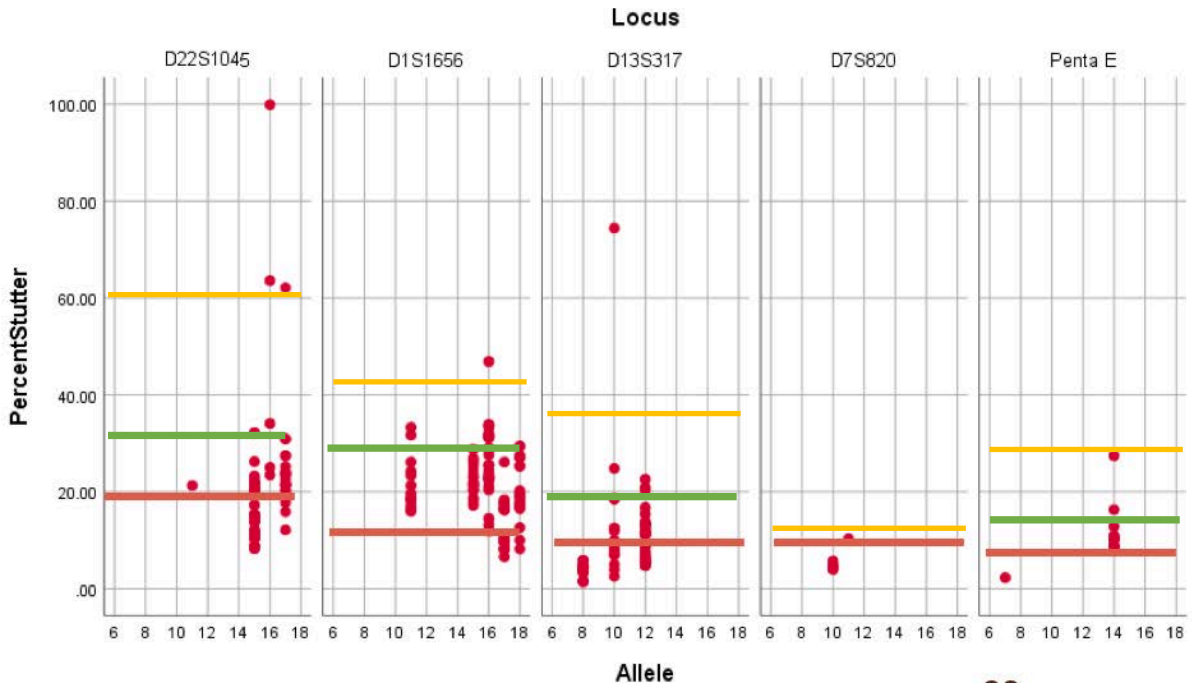

Figure 21: Stutter peaks (n-1) observed in TAZ™ (red channel) using VeriFiler™ Plus
The percent stutter height is shown for each instance of stutter. The percent stutter height is shown for each instance of stutter. Y-axis indicates percent stutter height; top x-axis shows locus; bottom x-axis indicates allele designation. Data points are indicated with red circles. The kit threshold is represented by the red line, the new calculated threshold using 1 standard deviation is indicated by the green line, the new calculated threshold using 3 standard deviations is indicated by the yellow line. Replicates per locus: D22S1045 n = 73, D1S1656 n = 107, D13S317 n = 73, D7S820 n = 6, Penta E n = 15.

For the red dye channel, an increase in stutter peak height was again observed, with the average stutter peak height between 5-20% of the allele height (Table12, Figure 21), with the presence of

stutter peaks up to 99% of our allele height (seen at locus D22S1045). With the development of new stutter thresholds, 80-95% of the stutter was filtered out when using a threshold with one standard deviation, and 96-100% of the stutter was filtered out when using a threshold with thee standard deviations. However, for D7S820 the calculated threshold with one standard deviation was lower was lower than the kit threshold, likely due to the low number of stutter instances observed, and with more replicates a more accurate threshold could likely be developed. In this instance, using the kit threshold or the threshold with three standard deviations would be recommended.

Table 13: Stutter peak (n-1) data for SID™ (purple channel) using VeriFiler™ Plus
The percent stutter average was calculated along with the standard deviation. A new threshold was calculated using 1 standard deviation and 3 standard deviations. The kit threshold is listed for reference**.** Replicates used to calculate the averages: Penta D n = 114, TH01 n = 50, D12S391 n = 61, D2S1336 n = 5, TPOX n = 8.

Locus	Percent Stutter Average	Std. Dev	Kit Threshold	New Threshold (+1 SD)	New Threshold (+3 SD)
Penta D	9.28	9.57	4.59	18.9	38.0
TH01	14.43	19.17	5.55	33.6	72.0
D12S391	27.95	13.52	15.61	41.5	68.5
D2S1336	53.50	18.06	16.69	72.1	N.A.
TPOX	10.68	3.88	5.82	14.6	22.3

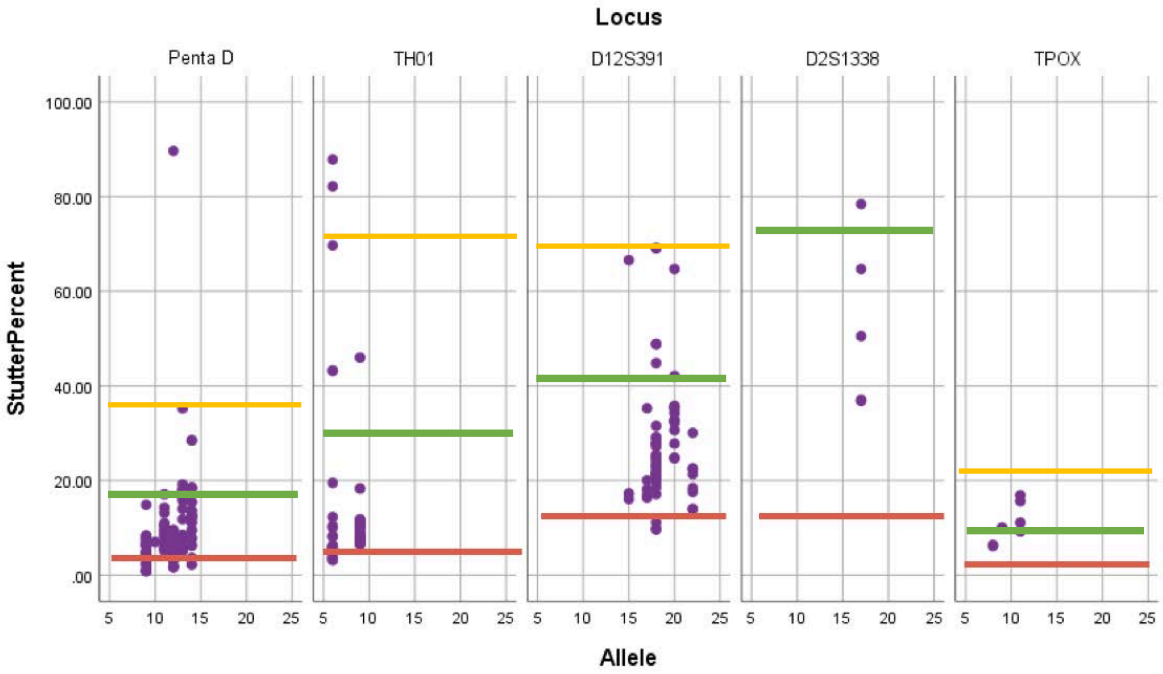

Figure 22: Stutter peaks (n-1) observed in SID™ (purple channel) using VeriFiler™ Plus
The percent stutter height is shown for each instance of stutter. The percent stutter height is shown for each instance of stutter. Y-axis indicates percent stutter height; top x-axis shows locus; bottom x-axis indicates allele designation. Data points are indicated with purple circles. The kit threshold is represented by the red line, the new calculated threshold using 1 standard deviation is indicated by the green line, the new calculated threshold using 3 standard deviations is indicated by the yellow line. Replicates per locus: Penta D n = 114, TH01 n = 50, D12S391 n = 61, D2S1336 n = 5, TPOX n = 8

In the purple dye channel, stutter peak heights were observed at an average between 9-53% of the allele height (Table 15, Figure 30). This value might be conflated due to the low number of data points for locus D2S1336. With the implementation of the new thresholds, 90-94% of stutter was filtered out for Penta D, TH01, and D12S391 when using the threshold with one standard deviation, and 96-100% was filtered out when using the threshold with three standard deviations. However, as locus D2S1228, only contained four data points, all of which did not seem have any consistency in peak height, impacting the accuracy of the threshold. The threshold calculated for this locus using one standard deviation was 72.1%, which is quite large and very likely to filter out any true alleles that might occur in the stutter position and only filtered out 75% of the stutter seen. The threshold calculated for this locus using three standard deviations was over 100%, and

therefore cannot be used. To calculate a more accurate threshold for these two loci, more data points would need to be evaluated.

When evaluating n-1 stutter, the results showed that at every single locus, stutter was observed well above the suggested stutter threshold for a typical amplification with VeriFiler™ Plus, thus implementation of a new threshold is necessary to filter stutter out when analyzing profiles produced from micromanipulated sperm cells. The method used to calculate the stutter height and produce a new threshold proved useful in removing the majority of all stutter for most loci, except those in which more data is needed. Further research should be done to collect more replicates for loci not typically amplified in LCN samples to determine more accurate thresholds.

Other types of stutter were observed in the profiles obtained using micromanipulation, including n-2, n-3, and n-4 as well as n+1 stutter. While these stutter types were present in the samples and could likely lead to the false attribution of alleles, they were not observed with enough frequency at enough loci, thus more samples should be obtained before evaluating the impact of these types of stutter on the value of the profiles.

4.2.2. Drop-In Analysis

After completing stutter analysis, the profiles were then analyzed for allele drop-in. Allele drop-in is the presence of an allele that is not attributable to the donor profile. In LCN samples, allele drop-in can be attributed to either contamination, or the development of spurious alleles due to stochastic effects. For our analysis purposes, peaks located in stutter positions were not counted as drop-in alleles, since an increased stutter presence due to our amplification methods was already recognized. To evaluate the impact of drop-in on our profiles, the total number of drop-in alleles as well as the total number of alleles calls per sample were recorded. The number of drop-in alleles was divided by the total number of allele calls to obtain the

percentage of overall alleles that were attributed to drop-in (Table 14). Furthermore, the number of samples that had drop-in alleles present was divided by the total number of profiles analyzed to determine the percentage with which drop-in effects our profiles (Table 15).

Table 14: Drop-In Analysis Results.
Total number of drop-in alleles calls and total number of allele calls is listed and the percent drop-in is listed. For 1-cell collections n=20, for 2-, 5-, 10-, and 20-cell collections n=8.

Cell Number	Number of Drop-in Alleles Observed	Total Number of Allele Calls	Percent Drop In
1	20	257	7.78
2	14	109	12.84
5	66	241	27.39
10	87	282	30.85
20	81	260	31.15

Table 15: Profiles with drop-in.
The number of profiles that contained at least one instance of drop-in, and the total number of profiles analyzed. The percentage of profiles containing drop-in was calculated. For 1-cell collections n=20, for 2-, 5-, 10-, and 20-cell collections n=8.

Cell Number	Number of Profiles with Drop-In Observed	Total Number of Profiles	Percent Profiles with Drop-In
1	9	20	45
2	3	8	37.5
5	7	8	87.5
10	6	8	75
20	5	8	62.5

For 1-cell samples, the percentage of drop-in alleles was low at ~8% while 45% of the profiles showed at least one instance of drop in. For 2-cell samples, the number of drop-in alleles increased to ~13% of total allele calls, with ~ 38% of our samples showing at least one instance of drop in. For 5-cell samples, the number of drop-in alleles increased to ~27% of allele calls, with 87.5% of all profiles showing at least one instance of drop-in. For 10-cells, the number of drop-in alleles increased slightly to ~31% of the alleles recovered, with 75% of profiles showing at least one instance of drop-in. For 20-cells, the number of drop-in alleles obtained did not increase, while the number of samples showing at least one instance of drop-in decreased. These

results show that while the presence of drop-in was significant it did not account for higher than ~31% percent in any cell number. Drop-in alleles had the least amount of impact on the 1-cell profiles, however the impact of drop-in increased significantly between 2- and 5-cells. 5-, 10- and 20-cell samples all had similar instances of drop-in, with over half the profiles obtained showing at least one drop-in allele. Therefore, the next step was to determine a method to filter out potential stutter while retaining the true allele calls in our profiles.

4.2.3 Overcoming drop-in

To overcome drop-in that could affect the quality of the DNA profiles obtained using micromanipulation, a new analysis threshold was created. During analysis of DNA profiles using the GeneMapper ID-X software, an analysis threshold is used to filter out any baseline noise from being called as an allele by the software. The analysis threshold is determined through internal validation studies performed by each lab prior to use. A universal default threshold of 150 RFUs is used prior to empirical study-specific threshold determinations. This is typically created using blank samples and determining the intensity difference between the highest peak and the lowest trough. However, according to SWADGM guidelines, a threshold can be created using samples with alleles present [65], as the presence of alleles amplifies stochastic effects and this increase could create spurious allele calls in the GeneMapper ID-X Software. Therefore, the analysis threshold was re-calculated using positive samples. To do this, the difference in intensity of the highest peak observed (not including allele calls) and the lowest trough was observed at each dye channel per cell number. These values were then combined to determine the average height of the stochastic threshold as well as the standard deviation. To calculate the new analysis, the average RFU height plus one standard deviation was used to determine a new value that would filter out most stochastic amplifications (Table 16).

Table 16: Stochastic threshold calculation results.
The stochastic threshold was calculated to overcome stochastic effects produced with LCN profiles. The difference in height from the highest non-allele peak and the lowest trough for each dye channel in each sample was calculated. The average of these values is shown in the table, as well as the standard deviation. These two values were added together and rounded to the nearest replicate of 5 to create a stochastic threshold for each dye channel for each cell number. For 1-cell samples, n=25. For 2-, 5-, 10-, and 20-cell samples, n=10.

Cell Number	Dye Channel	Average RFUs	Standard Deviation	Suggested Threshold
1-Cell		60	35.84	150
		165	115.77	280
		97	79.34	175
		126	90.39	220
		196	268.59	465
2-Cell		60	35.62	150
		175	163.92	340
		103	70.29	170
		140	83.03	225
		242	210.78	455
5-Cell		104	63.85	170
		290	153.59	445
		180	93.04	275
		186	97.67	285
		471	247.21	720
10-Cell		191	136.99	330
		272	175.63	450
		250	192.02	445
		239	164.60	405
		471	243.05	285
20-Cell		219	126.84	345
		405	182.03	590
		313	186.06	500
		273	185.79	460
		626	310.42	940

After implementation of the stochastic threshold, the instances of drop-in decreased for each cell number (Table 17). For 1-cell samples, drop-in alleles contributed to ~3% of all allele calls, compared to ~7% without the threshold. Additionally, the number of profiles observed with drop-in present decreased from 45% to 30% with the new threshold (Table 18). For 2-cells, the incidence of drop-in decreased from ~13% to ~8% with the stochastic threshold, and while the

number of profiles with drop-in present did not change. For 5-cells, the number of drop-in alleles decreased from ~27% to ~18%, and the number of profiles with drop-in present decreased from ~88% to ~62%. For 10-cells, the incidence of drop-in decreased from ~31% to ~21% and the number of samples with drop-in present did not change. For 20-cells, the incidence of drop-in decreased from ~31% to ~12% and the number of profiles with drop-in present did not change. Overall, the implementation of the stochastic threshold was successful in decreasing the incidence of drop-in seen in each cell numbers by 28-50%.

Table 17: Drop-In analysis after implementation of a new stochastic threshold
Total number of drop-in alleles calls, and total number of allele calls is listed, and the percent drop-in is listed. For 1-cell collections n=20, for 2-, 5-, 10-, and 20-cell collections n=8.

Cell Number	Number of Drop-in Alleles Observed	Total Number of Allele Calls	Percent Drop In
1	6	231	2.60
2	8	99	8.08
5	34	193	17.62
10	41	196	20.92
20	24	194	12.37

Table 18: Analysis of profiles containing drop-in
The number of profiles that contained at least one instance of drop-in, and the total number of profiles analyzed. The percentage of profiles containing drop-in was calculated.

Cell Number	Samples with Drop-in Present	Total Number of Samples	Percent of Samples with Drop-In
1	6	20	30
2	3	8	37.5
5	5	8	62.5
10	6	8	75
20	5	8	62.5

To ensure that the new stochastic threshold did not significantly decrease the number of true alleles recovered, percent difference in allele recovery before and after implementation of our threshold was calculated (Table 19). For 1-, 2- and 20-cell samples, the number of alleles recovered decreased by only 4-5%. For 5-cell samples, the number of alleles recovered decreased by ~9%. For 10-cell samples, the number of alleles recovered decreased by 20%. There was a higher decrease in allele recovery for 5- and 10-cells samples after implementation

of the stochastic threshold, however it did not significantly impact the allele recovery for our 1-, 2- and 20-cell samples. Further testing should be done using all cell numbers to ensure the reliability of the stochastic threshold and its effect on allelic recovery. However, the data here shows that the presence of drop-in alleles can be significantly decreased using the stochastic threshold without effecting recovery of alleles in our 1-, 2- and 20-cell samples.

Table 19: Impact of stochastic threshold of allele recovery
True allele calls were counted in total both before and after application of the stochastic threshold to determine the impact of the threshold to allele recovery. The percent difference of these values was also calculated**.** For 1-cell collections n=20, for 2-, 5-, 10-, and 20-cell collections n=8.

Cell Number	True Allele Calls before Threshold	True Allele Calls after Threshold	Percent Difference
1	237	225	5.06
2	95	91	4.21
5	175	159	9.14
10	195	155	20.5
20	179	170	5.02

.

CHAPTER 5: CONCLUSION

The main goal of this work was to develop methods that would permit the recovery and aSTR analysis of single or few sperm cells from late reported sexual assault evidence. While a few sperm cell recovery methods have been previously described, they each have disadvantages and have failed to be successfully integrated into U.S. crime laboratories. Therefore, this project focused on the evaluation of simplified micromanipulation as well as enhanced amplification strategies to produce a probative aSTR profile from single- or few-sperm cell samples.

The first aim of this study was to determine the most efficient lysis and amplification method for use with micromanipulated sperm cells. Four aSTR amplification kits were evaluated, AmpFLSTR™ IdentiFiler™ Plus, GlobalFiler™, PowerPlex® Fusion 6C, and VeriFiler™ Plus and two direct lysis methods, forensicGEM Sperm and LysePrep. After each collection, the allele recovery of each profiles was determined, and the percent recovery of each profile was calculated. Our results concluded that the LysePrep method was the advantageous method, as it allowed for higher allelic recovery at all cell numbers with each amplification kit. Our results also determined that the best performing amplification kits were PowerPlex® Fusion 6C and VeriFiler™ Plus, which both kits amplifying up to 46 autosomal loci. When comparing the performance of both amplification kits using the LysePrep method, our results showed that while PowerPlex® Fusion 6C recovered the greatest maximum number of alleles, VeriFiler™ Plus had a higher average allele recovery at all cell numbers and thus was deemed to be the more routinely successful kit. Therefore, the optimized method that produced the highest recovery of alleles from micromanipulated sperm was lysis using the LysePrep method and amplification using VeriFiler™ Plus.

The second aim of the study was to determine a limit of detection for the optimized method. This was done by determining the probative values of the samples obtained by calculating the random match probability of each profile. These results showed that even 1-cell profiles were able to obtain a 10^6 (1 in a million) power of discrimination, although this was not routinely achieved in all 1-cell samples. Our results showed that to routinely achieve a probative profile 5-sperm cells should be used. However, consensus profiling with multiple 1-cell replicates (e.g.10 replicates) could be used to improve the recovery of more probative profiles from samples containing less than 5-sperm cells. Furthermore, the quality of the profiles produced using the optimized method was evaluated by analyzing the impact of stochastic effects seen, specifically that of increased n-1 stutter peaks drop-in alleles. The results showed that there was a significant increase in the height of n-1 stutter peaks formed using the optimized method, and new, higher thresholds were calculated that successfully filtered out the majority of these peaks. The presence of drop-in alleles was evaluated and shown to be present in our LCN samples, however an analysis threshold was developed that helped to eliminate 28-50% of spurious allele formations seen in the profiles.

While this project shows that probative, reliable aSTR profiles can be retrieved from 1 to 20 sperm cells, further research needs to be conducted to verify the accuracy of the thresholds used. Additionally, further research should be conducted to allow for the optimal recovery of a full aSTR profile. This method should also be further tested on simulated mixtures and genuine post-coital samples. While work still needs to be conducted, this project resulted in the development of a novel simplified method for recovery of sperm cells that could prove to be beneficial in sexual assault casework.

APPENDIX: IRB MEMORANDUM

UNIVERSITY OF CENTRAL FLORIDA

Institutional Review Board
FWA00000351
IRB00001138, IRB00012110
Office of Research
12201 Research Parkway
Orlando, FL 32826-3246

Memorandum

To: Haley Hardin

From: UCF Institutional Review Board (IRB)

Date: June 5, 2020

Re: Request for IRB Determination

The IRB reviewed the information related to your dissertation *Single Cell Forensic Genomics: DNA Profiling of Micromanipulated Single Spermatozoa*.

As you know, the IRB cannot provide an official determination letter for your research because it was not submitted into our electronic submission system.

However, if you had completed a Huron submission, the IRB could make one of the following research determinations: "Not Human Subjects Research," "Exempt," "Expedited" or "Full Board.

Based on the information you provided, this study would have been issued a Not Human Subjects Research determination outcome letter had a request for a formal determination been submitted to the UCF IRB through Huron IRB system.

If you have any questions, please contact the UCF IRB irb@ucf.edu.

Sincerely,

Renea Carver
IRB Manager

REFERENCES

1. Butler, J.M., *Chapter 1 - Sample Collection, Storage, and Characterization, in Advanced Topics in Forensic DNA Typing: Methodology*, J.M. Butler, Editor. 2012, Academic Press: San Diego. p. 1-27.

2. Sharpe, N., *The Significance of Spermatozoa in Victims of Sexual Offences*, Canadian Medical Association J ournal, 1963. 89(10): p. 513-514

3. Wiegand, P., M. Schürenkamp, and U. Schütte, *DNA extraction from mixtures of body fluid using mild preferential lysis.* International Journal of Legal Medicine, 1992. 104(6): p. 359-360. doi: 10.1007/BF01369558

4. Yoshida, K., Sekiguchi, K., Mizuno, N., Kasia, K., Sato, H., & Seta, S.,*The modified method of two-step differential extraction of sperm and vaginal epithelial cell DNA from vaginal fluid mixed with semen*.Forensic Science International, 1995. 72(1): p. 25-33. doi: 10.1016/0379-0738(94)01668-u

5. McKiernan, H.E., and Danielson, P.B., Chapter 21 - *Molecular Diagnostic Applications in Forensic Science*, in *Molecular Diagnostics* (Third Edition), G.P. Patrinos, Editor. 2017, Academic Press. p. 371-394

6. Vuichard, S., Borer, U., Bottinelli, M., Cossu, C., Malik, N., Meier, V., Gehrig, C., Sulzer, A., Morerod, M., & Castella, V., *Differential DNA extraction of challenging simulated sexual-assault samples: a Swiss collaborative study*.Investigative Genetics, 2011. 2(1): p. 11.doi: 10.1186/2041-2223-2- 1.

7. Christian, C.W., Lavelle, J M., De Jon, A R., Loiselle, J., Brenner, L., & Joffe, M., *Forensic evidence findings in prepubertal victims of sexual assault.* Pediatrics, 2000. 106(1): p. 100-104. doi: 10.1542/peds.106.1.100

8. *National Best Practices for Sexual Assault Kits: A Multidisciplinary Approach.* National Institute of Justice. 2017.

9. Hanson, E.K. and Ballantyne, J., *Enhanced DNA profiling of the semen donor in late reported sexual assaults: use of Y-chromosome-targeted pre-amplification and next generation Y-STR amplification systems*, in Forensic DNA Typing Protocols. 2016, Springer. p. 185-200. doi: 10.1007/978-1- 4939-3597-0_15

10. Eungprabhanth, V., *Finding of the spermatozoa in the vagina related to elapsed time of coitus.* Zeitschrift für Rechtsmedizin, 1974. 74(4): p. 301-304. doi: 10.1007/BF02114842

11. Willott, G. and Allard, J., *Spermatozoa—their persistence after sexual intercourse*.Forensic Science International, 1982. 19(2): p. 135-154. 10.1016/0379-0738(82)90040-8

12. Perloff, W.H. and Steinberger, E.,*In vivo survival of spermatozoa in cervical mucus.*American Journal of Obstetrics and Gynecology, 1964. 88(4): p. 439-442. doi: 10.1016/0002-9378(64)90499-5

13. Morrison, A.I., *Persistence of spermatozoa in the vagina and cervix.* The British Journal of Venereal Diseases, 1972. 48(2): p. 141-143. doi: 10.1136/sti.48.2.141

14. Sanders, C.T., Sanchez, N., Ballantyne, J., Peterson, D., *Laser Microdissection Separation of Pure Spermatozoa from Epithelial Cells for Short Tandem Repeat Analysis.* Journal of Forensic Sciences, 2006. 51(4): p. 748-757. doi: 10.1111/j.1556-4029.2006.00180.x

15. Kayser, M., *Forensic use of Y-chromosome DNA: a general overview*. Human Genetics, 2017. 136(5): p. 621-635. doi: 10.1007/s00439-017-1776-9

16. Berger, B., Niedersättter, H., Köchl, S., Steinlechner, M., Parson, W., *Male/female DNA mixtures: a challenge for Y-STR analysis.* Progress in Forensic Genetics, 2003. 9: p. 295-299.doi: 10.1016/S0531-5131(02)00593-9

17. Cerri, N., Ricci, U., Sani, I., Verzeletti, A., De Ferrari, F., *Mixed stains from sexual assault cases: autosomal or Y-chromosome short tandem repeats?* Croatian MedicalJournal, 2003. 44(3): p. 289-292.

18. Hall, A. and Ballantyne, J., *The development of an 18-locus Y-STR system for forensic casework.* Analytical and Bioanalytical Chemistry, 2003. 376(8): p. 1234-1246. doi: 10.1007/s00216-003-2039-2

19. Hanson, E. and Ballantyne, J. , *A Y-short tandem repeat specific DNA enhancement strategy to aid the analysisof late reported (≥ 6 days) sexual assault cases.* Medicine, Science and the Law, 2014. 54(4): p. 209-218. doi: 10.1177/0025802413519761

20. Hall, A. and Ballantyne, J., N*ovel Y-STR typing strategies reveal the genetic profile of the semen donor in extended interval post-coital cervicovaginal samples.* Forensic Science International, 2003. 136(1): p. 58-72. doi: 10.1016/s0379-0738(03)00258-5

21. Mayntz-Press, K.A., Sims, L., Hall, A., Ballantyne, J., *Y-STR Profiling in Extended Interval (≥ 3 days) Postcoital Cervicovaginal Samples.* Journal of Forensic Sciences, 2008. 53(2): p. 342-348. doi:10.1111/j.1556-4029.2008.00672.x

22. Daniels, D.L., Hall, A. andBallantyne, J. *SWGDAM developmental validation of a 19-locus Y-STR system for forensic casework*. Journal of Forensic Science, 2004. 49(4): p. 1-16

23. Butler, J.M., *Forensic DNA typing: biology, technology, and genetics of STR markers.* 2005: Elsevier

24. Butler, J.M., *Genetics and genomics of core short tandem repeat loci used in human identity testing.* Journal of Forensic Sciences, 2006. 51(2): p. 253-265. doi: 10.1111/j.1556-4029.2006.00046.x

25. 6824. Butler, J.M., *Genetics and genomics of core short tandem repeat loci used in human identity testing*. Journal of Forensic Sciences, 2006. 51(2): p. 253-265. doi: 10.1111/j.1556-4029.2006.00046.x 25.

26. Shriver, M.D., Chapter 1 - *Forensic DNA Analysis: From Modest Beginnings to Molecular Photofitting, Genics, Genetics, Genomics, and the Pertinent Population Genetics Principles,* in *Molecular Photofitting,* T.N. Frudakis, Editor. 2008, Academic Press: Burlington. p. 1-34.

27. Budowle, B., Zoppis, S., & T.R. Moretti, *Forensics: analysis of short tandem repeat loci by multiplex PCR and real-time fluorescence detection during capillary electrophoresis, in DNA profiling and DNA fingerprinting.* 1999, Springer. p. 101-115. doi: 10.1007/978-1- 61779-461-2_1

28. Chakraborty, R., Stivers, D, N., 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(8): p. 1682-1696. doi:10.1002/(SICI)1522-2683(19990101)20:8<1682::AID-ELPS1682>3.0.CO;2-Z

29. Clark, C.P., Xu, K., Scott, O., Hickey, J., Tsuei, A., Jackson, K., & Landers, J, P., *Acoustic trapping of sperm cells from mock sexual assault samples.* Forensic Science International: Genetics, 2019. 41: p. 42-49. doi: 10.1016/j.fsigen.2019.03.012

30. Vandewoestyne, M. & Deforce, D., *Laser capture microdissection in forensic research: a review*. International Journal of Legal Medicine, 2010. 124(6): p. 513-521. doi:10.1007/s00414-010-0499-4

31. Norris, J.V., Evander, M., Horsman-Hall, K. M., Nilsson, J., Laurell, T., & Landers, J.P., A*coustic differential extraction for forensic analysis of sexual assault evidence.* Analytical Chemistry, 2009. 81(15): p. 6089-6095. doi: 10.1021/ac900439b

32. Elliott, K., Hill, D.S., Lambert, C., Burroughes, T.R., and Gill, P., *Use of laser microdissection greatly improves the recovery of DNA from sperm on microscope slides.* Forensic Science International, 2003. 137(1): p. 28-36. doi:10.1016/s0379-0738(03)00267-6

33. Kahn, M.C., *Chambers' micromanipulator for the isolation of a single bacterium.* The Journal of Infectious Diseases, 1922: p. 344-348.

34. Li, C.-x., Wang, G., Li, W., Huang, J., Ji, A., & Hu, L., *New cell separation technique for the isolation and analysis of cells from biological mixtures in forensic caseworks.* Croatian Medical Journal, 2011. 52(3): p. 293-298. doi: 10.3325/cmj.2011.52.293

35. Ashida, N., Ishii, S., Hayano,S., Tago, K., Tsuji, T., Yoshimure, Y., Otsuka, S., Senoo, K., *Isolation of functional single cells from environments using a micromanipulator: application to study denitrifying bacteria*. Applied Microbiology and Biotechnology, 2010. 85(4): p. 1211-1217. doi:10.1007/s00253-009-2330-z

36. Ng, S., Bongso, A., Sathananthan, H., & Ratnam, S.S., *Micromanipulation: its relevance to human in vitro fertilization*. Fertility and Sterility, Wallach, E.E., Editor.1990. 53(2): p. 203-219. doi: 10.1016/s0015-0282(16)53267-1

37. Farash, K., E.K. Hanson, & Ballantyne, J.,*Single source DNA profile recovery from single cells isolated from skin and fabric from touch DNA mixtures in mock physical assaults.* Science & Justice, 2018. 58(3): p. 191-199. doi: 10.1016/j.scijus.2017.12.006

38. Hanson, E.K. & Ballantyne, J.,*"Getting blood from a stone": ultrasensitive forensic DNA profiling of microscopic bio-particles recovered from "touch DNA" evidence*, in Nucleic Acid Detection. 2013, Springer. p. 3-17. doi:10.1007/978-1- 62703-535-4_1

39. Lounsbury, J.A., Coult, N., Miranian, D.C., Cronk, S.M., Haverstick, D.M., Kinnon, P., Saul, D.J., & Landers, J.P., *An enzyme-based DNA preparation method for application to forensic biological samples and degraded stains*. Forensic Science International: Genetics, 2012. 6(5): p. 607-615. doi: 10.1016/j.fsigen.2012.01.011

40. Moss, D., S.-A. Harbison, & Saul, D., *An easily automated, closed-tube forensic DNA extraction procedure using a thermostable proteinase.* International J ournal of Legal Medicine, 2003. 117(6): p. 340-349. doi: 10.1007/s00414-003-0400-9

41. Findlay, I., Taylor, A., Quirke, P., Frazier, R., & Urquhart, A., *DNA fingerprinting from single cells.* Nature, 1997. 389(6651): p. 555-556. doi: 10.1038/39225

42. Budowle, B., Hobson, D.L., Smerick, B., & Smith, J.A.L., *Low copy number—consideration and caution*. in *Twelfth International Symposium on Human Identification*. 2001. Promega Corporation.

43. Gill, P., Whitaker, J., Flaxman, C., Brown, N., Buckleton, J., *An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA.* Forensic Science International, 2000. 112(1): p. 17-40. doi:10.1016/s0379-0738(00)00158-4

44. Kloosterman, A.D. & Kersbergen, P., *Efficacy and limits of genotyping low copy number (LCN) DNA samples by multiplex PCR of STR loci.* J Soc Biol, 2003. 197(4): p. 351-9. doi: 10.1051/jbio/2003197040351

45. Budowle, B., Eisenberg, A.J., & Daal, A.V., *Validity of low copy number typing and applications to forensic science.* Croatian Medical Journal, 2009. 50(3): p. 207-217. doi: 10.3325/cmj.2009.50.207

46. Bessekri, M., Aggoune, A., Lazreg, S., Bucht, R., & Fuller, V., *Comparative study on the effects of reduced PCR reaction volumes and increased cycle number, on the sensitivity and the stochastic threshold of the AmpFlSTR Identifiler® Plus kit.* Forensic Science.

International: Genetics Supplement Series, 2013. 4(1): p. e306-e307. doi: 10.1016/j.fsigss.2013.10.156

47. McNevin, D., Edson, J., Robertson, J., Austin, J.J., *Reduced reaction volumes and increased Taq DNA polymerase concentration improve STR profiling outcomes from a real-world low template DNA source: telogen hairs.* Forensic Science, Medicine, and Pathology, 2015. 11(3): p. 326-338. doi: 10.1007/s12024-015-9679-3

48. Forster, L., Thomson, J., & Kutranov, S., *Direct comparison of post-28-cycle PCR purification and modified capillary electrophoresis methods with the 34-cycle "low copy number"(LCN) method for analysis of trace forensic DNA samples.* Forensic Science International: Genetics, 2008. 2(4): p. 318-328. doi:10.1016/j.fsigen.2008.04.005

49. Gaines, M.L., Wojtkiewicz, P.W., Valentine, J.A., & Brown, C.L., *Reduced volume PCR amplification reactions using the AmpFlSTR® Profiler Plus™ kit.* Journal of Forensic Science, 2002. 47(6): p. 1224-1237

50. Petricevic, S., Whitaker, J., Buckleton, J., Vintiner, S., Patel, J., Simon, P., Ferraby, H., Hermiz, W., & Russell, A., V*alidation and development of interpretation guidelines for low copy number (LCN) DNA profiling in New Zealand using the AmpFlSTR® SGM Plus™ multiplex.* Forensic Science International: Genetics, 2010. 4(5): p. 305-310. doi:10.1016/j.fsigen.2009.11.003

51. Whitaker, J.P., E.A. Cotton, and P. Gill, *A comparison of the characteristics of profiles produced with the AMPFlSTR® SGM Plus™ multiplex system for both standard and low copy number (LCN) STR DNA analysis.* Forensic Science International, 2001. **123**(2): p. 215-223.

52. Caragine, T., Mikulasovich, R., Tamariz, J., Bajda, E., Sebestyen, J., Baum, H., Prinz, M., *Validation of testing and interpretation protocols for low template DNA samples using AmpFlSTR® Identifiler®.* Croatian Medical Journal, 2009. 50(3): p. 250-267. doi:10.3325/cmj.2009.50.250

53. Benschop, C.C., van der Beek, C.P., Meiland, H.C., van Gorp, A.G.M., Westen, A.A., & Sijen, T., *Low template STR typing: effect of replicate number and consensus method on genotyping reliability and DNA database search results.* Forensic Science International: Genetics, 2011. 5(4): p. 316-328. doi: 10.1016/j.fsigen.2010.06.006

54. Kloosterman, A. & Kersbergen, P., *Efficacy and limits of genotyping low copy number DNA samples by multiplex PCR of STR loci.* in International Congress Series. 2003. Elsevier. doi: 10.1051/jbio/2003197040351

55. Brookes, C., Bright, J., Harbison, S., & Buckleton, J., *Characterising stutter in forensic STR multiplexes.* Forensic Science International: Genetics, 2012. 6(1): p. 58-63. doi:10.1016/j.fsigen.2011.02.001

56. Levinson, G. & Gutman,G.A.,*Slipped-strand mispairing: a major mechanism for DNA sequence evolution.* Molecular Biology and Evolution, 1987. 4(3): p. 203-221.doi:10.1093/oxfordjournals.molbev.a040442

57. Meldgaard, M. & Morling, N., *Detection and quantitative characterization of artificial extra peaks following polymerase chain reaction amplification of 14 short tandem repeat systems used in forensic investigations*. Electrophoresis, 1997. 18(11): p. 1928-1935. doi:10.1002/elps.1150181107

58. Grisedale, K.S. & van Daal, A.,*Comparison of STR profiling from low template DNA extracts with and without the consensus profiling method*. Investigative Genetics, 2012. 3(1): p. 14. doi: 10.1186/2041-2223-3- 1

59. Benschop, C.C.G., Wiebosch, D.C., Kloosterman, A.D., Sijen, T., *Post-coital vaginal sampling with nylon flocked swabs improves DNA typing*. Forensic Science International: Genetics, 2010. 4(2): p. 115-121. doi: 10.1016/j.fsigen.2009.07.003

60. *Guidelines for STR Enhanced Detection Methods, in Forensic Science Communications,* Scientific Working Group on DNA Analysis Methods (SWGDAM), Editor. 2014.

61. Buckleton, J. & Triggs, C., *Is the 2p rule always conservative?* Forensic Science International, 2006. 159(2-3): p. 206-209. doi: 10.1016/j.forsciint.2005.08.004

62. Manabe, S., Hamano, Y., Morimoto, C., Kawai, C., Fujimoto, S., & Tamaki, K., *New stutter ratio distribution for DNA mixture interpretation based on a continuous model.* Legal Medicine, 2016. 19: p. 16-21. doi: 10.1016/j.legalmed.2016.01.007

63. Bright, J.-A., Curran, J.M., & Buckleton, J.S., Investigation into the performance of different models for predicting stutter. Forensic Science International: Genetics, 2013. 7(4): p. 422-427. doi: 10.1016/j.fsigen.2013.04.008