FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

BODY FLUID IDENTIFICATION USING DNA EXTRACTION WASTE PRODUCT

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

in

FORENSIC SCIENCE

by

Rachel Tolces

To: Dean Michael R. Heithaus
College of Arts, Sciences and Education

This thesis, written by Rachel M. Tolces, and entitled Body Fluid Indentification Using DNA Extraction Waste Product, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

Sara Zapico

Bruce McCord, Co-Major Professor

Steven Lee, Co-Major Professor

Date of Defense: June 30, 2022

The thesis of Rachel M. Tolces is approved.

Dean Michael R. Heithaus
College of Arts, Sciences and Education

Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2022

ABSTRACT OF THE THESIS

BODY FLUID IDENTIFICATION USING DNA EXTRACTION WASTE PRODUCT

by

Rachel Tolces

Florida International University, 2022

Miami, Florida

Professor Steven Lee, Co-Major Professor

Professor Bruce McCord, Co-Major Professor

In many situations, the amount of DNA evidence recovered at a crime scene is at trace levels, limiting the amount of testing that can be done on the evidence. Because genetic profiling is generally considered the most imperative assay to be completed on a DNA sample, in situations where yield is low, DNA extracts may not be utilized for any other purpose. Fortunately, an alternate source of DNA may exist by utilizing the waste products resulting from DNA extraction.

The goal of this project was to develop and test a protocol for recovery of DNA from robotic extraction waste and utilizing this DNA, perform epigenetic based body fluid identification. Once extracted the sample will be analyzed using bisulfite conversion and pyrosequencing. In this process, specific loci were amplified and sequenced to detect the presence of buccal cells, blood, semen, or vaginal epithelia. By utilizing this method, information can be obtained about the origin of an unknown DNA sample that can provide crucial information in sexual assault and child abuse cases.

TABLE OF CONTENTS

CHAPTER PAGE

I. Introduction 1

II. Background 4
1. DNA Extraction 4
2. Waste Purification 8
3. Cell Counting 9
4. Epigenetics 10
5. Body Fluid Identification 14

III. Objectives 17

IV. Materials and Methods 18
1. Sample Collection 18
2. Cell Counting 18
3. Sample Preparation 19
4. DNA Extraction 20
5. Waste Purification 20
6. DNA Quantification 21
7. Bisulfite Conversion 21
8. Pyrosequencing 21

V. Results 22
1. Cell Counting 22
2. Waste and DNA Concentrations 23
3. Waste Quantification For Body Fluid Identification 24
4. Body Fluid Identification 25
5. Blood 25
6. Buccal 25
7. Semen 26
8. Vaginal Epithelia 26
9. Replication Study 28

VI. Discussion..31

VII. Conclusion..35

VIII. Future Work 36

References..38

TABLE OF TABLES

TABLE PAGE

Table 1: Current Body Fluid Identification Tests....................13

Table 2: Percent Methylation at CpG Sites for Each Body Fluid....................22

Table 3: Cell Counts Across Differing Donors....................22

Table 4: Normalized Cell Count and Extraction....................23

Table 5: DNA and Waste Average Yields....................23

Table 6: Concentration Ranges of Successful BFID....................24

TABLE OF FIGURES

FIGURE PAGE

Figure 1: EZ1 Cartridge Layout....6

Figure 2: Epigenetic Phenotyping....11

Figure 3: CpG Methylation....16

Figure 4: Methylation Comparisons....17

Figure 5: Blood Methylation CpG1....26

Figure 6: Buccal Methylation CpG3....27

Figure 7: Semen Methylation CpG2....27

Figure 8: Vaginal Epithelia Methylation CpG3....28

Figure 9: Blood Replication Results....29

Figure 10: Buccal Replication Results....30

Figure 11: Semen Replications Results....30

Figure 12: Vaginal Epithelia Replication Results....31

Introduction

Forensic DNA analysis is typically performed to generate a genetic profile to identify an offender or victim. However, there are many more purposes DNA analysis can serve. Epigenetic analysis on the DNA sample may provide additional information. One use of this epigenetic analysis is to perform body fluid identification to provide information about the origin of a DNA source, which may be crucial information in several cases.

One such scenario where body fluid identification is imperative is in suspected child abuse cases. Finding parental DNA on a child may not be uncommon, as parents may provide innocent, loving contact with their children. However, what if this DNA can be proven to be from a semen sample? This would bring to light new information regarding the case and possibly change a jury's verdict.

A second scenario where body fluid identification would prove helpful is if there is a suspected mixture of fluids. Semen and saliva are both translucent in appearance and may appear mixed in a sample or be masked completely if an opaquer and more colored sample is present such as blood. For this reason, crucial information regarding the sample can be missed without body fluid identification.

Body fluid identification can be performed at the crime scene using presumptive testing. Prior to capillary electrophoresis and epigenetic analysis, this method utilized enzymatic and antigen-antibody interactions to perform a color changing test to identify an unknown body fluid. Performing a presumptive test without having to wait for laboratory results can provide crucial information regarding how to move forward in the collection and analysis process, although it must be handled with care as these are chemical tests that will leave the remaining evidence unusable [46].

Following a presumptive test, a confirmatory test can be performed that may be less destructive to the remaining evidence. For example, a visual test of blood under a microscope to identify living blood cells is an easy and nondestructive way to confirm the presence of blood. A microscopic test can also be used to identify sperm cells in a semen sample. Other confirmatory tests involve UV spectroscopy and more labor-intensive dyes and stains. With these tests, while they may not destroy a sample, they may result in contamination that can compromise downstream DNA analyses [40]. For that reason, forensics has turned to epigenetics to answer some of these questions.

Studying the epigenetics of DNA is a novel field used to determine how the environment can impact the genome. Epigenetics demonstrates that while all cells contain the same genetic material, the way protein translation is completed depends on methylation patterns, histone wrapping, and non-coding RNAs, regions of RNA that are transcribed from the original DNA but do not get translated to form proteins. These mechanisms allow only certain parts of the DNA to be transcribed and thus translated, resulting in the different proteins each specific cell type is eligible to produce [1]. These epigenetic patterns can differentiate cell types, but can also be used to differentiate between identical twins, can determine age, and can provide dietary information.

The increased study of epigenetics has included the ability to differentiate between biological fluids based on their methylation patterns. DNA is methylated in specific ways to allow cell differentiation and protein production. By analyzing these specific differences in methylation, patterns can be utilized to identify critical body fluids, including blood, semen, vaginal epithelia, and buccal cells.

A common practice in forensic DNA labs is to prioritize the creation of a genetic profile with the limited DNA sample that may be found at a crime scene if they are attempting epigenetic body fluid identification. This sample must also be divided amongst all interested parties, so the amount of sample able to be utilized for genetic analysis is already minimal. For this reason, the chemical alteration of the sample required to perform confirmatory epigenetic body fluid identification is typically not performed on very low quantity samples.

Body fluid identification can be performed on unaltered DNA for presumptive testing [52]. This testing relies on antigen-antibody interactions or enzymatic reactions. Presumptive blood testing checks for catalytic activity of heme groups. Saliva assays test for a-amylase, and vaginal epithelial tests check for glycogen epithelial cells. Lastly, semen assays look for seminal acid phosphate [2]. These tests are considered to be presumptive as they are a good starting point for body fluid identification, however they are not confirmatory tests and therefore cannot be used for conclusive body fluid identification [54]. While there are confirmatory tests available, they are typically time and sample consuming, which makes it increasingly difficult to complete when there is a minute and finite amount of sample.

To perform body fluid identification with epigenetic markers, pyrosequencing is used. This is because pyrosequencing can detect SNPs which are where CpG sites for epigenetic analysis are located [3]. These CpG sites are ideal locations to focus on epigenetic analysis due to their location being right before transcription-regulating regions [17]. Pyrosequencing also occurs in real-time and 48 samples can be analyzed in under 2 hours.

This project aims to demonstrate that free-floating DNA remains in the extraction waste, and that it is possible to purify it, and utilize it for forensic body fluid identification.

At the same time, the original DNA extract can be used to create a genetic profile. This practice would ensure more information can be gained from a limited DNA sample, especially when this information must be retrieved.

Background

DNA Extraction

While DNA extraction has been around since its discovery in 1986, the protocols used to extract DNA have become more advanced, cleaner, and more efficient. The most commonly used manual method of DNA extraction is a phenol-chloroform extraction, which utilizes the chemical properties of the cell debris and negatively charged DNA to separate the molecules using phase extractions. However, this process is tedious, requires several hours to overnight processing and involves the use of dangerous chemicals that can harm both the scientist and the sample [42]. The most common commercial manual extraction kits utilize solid-phase extraction, relying on cell lysis and DNA separation and elution to provide replicable and reproducible results, but are still time consuming and labor intensive [4].

There has been an increase in demand for forensic sample processing that includes the need to batch DNA extractions. Therefore, the process needed to become automated. Automated extraction allows a larger number of samples to be run in a shorter amount of time with less user hands-on work, which is extremely beneficial for the time-sensitive nature of forensic science. For example, automated extraction on the EZ1 Robot takes 16 minutes from start to finish for 6-12 samples, compared to the over one hour it would take to perform extraction on one sample manually. Not only are automated instruments faster at performing DNA extraction, they generally result in higher yield and cleaner product, with less chance

for introducing contamination, helping to increase the yield and accuracy of downstream processes such as amplification and sequencing [5].

Several methods for performing an automated extraction include silica-based or magnetic bead extractions [6]. The EZ1 Robot (QIAGEN) utilizes magnetic bead extraction, which takes place over several manufactured cartridges. This extraction method relies on the chaotropic binding provided by the negative nature of DNA. The DNA is attached to the magnetic beads as salt concentrations increase. This step allows for the remaining waste to be washed away as the DNA is bound to the beads. Following the DNA binding, the salt concentration is lowered in the solution which releases the DNA to then be eluted in a separate buffer, purified and detached from any other waste or inhibitors [34].

A benefit of this extraction method is that it limits the amount of physical and chemical damage that can occur to the DNA sample itself [34]. For example, in manual extractions there are usually harsh and dangerous chemicals such as chloroform [35]. This method introduces chemicals that are not only dangerous to the sample, but to the technician as well. Similar manual extraction methods involve multiple vortex and centrifugation steps which may break the DNA and leave a sample compromised [34]. Automated, magnetic bead extraction removes the toxic chemicals and physically harsh conditions DNA is put under in manual extraction to potentially result in a cleaner extract.

Following the completion of extraction, these cartridges are typically discarded, despite the presence of biological sample remaining within the different waste cartridges. The different extraction protocols dispose of the waste in differing cartridges, which is where additional DNA, we hypothesize, remains and can then be used for purification. (Figure 1).

The benefit of using the EZ1 Robot is that the waste is discarded in individual cartridges that correlate with the sample processed. This allows for the waste to remain uncontaminated by other samples in the same run or by the instrumentation used. As seen in Figure 1, the waste remains in the cartridge containing the reagents for the extraction and each individual extraction relies on its own cartridge of reagents. Following the extraction protocol, this waste has already been isolated and can be separated to allow for purification.

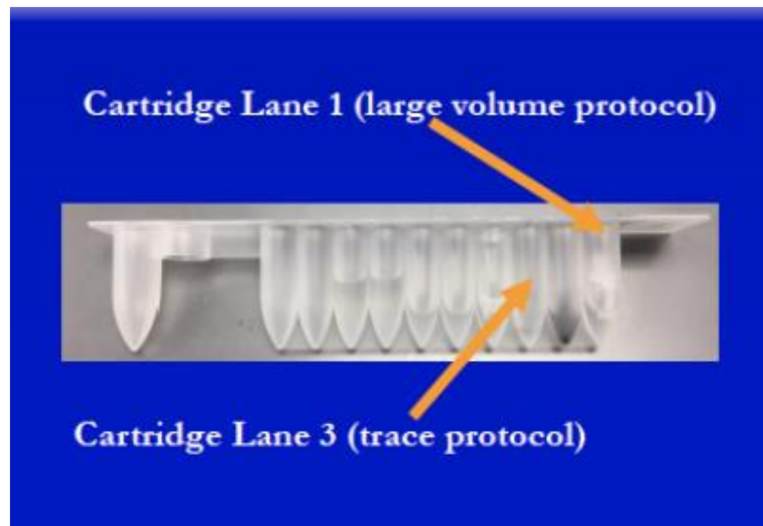


Figure 1: EZ1 Extraction Cartridges [7]

Advancements in forensic DNA analysis have allowed for trace amounts of DNA, known as "touch DNA" or "trace DNA" to be analyzed and used for forensic genotyping. This quantity of DNA is typically around 0.1-0.5ng [8]. It is extremely difficult to predict how much DNA is going to be deposited from a trace sample as the manner the sample is deposited and collected both have an impact on recovery, as well as the choice of downstream extraction and amplification processes. The force used to deposit the DNA, whether an aggressive grab or a light touch, the material being contacted and the mechanism and environmental storage conditions of the deposited sample can result in different amounts

of DNA originally deposited and differential recovery due to degradation or sample retention.

The amount of DNA left behind by a person differs depending on the individual [48]. Some people are "good shedders" while others are "bad shedders", meaning good shedders leave behind much more DNA from an identical touch than a bad shedder would. Even with some people being "better" shedders than others, individuals can shed different amounts of DNA on different days [46]. After DNA is deposited, the quantity remaining will begin to deplete as environmental factors and other contact is introduced to the surface of the material.

The difference in quantity of DNA remaining may be useful. It may assist in determining how much degradation has occurred since the initial contact was made, as well as how often that surface has also been touched to account for contamination [4]. This can also indicate the transfer of touch DNA if multiple people are involved and have DNA not belonging to them on their persons or property.

All off these differing factors make it extremely difficult, if not near impossible, to predict how much DNA remains of trace samples at a crime scene. Trace DNA may also be a mixture from a number of cell types, including a mixture of body fluids such as sweat, mucus, and saliva. Despite such a small amount of DNA being recovered, the abundance of cell types present renders it difficult to effectively and efficiently use the miniscule amount of sample provided by these cells [47].

That is why the majority of the sample when DNA is obtained is used for genetic profiling: the analysis of the DNA to ideally identify a person's identity. Especially when the amount of DNA found at a crime scene is so miniscule, additional testing is typically

bypassed as there is not enough sample to perform secondary assays. However, scientific assays typically result in the production of a waste product, that is discarded just as the name suggests.

Waste Purification

When analyzing such a minute amount of DNA, it is important to have as efficient of a recovery process as possible, and when discarding the waste, some of the sample will ultimately be discarded as well. From the same amount of sample, there may be vastly different amounts of DNA recovered. For example, when the same amount of sample was utilized for a set of 11 phenol-chloroform extractions, the concentrations of the DNA extracted had a standard deviation of 154 ng/uL. While this method successfully yielded DNA, these results demonstrated that there are varying degrees of success [48]. The FastDNA kit and the MoBio kits produced medians of 475ng/uL and 22ng/uL, respectively [49]. Different extraction methods produce differing amounts of DNA yield, and thus even more of that sample may be discarded or destroyed depending on the efficiency, damage and chemical modification imparted by the method used.

For this reason, it is important to analyze the waste product from these extraction methods, especially when performing a protocol designed to extract the maximum amount of DNA from a trace sample. The more DNA that can be recovered from a sample, the more information that can be acquired and therefore the more complete the information reported.

This biological material can be found still bound to the magnetic beads and in any liquids remaining from the extraction protocol [7]. Instead of being discarded, this waste can be purified by removing the magnetic beads and any other inhibitors utilized in the extraction process. Magnetic beads can be removed by placing the sample on a magnetic rack, thereby

separating the beads from the remainder of the sample. The magnetic rack will hold the beads in place while the remaining biological sample can be removed for purification. In preliminary testing, beads posed a problem for purification as they demonstrated properties that resulted in inhibition, and therefore obstructed the ability to purify the waste any further. For this reason, a magnetic rack was used to remove the beads from the supernatant as the supernatant was transferred to another tube.

Purification relies on cleansing buffers and a spin-column to trap the limited DNA sample remaining and discard the waste chemicals used in both the extraction and purification protocols. The spin-column has a small filter on the bottom that will not allow the DNA to pass through until it is bound in the elution buffer. This elution buffer is passed through the column twice to ensure any remaining DNA is brought into the final purified sample. Without an additional clean-up step, inhibitors may remain attached to the sample that would lead to PCR inhibition. This additional step is one of the most useful in reconcentrating the sample resulting in DNA purified from the waste [37].

Cell Counting

In an attempt to predict and optimize DNA extraction yield, by starting with a known number of cells in replicate samples, cell counting protocols were utilized to determine how much DNA could realistically be extracted from a sample. Cell counting would allow for the total number of cells to be observed with the trypan blue dye, allowing for calculations that can predict how many cells would be eligible for extraction [50]. Not all cell types will have the same amount of DNA, increasing the difficulty in attempting to predict how much DNA a sample reliably contains. For example, haploid sperm cells are predicted to have around

3.3pg of DNA per cell, and diploid epithelial cells are predicted to have around 6pg per cell [9]. With this information, bright fluorescent field cell counting can be used to count both live and dead cells within a sample.

Epigenetics

Epigenetics is a broad field that studies modifications to DNA that does not alter the physical base coding of the DNA itself [30]. This primarily involves gene expression, as every cell contains the same genetic content yet presents itself differently in different people. Epigenetics is most simply exemplified when looking at identical twins: despite being identical and having identical DNA, one may develop cancer while the other may not or have differing paces of aging [31]. This is due in part to changes in the way the genome expresses itself, rather than physical alterations to the coding of the genome itself.

Epigenetic analysis may be the key to understanding body fluid identification and performing it properly. This is due to epigenetics being used to study differences in methylation patterns across cells that contain the same genomic DNA but produce different proteins. Historically, epigenetics has been used in the medical field to help treat cancer as gene expression is altered to produce cancer cells [10]. The information provided by observing epigenetic patterns in cancer cells can be used to determine the stage of cancer and the best treatment plan to use moving forward.

More recently, epigenetic phenotyping has come into view as a possible addition to the field of forensic science. By utilizing epigenetic phenotyping, more information in regards to the DNA evidence discovered can be obtained, such as diet, smoking habits, age, and drug use (Figure 2). While this information may not be able to pinpoint a direct suspect, it can be used to gather more information about the sample or the person of interest that left a

sample at the crime scene. One of the benefits of epigenetic analysis is that it can be utilized to determine different tissue types, which allows for body fluid identification to be performed.

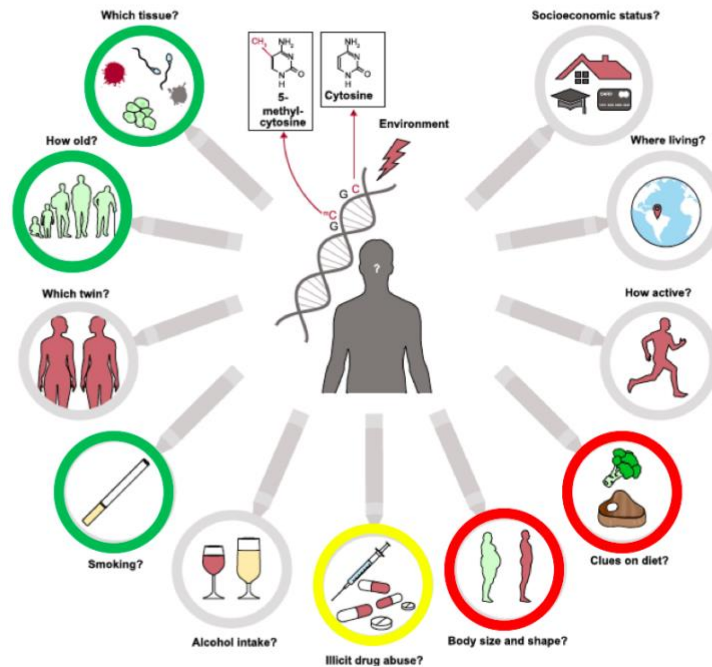


Figure 2: Epigenetic Phenotyping [11]

The information provided through epigenetic analysis can include behavioral characteristics about a person, including smoking, alcohol, and drug habits. These patterns can be identified using an assortment of visual, genetic, chemical, and biochemical tests. This information is extremely helpful in cases where DNA profiling may be limited due to the quantity and/or quality of the DNA recovered [51]. Specific epigenetic analysis relies on specialized tDMRs, located and determined to epigenetically identify a specific characteristic, such as body fluid expression [53].

There is also increased research to determine someone's chronological age using epigenetic markers. There are specific methylation markers that demonstrate unique changes as age increases, allowing age ranges to be determined based on how these genes are

expressed [55]. This is possible because aging results in changes of biological processes such as a change in metabolism. Certain epigenetic characteristics may not be able to individually pinpoint someone's age, but a unique combination of epigenetic traits can help narrow it down [56].

There are different modes of DNA epigenetic modifications. One way epigenetics presents itself is with chromatin modification and histone wrapping [32]. This process involves winding the DNA into tight segments that will not be transcribed as the enzymes move down the DNA strand. Different enzymes or acetyl groups can "open" a histone, which allows for the DNA to be unwound and thus transcribed. This is a way in which genes can be turned "on" and "off."

DNA methylation studies are the most common due in part to readily available technologies available to detect methylation. The methylation patterns found in cancer cells were the first kind of epigenetics studied and led the way for further studies to come [32]. Epigenetic methylation involves the addition of methyl groups to the 5' location on cytosines, which prevents transcription and translation of this region [33]. This alters the gene expression and therefore changes the way a cell functions.

Body fluids are unique for epigenetic analysis in that they are easily used to identify a number of different markers. By observing blood, semen, and other tissues, multiple epigenetic markers can be determined not only for body fluid identification but for diet, age, and drug habits as well [57].

Body Fluid Identification

Currently, body fluid identification is done with presumptive testing to test for enzymatic and antibody/antigen interactions. The first step of body fluid identification is generally visual analysis. This is most beneficial if there is an abundance of known sample that can be observed. Following a visual examination in many cases, presumptive testing would ensue [22]. Saliva and buccal testing rely on an amylase test that can be used as a presumptive test using either a spray or designated strip, and can have confirmatory testing done using DNA analysis via capillary electrophoresis [23]. Amylase is found in saliva to break down amylose into smaller sugars that are more easily digestible as it serves the function of breaking down carbohydrates [24].

Presumptive testing for semen is performed using colorimetric testing to detect acid-phosphate which is found in abundance in seminal fluid [22]. This test can be performed using a p30 card to study the antibody-antigen interactions between p30 antigens and the p30 protein [25] or using an SM test to observe the phosphate by using phosphatases to break down the phosphate resulting in a purple color [22].

Vaginal epithelial cells are the most difficult to identify using presumptive testing as they do not have any proteins solely found in its fluids. However, using Lugol's test, which is typically used for starches, has proven to be beneficial as vaginal cells are high in glycogen. This test does have limitations for cases where a mixture of vaginal cells and penile epithelial cells are suspected, such as in semen, due to penile cells also having a higher concentration of glycogen than other epithelial cells, but still less than vaginal cells [26].

Blood has many presumptive testing options, such as luminol and other commercial-made strips and reagents. Luminol has demonstrated the most success with presumptive

blood testing as it will glow a bright blue while under a blacklight, resulting in easily detected presence of blood [27]. While it may be easier to perform a presumptive test for blood, especially in large quantities, the opaque red coloration can possibly mask other samples, leading to the importance of specified, confirmatory testing.

Confirmatory body fluid identification involves studying and determining the locations of DNA segments used to regulate protein production which allows each cell type to have its individual and unique properties despite containing the same genomic DNA. This process of protein regulation is possibly due to CpG islands, sequences of DNA methylation in promoter regions of DNA [12]. These regions are some of the most stable loci for gene suppression which assists in making them consistent across different people displaying the same cell type, such as different semen donors [32].

Body Fluid	Presumptive Test	Confirmatory Test	Advantages	Limitations
Blood	Kastle-Meyer or Luminol	Crystal/Microscopy/UV Spectroscopy	Fast results	Renders sample unusable, some tests not performed in forensic labs
Buccal	Amylase	Spectroscopy/ELISA Test	Possible to identify many chemical components	Not specific; only used to match samples
Semen	P30 Card or PSA	RSID test for semenogelin	Does not require living semen-looks for seminal plasma	Not very sensitive

Vaginal Epithelia	Lugol's Test	Capillary Electrophoresis	Unique characteristics easy to discriminate for	Not commonly performed, lack of standardization

Table 1: Current Body Fluid Identification Tests [40]

While all of our cells contain the same genomic DNA, they all present differently to perform different functions, such as breaking down different nutrients or transporting different biological materials. This is due to the methylation that results from different portions of DNA containing additional methyl groups to prevent them from being transcribed and translated. In some cases, a decrease in methylation may cause a decrease in gene expression, especially if the gene being transcribed is an upstream regulator [45]. This process of gene silencing allows for the specific cell types to only produce proteins related to body-fluid specific functions.

In order to study these CpG sites and methylation patterns, a chemical alteration of the DNA must be completed to effectively analyze the methylation patterns of different biological fluids. This process, known as bisulfite conversion, converts unmethylated cytosine to uracil and leaves methylated cytosines intact [13]. This produces a new genetic sequence that can be amplified and analyzed to identify the body fluid being observed.

To analyze these new genetic sequences, a pyrosequencer is utilized. Pyrosequencing relies on the light emitted from the reaction of the enzyme luciferase with the dNTPs, or lack thereof, to produce a sequence able to read in a pyrogram [14]. CpG sites are regulation segments found prior to the promoter region that contain specific repeats of methylation and unmethylated cytosines. These sites have a distinct level of "covalent transfers of a methyl

group to the 5-carbon position of a CpG dinucleotide to form 5-methylcytosine" [15]. Figure 3 demonstrates how CpG methylation works [14]. By analyzing the amount of light emitted by the sequences at these regions, the overall methylation percentage can be calculated. Due to the unique nature of each body fluid's CpG site methylation, this data can be used to distinguish one body fluid from another.

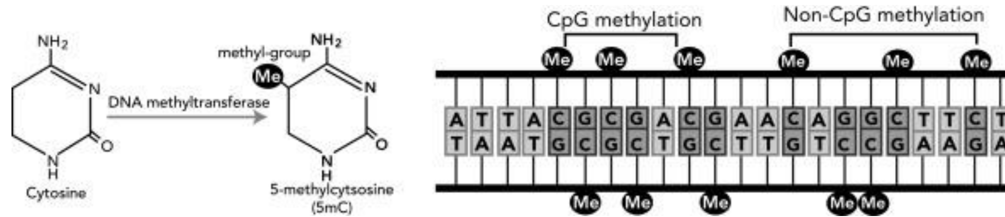


Figure 3: CpG Methylation [14]

CpG sites can be found with tissue-specific differentially methylated regions (tDMRs) [17]. These are regions found ahead of, or overlapping the transcription start site that assist in performing gene regulation and can be used to differentiate tissue cells that still contain the same genomic DNA. These loci have been determined to have specific methylation patterns that allow for expression within different cell types. For example, ZC3H12D has been utilized for detection of semen because it demonstrates extreme methylation by over 70% when comparing semen to other body fluids [19]. This marker targets a gene responsible for zinc production in semen cells [29]. Due to semen having an abundance of zinc present, this locus would thus be hypomethylated when compared to other different body fluids. On the other hand, VE08 in vaginal epithelial cells exhibited hypomethylation compared to other body fluids. BCAS4 for buccal cells shows hypermethylation, as does CG0679435 for blood, as shown in Figure 3. By analyzing these

tissue specific markers, and aiming to replicate the tDMRs, body fluid identification can be performed using pyrosequencing [44].

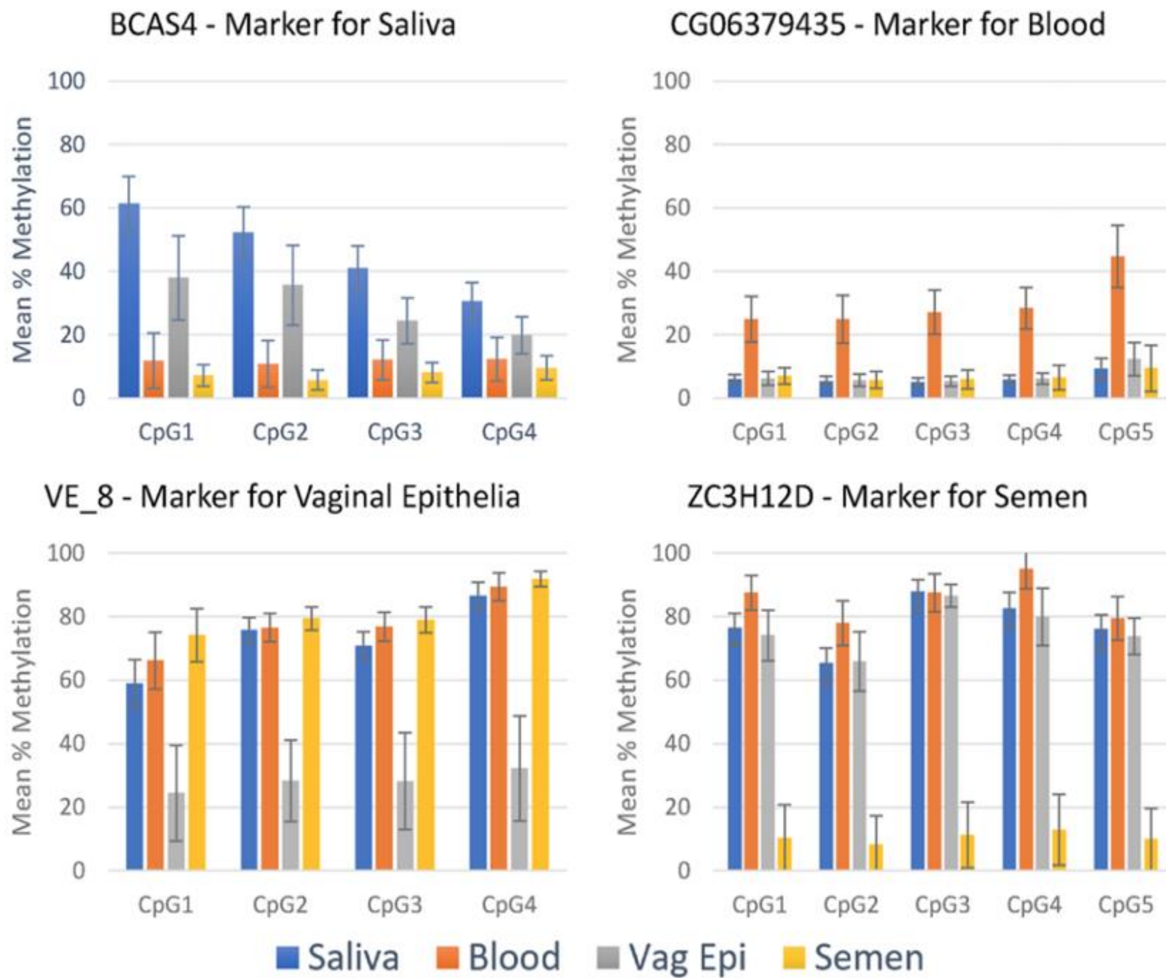


Figure 4: Methylation Comparisons [18]

When dealing with quantities of DNA this low, PCR bias may become a factor. PCR bias occurs when targets within a sample exhibit different amplification efficiency, and therefore the resulting PCR products are not representative of the starting target proportions in the sample [36]. This may occur when a sample template is too low or the cycle number for PCR is not optimized. This is one of the key challenges when performing epigenetic analysis with low quantities of DNA as the study relies on methylation percentages. If there

are more cells with increased methylation being amplified than those with decreased methylation, or vice versa, false positives and false negatives can be produced as the sample being amplified and sequenced no longer represents the percentage of methylation of the original sample.

Objectives

The overall aim of this study is to assess if DNA can be recovered from the waste product, and, after purification, used for body fluid identification based on epigenetic analysis. This main aim is divided into the following objectives:

1. If DNA can be purified from the extraction waste product, can it be used for epigenetic body fluid identification?
2. How much DNA is needed from the waste to successfully perform body fluid identification?
3. Can the amount of DNA found in a sample and its waste product be predicted through automated cell counting?

Materials and Methods

Sample Collection

Following the approval in IRB-21-0415, buccal, blood, vaginal epithelial, and semen samples were collected from willing participants. For the population study, 5 buccal samples and 6 vaginal epithelial samples were collecting on Covidien Curity Cotton Tipped Applicator Sterile Swabs (Dublin, Ireland). Additional buccal and vaginal epithelial samples were collected using the same IRB. The desired location was swabbed for 30 seconds and then processed. 5 blood samples were collected by a certified phlebotomist for the population study and 1 additional sample was collected and used for preliminary testing, deemed

"stock", and stored in an EDTA blood collection tube, and stored in a 4°C refrigerator. 5 semen samples were collected and stored in 2mL aliquots in a -20°C freezer along with one additional sample obtained earlier that was aliquoted and used as "stock". A total of 29 blood, 29 semen, 17 buccal and 25 vaginal epithelial samples were collected using the combination of these methods.

Cell Counting

The samples were centrifuged with 300uL 1x PBS and then pulse vortexed. 10uL of the vaginal epithelial and buccal samples were homogenized with 10uL of trypan blue, and 10uL of the mixture was placed on the slide and inserted into the LUNA-FL Dual Fluorescence Cell Counter (Annadale, VA). The Bright Field Cell Counting setting was utilized. Live, dead, and total cell counts were recorded according to the manufacturer's protocol[43].

Sample Preparation

Blood and semen samples at different volumes were transferred onto sterile swabs prior to pretreatment to test sensitivity. Blood, buccal, and vaginal epithelial samples were pretreated with 290uL G2 Buffer and 10uL ProK, and semen samples were pretreated with 270uL G2, 20uL DTT, and 10uL ProK (QIAGEN, Germantown, MD). The samples were heated on a shaking heat block at 56°C and 1400rpm for two hours and then transferred to a spin basket and spun down in a centrifuge at high speed for 1 minute. Samples were then transferred to tubes for the EZ1 robot (QIAGEN, Germantown, MD). If the Large Volume Protocol was utilized, 300uL MTL buffer was added following the completion of the

pretreatment. This protocol would be used for larger stains, such as on fabrics, that would require being soaked in a large volume of liquid prior to extraction.

DNA Extraction

DNA extraction was performed using the EZ1 Investigator Kit and EZ1 robot (QIAGEN, Germantown, MD). This robot utilizes magnetic bead extraction to perform the DNA extraction from the pretreated sample. Extraction utilized either the Trace Protocol or Large Volume Protocol and was eluted in 40uL TE Buffer according to the manufacturer's directions.

Waste Purification

The waste from the EZ1 cartridge was transferred to a 2mL tube placed on a magnetic rack. The sample was removed, leaving the remaining inhibitory magnetic beads from the extraction process behind. This sample was then treated using the MiniElute Reaction Cleanup Kit (QIAGEN, Germantown, MD)) and eluted in 40uL EB to be stored. Rather than splitting the sample into two spin columns for purification, one column was used, and the sample was run through the column in two cycles. The elution was also performed twice by running the already eluted sample back through the column to ensure all DNA was retrieved from the membrane.

DNA Quantification

Quantification was performed using the Rotor-Gene SYBR Green PCR Kit (QIAGEN, Germantown, MD) and conserved rRNA primers designed for vertebrate

quantification [20]. The samples were processed along with negative control and eight standards of a serial dilution as previously described [20]

Bisulfite Conversion

Bisulfite conversion was performed using the EpiTect Fast Bisulfite Conversion Kit (QIAGEN, Germantown, MD). This was followed by a cleanup process using the same kit and amplification using the PyroMark PCR Kit (QIAGEN, Germantown, MD) featuring CoralLoad Dye to amplify the tDMRs utilized for this work [16]. PCR was performed following the manufacturer's protocol.

Pyrosequencing

For pyrosequencing, four loci were utilized: BCAS4 for buccal at CpG site 1, ZC3H12D for semen at CpG site 2, VE_8 for vaginal epithelia at CpG site 3, and CG06379435 for blood at CpG site 1 as previously described [19]. The PyroMark Q48 (QIAGEN, Germantown MD) was used to perform all pyrosequencing with the Advanced CpG Reagents Kit (QIAGEN, Germantown, MD). All reagent calculations were performed using the PyroMark Q48 computer software. 10uL of sample was used to perform pyrosequencing.

Table 1 represents previously published data that was used to assess body fluid identification by observing methylation percentages and patterns across four different body fluids: blood, saliva (buccal), vaginal epithelia, and semen [19]. This data was used in this study to determine positive identification. Positive identification is defined as methylation percentages within two standard deviations of the primer designed for the tissue using a monoplex.

Table 2. Comparison of mean percent methylation and SDs for the five selected CpGs when analyzed in multiplex and in monoplex

	BCAS4_CpG1		CG06379435_CpG1		CG06379435_CpG2		VE_8_CpG3		ZC3H12D_CpG2	
	Multiplex	Monoplex	Multiplex	Monoplex	Multiplex	Monoplex	Multiplex	Monoplex	Multiplex	Monoplex
Saliva, $n = 38$	61.40 ± 8.49	68.5 ± 6.2	6.14 ± 1.40	8.7 ± 7.0	5.48 ± 1.51	2.6 ± 1.4	70.88 ± 4.32	57.67 ± 3.79	65.32 ± 4.83	78.0 ± 4.5
Blood, $n = 32$	11.86 ± 8.65	11.0 ± 4.2	24.98 ± 7.20	24 ± 7.8	24.91 ± 7.52	22.0 ± 6.7	76.83 ± 4.57	69.67 ± 3.78	77.90 ± 7.05	94.0 ± 2.7
Vaginal Epithelia, $n = 26$	37.97 ± 13.21	36.25 ± 9.38	6.27 ± 2.18	5.3 ± 3.2	5.70 ± 1.94	4.8 ± 1.6	28.23 ± 15.21	12.0 ± 9.54	65.87 ± 9.31	84.3 ± 8.48
Semen, $n = 28$	7.24 ± 3.45	5.60 ± 1.0	7.05 ± 2.56	3.4 ± 1.7	5.77 ± 2.69	2.4 ± 1.8	78.93 ± 4.07	68.33 ± 2.88	8.36 ± 8.87	6.1 ± 3.9

Table 2: Percent Methylation at CpG Sites for Each Body Fluid [19]

Results

Cell Counting

In order to determine the theoretical amount of starting nuclear DNA prior to extraction, automated cell counting was utilized. Cell counting demonstrated differences amongst specific donors while counting buccal cells, illustrating different amounts of shedding, as seen in Table 2. There is a magnitude of 100 difference between some live cell counts and a magnitude of 10 difference in the total cell counts.

Sample	Live	Dead	Total
1	1.13*10^4	9.2*10^5	9.31*10^5
2	3.3*10^5	2.59*10^6	2.92*10^6
3	1.97*10*4	8.06*10^5	8.25*10^5
4	7.86*10^3	3.14*10^5	3.22*10^5
5	1.97*10^4	7.19*10^5	7.39*10^5

Table 3: Cell Counts Across Differing Donors

A second set of cell counting was performed on one single donor. Following the cell counting protocol, a dilution was done in an attempt to “normalize” the number of cells. These samples were then extracted and quantified, along with their waste products, as seen in Table 3. Despite the attempt to ensure each sample had the same quantity of cells and thus the same quantity of DNA, there were very low quantities of DNA that did not correlate with

the lowest cell count. Both buccal samples and vaginal epithelial samples were studied this way and demonstrated automated cell counting was not a reliable method for predicting DNA and waste extraction quantities.

Sample	Total Cell Count	Dilution	DNA Concentration (ng/uL)	Waste Concentration (ng/uL)
Buccal 1	2.49*10^6	18.8uL sample/81.2uL PBS	0.009	0.001
Buccal 2	4.68*10^5	100uL sample/0uL PBS	0.041	0.005
Buccal 3	1.6*10^6	29.25uL sample/70.75uL PBS	0.017	0.002
Vaginal 1	1.27*10^6	67.8uL sample/32.2uL PBS	0.698	0.019
Vaginal 2	7.98*10^5	100uL sample/0uL PBS	1.638	0.05
Vaginal 3	1.58*10^6	50.5uL sample/49.5uL PBS	1.073	0.029

Table 4: Normalized Cell Count and Extraction

Waste and DNA Concentrations

Waste and DNA concentrations were determined using qPCR as cited above on 1uL of sample out of the 40uL of eluted product. The results for these quantifications can be found in Table 4. There were a number of samples that were extracted to quantify the DNA found in the waste product that were not then amplified for pyrosequencing.

Sample	Average DNA Conc (ng/uL)	Average DNA Yield (ng)	Average Waste Conc (ng/uL)	Average Waste Yield (ng)
Blood (n=29)	2.8	112	0.11	4.4

Buccal (n=17)	6.2	248	0.44	17.6
Semen (n=29)	2.9	116	0.32	12.8
Vaginal Epithelia (n=25)	14.3	572	1.7	68

Table 5: DNA and Waste Average Yields

Waste Quantification for Body Fluid Identification

Waste purification resulted in recovery of extraneous DNA remaining that would otherwise have been discarded. This DNA can be utilized to perform body fluid identification if there is a certain quantity of DNA remaining in the waste. This quantity recovered differed depending on based on the body fluid, and the resultant ranges for recovery are demonstrated in Table 5 for results on a total of 100 samples.

Sample	Low Conc (ng/uL)	High Conc (ng/uL)
Blood (n=29)	0.06	0.15
Buccal (n=17)	0.04	0.4
Semen (n=29)	0.01	1.12
Vaginal Epithelia (n=25)	0.8	1.07

Table 6: Concentration Ranges of Successful BFID

Body Fluid Identification

For the following tables, the blue column represents the DNA samples' methylation percentage at the specific CpG site. The orange, gray, and yellow columns represent the respective waste samples' methylation percentage at that same CpG site. Due to limiting reagents, not all DNA samples could be pyrosequenced, so the triplicate wastes are used from their corresponding triplicate extractions. The expected range of methylation was defined as

+/- two standard deviations from the average results obtained from the study of Gauthier et al. [19]

Blood (CG60639435)

In these studies we focused on CpG site 2, where blood is hypermethylated compared to all other body fluids. However, by comparing both CpG sites, as shown in Figure 3, a determination can be made for concentrations greater than 0.01 ng/uL.

Buccal

Buccal cells demonstrate hypermethylation, and therefore even samples that fall below or above the expected range shown in Table 3, could be determined as buccal. Figure 4 illustrates a range of results that differ in concentration but can be successfully defined as buccal and not confused with other body fluids.

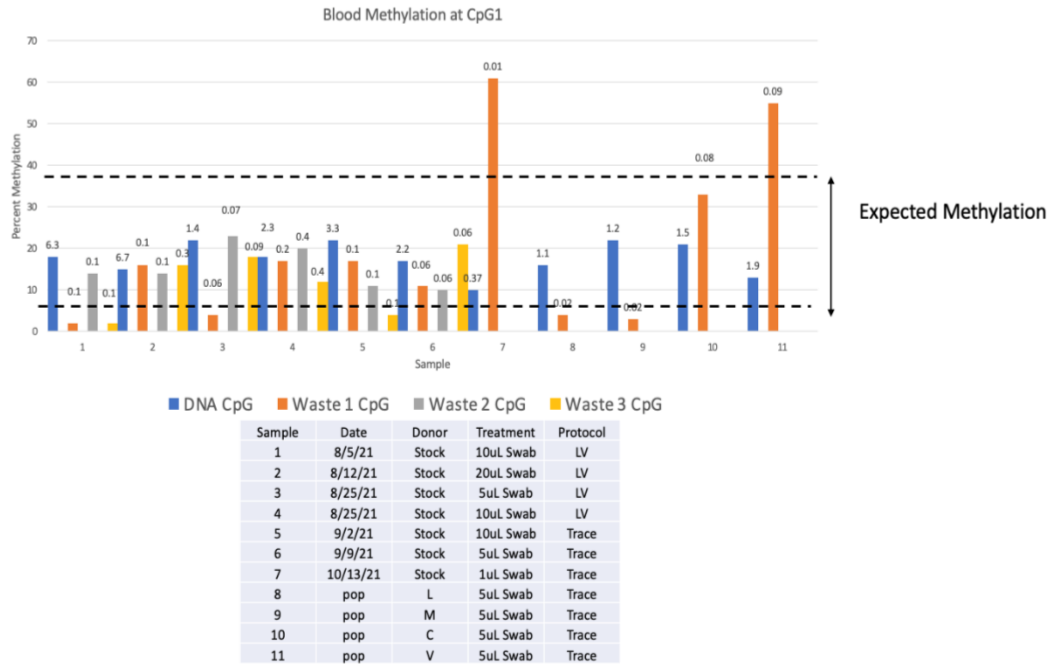
Semen

Due to semen being hypomethylated, it is straightforward to distinguish using this marker compared to all other body fluids. All samples of varying concentrations fell within the accepted range, seen in Figure 5, except for one. While it is unclear what caused the variance in this sample, other samples within the same concentration range were successfully determined.

Vaginal Epithelia

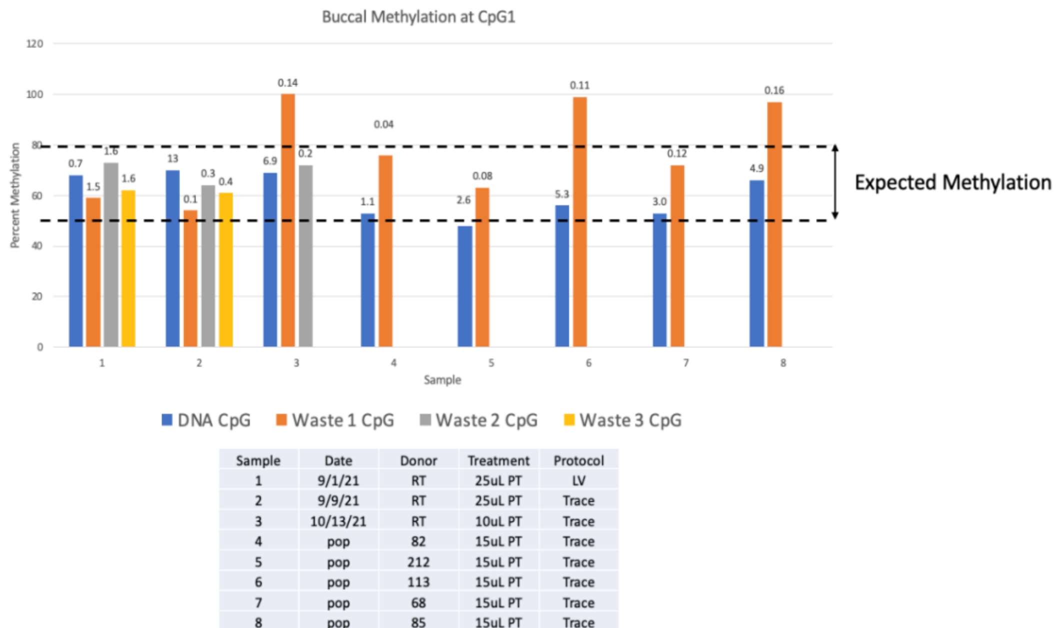
Vaginal epithelial samples demonstrated a wide range of methylation percentages. However, these samples were hypomethylated, whereas other body fluids using this marker would be much more hypermethylated. As demonstrated in Figure 6, while a majority of the results were within the expected range, two data points fell in a range that would overlap

differing body fluid showing methylation over 60%.   Such samples would need to be further evaluated with other body fluid specific markers as they may contain menstrual blood.



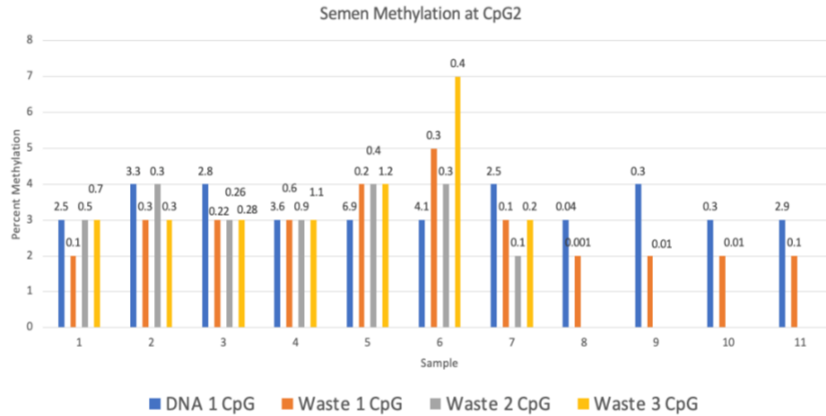
Sample	Date	Donor	Treatment	Protocol
1	8/5/21	Stock	10uL Swab	LV
2	8/12/21	Stock	20uL Swab	LV
3	8/25/21	Stock	5uL Swab	LV
4	8/25/21	Stock	10uL Swab	LV
5	9/2/21	Stock	10uL Swab	Trace
6	9/9/21	Stock	5uL Swab	Trace
7	10/13/21	Stock	1uL Swab	Trace
8	pop	L	5uL Swab	Trace
9	pop	M	5uL Swab	Trace
10	pop	C	5uL Swab	Trace
11	pop	V	5uL Swab	Trace

Figure 5: Blood Methylation Percentages at CpG Site 1



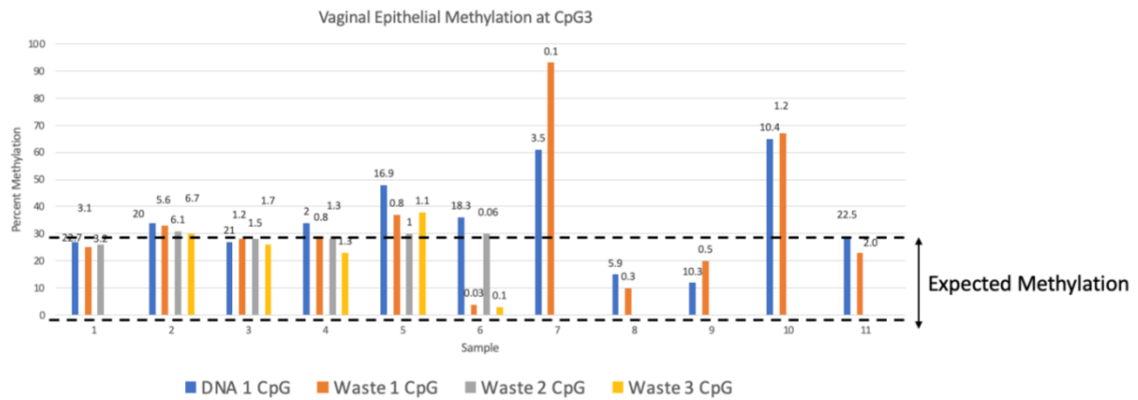
Sample	Date	Donor	Treatment	Protocol
1	9/1/21	RT	25uL PT	LV
2	9/9/21	RT	25uL PT	Trace
3	10/13/21	RT	10uL PT	Trace
4	pop	82	15uL PT	Trace
5	pop	212	15uL PT	Trace
6	pop	113	15uL PT	Trace
7	pop	68	15uL PT	Trace
8	pop	85	15uL PT	Trace

Figure 6: Buccal Methylation Percentages at CpG site 3



Sample	Date	Donor	Treatment	Protocol
1	8/5/21	Stock	10uL swab	LV
2	8/12/21	Stock	20uL swab	LV
3	8/25/21	Stock	5uL swab	LV
4	8/25/21	Stock	10uL swab	LV
5	9/2/21	Stock	10uL swab	Trace
6	9/9/21	Stock	5uL swab	Trace
7	10/13/21	Stock	1uL swab	Trace
8	pop	48	1uL swab	Trace
9	pop	11	1uL swab	Trace
10	pop	81	1uL swab	Trace
11	pop	166	1uL swab	Trace

Figure 7: Semen Methylation Percentages at CpG Site 2



Sample	Date	Donor	Treatment	Protocol
1	8/12/21	RT	150uL PT	LV
2	8/12/21	RT	100uL PT	LV
3	8/21/21	RT	50uL PT	LV
4	9/1/21	RT	10uL PT	LV
5	9/9/21	RT	25uL PT	Trace
6	10/13/21	RT	10uL PT	Trace
7	pop	67	15uL PT	Trace
8	pop	183	15uL PT	Trace
9	pop	62	15uL PT	Trace
10	pop	214	15uL PT	Trace
11	pop	207	15uL PT	Trace

Figure 8: Vaginal Epithelial Methylation Percentages at CpG Site 3

Replication Study

Triplicates of each sample were analyzed to determine method reproducibility and sensitivity. These samples were all run in parallel with each other to ensure the same treatment times were performed for each sample type.

As demonstrated by the blood replicates at CpG1 for marker CG06379435, all samples produced methylation percentages within the expected range. The concentrations of the waste, as shown as data labels above the bars, demonstrate a standard deviation of 0.04 and while there appears to be variability between each replicate, they all successfully reported methylation percentages associated with blood at CpG1, as seen in Figure 7.

Two out of three DNA/waste buccal pairs fell within the expected range of methylation at CpG1 using the marker BCAS4 as seen in Figure 8. While one waste sample demonstrated hypermethylation outside of the expected range, the buccal sample is already hypermethylated when compared to other body fluids and therefore can still be positively determined as a buccal sample despite its methylation percentage above the expected range.

Semen is the most hypomethylated sample and has produced the most stable results of all the body fluids processed. Figure 9 demonstrates the hypomethylation demonstrated by three semen replicates with only 5uL of starting sample being processed at CpG2 for marker ZC3H12D. The standard deviation of DNA methylation is 1.5 while for waste it is 0.6. For these amounts, the standard deviation of DNA is 0.25 and for waste it is 0.16. Semen has proven to be the most replicable and reliable body fluid when performing body fluid identification, which is crucial for distinguishing sex crimes.

Vaginal epithelial DNA exhibited hypermethylation at VE08 when compared to the expected range, seen in Figure 12, which may be due to the donor being a hypermethylated

individual, or the presence of menstrual blood. Despite this, all of these samples fall below the level of potentially interfering saliva samples and thus could be positively determined as vaginal epithelial fluid.

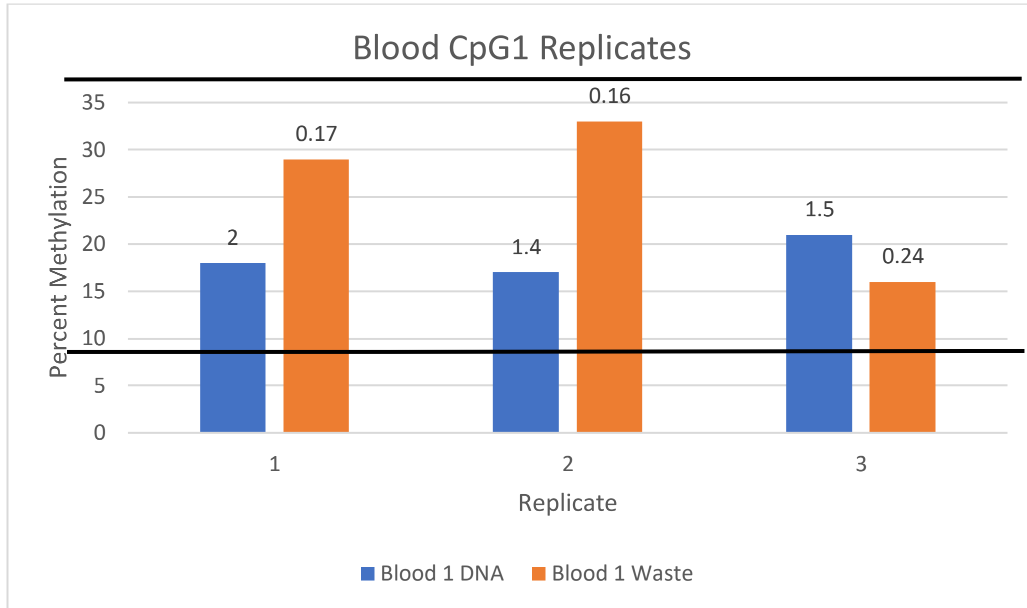


Figure 9: Blood Replication Results

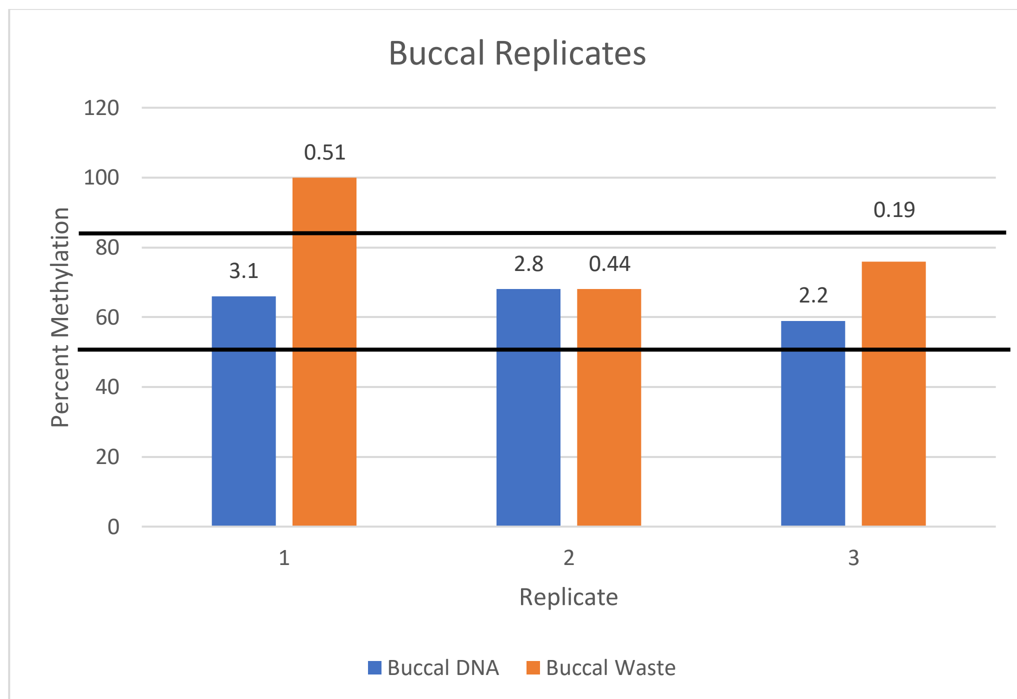


Figure 10: Buccal Replication Results

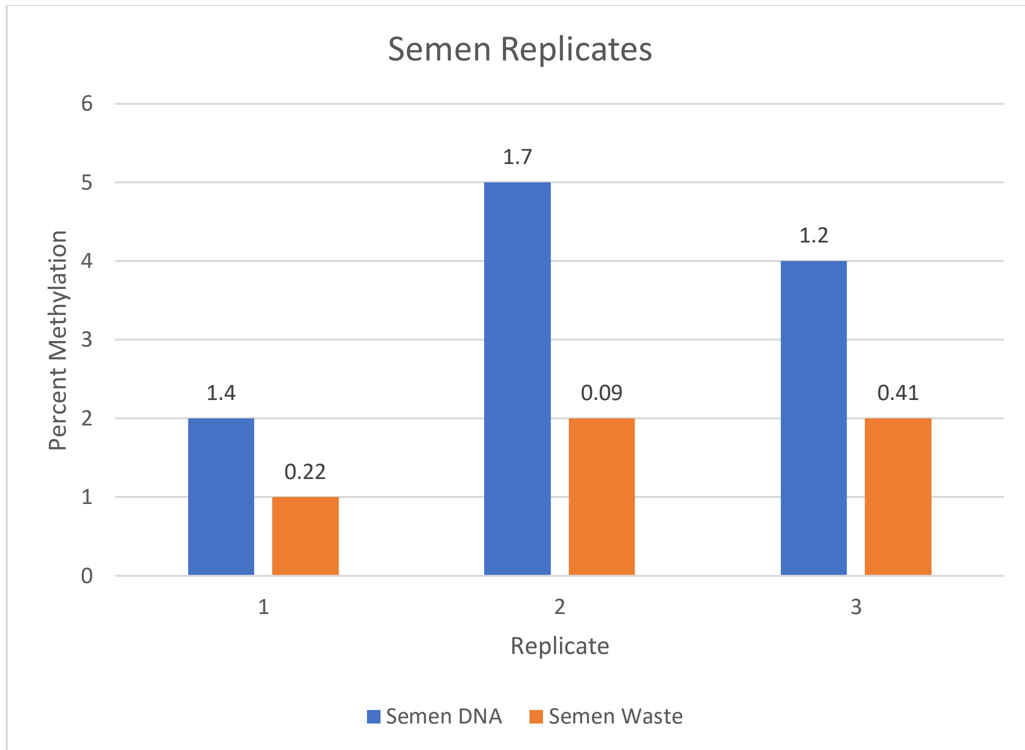


Figure 11: Semen Replication Results

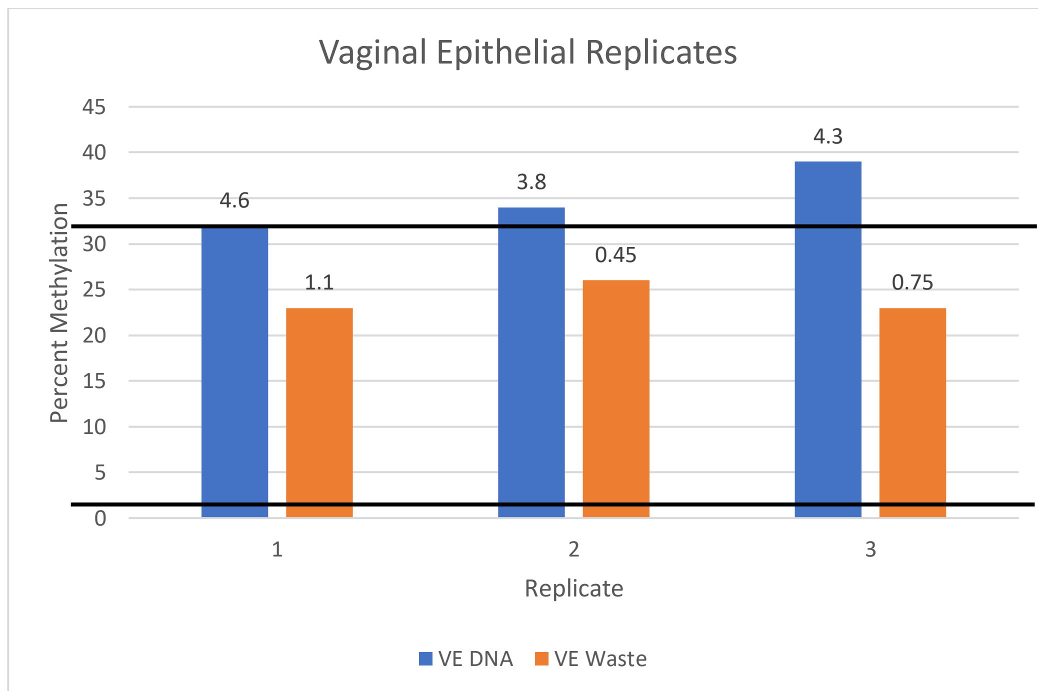


Figure 12: Vaginal Epithelial Replication Results

Discussion

One major difference between the methylation data in this study and that performed by Gauthier et al. was that a much smaller concentration of DNA was utilized here. [19] Original methylation patterns were studied using an average of 200ng of DNA, while the waste here produced an average of 4.4 ng per sample in blood, 17.6ng in buccal, 12.8ng in semen, and 68ng in vaginal epithelia, which is significantly less than in the previous study. In the blood samples specifically, when calculating for the effects of a 40% yield from bisulfite conversion and only 1.0 uL of that sample being used for PCR and pyrosequencing, the actual quantity amplified may have been as little as 0.15ng. This means the likelihood for PCR bias and stochastic effects are dramatically increased, which could result in unexpected additional cell, approximately 6pg of DNA, is methylated in a 15-cell sample, there would be an increase of 7% in the methylation percentage. When considering hypomethylated samples, this variation could put them within range of another body fluid. Similarly, the decrease of one methylated cell in an expected hypermethylated sample could lower the methylation to within the range of a more hypomethylated body fluid. such results could lead to possible false negatives when attempting to identify unknown body fluids.

This PCR bias may also explain why some expected hypermethylated samples are even more hypermethylated and fall above the range estimated. With a limited amount of sample, the probability of an increase in hypermethylation due to a nonproportional amplification is possible. However, if these samples remain out of range of a conflicting body fluid, as the replicates of the vaginal epithelia and buccal samples processed in this study have demonstrated, they can still be positively identified as their respective body fluid.

When comparing samples with different donors, the possibility that each individual may demonstrate unique methylation percentages and patterns for themselves also has to be considered. This could be due to competing epigenetic factors as well as genetic differences within the tDMRs which would lead to each individual possessing methylation percentages within a larger range than calculated prior.

Similarly, individuals shed different amounts of cells as seen in the cell counting experiment, that could also lead to different amounts of DNA being extracted from those cells as well. Some individuals produced six times as much DNA as others, which would lead to an increase in concentration in the waste as well (Figures 5-8). This may make their samples easier to analyze even when minute amounts of sample are being analyzed. The more DNA a sample contains, the more reliable the epigenetic analysis is through PCR, which is why it is important to determine how much DNA can successfully be used to perform body fluid identification.

As identified in a study relating to epithelial shedding in forensic cases, depending on the duration of contact and the force used for the contact, the overall number of cells and thus the amount of DNA will differ [39]. This results in additional challenges to predicting how much DNA a touch sample may theoretically yield as information regarding the way the sample was deposited and collected will differ and may be inconsistent. This can also become even more difficult if the sample is unexpectedly contaminated with additional cells.

The amount of DNA that can be used to positively identify a body fluid may be impacted by the type of body fluid being analyzed. For example, there appeared to be less DNA found in blood samples than all other samples. This would increase concern for PCR bias which may also cause false negatives or false positives. However, buccal cells contained

an abundance of DNA, and therefore demonstrated successful identification across all samples studied.

This becomes extremely important after bisulfite conversion, as only 4-14% of DNA is typically recovered [21]. By starting with a larger amount of sample, the possibility of ensuring a homogenized and representative sample after bisulfite conversion is more likely. For this reason, if blood is being analyzed, it may be crucial to ensure there is a larger concentration of sample in the waste, and if a mixture is analyzed, it would be imperative to recover as much DNA as possible to ensure that any body fluid being analyzed would be successfully converted, amplified, and sequenced.

Previous studies have demonstrated valid and consistent results down to 0.1ng of DNA concentrations across all four body fluids [38]. This study demonstrated that samples below 0.1ng/mL of DNA prior to bisulfite conversion with the significant loss of sample could successfully be identified for all four body fluids studied. This decrease in concentration can be used to assist in sensitivity studies and determining how much sample is realistically needed to obtain reliable and reproducible results.

A further way to enhance results and significance is to increase the number of samples that are tested. In this study, only five to six individuals were sampled, all of which are in the 18-30 age demographic. Epigenetic factors can possibly change as age increases, which is why it is important to increase the age diversity within this study [28]. Furthermore, evaluation of additional markers, may result in better body fluid discrimination as more samples with individuals of different backgrounds are analyzed. Age, diet, drug habits, and other life experiences can affect methylation patterns within cells and therefore may have an impact on the methylation pattern of body fluids as well [11].

A statistical analysis should also be performed on all of the data acquired to determine the significance of those samples which fall outside of the designated ranges previously published [19]. An ROC curve would provide information regarding the success of different body fluids in regards to their probability of identification.

When compared to previously used confirmatory tests, using the waste product has demonstrated to be most efficient in preventing the damage to the main sample. Crystal tests for confirmatory blood testing involve drying the sample with chemicals such as acetic acid and halide which would lead to the destruction of the sample [40]. In scenarios where there is minimal sample, epigenetic analysis may be more successful, and utilizing the waste ensures the pure extract can be used for future testing.

A previous method of epigenetic body fluid identification utilized much larger quantities of DNA, 3-50ng. While this testing proved successful, the quantity of DNA needed is much higher than the range determined in this study. Using MSRE-PCR, this testing also utilized one specific tDMR to perform epigenetic analysis, rather than the markers here that were specific for each body fluid [41]. Using the pyrosequencer and markers associated with the body fluids being tested for, a more accurate and efficient method of epigenetic analysis is being performed.

This method of utilizing the waste product to perform body fluid identification not only yields promising results in its ability to differentiate body fluids with such low quantities of DNA, but also provides for additional DNA that can be utilized for other genetic testing

Conclusion

The EZ1 Robot is not 100% efficient in delivering all extracted DNA into the final elution container. Using this remaining sample found in the waste product, body fluid identification can be performed if a significant quantity of DNA remains. Due to the unique properties of each body fluid, the quantity of DNA required to perform epigenetic body fluid identification differs.

It is also important to consider PCR bias and stochastic effects when working with low quantities of DNA that are typically found in the waste product. Low levels of DNA input can increase the variation in methylation percentages. Therefore, when attempting to analyze a mixture of body fluids, it is important to ensure sufficient DNA is present and that appropriate validation studies have been performed to properly determine the range of expected methylation percentages.

By utilizing the extraction waste product that can be used for body fluid identification, more information can be gained from the same amount of original sample, thus ensuring that the DNA can be utilized for genetic profiling, body fluid identification, and mixture analysis. Body fluid identification used to be only possible with an abundance of the sample. With this methodology, labs that utilize the EZ1 Robot can include the waste collection in their daily practice to determine what further testing can be conducted on the waste product, including body fluid identification.

The information gained using body fluid identification can be crucial for cases requiring mixture analysis or specific information regarding the type of DNA found, not just the donor. More information that can be gained from the same limited sample would benefit crime labs that are typically already working with low template DNA samples. The chemical

alterations involved in body fluid identification using bisulfite conversion makes it impossible to use the same sample for both genetic profiling and body fluid identification; however, by utilizing the waste product, it can be ensured that there is no damage is imparted to the critical DNA extract used for genetic profiling, and body fluid identification can still be performed.

Future Work

While this project is a big step towards performing body fluid identification regularly in labs, there are still a myriad of projects that can be completed to further this research. One such project is utilizing PCR product waste rather than extraction waste. In doing this, the initial DNA analysis workflow would remain the same until PCR is completed, at which step the PCR 'waste' containing the original template DNA, would be separated and conserved for methylation rather than utilizing DNA extraction waste.

A second experiment would entail performing both STR typing and body fluid identification on bisulfite converted DNA. This would involve a novel primer design that would allow genetic profiling to be performed on chemically altered DNA, therefore keeping the entire workflow for the DNA analysis the same until the sample is split to have both body fluid identification and genetic profiling done on the same sample. This method would ensure the least amount of DNA sample is lost due to it being utilized for the entirety of the analysis, rather than splitting up the sample earlier on.

There is also further research to be done on the loci used as markers to analyze body fluids. The range of expected methylation percentage ranges may increase when using low levels of DNA input. Due to this issue, it would also be beneficial to rework the bisulfite

conversion protocol in an attempt to increase the overall yield to ensure there is more DNA that survives the chemical alteration process.

A final study that can be completed would be to determine if the isolated waste can be recombined with the original extracted DNA to increase the overall yield of the extraction process. When working with evidence samples that are very small and minute, any and all recoverable DNA is crucial. If combining the original extracted DNA with its purified waste can increase the amount of DNA usable for genetic profiling and increase the efficiency and accuracy of the DNA analyzed, more complete data would be obtained that could increase the likelihood ratio and provide more statistical weight to the evidence.

There may also be work to be done in regards to the amount of template used in PCR. Only 1uL of template was utilized to perform PCR prior to pyrosequencing. If this template volume is increased, there may be an increase in PCR product which may yield more effective body fluid identification.

References

[1] Collins, L. J., Schönfeld, B., & Chen, X. S. (2011). The epigenetics of non-coding RNA. *Handbook of Epigenetics*, 49–61. https://doi.org/10.1016/b978-0-12-375709-8.00004-6

[2] Harbison, S. A., & Fleming, R. I. (2016, February 10). *Forensic body fluid identification: State of the art: RRFMS*. Research and Reports in Forensic Medical Science. Retrieved June 6, 2022, from https://www.dovepress.com/forensic-body-fluid-identification-state-of-the-art-peer-reviewed

[3] Delaney, C., Garg, S. K., & Yung, R. (2015). Analysis of DNA methylation by pyrosequencing. *Methods in Molecular Biology*, 249–264. https://doi.org/10.1007/978-1-4939-2963-4_19

[4] Gahlon, H. L. (2020). A brief history and practical applications in DNA extraction. *CHIMIA*, *74*(11), 907. https://doi.org/10.2533/chimia.2020.907

[5] Foley, C., O'Farrelly, C., & Meade, K. G. (2011, May 23). *Technical note: Comparative analyses of the quality and yield of genomic DNA from invasive and noninvasive, automated and manual extraction methods*. Journal of Dairy Science. Retrieved June 4, 2022, from https://www.sciencedirect.com/science/article/pii/S0022030211003031

[6] Fernandez-Tejero, N., Ghemrawi, M., McCord, B., Lee, S. B. (2022). DNA extraction methods in forensic analysis. *Encyclopedia of Analytical Chemistry*, 1–18. https://doi.org/10.1002/9780470027318.a1104m.pub2

[7] Hanson, E., Lopez, J., Ballantyne, J., "A Novel DNA/RNA Co-Extraction Method Using EZ1 DNA Investigator Extraction 'Waste'" (2017)

[8] van Oorschot, R. A., Ballantyne, K. N., & Mitchell, R. J. (2010). Forensic trace DNA: a review. *Investigative genetics*, *1*(1), 14. https://doi.org/10.1186/2041-2223-1-14

[9] Gillooly, J. F., Hein, A., & Damiani, R. (2015). Nuclear DNA content varies with cell size across human cell types. *Cold Spring Harbor Perspectives in Biology*, *7*(7). https://doi.org/10.1101/cshperspect.a019091

[10] Kristensen, L. S., Nielsen, H. M., & Hansen, L. L. (2009, October 15). *Epigenetics and cancer treatment*. European Journal of Pharmacology. Retrieved June 4, 2022

[11] Vidaki, A., & Kayser, M. (2017). From forensic epigenetics to forensic epigenomics: Broadening DNA investigative intelligence. *Genome Biology*, *18*(1). https://doi.org/10.1186/s13059-017-1373-1

[12] Lim, W.-J., Kim, K. H., Kim, J.-Y., Jeong, S., & Kim, N. (2019). Identification of DNA-methylated CPG islands associated with gene silencing in the adult body tissues of the

Ogye chicken using RNA-seq and reduced representation bisulfite sequencing. *Frontiers in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.00346

[13] Bocklandt, S., Lin, W., Sehl, M. E., Sánchez, F. J., Sinsheimer, J. S., Horvath, S., & Vilain, E. (2011). Epigenetic predictor of age. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0014821

[14] Ahmadian, A., Ehn, M., & Hober, S. (2006). Pyrosequencing: History, biochemistry and future. *Clinica Chimica Acta*, *363*(1-2), 83–94. https://doi.org/10.1016/j.cccn.2005.04.038

[15] Clark, J., & Rager, J. E. (2020). Epigenetics: An overview of CPG methylation, chromatin remodeling, and regulatory/noncoding RNAS. *Environmental Epigenetics in Toxicology and Public Health*, 3–32. https://doi.org/10.1016/b978-0-12-819968-8.00001-9" https://doi.org/10.1016/b978-0-12-819968-8.00001-9

[16] Qiagen. (2021, September). EZ1&2 Investigator Kit Handbook. USA; Qiagen.

[17] Antunes, J., Gauthier, Q., Aguiar-Pulido, V., Duncan, G., & McCord, B. (2021). A data-driven, high-throughput methodology to determine tissue-specific differentially methylated regions able to discriminate body fluids. *ELECTROPHORESIS*, *42*(9-10), 1168–1176. https://doi.org/10.1002/elps.202000217

[18] Antunes, J., Silva, D. S. B. S., Balamurugan, K., Duncan, G., Alho, C. S., & McCord, B. (2016). Forensic discrimination of vaginal epithelia by DNA methylation analysis through pyrosequencing. *ELECTROPHORESIS*, *37*(21), 2751–2758. https://doi.org/10.1002/elps.201600037

[19] Gauthier, Q. T., Cho, S., Carmel, J. H., & McCord, B. R. (2019). Development of a body fluid identification multiplex via DNA methylation analysis. *ELECTROPHORESIS*. https://doi.org/10.1002/elps.201900118

[20] Ghemrawi, M., Fischinger, F., Duncan, G., Dukes, M. J., Guilliano, M., & McCord, B. (2021). Developmental validation of SpeID: A pyrosequencing-based assay for species identification. *Forensic Science International: Genetics*, *55*, 102560. https://doi.org/10.1016/j.fsigen.2021.102560

[21] Darst, R. P., Pardo, C. E., Ai, L., Brown, K. D., & Kladde, M. P. (2010). Bisulfite sequencing of DNA. *Current Protocols in Molecular Biology*, *91*(1). https://doi.org/10.1002/0471142727.mb0709s91

[22] Sakurada, K., Watanabe, K., & Akutsu, T. (2020). Current Methods for Body Fluid Identification Related to Sexual Crime: Focusing on Saliva, Semen, and Vaginal Fluid. *Diagnostics (Basel, Switzerland)*, *10*(9), 693. https://doi.org/10.3390/diagnostics10090693

[23] Chatterjee S. (2019). Saliva as a forensic tool. *Journal of forensic dental sciences*, *11*(1), 1–4. https://doi.org/10.4103/jfo.jfds_69_18

[24] Libretexts. (2020, August 14). *4.2: Digestion and absorption of carbohydrates*. Medicine LibreTexts. Retrieved June 8, 2022, from https://med.libretexts.org/" https://med.libretexts.org/

[25] *Bureau of Criminal Apprehension*. Learn more about DPS, Divisions, Programs, Boards and Committees. (n.d.). Retrieved June 8, 2022, from https://dps.mn.gov/divisions/bca/bca-divisions/forensic-science/

[26] Hausmann R, Pregler C, Schellmann B. The value of the Lugol's iodine staining technique for the identification of vaginal epithelial cells. Int J Legal Med. 1994;106(6):298-301. doi: 10.1007/BF01224775. PMID: 7524629.

[27] Tobe, S. S., Watson, N., & Daéid, N. N. (2007). Evaluation of six presumptive tests for blood, their specificity, sensitivity, and effect on high molecular-weight DNA. *Journal of Forensic Sciences*, *52*(1), 102–109. https://doi.org/10.1111/j.1556-4029.2006.00324.x" https://doi.org/10.1111/j.1556-4029.2006.00324.x

[28] Pal, S., & Tyler, J. K. (2016). Epigenetics and aging. *Science advances*, *2*(7), e1600584. https://doi.org/10.1126/sciadv.1600584

[29] Madi, T., Balamurugan, K., Bombardi, R., Duncan, G., & McCord, B. (2012). The determination of tissue-specific DNA methylation patterns in forensic biofluids using bisulfite modification and pyrosequencing. *ELECTROPHORESIS*, *33*(12), 1736–1745. https://doi.org/10.1002/elps.201100711" https://doi.org/10.1002/elps.201100711

[30] Centers for Disease Control and Prevention. (2022, May 18). *What is epigenetics?* Centers for Disease Control and Prevention. Retrieved June 8, 2022, from https://www.cdc.gov/genomics/disease/epigenetics.htm

[31] Bell, J. T., & Spector, T. D. (2011). A twin approach to unraveling epigenetics. *Trends in Genetics : TIG*, *27*(3), 116–125. https://doi.org/10.1016/j.tig.2010.12.005" https://doi.org/10.1016/j.tig.2010.12.005

[32] Feinberg, A. P., & Tycko, B. (2004). The history of cancer epigenetics. *Nature Reviews Cancer*, *4*(2), 143–153. https://doi.org/10.1038/nrc1279

[33] Moore, L. D., Le, T., & Fan, G. (2012). DNA methylation and its basic function. *Neuropsychopharmacology*, *38*(1), 23–38. https://doi.org/10.1038/npp.2012.112

[34] Ali, N., Rampazzo, R., Costa, A., & Krieger, M. A. (2017). Current Nucleic Acid Extraction Methods and Their Implications to Point-of-Care Diagnostics. *BioMed research international*, *2017*, 9306564. https://doi.org/10.1155/2017/9306564

[35] Psifidi, A., Dovas, C. I., Bramis, G., Lazou, T., Russel, C. L., Arsenos, G., & Banos, G. (2015). Comparison of eleven methods for genomic DNA extraction suitable for large-scale whole-genome genotyping and long-term DNA banking using blood samples. *PLOS ONE*, *10*(1). https://doi.org/10.1371/journal.pone.0115960" https://doi.org/10.1371/journal.pone.0115960

[36] Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., & Polz, M. F. (2005). PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and environmental microbiology*, *71*(12), 8966–8969. https://doi.org/10.1128/AEM.71.12.8966-8969.2005" https://doi.org/10.1128/AEM.71.12.8966-8969.2005

[37] Lazarevic, V., Gaïa, N., Girard, M., François, P., & Schrenzel, J. (2013). Comparison of DNA extraction methods in analysis of salivary bacterial communities. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0067699

[38] McCord, B., Gauthier, Q., Alghanim, H., Antunes, J., Fernandez Tejero, N., Duncan, G., & Balamurugan, K. (2019). Applications of epigenetic methylation in body fluid identification, age determination and phenotyping. *Forensic Science International: Genetics Supplement Series*, *7*(1), 485–487. https://doi.org/10.1016/j.fsigss.2019.10.061

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

[40] Virkler, K., & Lednev, I. K. (2009). Analysis of body fluids for forensic purposes: From laboratory testing to non-destructive rapid confirmatory identification at a crime scene. *Forensic Science International*, *188*(1-3), 1–17. https://doi.org/10.1016/j.forsciint.2009.02.013

[41] An, J. H., Choi, A., Shin, K.-J., Yang, W. I., & Lee, H. Y. (2012). DNA methylation-specific multiplex assays for body fluid identification. *International Journal of Legal Medicine*, *127*(1), 35–43. https://doi.org/10.1007/s00414-012-0719-1

[42] Barnett, R., & Larson, G. (2011). A phenol–chloroform protocol for extracting DNA from ancient samples. *Methods in Molecular Biology*, 13–19. https://doi.org/10.1007/978-1-61779-516-9_2

[43] Crowley, L. C., Marfell, B. J., Christensen, M. E., & Waterhouse, N. J. (2016). Measuring cell death by Trypan Blue Uptake and light microscopy. *Cold Spring Harbor Protocols*, *2016*(7). https://doi.org/10.1101/pdb.prot087155

[44] Vidaki, A., Giangasparo, F., & Syndercombe Court, D. (2016). Discovery of potential DNA methylation markers for forensic tissue identification using bisulphite

pyrosequencing. *ELECTROPHORESIS*, *37*(21), 2767–2779. https://doi.org/10.1002/elps.201600261

[45] Attwood, J. T., Yung, R. L., & Richardson, B. C. (2002). DNA methylation and the regulation of Gene Transcription. *Cellular and Molecular Life Sciences CMLS*, *59*(2), 241–257. https://doi.org/10.1007/s00018-002-8420-z

[46] Ladd, C., Adamowicz, M. S., Bourke, M. T., Scherczinger, C. A., & Lee, H. C. (1999). A systematic analysis of secondary DNA transfer. *Journal of Forensic Sciences*, *44*(6). https://doi.org/10.1520/jfs14599j

[47] Burrill, J., Daniel, B., & Frascione, N. (2019). A review of Trace "touch DNA" deposits: Variability factors and an exploration of cellular composition. *Forensic Science International: Genetics*, *39*, 8–18. https://doi.org/10.1016/j.fsigen.2018.11.019

[48] Psifidi, A., Dovas, C. I., Bramis, G., Lazou, T., Russel, C. L., Arsenos, G., & Banos, G. (2015). Comparison of eleven methods for genomic DNA extraction suitable for large-scale whole-genome genotyping and long-term DNA banking using blood samples. *PLOS ONE*, *10*(1). https://doi.org/10.1371/journal.pone.0115960

[49] Kennedy, N. A., Walker, A. W., Berry, S. H., Duncan, S. H., Farquarson, F. M., Louis, P., Thomson, J. M., Satsangi, J., Flint, H. J., Parkhill, J., Lees, C. W., & Hold, G. L. (2014). The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rrna gene sequencing. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0088982

[50] Gervin, K., Page, C. M., Aass, H. C., Jansen, M. A., Fjeldstad, H. E., Andreassen, B. K., Duijts, L., van Meurs, J. B., van Zelm, M. C., Jaddoe, V. W., Nordeng, H., Knudsen, G. P., Magnus, P., Nystad, W., Staff, A. C., Felix, J. F., & Lyle, R. (2016). Cell type specific DNA methylation in cord blood: A 450K-reference data set and cell count-based validation of estimated cell type composition. *Epigenetics*, *11*(9), 690–698. https://doi.org/10.1080/15592294.2016.1214782

[51] Varela Morillas, Á., Suhling, K., & Frascione, N. (2022). Unlocking the potential of forensic traces: Analytical approaches to generate investigative leads. *Science & Justice*, *62*(3), 310–326. https://doi.org/10.1016/j.scijus.2022.03.005

[52] Jones, E. L. (2005). The identification of semen and other body fluids. *Forensic science handbook*, *2*, 329-99.

[53] Mishra, A., Gondhali, U., & Choudhary, S. (2022). Potential of DNA technique-based body fluid identification. *Handbook of DNA Profiling*, 321–337. https://doi.org/10.1007/978-981-16-4318-7_11

[54] Gaensslen, R. E. (2000). Forensic Analysis of Biological Evidence. *Forensic Sciences*, *1*(29).

[55] Haddrill, P. R. (2021). Developments in forensic DNA analysis. *Emerging Topics in Life Sciences*, *5*(3), 381–393. https://doi.org/10.1042/etls20200304

[56] Shabani, M., Borry, P., Smeers, I., & Bekaert, B. (2018). Forensic epigenetic age estimation and beyond: Ethical and legal considerations. *Trends in Genetics*, *34*(7), 489–491. https://doi.org/10.1016/j.tig.2018.03.006

[57] Freire-Aradas, A., Girón-Santamaría, L., Mosquera-Miguel, A., Ambroa-Conde, A., Phillips, C., Casares de Cal, M., Gómez-Tato, A., Álvarez-Dios, J., Pospiech, E., Aliferi, A., Syndercombe Court, D., Branicki, W., & Lareu, M. V. (2022). A common epigenetic clock from childhood to old age. *Forensic Science International: Genetics*, *60*, 102743. https://doi.org/10.1016/j.fsigen.2022.102743

Rachel Tolces is from Boca Raton, Florida and graduated with her B.S. in Biology from the University of Florida in Spring 2020. After graduating, she began her studies at Florida International University of obtain an M.S. in Forensic Science. Upon graduating with her Master Degree, Rachel will work at AFDIL as a Forensic Scientist in Dover, DE.