

2024

Evaluation of the Y-screening and EZ1 differential extraction utilization by a crime laboratory

<https://hdl.handle.net/2144/51851>

"Downloaded from OpenBU. Boston University's institutional repository."

BOSTON UNIVERSITY

ARAM V. CHOBANIAN & EDWARD AVEDISIAN SCHOOL OF MEDICINE

Thesis

**EVALUATION OF THE Y-SCREENING AND EZ1 DIFFERENTIAL
EXTRACTION UTILIZATION BY A CRIME LABORATORY**

by

XIAOYU ZHANG

B.A., Boston University, 2022

Submitted in partial fulfillment of the
requirements for the degree of
Master of Science

2024

Approved by

First Reader

Celeste M. Wareing, MS
Instructor, Program in Biomedical Forensic Science

Second Reader

Robin W. Cotton, Ph.D.
Associate Professor of Anatomy & Neurobiology
Program in Biomedical Forensic Sciences

Third Reader

Saman Saleem, M.S.
Forensic Scientist III, Massachusetts State Police Crime Laboratory –
DNA Unit

ACKNOWLEDGMENTS

I would like to extend my deepest gratitude to my research advisor, Dr. Robin Cotton, for her invaluable guidance and support throughout the development of this thesis. Her insights and expertise have been fundamental to my research journey. Special thanks also go to Prof. Celeste M. Wareing and Saman Saleem for their assistance and contributions to my thesis. Their input has significantly enriched my work. I am grateful to all the faculty within the Boston University Biomedical Forensic Sciences M.S. program for their educational support and commitment to fostering a conducive learning environment.

My appreciation also extends to my classmates and the members of my research group under Dr. Robin Cotton. Their perspectives, assistance, and feedback have been immensely helpful and have greatly contributed to my research experience.

Lastly, I would like to thank my family for their unwavering support and encouragement. Their belief in my abilities has been a constant source of motivation throughout my academic endeavors.

**EVALUATION OF THE Y-SCREENING AND EZ1 DIFFERENTIAL
EXTRACTION UTILIZATION BY A CRIME LABORATORY**

XIAOYU ZHANG

ABSTRACT

Due to the frequent occurrence of sexual assaults, crime laboratories often face resource and personnel shortages when attempting to process the large numbers of sexual assault cases submitted to the Laboratory in a timely manner. As a result, laboratories may accumulate a backlog of Sexual Assault Kits (SAEKs). To address this issue, some crime laboratories have adopted Y-screening as an alternative to traditional serological tests for semen detection. This study examines the effectiveness of Y-screening and the EZ1 differential extraction method in the analysis of SAEKs and attempts to determine whether Y-screening can accurately predict the presence and amount of male and female deoxyribonucleic acid (DNA) in relevant samples. The information can be used by the laboratory to evaluate the forensic process and is designed to alleviate SAEK backlogs. By analyzing real casework quantification values and looking at simulated samples, the precision of Y-screening in estimating DNA concentrations and the ability of the EZ1 system to separate male from female DNA were evaluated. The research identified significant variability in the distribution of DNA across samples and noted that there is DNA loss during the EZ1 extraction process, in part because a portion of male DNA remains on the swab post-extraction. While Y-screening effectively detected male DNA, relying on its quantitative polymerase chain reaction (qPCR) results to decide how much sample to use for short tandem repeat (STR) analysis was sometimes inaccurate. These

findings underscore the necessity for further understanding and refinement of forensic protocols and technologies to enhance the handling of sexual assault evidence.

TABLE OF CONTENTS

ACKNOWLEDGMENTS.....	iv
ABSTRACT	v
TABLE OF CONTENTS	vii
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xii
1. INTRODUCTION	1
1.1 Sexual Assault Kit	1
1.2 Serological Test for Semen Detection	4
1.3 Backlog.....	6
1.4 STR and Y-STR.....	7
1.5 Y-screening of SAEKs.....	9
1.6 ForensicGEM [®] Sperm Lysis.....	11
1.7 Differential Extraction Using EZ1.....	12
1.8 Purpose of Study.....	13
2. MATERIALS AND METHODS	15
2.1 Data Analysis of Laboratory Casework.....	15
2.2 Mock Evidence Preparation	16
2.2.1 Non-Casework Post-Coital Swabs.....	16
2.2.2 Sample Cutting Methods for Casework Direct System.....	17

2.3 Y-screening Study.....	18
2.3.1 Casework Direct System.....	18
2.3.2 ForensicGEM® Recovery Study.....	19
2.3.3 Quantification Using qPCR.....	20
2.4 Qiagen EZ1 Extraction.....	21
2.4.1 EZ1 DNA Investigator Kit.....	21
2.4.2 ForensicGEM® Recover Study.....	22
2.4.3 Quantification Using qPCR.....	23
2.5 Statistics.....	23
2.5.1 Crime Laboratory Sexual Assault Cases Data.....	24
2.5.2 Mock Case Data.....	26
3. RESULTS.....	28
3.1 Laboratory Casework Evaluation.....	28
3.1.1 Casework Direct Analysis.....	28
3.1.2 EZ1 Evaluation.....	30
3.2 Mock Case Study.....	40
3.2.1 Distribution of DNA on a Swab Using Casework Direct System.....	40
3.2.2 Efficiency of Casework Direct System.....	42
3.2.3 Efficiency of EZ1 Extraction.....	43
4. DISCUSSION	45
4.1 Laboratory Casework Analysis.....	45

4.1.1 Accuracy of Y-screening	45
4.1.2 Efficiency of EZ1	47
4.2 Mock Case Analysis.....	48
4.2.1 Distribution of DNA on a Swab.....	48
4.2.2 Efficiency of EZ1.....	50
5. CONCLUSIONS	52
5.1 Future Directions.....	53
LIST OF JOURNAL ABBREVIATIONS	55
BIBLIOGRAPHY	56
CURRICULUM VITAE	60

LIST OF TABLES

Table 1. Thermocycler temperature cycle.....	19
Table 2. Thermocycler temperature cycle.....	23
Table 3. Distribution of total mass of female DNA in 1/2 Y-screen swabs.....	29
Table 4. Distribution of total mass of male DNA in 1/2 Y-screen swabs.....	30
Table 5. Comparative distribution of female DNA mass in 1/2 Y-Screen swab and F1 of 1/2 EZ1 Swab	35
Table 6. Distribution of male DNA mass in F1 (1/2 EZ1 swab).....	36
Table 7. Distribution of female DNA mass in F2 (1/2 EZ1 swab).....	36
Table 8. Comparative distribution of male DNA mass in 1/2 Y-Screen swab and F2 of 1/2 EZ1 Swab.....	37
Table 9. Total mass of male and female DNA on a Y-screen swab.....	40

LIST OF FIGURES

Figure 1. Comparative illustration of sample swab cutting methods for analysis	18
Figure 2. Comparison of total female DNA in 1/2 Y-screen swabs vs. 1/2 EZ1 swabs	31
Figure 3. Comparison of total female DNA in 1/2 Y-screen swabs vs. 1/2 EZ1 swabs	32
Figure 4. Comparison of total male DNA in 1/2 Y-screen swabs vs. 1/2 EZ1 swabs.....	33
Figure 5. Comparison of total male DNA in 1/2 Y-screen swabs vs. 1/2 EZ1 swabs.....	34
Figure 6. Direct comparison of male DNA in F1 vs. F2 (1/2 EZ1 swab)	38
Figure 7. Detailed comparison of male DNA in F1 vs. F2 (1/2 EZ1 swab)	38
Figure 8. The proportion of female and male DNA in F2 (1/2 EZ1 swab)	39
Figure 9. Regional distribution of male and female DNA on post-coital swabs from volunteers using method 1.....	41
Figure 10. Regional distribution of male and female DNA on post-coital swabs from volunteers using method 2.....	42
Figure 11. Total female and male DNA in F1, F2, and material section of a swab.	43
Figure 12. Proportion of female and male DNA in F1 and F2 (1/2 EZ1 swab)	44

LIST OF ABBREVIATIONS

AP	Acid Phosphatase
BP	Base Pairs
CODIS	Combined DNA Index System
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
IRB	Institutional Review Board
NIJ	National Institute of Justice
PSA	Prostate-specific Antigen
qPCR	Quantitative Polymerase Chain Reaction
SAEKs	Sexual Assault Evidence Kits
SANE	Sexual Assault Nurse Examiner
SAFER	Sexual Assault Forensic Evidence Reporting
STRs	Short Tandem Repeats
SDS	Sodium Dodecyl Sulfate
TAT	Turnaround Times
U.S.	United States
μ L	Microliter
mL	Milliliter
ng	Nanogram

1. INTRODUCTION

1.1 Sexual Assault Kit

Sexual assault is a prevalent and severe crime that has persisted in the United States (U.S.). The impact of this violence can bring devastating and long-term physical and/or psychological damage to victims of sexual assault. According to the National Crime Survey, around 1.5 million instances of rape or attempted rape occurred over a 10-year span starting from 1973 (1). Recent data from the U.S. Department of Justice's Bureau of Justice Statistics, covering the period of 2018 to 2022, reveals a concerning escalation in the number of reported rape and sexual assault cases, totaling 2.4 million incidents over 5 years (2).

The victims of sexual assault are advised to seek a medical forensic exam. The exam provides them with healthcare support for treating injuries, recommending and providing prophylactic treatment for sexually transmitted infections, assessing reproductive health concerns, and coordinating follow-up care and communication (3).

The initial rape kit in the U.S. was developed during the period of 1973 to 1975 through the collaborative efforts of Marty Goddard, a victim advocate based in Chicago, and Dr. Cynthia Porter Erie, an assistant professor at the University of Illinois at Chicago (4). Through their investigation of the problems encountered in rape prosecutions by law enforcement, state attorneys' offices, and crime laboratories across the country, they became aware of an imperative but solvable problem, the quality of collected evidence. Accordingly, they designed a kit for evidence collection and preservation, known as the

Vitullo[®] Kit (4). This kit includes standard supplies for evidence collection, such as tubes, swabs, labeled paper envelopes, glass slides, protocols, etc. The evolution of this rape kit was instrumental in establishing the framework for the Sexual Assault Evidence Kits (SAEKs), which is now commonly recognized and used in forensic investigations. The primary purpose of such kits is to meticulously collect evidence from victims while ensuring the preservation of biological materials through the use of packaging methods that prevent contamination or loss of biological materials. The components of the kit may vary according to jurisdictional guidelines. Generally, these kits include a comb, vaginal or penile swabs, rectal swabs, oral swabs, extracorporeal swabs, tools for fingernail scrapings, and containers for blood, hair, urine, and other bodily fluids collected from the victim. In addition, in some cases, the victim's undergarments are collected and packaged (5). While this comprehensive approach is designed to secure critical evidence in a manner that maintains its integrity and reliability, it is a lengthy process and significantly intrusive for the victims involved.

The SAEKs are collected by medical professionals, such as the Sexual Assault Nurse Examiner (SANE). SANEs are specially trained forensic nurses who provide post-rape medical care, support victims' psychological recovery, and document forensic evidence correctly (6). Following the collection of SAEKs, law enforcement officers take the evidence into custody and then submit it to a forensic laboratory for further examination. The samples in the kits are examined for biological evidence, including bodily fluids; if present, Deoxyribonucleic acid (DNA) can be extracted from these samples for analysis and interpretation. The completed DNA results excluding victim

profiles are uploaded to the Combined DNA Index System (CODIS), a national DNA database that connects labs by storing DNA data at the local, state, and national levels using hardware and software. The majority of this data is composed of reference DNA profiles from convicted offenders/arrestees and from crime scene evidence without a suspect (7). The DNA profiles obtained from SAEKs can be compared against these reference samples, and if they match, this may provide law enforcement with leads to identify the offender or uncover patterns of repeated offenses (8).

For most sexual assault cases, the most significant element that leads to conviction is the detection of male DNA. Thus, locating and collecting bodily fluids and DNA that have been deposited on the victims of sexual assault is imperative. While the fundamental process of identifying and characterizing bodily fluids and DNA remains pivotal in suspect identification, it has seen limited changes over time. However, it is important to highlight the substantial evolution in the sensitivity of DNA testing over the past years (5). The increased sensitivity allows for a longer time frame for the collection of biological specimens, given that semen can be collected from the surface of the body and in a body cavity for multiple days after a sexual assault (5). The purpose of examining SAEKs and increasing the sensitivity of DNA tests is to generate a CODIS eligible DNA profile. To enhance the generation of evidentiary DNA profiles, high throughput processing is a standardized approach that incorporates laboratory optimization relying on the fact that the majority of SAEKs have one common collecting substrate made up of swabs taken from different bodily fluids (5).

1.2 Serological Test for Semen Detection

Following SAEKs collection, law enforcement agencies send some kits directly to forensic labs for DNA testing, while others are retained in custody for a period of time before being forwarded to a forensic laboratory (9). Depending upon the laboratory, there are different methods used for screening and analyzing the evidence in the kits as well as different turnaround times (TAT) for testing (10). Unfortunately, many state and municipal forensic laboratories in the U.S. lack laboratory-specific TAT listed on any publicly accessible website. A 2010 study collected TAT data for the entire U.S., and only 14 laboratories provided times that met the standard (the average TAT for DNA analysis in criminal cases was 152 days) (11). More recently, the average TAT for DNA analysis in Massachusetts for the months of April through June 2023 was 90 days (12).

The process of examining a SAEK starts with creating a list and detailing all the items gathered in the kit. Samples are placed on slides, stained, and microscopically examined for the presence of spermatozoa cells. The most popular presumptive test for semen in a forensic laboratory is dependent on the identification of a semen stain on the evidence, specifically, testing for the enzyme called acid phosphatase (AP) which, in comparison to other bodily fluids, is contained in seminal plasma at a high quantity (13). Two formats of AP tests are utilized in the laboratory, including indirect and direct AP tests. The indirect AP test begins by moistening the evidence samples, which are often textiles, and applying pressure on the item surface with filter paper to transfer any possible semen stains onto the paper (14). Then, the AP reagent is applied to the filter

paper to locate the suspect semen stain by observing a color change (14). Directly applying the AP reagent to the samples, a direct AP test, is more useful when diluted semen stains are present and are then confirmed with a second AP positive stain (15).

The other presumptive test of semen is the detection of prostate-specific antigen (PSA, also known as p30), which is a glycoprotein synthesized by the prostatic gland and released into seminal plasma and serves as a reliable indicator for identifying semen in forensic evidence, even in cases involving samples from individuals who have had a vasectomy or are azoospermic (13). The test is a lateral flow immunochromatographic test that uses ABACard P30 to detect antigens in semen. If the results of these presumptive tests are positive, indicating the presence of semen, the sample is then confirmed by microscopic identification for sperm, which is a confirmatory test.

The microscopic identification of sperm is the only confirmatory test for semen because it detects the presence of sperm cells even if AP and PSA are inactive or absent. After the evidence shows a positive result of AP or PSA, it is sampled on a slide and is stained using Christmas Tree staining or Hemotoxlyin and Eosin staining to distinguish sperm cells from other cellular materials (16).

One major drawback of utilizing both preliminary and confirmatory tests to identify semen is that these methods are time-consuming and labor-intensive in forensic laboratories. Each swab from potential semen-containing evidence needs to be individually tested on a P30 card, and this can be expensive in each case since a SAEK may contain up to 10 swabs. Performing staining methods and microscopic

identification of sperm takes a considerable amount of time because each swab that shows a positive presumptive test must be sampled and placed on a slide, stained, and observed under a microscope. When the sample contains a low quantity of sperm cells, the microscopic search for sperm cells may take longer. Due to these inefficiencies use of traditional serological testing methods can contribute to backlogs.

1.3 Backlog

For a variety of reasons, some law enforcement agencies may wait to send SAEKs for varying lengths of time and reasons to the forensic laboratory. As the number of untested SAEKs increases and accumulates, a backlog of SAEKs follows. The definition of the SAEK backlog varies across jurisdictions, but it commonly refers to kits that have not been tested and are either held by law enforcement without being sent for lab analysis or have been at a crime lab for over 30 days without being processed (10). A 2021 study analyzed data from 911 counties across 15 states, including various county-level variables, estimated that between 300,000 and 400,000 SAEKs went unsubmitted across the U.S. during the years 2014 to 2018 (17). Such a large number of untested kits are caused by limited laboratory and law enforcement resources or police discretion.

The rapid development and advancement of DNA testing has improved awareness of the overall value of DNA evidence aiding in solving crimes. This has also led to an increased demand for DNA analysis. However, this huge demand may exceed the analytical capacity of the labs due to the limited resources and staffing. This shortfall is one of the reasons why law enforcement may choose not to submit SAEKs for testing. In

addition, some studies suggest that the reasons SAEKs are not tested are not only due to a shortage of resources but are more significantly influenced by concerns over victim's credibility and cooperation (18).

To address these problems, the federal government provides some grant programs to support state and local law enforcement and crime labs. With regard to sexual assault cases in particular, the National Institute of Justice (NIJ) is empowered to create procedures for effective DNA collection and processing under the Sexual Assault Forensic Evidence Reporting Act (SAFER Act). To support NIJ, a specialized SAFER Working Group was convened to focus on enhancing the laboratory processing of sexual assault evidence kits (5). They have suggested that a technique known as the Direct to DNA Y-screening approach, which is a more sensitive method for male DNA screening through quantification, be employed to assess the potential consumption of the sample (5). The high sensitivity and specificity of DNA quantification can help laboratories conserve time and resources by avoiding unnecessary testing on samples without human or male DNA, thereby effectively reducing backlogs.

1.4 STR and Y-STR

DNA contains highly repeated sequences that are prevalent throughout the human genome, making up over half of its content and are also highly prevalent in eukaryotes (19). Satellite DNA is one of the types of tandem repeats, and it can be subdivided into satellite DNA, minisatellite DNA, and microsatellite DNA (19). Autosomal short tandem repeats (STRs) are also known as microsatellite DNA, the most

used sequences for analyzing DNA in criminal cases. They involve a repetitive unit of 2-6 base pairs (bp) and account for approximately 3% of the genome (20). Most STRs are found in non-coding regions but a small amount is located within coding regions, and importantly, duplicative mutations in these regions occur frequently, with rates ranging from 1×10^{-3} to 1×10^{-4} per generation (21, 22). Such a high mutation rate leads to a great variety of STRs in humans, resulting in a genetic variation that helps to differentiate each individual for forensic purposes.

“Y-STR” refers to STRs are located on the Y chromosome, which is a single haplotype that is passed down directly from father to son and carries the male determining loci (23). The Y chromosome has fewer genes compared than other chromosomes, as more than 50% of its sequences are repetitive. When standard serology methods are unable to detect sperm or seminal fluid, Y-STR analysis is used, as it may be more sensitive than standard serology methods in some cases. The use of Y-STR testing is beneficial in sexual assault cases involving female victims and male perpetrators because of its specific amplification of male DNA. Situations that can reduce the effectiveness of male DNA detection occur when a rape victim reports the incident more than three days post incident. In that case, the collected sample is from an "extended" post-coital interval (24). The failure to obtain male DNA is because of sperm loss and lysis, which is due to the vaginal lavage and drainage, menstruation, and the normal intra-cervicovaginal sperm degradative changes that take place throughout time (25). In many sexual assault cases, the overwhelming amount of female DNA in the sample competes with male DNA, leading to a low quality of male DNA profile or

no male STR results (26). Resolutions of these problems, utilizes the Y chromosome-specific primer as a way to increase the chances of detecting a low level of a male DNA in the presence of excess female victim's DNA (25).

For forensic use, autosomal STR profiles are preferred because autosomal DNA provides stronger discrimination power than Y-STR profiles. This is due to the property of random recombination of autosomal DNA. Although Y-STR profiles are statistically weak due to haploidy and patrilineal inheritance, they are a valuable complement to autosomal STR profiles and can help to distinguish between individuals, especially when mixed samples are involved (27).

1.5 Y-screening of SAEKs

The Y-screening approach is a rapid test that is directly applied to swab evidence in the SAEKs to quickly determine the presence of male DNA, using a quantification kit without the need for a purification step. The Y-screening approach is primarily employed for samples with limited DNA, where male DNA quantities are undetectable by traditional serological methods. This method shifts the focus from conventional serological methods to screening cases using DNA quantification. This shift is significant because serology testing is usually time-consuming and labor-intensive.

The Y-screening assay, which employs a non-differential extraction method, utilizes approximately one-eighth of a swab sample. Following this, quantification is carried out to identify the amount of total human and total male DNA (5). The Y screening is either a one-extract method or part of a two-extract workflow, and both methods involve

selecting samples from the kit, extracting DNA, and quantifying it to decide which samples are suitable for STR analysis, either autosomal or Y-STR analysis, based on the quantification results (28). The one-extract workflow uses the same extract for Y-screening method and STR analysis, while the two-extract workflow uses a separate initial sample for Y-screening method and a second separate sample that undergoes its own extraction and purification process to prepare for STR analysis (28). For some laboratories, Y-screening is conducted by personnel from the Criminalistics unit, who sample from the entirety of the evidence present in SAEKs. Subsequently, the laboratory assesses the quantification results of male DNA to determine the appropriate number of samples requiring DNA analysis using techniques such as differential extraction, quantification, and amplification. It might not be necessary to process the sample any further if no male DNA is detected. The underlying assumption for using this sampling technique is that the swabs submitted are homogeneous, which theoretically enables an accurate estimation of the DNA quantity from a swab. However, it is important to recognize that this assumption may not be accurate (28). Different sampling techniques and varying amounts of samples deposited on a swab can affect the homogeneity of a swab.

One Y-screening method was developed by Promega which is known as the Casework Direct system. This is designed to quickly produce DNA lysate from forensic casework samples of different sizes and types by using a Casework Direct Reagent and 1-Thioglycerol as a reducing agent (29). The lysate from Casework Direct is compatible with multiple quantification and amplification systems, such as PowerQuant[®] and PowerPlex[®], and its non-binding or non-washing conditions reduce the possibility of DNA

loss during processing (30). The Quantifiler™ Trio DNA Quantification Kit from Applied Biosystems, a highly sensitive Real-Time Polymerase Chain Reaction (Real-Time PCR) kit, can also be utilized for quantitation of the Casework Direct lysates, and it is designed to quantify the total amount of human DNA and male DNA in the samples. The kit has four targets, including a small autosomal target, a large autosomal target, Y chromosome target, and an internal PCR control (31).

The high sensitivity of Y-screening can aid in processing SAEKs and may serve as a replacement for conventional serological methods. Y-screening also has benefits such as increased capacity, reproducibility, and objectivity; furthermore, it may require that laboratories implementing it undertake validation to determine the most effective approaches for case analysis (28).

1.6 ForensicGEM® Sperm Lysis

The forensicGEM® Sperm Lysis is an alternative Y-screening assay for DNA extraction. The forensicGEM® Sperm Lysis Kit employs MicroGEM's Acrosolv reagent to lyse sperm cells effectively and remove nucleoproteins from the DNA, avoiding the use of qPCR-inhibiting chemicals like sodium dodecyl sulfate (SDS), mercaptoethanol, and dithiothreitol (DTT) (32). A unique combination of mesophilic and thermophilic enzymes, which are activated at varying temperatures, are included in the extraction kit. At lower temperatures, mesophilic enzymes that break down cell walls are activated. A subsequent step at 75°C activates thermophilic proteinases, which are responsible for destroying nucleases, and removing nucleosomes from DNA. Finally, a 95°C step is used to

deactivate these thermophilic proteinases (33). In this direct lysis procedure, the substrate and reaction components are placed in one tube, thus theoretically allowing efficient extraction and significant preservation of DNA on the evidence. This method was used to assess the extraction efficiency of the EZ1 procedure and to determine if any DNA was left on the swab after extraction by EZ1.

1.7 Differential Extraction Using EZ1

The use of SAEKs is aimed at recovering male suspect's DNA from a female victim, with the expectation that the samples will contain both female and male DNA. To generate a high-quality DNA profile, it is critical to separate male DNA from a sample containing an overwhelming presence of female DNA. This process is achieved by performing a differential extraction.

The first protocol of differential extraction was developed in a 1985 study to separate sperm cells from vaginal epithelial cells in a cotton matrix. The process involves lysing the epithelial cells through incubation in a solution of SDS, an ionic denaturing detergent, combined with proteinase K, while sperm cells are then isolated intact via centrifugation. The supernatant, containing the female DNA, is carefully removed. Following this, the sperm pellets are lysed in a mixture of SDS, proteinase K, and DTT, which is a substance that effectively cleaves the disulfide bonds that are notably abundant in the head region of sperm cells (34). The differential extraction method successfully separates DNA from the sample into two distinct portions: a non-sperm-cell fraction (F1) and a sperm-cell fraction (F2), yet DNA carryover between these fractions is observed.

Moreover, this method involves numerous incubation and washing steps, making it quite time intensive.

To produce an optimal DNA profile from biological samples, the extraction technique must effectively separate male DNA from substantial amounts of female DNA and also isolate human DNA from other cellular materials and inhibitors. It is important to remove PCR inhibitors present in the samples while preserving the DNA. These inhibitors can interfere with DNA polymerases which can potentially lead to PCR failure. EZ1 Advanced or EZ1 Advanced XL instruments developed by Qiagen are designed to automatically purify genomic DNA from various sample types in approximately 20 minutes. The EZ1 cartridges come pre-filled with all necessary reagents, minimizing manual handling and the risk of contamination (35). The EZ1 DNA Investigator Kit utilizes magnetic-particle technology to bind DNA and incorporates a wash buffer to remove inhibitors (35). The purified DNA is then ready for direct use in downstream STR analysis.

1.8 Purpose of Study

The purpose of employing Y-screening in crime laboratories is to save time traditionally spent on serological tests for examining SAEKs, thereby reducing backlogs. Some crime labs use Y-screening to determine the amount of male DNA in a sample, and then use the results from qPCR to decide the appropriate quantity of the sample to send to the DNA unit for further analysis. This study evaluates real sexual assault case data analyzed by a crime laboratory to determine if Y-screening results were able to predict the

amounts of male and female DNA present in the samples. Additionally, the differential extraction efficiency of the EZ1 system was assessed. To investigate the lysis efficiency of Y-screening, mock case samples were prepared. Casework Direct system and Acrosolv were used to check for any DNA residue left on swabs. Moreover, Y-screening was used to determine the DNA distribution on a sample. The extraction efficiency of EZ1 was also examined to ascertain whether any DNA was still present on the used swabs.

2. MATERIALS AND METHODS

2.1 Data Analysis of Laboratory Casework

The initial step of the study involved evaluating quantitative PCR data (qPCR) that was devoid of any sample identification data from a crime laboratory's sexual assault cases. This data encompassed Y-screening information using the Casework Direct system, EZ1 information using the Qiagen EZ1 Advanced XL Robot, and corresponding qPCR outcomes obtained with the PowerQuant[®] System via the ABI 7500 Real-Time PCR system. Specifically, the data presents the concentration values of autosomal and male DNA found on the swab using the Y-screening assay. In the process of Y-screening, analysts initially examined all swabs from SAEKs. Each sample within these kits may contain multiple swabs. For the screening, analysts cut $\frac{1}{8}$ of each swab from the same sample and combined these portions into a single tube (designated as the Y-screen swab). This combined sample undergoes Y-screening to identify the sample with the highest concentration of male DNA. Once determined, a consistent portion from all swabs of this sample (for example consistent portions of 4 vaginal swabs) was cut and put into another tube, which was then sent to the DNA unit for further analysis. The data also includes the DNA concentration values of autosomal and male DNA from fraction 1 (F1) and fraction 2 (F2), and the corresponding amounts of the swab utilized for testing (designated as the EZ1 swab). Before quantifying the EZ1 extracts, the sample volumes were concentrated to a final volume either 13 μ L or 14 μ L using Vivacon 500 Concentrator. Quantification was done using these concentrated extracts. Validation studies have demonstrated that using the Vivacon 500 Concentrator may result in a loss of approximately 20% of the

sample (36). For the purposes of the study, this 20% loss was compensated for by adding an equivalent amount back to the DNA total. The aim of analyzing this data was to compare the initial amount of DNA present on the swab using the Y-screening with the amount of DNA recovered via the EZ1 instrument. To ensure the data sets were comparable, DNA amounts obtained from the swab used for both Y-screening and EZ1 were normalized to $\frac{1}{2}$ swab.

In this dataset, it was observed that in the F2 from 13 of 90 cases, the concentration of male DNA exceeded that of small autosomal DNA. To address this issue, validation data from a virtual standard curve was employed. The average and standard deviation (SD) of differences between varying concentrations of small autosomal and male DNA were calculated. Based on these calculations, F2 samples were excluded from calculations if the difference between the small autosomal DNA and male DNA concentrations exceeded two SD from the established range. This difference affected 4 cases. If the difference did not exceed this range, the concentration of male DNA was adjusted to match that of the autosomal DNA.

2.2 Mock Evidence Preparation

2.2.1 Non-Casework Post-Coital Swabs

To compare the effectiveness of Y-screening and use of differential extraction on the EZ1, post-coital samples were obtained from unidentified volunteers following the laboratory institutional review board (IRB) procedure. The mock evidence samples were used to mimic vaginal swabs collected from SAEKs. Two unidentified female donors

(Donor A and Donor B) self-collected the samples within 12 hours after intercourse, using four cotton swabs each, resulting in a total of eight post-coital swabs. All samples were then dried and stored in envelopes until used.

2.2.2 Sample Cutting Methods for Casework Direct System

Two cutting methods were used for two swabs per each donor to determine the distribution of female DNA and male DNA located on the swab (Figure 1): Method #1 involved sectioning a swab into six pieces. Initially, the tip of the swab was sliced, and this portion was divided into two parts, labeled *1-1* and *1-2*. Subsequently, the remaining shaft of the swab was vertically sliced into four equal sections, resulting in a total of six pieces. These additional segments were designated as *2-1*, *2-2*, *2-3*, and *2-4*. Method #2 also partitioned a swab into six distinct portions but with a specialized approach for separating materials. Initially, the tip of the swab was cut off similarly to Method #1. All cotton from this section was then meticulously removed, leaving behind only the outermost layer which counted as two separate parts. The remainder of the swab was subsequently divided vertically into two halves. From these sections, the inner cotton was carefully removed, leaving the outer layers. The segments from the top were designated as *1-1 O* for the outer layer and *1-2 In* for the inner cotton. Similarly, one of the bottom halves were labeled as *2-1 O* for the outer layer and *2-2 In* for the inner cotton, with *3-1 O* for the outer layer and *3-2 In* for the inner cotton of the second half.

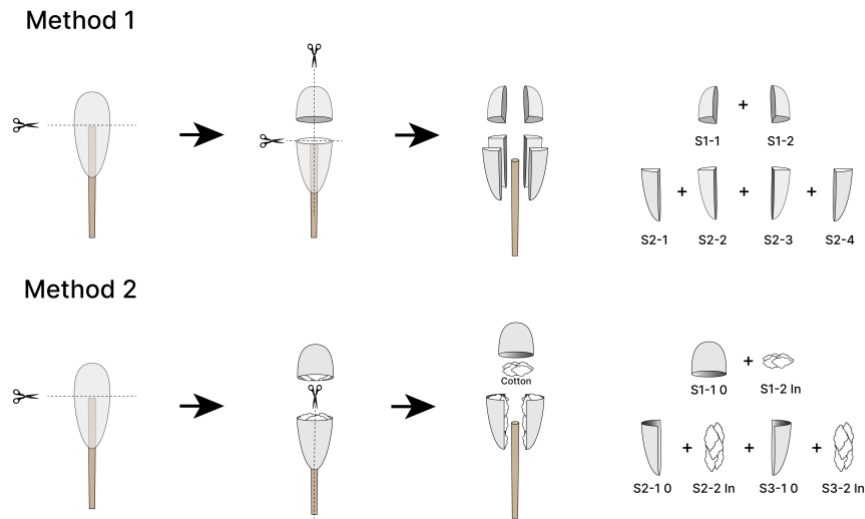


Figure 1. Comparative illustration of sample swab cutting methods for analysis. Method 1 analyzed the DNA distribution across different sections of a swab. Method 2 examined the distribution of DNA between the outer layer and the cotton interior.

2.3 Y-screening Study

2.3.1 Casework Direct System

Each cutting of the post-coital swabs, using the two methods, was placed inside an Investigator Lyse&Spin Basket with a 2 mL collection tube (Qiagen, Hilden, Germany). To prepare the Casework Direct solution, 1 μ L of tenfold diluted 1-Thioglycerol was added to 200 μ L of Casework Direct Reagent for each cutting, and subsequently, 200 μ L of the solution was added to each basket, according to the manufacturer's recommended protocol (29). Those tubes were incubated in a shaking heat block (Thermo Fisher Scientific, Waltham Massachusetts) at 70°C for 30 minutes. After incubation, the tubes were vortexed on a vortex mixer (Thermo Fisher Scientific, Waltham Massachusetts) for 10 seconds and then centrifuged in a Centrifuge 5418 (Eppendorf, Hamburg, Germany) at room temperature for 5 minutes at 14,000 rpm. Some of the cuttings were centrifuged for an

additional 5 minutes to move all the liquid to the tube. The lysate was directly used for quantification. All cuttings were preserved in different tubes in a freezer at -18 °C for the forensicGEM/Acrosolv test.

2.3.2 ForensicGEM® Recovery Study

The cuttings used for the Casework Direct lysis described above were thawed and were placed into 0.5 mL reaction tubes (Thermo Fisher Scientific, Waltham Massachusetts). To digest cells remaining on the cotton, 10 µL of Acrosolv, 10 µL of 10X RED+ buffer, 2 µL of ForensicGEM from the ForensicGEM® Sperm kit (MicroGEM, Charlottesville Virginia), and 78 µL of ultrapure water were added to the tubes, bringing the total volume to 100 µL to adequately cover the cotton. The tubes were then placed in a Veriti™ 60 Well Thermal Cycler (Thermo Fisher Scientific, Waltham Massachusetts) and subjected to an 11-minute temperature cycle as prescribed in the manufacturer's protocol (Table 1) (33). After incubation, the tubes were briefly centrifuged, and all cotton was transferred to filterless baskets in 2 mL microcentrifuge tubes. These tubes were then centrifuged at 14,000 rpm for 3 minutes. The baskets and cotton were discarded, and the liquid was transferred back into the original 0.5 mL microcentrifuge tubes used for the forensicGEM digest. The digested samples were ready for quantification using qPCR.

Table 1. Thermocycler temperature cycle. The recommended temperature cycle for the effective use of the forensicGEM® Sperm kit, developed by MicroGEM.

Step	Time (min)	Temperature (°C)
1	5	52
2	3	75
3	3	95

2.3.3 Quantification Using qPCR

Casework Direct lysates were diluted to a 1:10 ratio. From the lysates, 10 μL was pipetted into 90 μL of TE buffer to achieve a 1:10 dilution. The TE buffer, prepared in the DNA laboratory of the Biomedical Forensic Sciences program, consisted of 10 mM Tris at pH 8.0 and 0.1 mM Ethylenediaminetetraacetic acid (EDTA). The tubes containing lysates were briefly vortexed and centrifuged before being used. For quantification, 2 μL of the diluted extracts were added to a well containing 18 μL of the qPCR master mix that was prepared using the Quantifiler™ Trio DNA Quantification Kit (Applied Biosystems™, Thermo Fisher Scientific, Waltham Massachusetts), in accordance with the manufacturer's suggested protocol (31). Two wells of the positive control and one well of a negative control, using TE buffer, were run with the samples for quality control. The qPCR was conducted on a 7500 Real Time PCR System (Applied Biosystems™, Thermo Fisher Scientific, Waltham Massachusetts), using a virtual standard curve for assistance (37).

The concentration results obtained from the qPCR for Casework Direct lysates were multiplied by the total volume of extraction (200 μL) and compensated for the 1:10 dilution to convert to nanograms (ng). The concentration results obtained from qPCR for forensicGEM lysates were also multiplied by the total volume of extraction (100 μL) to convert to ng.

2.4 Qiagen EZ1 Extraction

2.4.1 EZ1 DNA Investigator Kit

Two swabs from donor B were used for the EZ1 study. Each swab was sliced in half in preparation for the EZ1 pretreatment of mixture samples involving epithelial and sperm cells, according to the manufacturer's suggested protocol (35). Each half of the sample was placed in a 2 mL microcentrifuge tube (Eppendorf, Hamburg, Germany), and 475 μ L of buffer G2 along with 25 μ L of Proteinase K were added to the tube. The tubes were then vortexed for 10 seconds before being incubated at 56°C for 1 hour and 10 minutes at 900 rpm in a thermomixer. After incubation, the tubes were briefly centrifuged, and the solid samples were transferred to filterless spin baskets (Costar®, Salt Lake City Utah). The baskets were placed back in the sample tubes and then centrifuged at 14,000 rpm for 5 minutes. All the cotton was preserved in another 2 mL microcentrifuge tube for the forensicGEM One-Step Lysis study. The supernatant was then transferred to a new 2 mL tube as the epithelial-cell fraction (F1), ready to be purified. The sperm cell pellet in the tubes (F2) was washed by adding 500 μ L of Buffer G2 and subsequently centrifuged at 14,000 rpm for 5 minutes before discarding the supernatant. The sperm cell pellet was washed only once to preserve the quantity of male DNA in the sample. After the washing step, 160 μ L of Buffer G2, 10 μ L of Proteinase K, and 40 μ L of 1 M DTT were added to the tubes and vortexed for 10 seconds. Then, the tubes were incubated at 70°C for 10 minutes at 900 rpm in a thermomixer. After incubation, the tubes were briefly centrifuged. For the F1, 400 μ L of Buffer MTL was added to each sample tube. Then, 1 μ L of carrier RNA (cRNA) was added to all samples.

For the purification of samples, the Trace Protocol was performed on F2, and the Large-Volume Protocol was used for F1 using the EZ1[®] DNA Investigator[®] Kit (Qiagen, Hilden, Germany) according to the manufacturer's suggested protocol (35). The EZ1[®] DNA Investigator[®] Card was inserted into the card slot, and the EZ1 instrument was turned on. On the instrument's screen, "START" was pressed to initiate the protocol setup. Button "1" was selected for the Trace Protocol for F2, and 100 μ L of TE buffer was chosen as the elution buffer. Button "3" was selected for the Large-Volume Protocol for F1, and 100 μ L of TE buffer was chosen as the elution buffer. The reagent cartridges were taken out from the box and inverted to mix the magnetic particles until completely resuspended. The cartridges were then loaded into the cartridge rack. Opened elution tubes, tip holders containing filter-tips, and opened sample tubes were placed properly onto the tip rack. The instrument door was closed, and the purification procedure was started. After purification, the elution tubes containing purified DNA were ready for quantification using qPCR.

2.4.2 ForensicGEM[®] Recovery Study

Each cotton piece recovered from the EZ1 extraction was re-extracted using the ForensicGEM[®] Sperm kit (MicroGEM, Charlottesville Virginia). Because the cotton pieces were relatively larger, 10 μ L of Acrosolv, 20 μ L of 10X RED+ buffer, 2 μ L of ForensicGEM, and 168 μ L of ultrapure water were added to the tubes, bringing the total volume to 200 μ L to adequately cover the cotton. The tubes were then placed in a Veriti[™] 60 Well Thermal Cycler and subjected to a 22-minute temperature cycle. Given that the volume (200 μ L) used to cover the cotton was twice the recommended volume (100 μ L),

the incubation time in the thermal cycler was also doubled to a total of 22 minutes (Table 2).

Table 2. Thermocycler Temperature Cycle. This temperature cycle is twice the recommended temperature for use with the forensicGEM[®] Sperm kit by MicroGEM.

Step	Time (min)	Temperature (°C)
1	10	52
2	6	75
3	6	95

2.4.3 Quantification Using qPCR

To determine the amount of human DNA in the samples after EZ1 extraction and purification, as well as the amount remaining on the cotton after forensicGEM treatment, the DNA extracts from both experiments were quantified using the Quantifiler[™] Trio DNA Quantification Kit (Applied Biosystems[™], Thermo Fisher Scientific, Waltham Massachusetts), in accordance with the manufacturer's suggested protocol (31). The qPCR was conducted on a 7500 Real Time PCR System (Applied Biosystems[™], Thermo Fisher Scientific, Waltham Massachusetts), using a virtual standard curve for assistance (37).

2.5 Statistics

The statistical analysis of this study was conducted by using JMP[®] statistical software (JMP Statistical Discovery LLC, Cary North Carolina) and Microsoft Excel (Microsoft, Redmond Washington).

2.5.1 Crime Laboratory Sexual Assault Cases Data

The first step was to calculate the total nanograms (ng) of autosomal and male DNA present in the Y-screen sample based on the qPCR results. The concentration values were multiplied by the volume of Casework Direct reagent (200 μ L) added to the sample tubes:

$$(1) \text{ total ng of DNA in Y screen swab} = \text{concentration value} \left(\frac{\text{ng}}{\mu\text{L}} \right) \times \text{Lysate volume} (\mu\text{L})$$

As outlined in Section 2.1, the analyst cuts 1/8 of each swab from the sample and places all these portions into one tube. The total of these portions equals the number of swabs used for the Y-screening assay. For instance, if there are four vaginal swabs, $\frac{1}{8}$ is cut from each, and these portions are combined to equal half a swab. This amount is labeled as the Y-screen swab. Similarly, EZ1 swab also indicates the quantity of swab used in the EZ1 process. The Y-screen swab was normalized to $\frac{1}{2}$ swab to be comparable:

$$(2) \text{ total ng of DNA in Y screen swab (normalized)} = \frac{0.5}{\text{number of swabs used for Y screening}} \times \text{total ng of DNA in Y screen swab}$$

All calculations were normalized. The total ng of female DNA in the Y-screen swab was calculated by subtracting the total ng of male DNA from the total ng of small autosomal DNA. There were a few cases where the male DNA concentrations were similar to the small autosomal DNA concentrations, except that the small autosomal concentrations were slightly higher. This may be due to variation in quantification results between the small autosomal values and the Y values. These discrepancies were not accounted for in this study. We attributed these differences to be female DNA, understanding that the actual values of male DNA might be slightly higher due to variation between the small autosomal and male DNA concentration values. In instances where the DNA concentrations

measured by the small autosomal probe are greater than those measured by the Y probe, the true concentration of male DNA may be as large as the small autosomal value if the solution contains only male DNA. If this information is needed in a case situation, the determination could be found by examination of DNA profiles.

The second step was to calculate the total ng of autosomal and male DNA present in the EZ1 swab by multiplying the concentrated volume:

$$(3) \text{ total ng of DNA in EZ1 swab} = \text{concentration value} \left(\frac{\text{ng}}{\mu\text{L}} \right) \times \text{concentrated volume} (\mu\text{L})$$

The EZ1 swab was normalized to $\frac{1}{2}$ swab to be comparable:

$$(4) \text{ total ng of DNA in EZ1 swab (normalized)} = \frac{0.5}{\text{number of swabs used for EZ1}} \times \text{total ng of DNA in EZ1 swab}$$

The 20% of DNA that was originally lost from the sample due to the concentration step was added back in for the purposes of the study:

$$(5) \text{ total ng of DNA in EZ1 swab without loss (normalized)} = \text{total ng of DNA in EZ1 swab (normalized)} \times 1.2$$

To compare the proportions of female and male DNA in EZ1 swab, the ng of female and male DNA was each divided by the ng of autosomal DNA in the EZ1 swab:

$$(6) \text{ proportion of female or male DNA in EZ1 snippet without loss (normalized)} = \frac{\text{total ng of female or male DNA in EZ1 swab (normalized)}}{\text{total ng of autosomal DNA in EZ1 swab}} \times 100\%$$

The percentage distribution of female and male DNA in fractions F1 and F2 was calculated by dividing the ng of female or male DNA in each fraction by the total ng of female or male DNA in the EZ1 swab, respectively:

$$(7) \% \text{ of distribution} = \frac{\text{ng of female or male DNA in F1 or F2}}{\text{Total ng of female or male DNA in EZ1 swab (F1+F2)}} \times 100\%$$

The analysis comparing the differences in DNA between females and males was conducted by generating descriptive statistics, which included means (\bar{x}), SD ($SD = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \mu)^2}$), minimum and maximum values, and first (Q1) and third quartile (Q3).

2.5.2 Mock Case Data

The first goal of using qPCR was to determine the initial concentration of human DNA present on each section of the volunteers' sample swabs using the Casework Direct system. These concentrations were then converted to ng to analyze the distribution of human DNA on the swabs, distinguishing between female and male DNA. To calculate the undiluted mass of autosomal and male DNA on each swab section, the concentration value obtained from qPCR was multiplied by the volume of Casework Direct reagent (200 μ L) and further multiplied by a dilution factor of 10:

$$(8) \text{ ng of DNA in each section of a swab} = \text{concentration value} \left(\frac{\text{ng}}{\mu\text{L}} \right) \times \text{Lysate volume} (\mu\text{L}) \times 10$$

The amount of female DNA in each swab section was calculated by subtracting the ng of male DNA from the ng of autosomal DNA. The total ng of female and male DNA for each swab was then obtained by summing the ng of female and male DNA across all sections. To compare the proportions of female and male DNA in each section of a swab, the amounts of female and male DNA were each divided by the ng of small autosomal DNA in the same section:

$$(9) \text{ proportion of male DNA in the section} = \frac{\text{ng of male DNA in the section}}{\text{ng of small autosomal DNA in the section}} \times 100\%$$

The next step was to compare the regional distribution of female and male DNA on the swab. The amounts of female and male DNA in each section were divided by the total ng of female and male DNA on the swab, respectively:

$$(10) \textit{ regional distribution of male DNA in the section} = \frac{\textit{ng of male DNA in the section}}{\textit{total ng of male DNA on the swab}} \times 100\%$$

The second goal of using qPCR was to determine the efficiency of the EZ1 instrument by comparing the concentration values obtained from qPCR after purification by EZ1 with those obtained after forensicGEM lysis. To calculate the ng of autosomal and male DNA present on each half of the swab and to determine how much DNA was left on the swab after forensicGEM lysis, the concentration values obtained from qPCR were multiplied by the elution volume (100 μ L) or the lysate volume (200 μ L).

$$(11) \textit{ ng of DNA in half of a swab} = \textit{concentration value} \left(\frac{\textit{ng}}{\mu\textit{L}} \right) \times \textit{elution or lysate volume} (\mu\textit{L})$$

The proportions of female and male DNA in F1 and F2 were also calculated. These proportions from the EZ1 study were then compared with those found in the Casework Direct study and actual casework data to assess the correlation between volunteer samples and laboratory samples from real cases. To determine the percentage of DNA remaining on the swab in relation to the total ng of DNA present, the human DNA in the material was divided by the total human DNA on the swab, which includes the DNA found in F1, F2, and the material:

$$(12) \textit{ percentage of DNA remaining on the swab} = \frac{\textit{ng of DNA in M}}{\textit{total ng of DNA in F1+F2+M}} \times 100\%$$

3. RESULTS

3.1 Laboratory Casework Evaluation

3.1.1 Casework Direct Analysis

In this study, the quantification results of female and male DNA from a half of a swab from laboratory samples were analyzed across 90 sexual assault cases. The Casework Direct system was used to obtain the total ng of both female and male DNA per sample. Table 2 presents a distribution of the total mass of female DNA detected in $\frac{1}{2}$ swab of the Y-screen data. It categorizes the female DNA mass into defined ranges and lists the proportion of cases corresponding to each range. The data demonstrated a wide range of total ng of female DNA, with a maximum value of 44072.16 ng and a minimum value of 0.62 ng. The mean ng of total female DNA across the cases was 6113.19 ng, with a standard deviation of 9077.55 ng, indicating substantial variability. The Q1 was at 1326.48 ng, and the Q3 was at 6226.57 ng. The data indicated that a majority of samples have total female DNA mass ranging from 1000 ng to 5000 ng, accounting for 53.3% of the total samples. This data represented a significant variability in the amount of female DNA collected from real cases.

Table 3. Distribution of total mass of female DNA in 1/2 Y-screen swabs. The female DNA was categorized into five ranges, and each category demonstrated the percentage of cases falling within that specific DNA mass range.

Total Female DNA, ng (1/2 Y-screen Swab)	Percentage of Cases
0 ng - 1 ng	1.1%
1 ng - 100 ng	0%
100 ng - 1000 ng	16.7%
1000 ng - 5000 ng	53.3%
5000 ng - 10000 ng	15.6%
> 10000 ng	13.3%

A distribution of the total ng of male DNA present in the Y-screen for these same $\frac{1}{2}$ swabs is depicted in Table 3. The data reveal a narrower range of total ng of male DNA compared to female DNA, with a maximum of 675.72 ng and a minimum of 0.06 ng. The mean total ng of male DNA across the cases is 39.03 ng, with a standard deviation of 117.27 ng, indicating a high degree of dispersion. The Q1 is at 0.39 ng, with approximately 40% of the cases having less than 1 ng of male DNA. The Q3 is at 13.85 ng, and approximately 8.9% of the cases have more than 100 ng of male DNA. The data indicates that the majority of male DNA quantities obtained from half of a Y-screen swab were less than 10 ng. Furthermore, the lowest quantity of male DNA obtained from half of a swab was only about 0.06 ng. This suggests that the amount of swab material sent to the DNA analysis unit should exceed half a swab to achieve optimal amplification target, according to the Y-screening data.

Table 4. Distribution of total mass of male DNA in 1/2 Y-screen swabs. The male DNA was categorized into five ranges, and each category demonstrated the percentage of cases falling within that specific DNA mass range.

Total Male DNA, ng (1/2 Y-screen Swab)	Percentage of Cases
0 ng - 0.1 ng	7.8%
0.1 ng - 1 ng	32.2%
1 ng - 10 ng	32.2%
10 ng - 100 ng	18.9%
> 100 ng	8.9%

3.1.2 EZ1 Evaluation

The quantity of EZ1 swab used for EZ1 extraction was adjusted by the laboratory based on the initial Y-screening results. Specifically, a greater amount of swab was utilized for the EZ1 extraction when low DNA concentrations were indicated by the Y-screening. For the purposes of this study, the EZ1 swab size was standardized to half a swab.

To effectively compare the total ng of female DNA in half of the EZ1 swab with that detected in half of the Y-screen swab, Figures 2 and 3 are organized in ascending order based on the total ng of female DNA found in the Y-screen swab. Figure 2 displays the right half of the figure, while figure 3 shows the left half, highlighting the contrasts in DNA quantities between the two treatments. The maximum observed ng of female DNA in the EZ1 swab was 2886.09 ng, with the minimum being 0.34 ng. Across all cases, the average ng of female DNA in the EZ1 swab was 615.67 ng with a SD of 600.51 ng. This mean is approximately 10 times lower than that of the Y-screen swab, which also demonstrated a significantly lower SD, indicating less variability in the DNA quantities collected through the EZ1 extraction and purification process. The Q1 and Q3 are marked at 49.79 ng and

900.87 ng, respectively. Within the EZ1 study data, 77.8% of cases had female DNA mass below 1000 ng, in contrast to only 17.8% in the Y-screen study. Notably, 6 out of 90 cases exhibited higher mass of female DNA in the EZ1 swab compared to the Y-screen swab, occurring exclusively when the ng of female DNA in the Y-screen swab was below 1000 ng (Figure 3). The higher amount of female DNA collected in the Y-screen swab does not necessarily correlate with higher quantities in the EZ1 swab (Figure 2), indicating that the Y-screening assay does not accurately predict the ng of female DNA determined by the EZ1 method. It can be concluded that some female DNA was lost during the extraction and purification process using the EZ1 instrument.

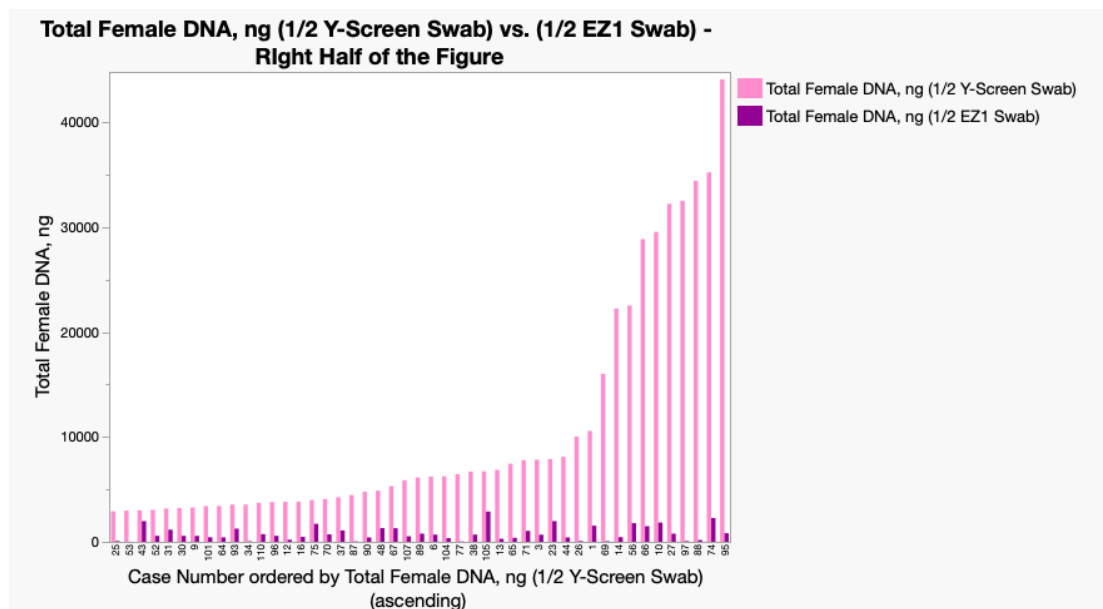


Figure 2. Comparison of total female DNA in 1/2 Y-screen swabs vs. 1/2 EZ1 swabs. This figure depicts the half of the cases that contained the highest amount of total female DNA in Y-screen swabs. The pink color indicates female DNA in the Y-screen swabs, and the purple color indicates female DNA in the EZ1 swabs.

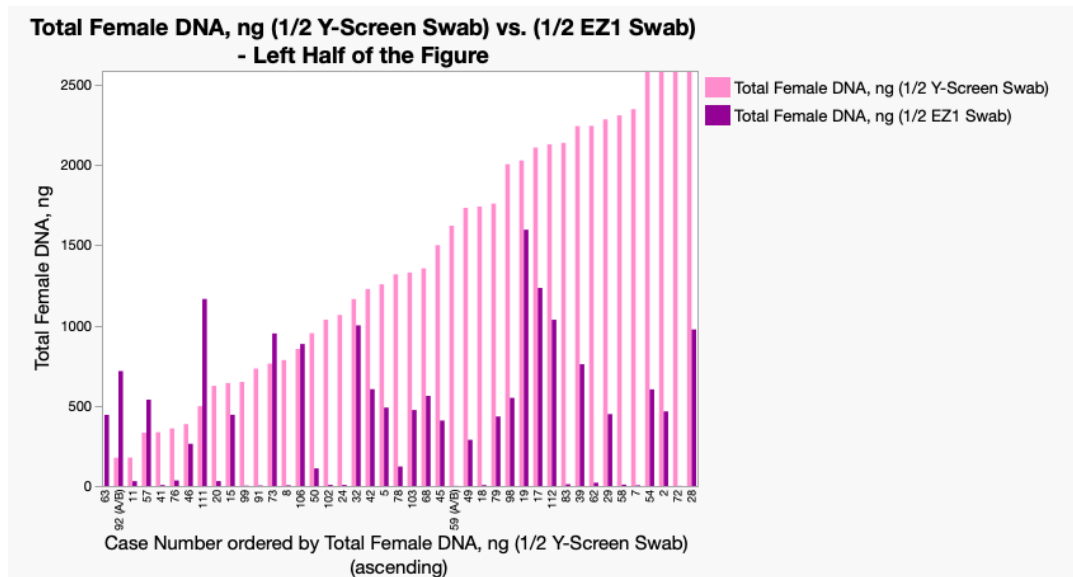


Figure 3. Comparison of total female DNA in ½ Y-screen swabs vs. ½ EZ1 swabs. This figure depicts the half of the cases that contained the lowest amount of total female DNA in Y-screen swabs. The pink color indicates female DNA in the Y-screen swabs, and the purple color indicates female DNA in the EZ1 swabs.

To effectively compare the total ng of male DNA in half of EZ1 swab with that in half of Y-screen swab, Figures 4 and 5 were organized in ascending order based on the total ng of male DNA found in half of Y-screen swab. Figure 4 displays the right half of the figure, and figure 5 shows the left half. The maximum ng of male DNA observed in half of EZ1 swab was 507.38 ng, while the minimum was 0 ng. The average ng of male DNA and SD across all cases was 21.29 ± 74.75 ng. The Q1 and Q3 were marked at 0.040 ng and 4.29 ng, respectively. Furthermore, 61.1% of cases exhibited total ng of male DNA below 1 ng, with 35.6% of these having less than 0.1 ng (Figure 5). Notably, 5 cases showed no detectable male DNA found in half of the EZ1 swabs. The average total ng of male DNA observed in half of the Y-screen swab was 39.03 ng, approximately 2 times the amount found in half of the EZ1 swab. Significantly, within the dataset, 11 out of 90 cases

exhibited higher amount of male DNA in half of the EZ1 swab compared to those in half of the Y-screen swabs. However, for the majority of cases, the male DNA quantities obtained by the Y-screening assay did not accurately predict those determined by the EZ1 method. The loss of male DNA can be concluded.

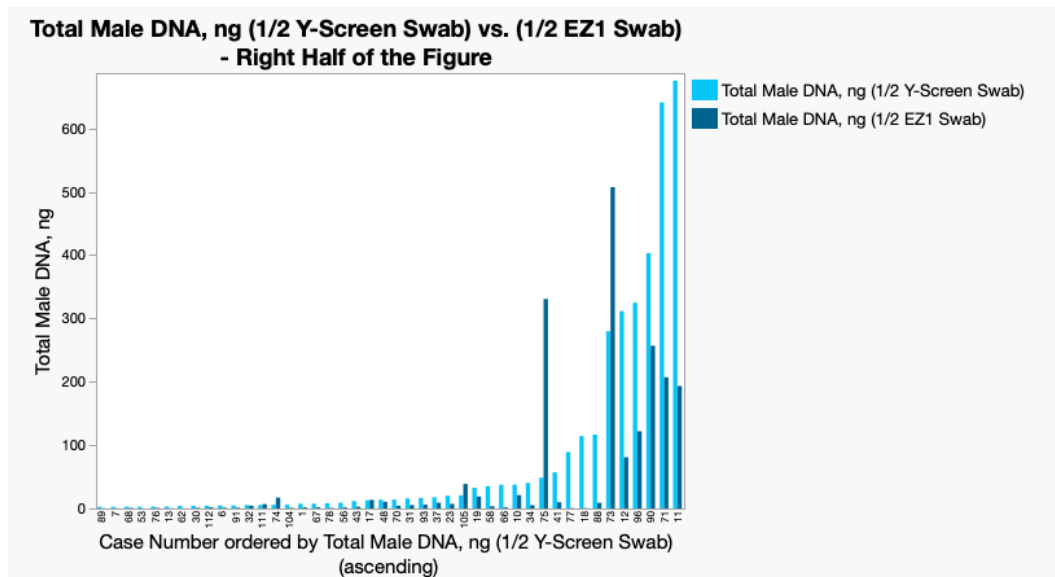


Figure 4. Comparison of total male DNA in 1/2 Y-screen swabs vs. 1/2 EZ1 swabs. This figure depicts the half of the cases that contained the highest amount of total male DNA in Y-screen swabs. The light blue color indicates male DNA in the Y-screen swabs, and the dark blue color indicates male DNA in the EZ1 swabs.

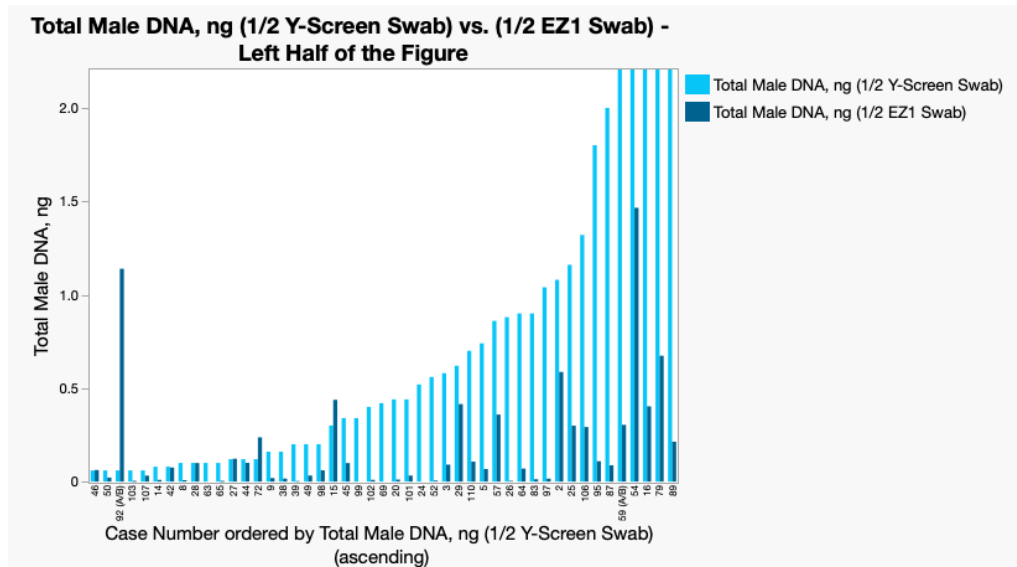


Figure 5. Comparison of total male DNA in ½ Y-screen swabs vs. ½ EZ1 swabs. This figure depicts the half of the cases that contained the lowest amount of total male DNA in Y-screen swabs. The light blue color indicates male DNA in the Y-screen swabs, and the dark blue color indicates male DNA in the EZ1 swabs.

Female and male DNA in fractions F1 and F2 were separately evaluated, theoretically ensuring the majority of female DNA was present in F1 and the majority of male DNA in F2 for effective STR profile analysis.

To assess the differential extraction procedure of the EZ1 system, the distribution of ng of female DNA detected in F1 from half an EZ1 swab is summarized and compared to that of female DNA detected in half a Y-screen swab (Table 5). It categorizes the female DNA mass into defined ranges and lists the proportion of cases corresponding to each range. As expected, F1 contained large quantities of female DNA, with an average of 612.84 ng and a standard deviation of 599.21 ng. This average closely corresponds to the total average ng of female DNA in the entire EZ1 swab, confirming that the majority of female DNA was located in F1. The maximum amount of female DNA observed in F1 was

2878.51 ng, with the minimum at 0 ng, including two cases with no detectable female DNA.

The Q1 and Q3 were marked at 43.07 ng and 900.54 ng, respectively. Notably, almost 13.3% of cases in F1 contained less than 10 ng of female DNA.

Table 5. Comparative distribution of female DNA mass in 1/2 Y-Screen swab and F1 of 1/2 EZ1 Swab. A side-by-side comparison of the distribution of female DNA mass measured through two distinct methods: the Y-screening and EZ1. The female DNA is categorized into different ranges, and each category depicts the percentage of cases falling within that specific DNA mass range.

Total Female DNA, ng (1/2 Y-screen Swab)	Percentage of Cases	Female DNA in F1, ng (1/2 EZ1 Swab)	Percentage of Cases
0 ng - 100 ng	1.1%	0 ng - 1 ng	2.2%
100 ng - 1000 ng	16.7%	1 ng - 10 ng	11.1%
1000 ng - 5000 ng	53.3%	10 ng -100 ng	13.3%
5000 ng - 10000 ng	15.6%	100 ng - 1000 ng	52.2%
> 10000 ng	13.3%	1000 ng - 3000 ng	21.1%

To further assess the differential extraction procedure using the EZ1 system, the distribution of male DNA in F1 of half of the EZ1 swab was analyzed (Table 6). The average amount of male DNA in F2 was notably low, recorded at just 1.83 ng, with a standard deviation of 6.05 ng. This result is consistent with expectations given the typically small quantity of male DNA in total DNA samples. The maximum ng of male DNA observed in F1 was 51.89 ng, with a minimum of 0 ng. The Q1 and Q3 were recorded at 0 ng and 1.00 ng, respectively. Notably, 27.8% of the cases had no detectable male DNA in F1. Within the dataset, 20.0% of cases contained male DNA ranging from 1 ng to 10 ng, and 4.4% had ng of male DNA that were higher than 10 ng.

Table 6. Distribution of male DNA mass in F1 (1/2 EZ1 swab). The male DNA is categorized into seven ranges, and each category depicts the percentage of cases falling within that specific DNA mass range.

Male DNA in F1, ng (1/2 EZ1 Swab)	Percentage of Cases
0	27.8%
0 ng - 0.01 ng	5.6%
0.01 ng - 0.1 ng	20.0%
0.1 ng - 0.5 ng	14.4%
0.5 ng - 1 ng	7.8%
1 ng - 10 ng	20.0%
> 10 ng	4.4%

The distribution of female DNA in F2 for half of an EZ1 swab was also analyzed (Table 7). In contrast to the female DNA found in F1, the average amount of female DNA in F2 was 2.83 ng, with a standard deviation of 8.58 ng. The maximum amount recorded was 60.82 ng, while the minimum was 0 ng. The Q1 and Q3 were noted at 0.16 ng and 1.21 ng, respectively. The majority of cases displayed female DNA amounts ranging from 0.1 ng to 1 ng, although, 27.8% of the cases had female DNA amounts exceeding 1 ng.

Table 7. Distribution of female DNA mass in F2 (1/2 EZ1 swab). The female DNA is categorized into five ranges, and each category depicts the percentage of cases falling within that specific DNA mass range.

Female DNA in F2, ng (1/2 EZ1 Swab)	Percentage of Cases
0 ng	8.9%
0 ng – 0.1 ng	10.0%
0.1 ng – 1 ng	53.3%
1 ng – 10 ng	22.2%
> 10 ng	5.6%

The distribution of male DNA in F2 of half of an EZ1 swab is presented and compared to that of male DNA detected in half a Y-screen swab (Table 8). Contrary to F1,

a substantial amount of male DNA was expected to be found in F2. The average amount of male DNA was 19.46 ng, with a range from 0 to 492.74 ng and a standard deviation of 71.61 ng. The Q1 and Q3 were recorded at 0.0013 ng and 1.19 ng, respectively. Although the average mass of male DNA in F2 was much higher than that of male DNA in F1, excluding the six cases containing more than 100 ng of male DNA reduces the average to 2.70 ng, which is only double that found in F1. The majority of cases, approximately 72.2%, had less than 0.75 ng of male DNA in F2, which was unexpected based on the Y-screen results.

Table 8. Comparative distribution of male DNA mass in 1/2 Y-Screen swab and F2 of 1/2 EZ1 Swab. A side-by-side comparison of the distribution of male DNA mass measured through two distinct methods: the Y-screening and EZ1. The male DNA is categorized into different ranges, and each category depicts the percentage of cases falling within that specific DNA mass range.

Total Male DNA, ng (1/2 Y-screen Swab)	Percentage of Cases	Male DNA in F2, ng (1/2 EZ1 Swab)	Percentage of Cases
0 ng - 0.1 ng	7.8%	0 ng	24.4%
0.1 ng - 1 ng	32.2%	0 ng - 0.01 ng	11.1%
1 ng - 10 ng	32.2%	0.01 ng - 0.1 ng	15.6%
10 ng - 100 ng	18.9%	0.1 ng - 1 ng	22.2%
> 100 ng	8.9%	1 ng - 10 ng	14.4%
		10 ng - 100 ng	5.6%
		> 100 ng	6.7%

Surprisingly, a comparison of male DNA in F1 to that in F2 (Figures 6 & 7) revealed that although the average and maximum mass of male DNA in F2 were higher than in F1, 37.8% of cases had more male DNA in F1, with differences ranging from 0.0042 ng to 13.52 ng. Notably, within this 37.8% of cases, half of them had male DNA present only in F1 and not in F2 (Figure 7). In addition, 5 cases detected no male DNA present in

F1 and F2 while those corresponding Y-screen swabs had male DNA ranging from 0.1 ng to 2.78 ng.

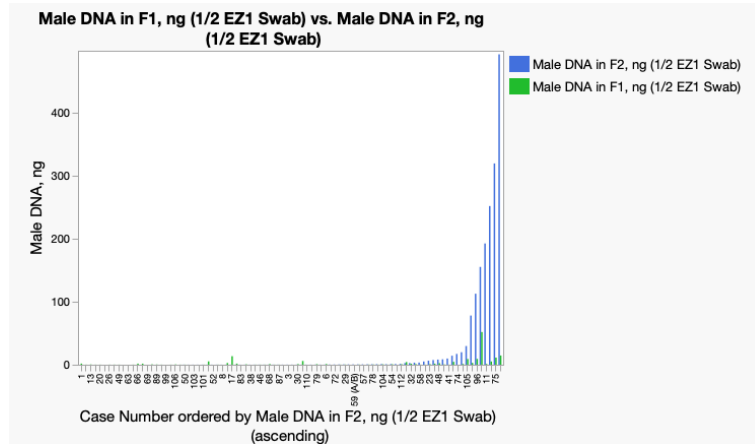


Figure 6. Direct comparison of male DNA in F1 vs. F2 (1/2 EZ1 swab). This figure illustrates the cases with the highest amounts of male DNA in F2. The blue color indicates male DNA in F2, and the green color indicates male DNA in F1.

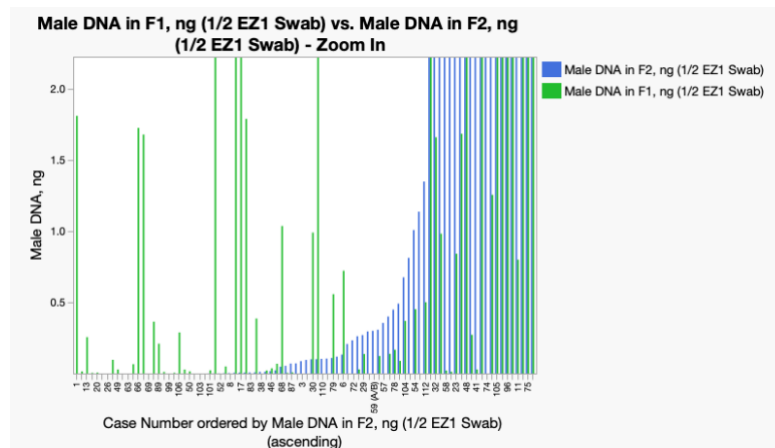


Figure 7. Detailed comparison of male DNA in F1 vs. F2 (1/2 EZ1 swab). This figure displays the cases with the lowest amount of male DNA in F2 alongside cases with higher male DNA amount in F1. The blue color indicates male DNA in F2, and the green color indicates male DNA in F1.

The effectiveness of the EZ1 procedure in isolating male DNA from female DNA within forensic samples was evaluated by calculating the proportion of female and male

DNA in F2 (Figure 8). The design of F2 was specifically intended to mitigate the issue of an excessive amount of female DNA overshadowing male DNA, which could compromise the integrity of male DNA profiles. However, the data indicates that the separation was less effective than anticipated. Specifically, in F2 fractions, female DNA constituted over 50% of the total DNA in 66.7% of the cases. Furthermore, in the same set of cases, female DNA accounted for more than 90% of the total DNA content in 51.1% of the cases, including 30% of cases where only female DNA was detected. These profiles would yield little genetic information from the male contributor. For case samples where the proportion of male DNA was below 10% in F2, the average male DNA detected by Y-screening was about 4.83 ng, ranging from 0.06 to 89.24 ng. Conversely, for case samples where the male DNA in F2 exceeded 50%, the average male DNA in Y-screen was about 108.25 ng. This data indicates that separation is linked to the original quantity of male DNA present, where a higher quantification of male DNA generally results in better overall separation.

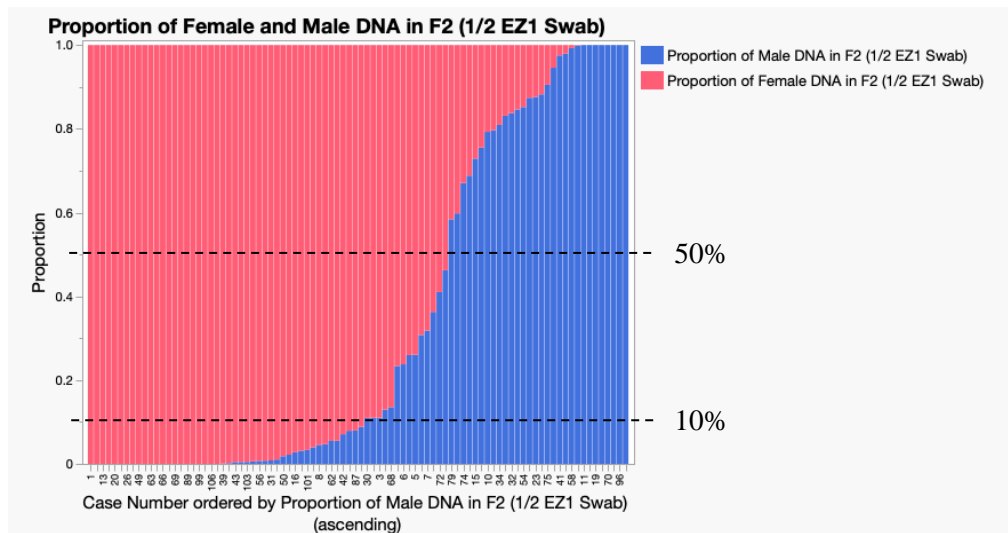


Figure 8. The proportion of female and male DNA in F2 (1/2 EZ1 swab). This figure is marked with two dashed lines indicating the percentage of cases where female DNA exceeded 50% and 90% of the total DNA.

3.2 Mock Case Study

3.2.1 Distribution of DNA on a Swab Using Casework Direct System

The mock case samples were designed to simulate the processing of sexual assault cases using Y-screening and EZ1 assays. The total amounts of female and male DNA from four swabs across two donors were determined (Table 9). The male DNA levels from both donors were significantly higher than those of female DNA, a result attributed to the brief interval between intercourse and sample collection. Specifically, samples from donor A were collected approximately 5 minutes post-intercourse, while those from donor B were obtained around 12 hours later.

Table 9. Total mass of male and female DNA on a Y-screen swab. This table displays the quantity of DNA recovered from two swabs for each donor.

Sample Name	Female (ng)	Male (ng)
Donor A S1	3072.78	3616.17
Donor A S2	2115.18	3383.81
Donor B S1	1338.08	2614.13
Donor B S2	1529.11	3111.56

To determine the DNA distribution on a swab, the swabs were cut into 6 pieces using either method 1 or method 2 (Figures 9 & 10). The Y-screening assay was employed based on the supposition that the swabs would exhibit a uniform distribution of samples (42). However, the distribution of male and female DNA on the swabs was quite uneven. It is evident that for donor A using method 1, the DNA distribution was relatively even at the tip, but varied significantly along the shaft; one quarter of the shaft contained 40% of the total female DNA, while another quarter had only 1.1% of the total female DNA (Figure 9). For donor B using method 1, the tip of the swab had a higher amount of DNA,

accounting for 34% of the total male DNA, but the remainder of the shaft indicated a range from a high of 21.4% to a low of 3.7% of the total male DNA (Figure 9). The distribution of female DNA was similarly uneven. This clearly indicates that the DNA distribution on the swabs was highly uneven.

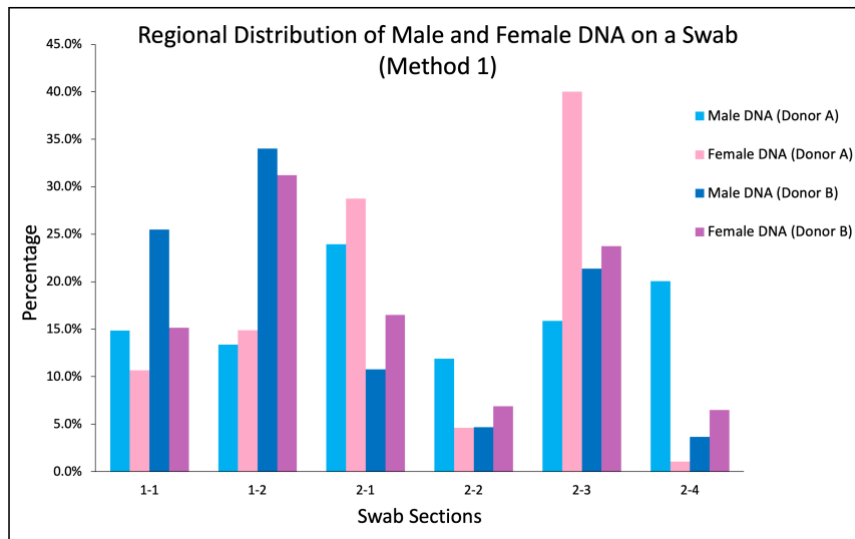


Figure 9. Regional distribution of male and female DNA on post-coital swabs from volunteers using method 1. The colors light blue, pink, dark blue, and purple correspond to male DNA from Donor A, female DNA from Donor A, male DNA from Donor B, and female DNA from Donor B, respectively.

Method 2 was employed to test the distribution differences between the DNA on the outer layer and the internal cotton of the swab. Unlike the distribution observed with method 1, both female and male DNA were predominantly present in the outer layer of the swab (Figure 10). For both donors, the outer layer of the tip, designated as *I-I O*, contained significantly more DNA, with up to 46.2% of the total female DNA and a minimum of 29.2%. The male DNA was found in a similar range. The remaining two halves of the shaft had slightly less DNA than the tip, except for donor B's female DNA, which

constituted 53.5% of the total in section 2-1 O. DNA was least present in all inner cotton areas. For donor A, neither of the two inner cotton sections contained any female DNA, and the maximum male DNA found accounted for only 7.8% of the total. For donor B, the content of female and male DNA in the cotton was less than 6% of the total. This suggests that for a cotton swab, the majority of both female and male DNA is located in the outer layer.

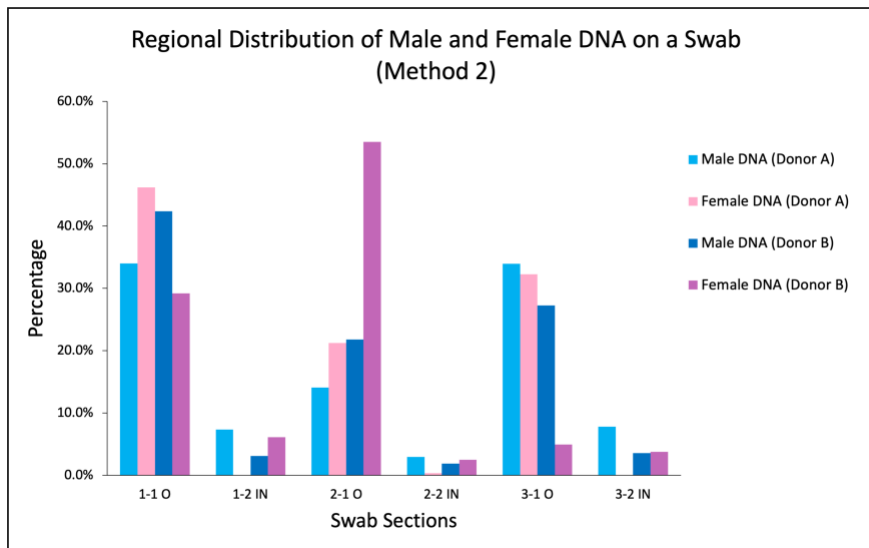


Figure 10. Regional distribution of male and female DNA on post-coital swabs from volunteers using method 2. The colors light blue, pink, dark blue, and purple correspond to male DNA from Donor A, female DNA from Donor A, male DNA from Donor B, and female DNA from Donor B, respectively.

3.2.2 Efficiency of Casework Direct System

The forensicGEM Sperm was utilized to determine the amount of DNA remaining on the Y-screen swabs. According to the data, only 4.2% of the total DNA was left on the cuttings of donor A. Consequently, the Casework Direct system proved effective in lysing DNA from a cotton swab.

3.2.3 Efficiency of EZ1 Extraction

Two swabs from donor B were analyzed to evaluate the efficiency of the EZ1 extraction process (Figure 11). The amounts of female and male DNA extracted from swab 1 in the F1 and F2 fractions were 974.6 ng and 360.1 ng, respectively. Similarly, from swab 2, the amounts of female and male DNA extracted in the F1 and F2 were 1533.1 ng and 450.9 ng, respectively. Contrasting with the DNA distributions noted on the Y-screen swabs, the EZ1 swabs revealed a higher presence of female DNA compared to male DNA. Female DNA was most prevalent in the F1 fraction for both swabs. In comparison, male DNA was more abundant in the F2 fraction. However, it was surprising to find that on swab 2, the male DNA remaining in the M fraction constituted 71% of the total male DNA on the swab, encompassing the F1, F2, and M fractions. For swab 1, up to 48.8% of the total male DNA remained on the swab. The extraction performance of female DNA using EZ1 was significantly better than that of male DNA, with only 4.7% and 14.2% of the total female DNA remaining on swabs 1 and 2, respectively.

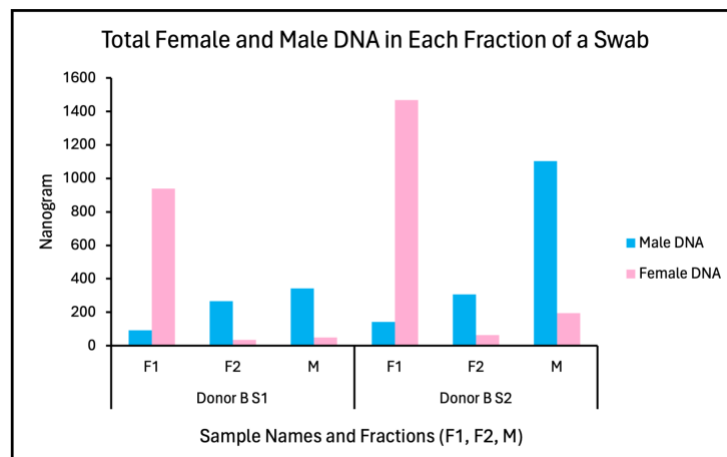


Figure 11. Total female and male DNA in F1, F2, and material section of a swab.

The figure illustrates female and male DNA in each section of swabs 1 and 2. Blue represents male DNA, while pink denotes female DNA.

The efficiency of separating sperm cells from female cells using EZ1 was determined (Figure 12). Because of the sample size, the phenomenon observed in the laboratory data was different from volunteer's samples that the highest proportion of female DNA in F2 was 18.2%. The proportion male DNA in F1 performed normally.

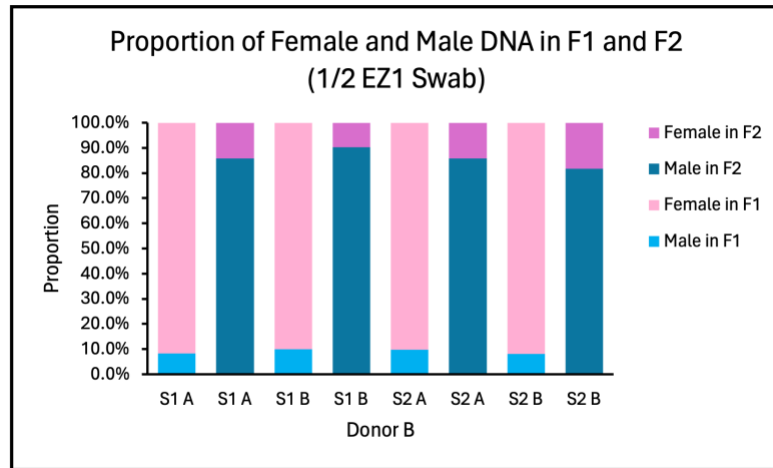


Figure 12. Proportion of female and male DNA in F1 and F2 (1/2 EZ1 swab). This figure displays the proportions of female and male DNA in two halves of two swabs, identified as Section A and Section B. Purple and dark blue represent female and male DNA in F2, respectively, while pink and light blue represent female and male DNA in F1.

4. DISCUSSION

4.1 Laboratory Casework Analysis

4.1.1 Accuracy of Y-screening

The primary goal of using Y-screening in crime laboratories is to test for the presence of male DNA in samples from sexual assault kits. This method also helps determine the number of samples to be forwarded to the DNA unit based on the quantity of male DNA detected. The Casework Direct system is specifically designed for Y-screening and ensures efficient DNA lysis directly from the swab, significantly reducing the possibility of DNA loss and saving time. For the purpose of this study, it is assumed that lysates produced using the Casework Direct system contain all the DNA present in the sample.

Analysis of real casework data from the crime laboratory demonstrates that the sample lysates produced by the Casework Direct system show high variability in DNA amounts, particularly for female DNA. The content of female DNA on one half of a Y-screen swab ranged from a minimum of 0.62 ng to a maximum of 44,072.16 ng. Approximately 82.2% of the samples contained more than 1000 ng of female DNA, with an average of 6113.19 ng, indicating a substantial presence of female DNA. Similarly, male DNA results ranged from 0.06 ng to 675.72 ng, with an average of 39.03 ng. 60% of samples contained more than 1 ng of male DNA, which if isolated away from female DNA is suitable for subsequent STR analysis.

Casework data is used to estimate the number of swabs used on the EZ1. However, all swab counts were normalized to half a swab for the data comparisons made here. The

maximum amount of female DNA on an EZ1 swab was only 2886.09 ng, about $\frac{1}{15}$ of that on a Y-screen swab. The average amount of female DNA was also significantly lower on the EZ1 swabs, with 615.67 ng compared to 6113.19 ng on Y-screen swabs, which was only about $\frac{1}{10}$. Approximately 77.8% of the samples had less than 1000 ng of female DNA after EZ1 extraction, suggesting that the amount of female DNA obtained from EZ1 swabs was only about one-tenth of that obtained from Y-screen swabs.

Furthermore, obtaining more DNA from a Y-screen swab does not guarantee more DNA on an EZ1 swab. Of the nine samples that had over 100 ng/ μ L of female DNA on a Y-screen swab, the amount of female DNA on the EZ1 swab was only 3.4% of that measured on the Y-screen swab. For samples with less than 100 ng/ μ L on a Y-screen, about 17% of the Y-screen female DNA amount was found on EZ1 swabs. It is advisable to re-quant when Y-screen results show DNA concentrations over 100 ng/ μ L to obtain a more accurate concentration. While male DNA is crucial for sexual assault cases, the average amount of male DNA obtained from EZ1 swabs was only half of the average DNA that from Y-screen swabs. On Y-screen swabs, most samples had more than 1 ng of male DNA, yet 61.1% of the samples tested on EZ1 swabs indicated total male DNA below 1 ng, and 35.6% had even less than 0.1 ng. Although 12.2% of the cases showed more male DNA on EZ1 swabs, these were exceptional cases. The results presented here suggest that it is inaccurate to predict that a similar amount of DNA or total amount of DNA will be obtained from EZ1 swabs based solely on Y-screening results. There is a significant possibility that some of the DNA is lost during the EZ1 process. However, Y-screening is

appropriate for determining the presence of male DNA since very little DNA is lost in the Y-screening process.

4.1.2 Efficiency of EZ1

The sperm cells were isolated from the female DNA by EZ1 pretreatment of the sample, and two fractions F1 and F2 were obtained. These two fractions are separately purified using the EZ1 instrument. As expected, most of the DNA in the F1 was female, with male DNA being significantly less, averaging just 1.83 ng. About a quarter of the case samples showed no male DNA in F1, although an equal proportion contained more than 1ng of male DNA. The amount of female DNA in F2 was also low, averaging 2.83 ng, and more than half of the F2 samples contained less than 1 ng female DNA, which was beneficial.

Male DNA was found to be more abundant in F2 than in F1. While the average amount of male DNA in F2, 19.46 ng, was 11 times higher than the ng of male DNA in F1, this average was high as a result of a few cases with unusually high levels of male DNA. When these outlier values are excluded, the average amount of male DNA in F2 drops to 2.70 ng, which is only 1.5 times the average amount found in F1. Remarkably, about one-quarter of the cases also lacked any male DNA in F2. The separate comparisons of male DNA amounts in F1 and F2 indicated that in 37.8% of the cases, there was more male DNA in F1 than in F2, and half of these cases only detected male DNA in F1. Because Y-screening has replaced serology testing, it is uncertain whether all the male DNA in a sample originates from sperm cells; some may also come from male epithelial cells. Male

epithelial cells or other male cell types found in semen, which belong to the category of round cells found in semen, are types of non-sperm cells. It has been reported that the concentration of round cells in semen from healthy volunteers is about $2.6 \times 10^3/\mu\text{L}$ (38). Therefore, this occurrence was expected. The detection of male DNA exclusively in F1 but not in F2 might be due to the absence of sperm cells, suggesting that the detected male DNA could stem from non-sperm male cells. Additionally, five cases had no male DNA in either fraction. A high amount of female DNA in F1 could make it difficult to detect male DNA in F1 or to obtain complete male STR profiles. Regarding the proportion of female to male DNA in F2, although the average female DNA content was low, about half of the cases have female DNA which is over 90% of the total DNA in this fraction, with the male DNA making up less than 10%. As mentioned in the Materials and Methods, if the ng of male DNA is artificially low due to quantification discrepancies, the blue portion of this graph might shift slightly to the right, indicating that the ng of male DNA may be a bit higher. This high proportion of female DNA can lead to low-quality male DNA profiles or no male STR results at all, and this increases the difficulty in the analysis significantly.

4.2 Mock Case Analysis

4.2.1 Distribution of DNA on a Swab

The preparation of mock case samples was intended to replicate the laboratory testing of actual sexual assault samples and to investigate whether similar issues could arise with the volunteer samples. This preparation allowed for the exploration of DNA distribution on swabs, examined the efficiency of DNA extraction by the Casework Direct

system and looked at the efficiency of DNA extraction and purification using the EZ1 bio-robot. The use of forensicGEM to measure the amount of DNA remaining on used swabs was based on the one-tube extraction method, which aims to maximize DNA recovery. Using forensicGEM and Acrosolv to analyze DNA left on Y-screen swabs resulted in minimal DNA detection. This outcome confirmed the effectiveness of the Casework Direct system for DNA extraction from swabs. This study also provided a plausible explanation for the significant amount of DNA lost during the EZ1 process.

A key assumption of Y-screening is that the DNA on the swab is evenly distributed; however, this assumption may be inaccurate (28). To study the distribution of DNA on cotton swabs, two methods were used. The first cutting method revealed significant differences in the distribution of DNA between the tip and bottom of the swab, with the percentage of male DNA ranging from a low of 3.7% to a high of 34%. Similarly, the distribution of a female DNA varies greatly across different sections. These findings confirm that the homogeneity hypothesis is incorrect. Using only one-eighth or half of a swab to determine the total DNA content of a sample creates a significant risk of underestimating or overestimating the actual DNA content.

The second cutting method was employed to assess whether DNA is predominantly located on the outer layer or within the internal cotton of the swab. The data revealed that the majority of both female and male DNA was present on the swab's outer layer. However, the distribution of DNA in the outer layer was also found to be highly uneven for both female and male DNA. The internal cotton contained very little DNA, and for one of the donors, no female DNA was detected in two of the cotton sections at all. These findings

further confirm that estimation of the total DNA content and amount of DNA recovery based solely on qPCR results from a portion of the sample swab can often be inaccurate. The Y-screening procedure is very accurate in the detection of male DNA but lacks accuracy in predicting the success of the differential extraction or the amount of male DNA that will be recovered.

4.2.2 Efficiency of EZ1

The purpose of using EZ1 was to separate the male DNA from the female DNA and then purify the two parts. Analysis of real-world case data shows that there was a significant amount of DNA loss during this process when compared to Y-screening results. Therefore, the forensicGEM Acrosolv component was used to analyze any DNA remaining on the EZ1 swab from the small sample set of donor-provided vaginal swabs. This digest also provides insight into the efficiency of male DNA isolation. Due to the small volunteer sample size, there were differences between the results from the volunteer samples and from real casework samples.

Differential extraction with the EZ1 instrument was effective in processing volunteer samples. It resulted in a high amount of female DNA in Fraction F1, where most of the female DNA was localized. In contrast, the majority of male DNA was found in F2, although a small portion was not differentiated and remained in F1. Importantly, the amount of male DNA in F2 was adequate. However, when using forensicGEM and Acrosolv to assess residual DNA, a substantial amount of male DNA and a smaller amount of female DNA were detected in the material fraction, with about 71% of the male DNA

not being extracted from the swab. Previous research has shown that the efficiency of cotton swabs for pure DNA extraction and recovery was below 50% (39). The amount of DNA residue might be attributed to the extraction's incubation time being only 1 hour and 10 minutes, which was shorter than the recommended maximum of 2 hours. Although only a small number of samples were tested, this data suggests that a large proportion of male DNA remained on the swab after EZ1 processing, which could explain the discrepancy in male DNA quantities obtained compared to those from Y-screening. Further studies are required to explore the causes of this DNA loss.

5. CONCLUSIONS

Due to the growing backlog of SAEKs, crime laboratories must find ways to mitigate this issue. One such method is the use of Y-screening, which eliminates the need for time-consuming and labor-intensive traditional serological testing. The Casework Direct system not only saves time but also maximizes DNA extraction and minimizes DNA loss, allowing for a more accurate determination of DNA concentration in a sample. However, it may not be advisable for crime lab analysts to test just a one-eighth or half of a swab for DNA content and then assume the entire swab has a uniform DNA concentration to determine the appropriate quantity of swab to send for STR analysis. This is because the distribution of DNA on cotton swabs is not homogeneous and can vary significantly, with some parts containing up to 30% more DNA than others. Using only a part of the swab to estimate the total DNA content leads to a significant risk of underestimating or overestimating the actual DNA content, potentially resulting in sending an inadequate number of swabs for analysis.

Besides the variable distribution of DNA on swabs, another issue is the substantial DNA loss during the differential extraction and purification processes of EZ1. This loss can lead to samples that originally had sufficient DNA yielding too little for a full STR profile, resulting in no or only partial profiles. This is why Y-screening assays are better suited for determining the presence of male DNA in a sample. For male DNA, a significant reason for loss is that a considerable amount remains on the swab as EZ1 pretreatment may not effectively extract all the male DNA. In contrast, only a small fraction of female DNA

remained on the swab, so a large loss of female DNA cannot be explained for this reason. The result of lost DNA remains a concern.

Moreover, the ability of EZ1 pretreatment to effectively separate male from female DNA is limited for some pieces of evidence. Sometimes, the non-sperm fraction contained more male DNA than the sperm fraction. More critically, in the sperm fraction, the likelihood of female DNA constituting up to 90% of total DNA can be high. The excess of female DNA can result in no male DNA profile, low-quality or partial profiles, increasing the difficulty of analysis.

In conclusion, the data presented in this study confirmed that in most cases, the utilization of Y-screening and EZ1 pretreatment will assist in reducing the current backlog of SAEKs, saving time and resources in crime labs.

5.1 Future Directions

In this study, many factors were not considered. The data from actual sexual assault cases included only Y-screening and EZ1 results; there was no information on amplification, CE, STR profiles, or CODIS. Therefore, any conclusions about the impact on subsequent steps are inferred from limited quantification results. The ultimate consequences of DNA loss during the EZ1 process needs to be evaluated based on whether the remaining DNA meets amplification targets, can create high-quality STR profiles, and is suitable for uploading to CODIS. It's possible that even if a significant amount of DNA is lost, the amount of DNA obtained from the sample might be high enough to avoid adversely affecting further analysis. This depends on the initial amount of DNA in the

sample. The amount of female DNA in the sperm fraction of some samples may affect the quality of the STR profile. More information from follow-up results is needed to analyze whether samples containing more than 90% female DNA are adversely affected. In this study, only cotton swabs were used, and there was no testing to see if the EZ1 system had better extraction performance with other types of swabs. Future studies could test various swab types, such as cotton, nylon, and foam, to determine if the EZ1 system achieves better recovery with certain swab materials.

Additionally, due to the limited number of post-coital samples collected, the data derived from this small sample size may not be representative of all scenarios. More samples are needed to test DNA distribution and the amount of DNA remaining.

LIST OF JOURNAL ABBREVIATIONS

Arch Androl	Archives of Andrology
Anal Bioanal Chem.	Analytical and Bioanalytical Chemistry
Am J Hum Genet.	American Journal of Human Genetics
Forensic Sci Int	Forensic Science International
Forensic Sci Med Pathol	Forensic Science, Medicine and Pathology
J Forensic Sci	Journal of Forensic Sciences
Microbiol Mol Biol Rev	Microbiology and Molecular Biology Reviews
Med Sci Law	Medicine, Science and the Law
Trends Genet	Trend in Genetics

BIBLIOGRAPHY

1. Patsy A. Klaus, Marshall M. DeBerry. The crime of rape. 1985 March.
2. Thompson A, Tapp SN. Criminal victimization, 2022. Bureau of Justice Statistics. <https://bjs.ojp.gov/document/cv22.pdf> (accessed May 13, 2024).
3. Little K. A national protocol for sexual assault medical forensic examinations: Adults/adolescents. US Department of Justice, Office on Violence Against Women; 2013. <https://www.ojp.gov/pdffiles1/ovw/228119.pdf> (accessed May 13, 2024).
4. Shelby R. Whose rape kit? stabilizing the vitullo® kit through positivist criminology and protocol feminism. *Theoretical Criminology*. 2020;24(4):669-88. <https://doi.org/10.1177/1362480618819805>.
5. Recommendations for the efficient DNA processing of sexual assault evidence kits. Scientific Working Group on DNA Analysis Methods; 2016. https://www.swgdam.org/_files/ugd/4344b0_4daf2bb5512b4e2582f895c4a133a0ed.pdf (accessed May 13, 2024).
6. Campbell R, Patterson D, Lichty LF. The effectiveness of sexual assault nurse examiner (SANE) programs: a review of psychological, medical, legal, and community outcomes. *Trauma Violence Abuse*. 2005;6(4):313-329. <https://doi.org/10.1177/1524838005280328>.
7. Butler JM. Advanced topics in forensic DNA typing: Methodology. In: *Advanced Topics in Forensic DNA Typing: Methodology*. United States: Elsevier Science & Technology; 2011.
8. Campbell R, Fehler-Cabral G, Pierce SJ, Sharma DB, Bybee D, Shaw J, et al. The Detroit Sexual Assault Kit (SAK) Action Research Project (ARP), Final Report; 2015. <https://www.ojp.gov/pdffiles1/nij/grants/248680.pdf> (accessed May 13, 2024).
9. Hanson EJ. Sexual assault kits (SAKs) and the backlog of untested sexual assault evidence: In brief. Congressional Research Service; 2022. <https://crsreports.congress.gov/product/pdf/R/R44237/5> (accessed May 13, 2024).
10. Sexual Assault Kits using science to find solutions. National Institute of Justice. <https://nij.ojp.gov/sites/g/files/xyckuh171/files/media/document/unsubmitted-kits.pdf> (accessed May 13, 2024).
11. Hayes J, et al. Forensic testing turnaround times in 50 states. OLR research Report; 2010. <https://www.cga.ct.gov/2010/rpt/2010-R-0086.htm> (accessed May 13, 2024).

12. Massachusetts State Police Crime Laboratory Quarterly Report 2023. <https://malegislature.gov/Bills/193/HD4948.pdf> (accessed May 13, 2024).
13. Sensabaugh GF, Crim D. Isolation and Characterization of a Semen-Specific Protein from Human Seminal Plasma: A Potential New Marker for Semen Identification. *J. Forensic Sci.* 1971;23:106–115. <https://doi.org/10.1520/JFS10659J>.
14. Herman Y, Feine I, Gafny R. Acid phosphatase test on Phadebas[®] sheets - An optimized method for presumptive saliva and semen detection. *Forensic Sci Int.* 2018 Jul;288:218-222. <https://doi.org/10.1016/j.forsciint.2018.04.047>.
15. Lewis J, Baird A, Mcalister C, Siemieniuk A, Blackmore L, McCabe B, et al. Improved detection of semen by use of direct acid phosphatase testing. *Science & Justice.* 2013 -12;53(4):385. <https://doi.org/10.1016/j.scijus.2013.04.009>.
16. Allery JP, Telmon N, Mieusset R, Blanc A, Rougé D. Cytological detection of spermatozoa: comparison of three staining methods. *J Forensic Sci.* 2001;46(2):349-351.
17. Strom K, Scott T, Feeney H, Young A, Couzens L, Berzofsky M. How much justice is denied? An estimate of unsubmitted sexual assault kits in the United States. *Journal of Criminal Justice.* 2020-10-20;73. <https://doi.org/10.1016/j.jcrimjus.2020.101746>.
18. Campbell R, Fehler-Cabral G. Why Police “Couldn’t or Wouldn’t” Submit Sexual Assault Kits for Forensic DNA Testing: A Focal Concerns Theory Analysis of Untested Rape Kits. *Law & Society Review.* 2018;52(1):73-105. <https://doi.org/10.1111/lasr.12310>.
19. Richard GF, Kerrest A, Dujon B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol Mol Biol Rev.* 2008 Dec;72(4):686-727. <https://doi.org/10.1128/MMBR.00011-08>.
20. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature,* 409(6822), 860–921. <https://doi.org/10.1038/35057062>.
21. Herrera RJ, Garcia-Bertrand R. The nature of evolution. Ancestral DNA, Human Origins, and Migrations. 2018:1. <https://doi.org/10.1016/b978-0-12-804124-6.00001-x>.
22. Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am J Hum Genet.* 1998;62(6):1408-1415. <https://doi.org/10.1086/301869>.

23. Jobling MA, Tyler-Smith C. Fathers and sons: the Y chromosome and human evolution. *Trends Genet.* 1995;11(11):449-456. [https://doi.org/10.1016/s0168-9525\(00\)89144-1](https://doi.org/10.1016/s0168-9525(00)89144-1).
24. Hanson EK, Ballantyne J. A Y-short tandem repeat specific DNA enhancement strategy to aid the analysis of late reported (≥ 6 days) sexual assault cases. *Med Sci Law.* 2014;54(4):209-218. <https://doi.org/10.1177/0025802413519761>.
25. Hall A, Ballantyne J. Novel Y-STR typing strategies reveal the genetic profile of the semen donor in extended interval post-coital cervicovaginal samples. *Forensic Sci Int.* 2003;136(1-3):58-72. [https://doi.org/10.1016/s0379-0738\(03\)00258-5](https://doi.org/10.1016/s0379-0738(03)00258-5).
26. Hall A, Ballantyne J. The development of an 18-locus Y-STR system for forensic casework. *Anal Bioanal Chem.* 2003;376(8):1234-1246. <https://doi.org/10.1007/s00216-003-2039-2>.
27. Roewer L. Y chromosome STR typing in crime casework. *Forensic Sci Med Pathol.* 2009;5(2):77-84. <https://doi.org/10.1007/s12024-009-9089-5>.
28. Report on Y-Screening of Sexual Assault Evidence Kits (SAEKs). Scientific Working Group on DNA Analysis Methods. 2020. https://www.swgdam.org/_files/ugd/4344b0_e8334cb799704a1dabbd8d41f58b979d.pdf (accessed May 13, 2024).
29. Casework Direct system Technical Manual. 2020.
30. Graham EK, Loten M, Thompson JM, Drobac J, Gopalakrishnan A. Developmental Validation of the Casework Direct system, Custom: A Method for the Rapid Processing of Casework Samples WP108. 2018.
31. Quantifiler™ HP and Trio DNA Quantification Kits User Guide. 2018.
32. Hickey J. forensicGEM® sperm lysis: Enzymatic method achieves better results than chemical method. https://www.technosaurus.co.jp/files/forensicGEM_YScreening_AppNote_20191007.pdf (accessed May 13, 2024).
33. Roberts AL. Optimization of a Novel Temperature Controlled Differential Extraction Procedure for Aged Sample Analysis. 2019.
34. Gill P, Jeffreys AJ, Werrett DJ. Forensic application of DNA 'fingerprints'. *Nature.* 1985;318(6046):577-579. <https://doi.org/10.1038/318577a0>.

35. EZ1&2[®] DNA investigator[®] Kit Handbook for automated purification of DNA from forensic and human ID samples using EZ1[®] instruments. 2022.
36. Instructions for Use Vivacon[®] 500. 2021.
37. Grgicak CM, Urban ZM, Cotton RW. Investigation of Reproducibility and Error Associated with qPCR Methods using Quantifiler[®] Duo DNA Quantification Kit. *Journal of Forensic Sciences* 2010;55(5):1331–9. <https://doi.org/10.1111/j.1556-4029.2010.01460.x>.
38. Arata de Bellabarba G, Tortolero I, Villarroel V, Molina CZ, Bellabarba C, Velazquez E. Nonsperm cells in human semen and their relationship with semen parameters. *Arch Androl.* 2000 Nov-Dec;45(3):131-6. <https://doi.org/10.1080/01485010050193896>.
39. Bruijns BB, Tiggelaar RM, Gardeniers H. The Extraction and Recovery Efficiency of Pure DNA for Different Types of Swabs. *J Forensic Sci.* 2018;63(5):1492-1499. <https://doi.org/10.1111/1556-4029.13837>.

