

Spring 2019

# Methylation-Specific Differentiation of Vaginal Epithelial Cells for Forensic Tissue Typing by Bisulfite Conversion and Pyrosequencing

Elise Pood

Follow this and additional works at: [https://aquila.usm.edu/masters\\_theses](https://aquila.usm.edu/masters_theses)

Part of the [Biology Commons](#), and the [Genetics and Genomics Commons](#)

---

## Recommended Citation

Pood, Elise, "Methylation-Specific Differentiation of Vaginal Epithelial Cells for Forensic Tissue Typing by Bisulfite Conversion and Pyrosequencing" (2019). *Master's Theses*. 641.  
[https://aquila.usm.edu/masters\\_theses/641](https://aquila.usm.edu/masters_theses/641)

This Masters Thesis is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Master's Theses by an authorized administrator of The Aquila Digital Community. For more information, please contact [Joshua.Cromwell@usm.edu](mailto:Joshua.Cromwell@usm.edu).

METHYLATION-SPECIFIC DIFFERENTIATION OF VAGINAL EPITHELIAL  
CELLS FOR FORENSIC TISSUE TYPING BY BISULFITE CONVERSION AND  
PYROSEQUENCING

by

Elise Nicole Pood

A Thesis  
Submitted to the Graduate School,  
the College of Arts and Sciences  
and the School of Criminal Justice, Forensic Science, and Security  
at The University of Southern Mississippi  
in Partial Fulfillment of the Requirements  
for the Degree of Master of Science

Approved by:

Dr. Kuppareddi Balamurugan, Committee Chair  
Dr. Philip Carlan  
Dr. Dean Bertram

---

Dr. Kuppareddi  
Balamurugan  
Committee Chair

---

Dr. Lisa Nored  
Director of School

---

Dr. Karen S. Coats  
Dean of the Graduate School

May 2019

COPYRIGHT BY

Elise Nicole Pood

2019

*Published by the Graduate School*



## ABSTRACT

The identification of bodily fluids and tissues is often applied to criminal investigations to clarify events that may or may not have taken place. Current forensic techniques can identify blood, saliva, seminal fluid, and spermatozoa, but there is a clear absence of reliable testing to identify vaginal epithelial tissue. Though there are serological tests available for this purpose, tissue-specific methylation markers have recently been investigated as a candidate for the identification of blood, saliva, and spermatozoa.

In this study, tissue-specific methylation markers were analyzed to identify a set of markers for the differentiation of vaginal fluid from blood, saliva, and semen. From the four tissue types collected, genomic DNA was extracted, quantitated, and bisulfite-modified to preserve the methylation information. Candidate markers were amplified then pyrosequenced to determine the percent methylation of specific CpG sites. The level of significance between tissues was determined using one way ANOVA Tukey's posthoc test by SPSS statistical package.

Three markers, cg4739647, cg6266993, and cg9323727 were found to be hypermethylated in vaginal fluid compared to blood, saliva, and semen. The differences between methylation levels at nearly all analyzed CpG sites were found to be significant, suggesting that these markers may be used to identify vaginal epithelial tissue for forensic purposes. Pyrosequencing has several advantages over conventional serological analysis, and the development of a multiplex kit using these markers will aid in the conservation of precious DNA samples that can be used for other forensic purposes.

## ACKNOWLEDGMENTS

Above all, I am grateful to Dr. Kuppareddi Balamurugan for his unequivocal support, encouragement, and guidance through my coursework and research, and for imparting to his students a lifetime's worth of education and experience pertaining to the Biology and Forensic science fields. Likewise, I am grateful to the members of my committee for providing their time and expertise to ensure the quality of this study.

I would also like to recognize Dr. Lisa Nored, Chair of the School of Criminal Justice, Forensic Science, and Security, and other university faculty and staff members for their constant support and advisement during my academic career at the University of Southern Mississippi.

## DEDICATION

This research is dedicated to my late father, Dr. Elliott A. Pood, who inspired in me a respect for higher education and an interest in the criminal justice process.

TABLE OF CONTENTS

ABSTRACT ..... ii

ACKNOWLEDGMENTS ..... iii

DEDICATION ..... iv

LIST OF TABLES ..... viii

LIST OF ILLUSTRATIONS ..... ix

CHAPTER I - INTRODUCTION ..... 1

    Current DNA Technology..... 1

    Current Tissue ID Methods..... 2

    DNA Methylation ..... 4

    Bisulfite Modification..... 6

    Methods to Measure DNA Methylation ..... 7

    Pyrosequencing ..... 8

    Aims and Objectives..... 9

CHAPTER II – REVIEW OF LITERATURE ..... 11

    Current Presumptive and Confirmatory Tissue ID Methods ..... 11

        RSID ..... 11

        Blood..... 11

        Saliva..... 14

        Semen..... 15

Vaginal epithelial cells.....	16
mRNA profiling.....	16
Microbiome Differentiation.....	17
Methylation-Based Tissue Differentiation.....	17
CHAPTER III - MATERIALS AND METHODS.....	20
Sample Collection.....	20
Blood.....	20
Buccal.....	20
Semen.....	20
Vaginal Epithelial.....	20
DNA Extraction and Quantitation.....	21
Bisulfite Conversion.....	22
Bisulfite Conversion Cleanup Process.....	23
CpG Marker Selection and Primer Design.....	24
Polymerase Chain Reaction (PCR).....	25
Pyrosequencing.....	26
Data Analysis.....	27
Bisulfite Controls.....	27
CHAPTER IV – RESULTS.....	29
Vaginal Epithelial Marker 4739647.....	32



Vaginal Epithelial Marker 6266993.....	35
Vaginal Epithelial Marker 9323727.....	37
CHAPTER V – DISCUSSION.....	39
APPENDIX A – Primer Specifications .....	45
REFERENCES .....	46

## LIST OF TABLES

Table 1.	Bisulfite reaction components.....	22
Table 2.	PCR components per sample for sample amplification.....	25
Table 3.	PCR cycling conditions for sample amplification. ....	26
Table 4.	Percent Methylation Averages and Standard Deviations-Marker 4739647 .	33
Table 5.	Levels of significance (p) for 4739647 .....	34
Table 6.	Percent Methylation Averages and Standard Deviations-Marker 6266993 .	35
Table 7.	Levels of significance (p) for 6266993 .....	36
Table 8.	Percent Methylation Averages and Standard Deviations-Marker 9323727 .	37
Table 9.	Levels of significance (p) for 9323727 .....	38

## LIST OF ILLUSTRATIONS

Figure 1.	Methylation of a Cytosine.....	4
Figure 2.	Activation and inactivation of a gene via methylation. ....	5
Figure 3.	Bisulfite Modification of Unmethylated vs. Methylated DNA .....	7
Figure 4.	Pink colored, feather shaped Takayama hemochromogen crystals. ....	13
Figure 5.	1% agarose gel quantitation of genomic DNA. ....	29
Figure 6.	2% agarose gel quantitation of PCR products .....	29
Figure 7.	Hypermethylated Pyrogram .....	30
Figure 8.	Hypomethylated Pyrogram .....	31
Figure 9.	Reagent Blank Pyrogram .....	32
Figure 10.	Histogram for 4739647 .....	34
Figure 11.	Histogram for 6266993 .....	36
Figure 12.	Histogram for 9323727 .....	38

## CHAPTER I - INTRODUCTION

### **Current DNA Technology**

Within the scope of scientific advancement, the application of DNA technology to human identification is still relatively young. There is recent, rapid progress in many facets of human identification using DNA technology. Prior to 1985, Forensic Biology essentially relied on ABO blood typing, a method of discriminating between individuals based on the antigens present on their red blood cells. This method was developed by Karl Landsteiner in the early 1900s (Landsteiner, 1901), and improvements upon Landsteiner's blood typing method included MN typing and Rh factor identification (Landsteiner and Wiener, 1940). Comparison of red and white blood cell enzymes were later found to have a higher rate of discrimination between individuals than blood typing alone and were often used for clarification in paternity cases. The first Restriction Fragment Length Polymorphism (RFLP) was discussed by Ray White and colleagues in 1980 (Botstein et al., 1980), though it was not until 1985 that Alec Jeffreys released his discovery of the first probe for human identification (Gill et al., 1985). Alec Jeffreys was also credited for using RFLP DNA technology for the first time in a criminal case to convict Colin Pitchfork of sexual assault and murder (Butler, J. M. 2010). The Variable Number Tandem Repeat (VNTR) analysis was considered the gold standard in forensic analysis but this process had its own disadvantages as well. The need for a larger quantity of DNA (100+ ng) as well as high molecular weight DNA complicated the analysis in cases where there was limited DNA available (10ng or less). Often the DNA available for casework is severely degraded and is not suitable for VNTR analysis. These two disadvantages associated with RFLP technology generated the need for a more robust

technology where small amounts of DNA could be used to obtain a complete DNA profile. The identification of Short Tandem Repeat (STR) polymorphism (Edwards et al., 1991) and the invention of Polymerase Chain Reaction (PCR) (Saiki RK, 1985) have overcome the two major disadvantages with RFLP analysis. This transition of technologies happened around 1995 and by 2000 nearly all labs in the country had validated the use of STR markers for forensic casework.

In 1985 Kary Mullis, a scientist at Cetus Corporation, released his development of PCR technology (Saiki RK, 1985) which allows small quantities of genomic DNA to be amplified quickly and easily. This advancement circumvented the previous need for high-quality, undegraded DNA for VNTR analysis, and allowed even relatively-degraded samples in low quantities to be used without deleterious effect to results. The PCR technique revolutionized not only the field of human identification but impacted most biological disciplines that were reliant upon DNA for advancement.

In criminal cases involving biological evidence, the presence or absence of tissue types such as blood, semen, saliva, or vaginal secretions can be a crucial component for understanding the events that transpired during the reported crime. Unfortunately, methods for tissue identification have not seen similarly rapid advancements as those pertaining to human identification (Patzelt, 2004).

### **Current Tissue ID Methods**

Several serological tests exist for the detection of human body fluid. These tests are employed by forensic professionals in the field and in the laboratory and are generally delineated into two groups—presumptive and confirmatory tests. Presumptive tests are often used in the field to suggest the presence of a bodily fluid, and they tend to have a

high level of sensitivity but lack the species-specificity. Confirmatory tests are used to authenticate the presumed presence of a bodily fluid but can tend to be less sensitive than presumptive tests and may offer false negative results in diluted stains. For example, the presumptive Phenolphthalein test employed in the identification of blood is highly sensitive and can detect more than 1000-fold dilution of blood, while the confirmatory test for blood, the Takayama crystal test is sensitive only up to a 50-fold dilution of blood.

The presumptive tests for blood identification include the Kastle-Meyer test and the Hemastix test while the confirmatory test relies on the identification of hemoglobin crystals in the sample. The presumptive test for seminal fluid depends on the presence of excessive amounts of Acid Phosphatase enzyme while the confirmatory tests depend on microscopic identification of spermatozoa or detection of Prostate Specific Antigens (PSA) in the sample. The identification of saliva is based on the detection of excessive amounts of the enzyme Amylase that is also found in other tissues as well in smaller quantities. Current tests for presumptive and confirmatory tissue identification possess numerous disadvantages such as expending exorbitant amounts of the available sample, possessing high rates of false positives or negatives under such common circumstances as dilutions, and lacking specificity for species identification or even differentiation between similar tissues. Additionally, no standard protocol exists for the differentiation of vaginal epithelial cells or vaginal fluid from other tissues commonly encountered in forensic casework. Development of a simultaneous method for accurate and reliable differentiation of forensically-relevant tissues would reduce the time necessary for serological processes during casework and minimize the expenditure of available

evidentiary samples. So, a more reliable method of tissue identification becomes necessary for serological analysis of evidence.

### DNA Methylation

All nucleated cells in humans contain identical genetic sequences, but the chemical changes to the DNA nucleotides within each cell, referred to as the epigenetic modifications, differ widely (Ng & Gurdon, 2008). Epigenetic modifications include histone modification, chromatin remodeling, gene silencing or activation, and carcinogenesis, but one of the most common and well-studied epigenetic modifications is methylation. Methylation is the addition of a methyl group (CH<sub>3</sub>) at the 5-carbon of the Cytosine (C) ring resulting in 5-methylcytosine (5mC) (Kader & Ghai, 2015). (Figure 1).

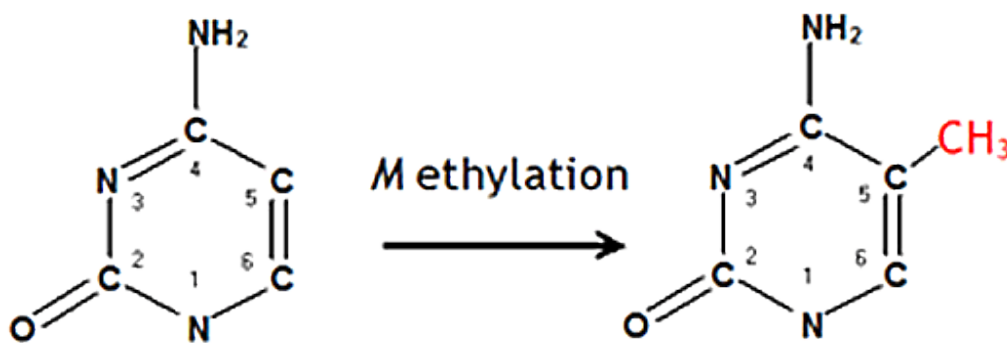


Figure 1. Methylation of a Cytosine

Addition of a methyl group to the 5-C position of cytosine (Chakarov et al., 2014).

Since the methylation occurs in the Cytosine that is immediately followed by a Guanine (G), they are usually referred to as CpG sites. In human DNA, 5-methylcytosine is found in approximately 1.5% of genomic DNA (Lister et al., 2009). While methylation typically occurs at a CpG site, some non-CpG methylation has been reported in humans (Kader & Ghai, 2015). The body employs methylation often in the promoter regions of

genes (Figure 2) to regulate expression in different areas of the genome in different cells, as necessary (Patterson et al., 2011).

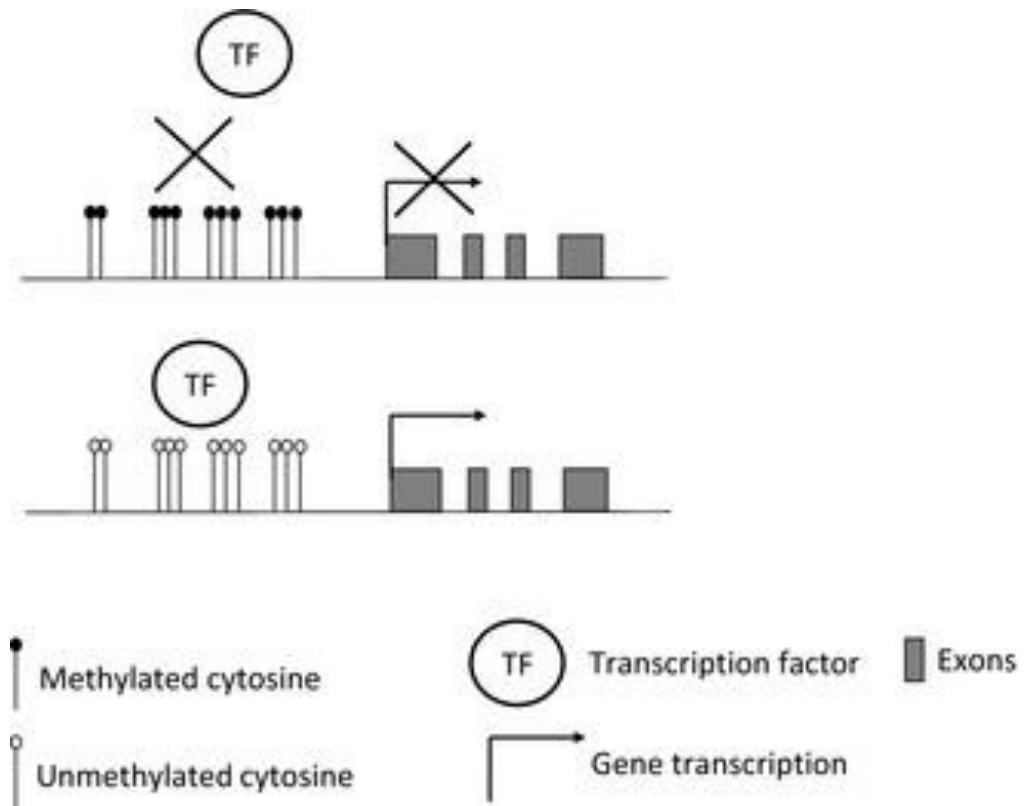


Figure 2. Activation and inactivation of a gene via methylation.

When a CpG island in the promoter region of a gene is methylated, expression of the gene is repressed (Reynolds, Jacobson, & Drake, 2013).

Methylation can occur in response to environmental stimuli in utero and throughout the individual's lifetime (Lee et al., 2012), but is often associated with X-inactivation, genetic memory, and cell fate (Madi et al., 2012). Methylation is often found in areas known as CpG islands, which are clusters of CpG sites that typically exist within the promoter region of a gene and contain high GC percentages. The expression of genes is regulated by methylation interfering with the binding of ribosomes to the promoter region of genes (Gardiner-Garden and Frommer, 1987). Numerous genome-



wide studies have indicated a relationship between DNA methylation and certain regions of the genome of a specific tissue known as tissue-specific differentially methylated regions (tDMRs) (Vidaki et al., 2016; Lee et al., 2012). Earlier Eckhardt et al. (2006) reported the high-resolution DNA methylation profile of human chromosomes 6, 20, and 22 and provided a resource of DNA methylation for 12 different tissues. The use of DNA methylation-based forensic tissue identification using methylation-sensitive restriction enzymes was reported by Frumkin et. al. (2011). In this study, the authors differentiated blood, saliva, semen, and skin cells based on the methylation pattern of the tissues. Lee et al. (2012) reported the identification of CpG site-specific methylation information for semen, blood, and vaginal fluid using the Illumina HumanMethylation 450K BeadChip array. Other studies have reported the identification of semen-specific methylation patterns (Wasserstrom et. al., 2013).

### **Bisulfite Modification**

Direct genomic sequencing of post-PCR samples does not reveal the methylation status of CpG sites, since the methylation information is lost during PCR; conversions such as bisulfite modification may be used to protect and reveal methylation status of CpG sites. Bisulfite modification converts unmethylated cytosine nucleotides to uracil, which are converted to thymine during the subsequent PCR process. The methylated cytosines are protected from such conversion and remain as cytosine (Figure 3). The bisulfite-modified DNA is then PCR amplified and pyrosequenced at single-nucleotide resolution to determine the methylation status of specific CpG sites. The percentage of methylation for an unknown tissue is calculated by comparing the percentage of cytosine

nucleotides that have been converted to thymine nucleotides in the sample (Patterson et. al. 2011).

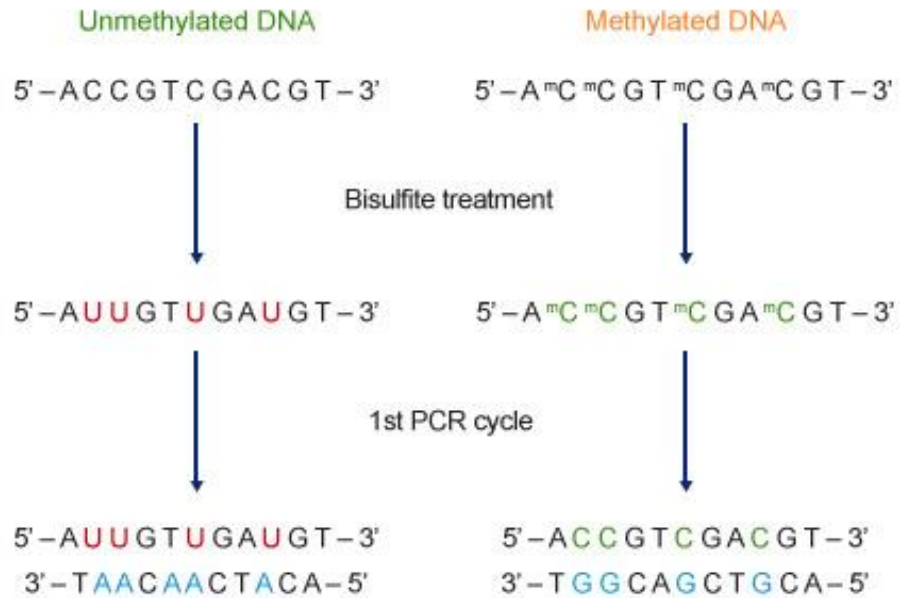


Figure 3. Bisulfite Modification of Unmethylated vs. Methylated DNA

(DNA methylation: a guide, 2017).

### Methods to Measure DNA Methylation

Several technologies have been utilized to quantitate the percentage of methylation within certain regions of the genome. One of these methods is ligation-mediated PCR, which utilizes a pair of non-isoschizomeric restriction enzymes, one of which is methylation-sensitive. The cleavage by the methylation sensitive enzyme is quantitated by amplifying the restricted products with ligation-mediated PCR and radioactive labeling. The amplified product ratios of the two digested products are compared and correlate directly to the amount of methylation present in the restriction site (McGrew & Rosenthal, 1993).

The single-nucleotide resolution of capillary electrophoresis sequencing can be applied to methylation analysis studies as well. Small amounts of DNA undergo enzymatic hydrolyzation to single nucleotides which are then derivatized with a fluorescent marker, followed by capillary electrophoresis. Corresponding strands known to be fully-methylated and fully-unmethylated are analyzed separately to establish a reference for methylation, then the DNA in question is compared with the reference to derive its percent methylation (Stach et al., 2003).

Other studies for DNA methylation analysis include digestion of genomic DNA using methylation-sensitive restriction enzymes. Many restriction enzymes are sensitive to DNA methylation and cleavage of the restriction site may be impaired if a base involved in the restriction site is modified. In this experiment, the genomic DNA is digested with methylation-sensitive restriction enzymes and amplified using fluorescently-labeled primers that flank the restriction site. If the DNA fragment is cleaved by the enzyme (non-methylated site), there is no amplification. If the DNA fragment is not cleaved by the enzyme, the target segment is amplified (methylated site) (Frumkin et. al. 2011).

### **Pyrosequencing**

Pyrosequencing is one of the many sequencing technologies that offers single-nucleotide resolution necessary for measuring methylation. Pyrosequencing is a “sequence-by-synthesis” technique. Genomic DNA undergoes bisulfite conversion, followed by PCR amplification of the target region containing CpGs for which information about methylation levels is desired. Biotinylated PCR products are generated by tagging one of the PCR primers with Biotin, a necessary step before pyrosequencing.

The amplified products are bound to streptavidin beads and treated with an alkali to result in single-stranded DNA. To initiate a pyrosequencing reaction, a sequencing primer is hybridized to a single-stranded template and incubated with appropriate enzymes. Nucleotides are dispensed according to each assay's individual dispensation order. In the presence of a complimentary base, a dNTP is ligated onto the template strand by DNA polymerase and the incorporation results in a flash of light from the photoluminescent chemicals. The pyrosequencer records which nucleotide was incorporated, and the amount of light emitted before degrading the unused dNTP's and moving to the next nucleotide on the template strand. The results of each sample are displayed as a pyrogram with the relative methylation levels of each CpG site in the target segment. Pyrosequencing is a reliable, sensitive method for methylation analysis in both hyper-and hypomethylated regions (Tost & Gut, 2007; Madi et. al. 2012). Pyrosequencing technology is advantageous in that the methylation information is quantified for each CpG site. The differential methylation pattern of specific sites can be determined for a variety of tissues and the resulting percent methylation can be used to identify the tissue source of a DNA sample. Once a set of epigenetic markers is identified, this technology can be routinely used in crime labs to differentiate one tissue from others.

### **Aims and Objectives**

The primary purpose of this study is to identify a set of DNA methylation markers that can be used to differentiate vaginal epithelial cells from three other forensically-relevant tissues (blood, saliva, and semen) for forensic human identification, based on the relative percent methylation of the target CpG sites. This will be accompanied by genomic DNA extraction, bisulfite conversion of genomic DNA, PCR amplification of

target segments using site-specific PCR primers, and finding the relative percent methylation of the target segments in different tissues by quantitative pyrosequencing. The development of a set of markers for vaginal epithelial cell identification could be combined with the known markers for several forensically-relevant tissues to develop a multiplex kit for serological analysis. A multiplex tissue identification kit could reduce the time and resources currently being applied to the serological process in forensic laboratories.

## CHAPTER II – REVIEW OF LITERATURE

### **Current Presumptive and Confirmatory Tissue ID Methods**

#### *RSID*

Discrimination between forensically-relevant body fluids such as blood, saliva, and semen has seen some recent advancement with the development of the Rapid Stain Identification Series (RSID) confirmatory tests, which utilize secretion-specific antigens to confirm the presence of specific tissues or bodily fluids. The principle of the RSID test strips is similar to that of the ABACard Hematrace test, where an immunochromatographic strip is impregnated with antihuman antibodies that are specific to each bodily fluid and have no cross-specificity with each other (Casey & Price, 2010).

#### *Blood*

One commonly used test for the presumptive detection of blood is the phenolphthalein, or Kastle-Meyer test. The colorless phenolphthalein reagent, in a reduced, alkaline state, when mixed with blood and hydrogen peroxide, turns pink to indicate the possible presence of blood. The pink color is a result of the oxidation reaction carried out by the peroxidase-like activity of hemoglobin in presence of hydrogen peroxide. Though the Kastle-Meyer test is useful in the field for rapid presumptive identification of blood, false positive reactions have been reported for a myriad of other substances, including vegetable extracts, nasal mucus, saliva, and milk (Gaensslen, 1983). The sensitivity of the phenolphthalein test makes it one of the preferred reagents among a plethora of other tests.

Chemiluminescent tests such as Luminol are frequently used in field for the presumptive identification of blood. Luminescence is achieved by the oxidation of the

luminol compound by the hemoglobin in blood. Luminol possesses an advantage of fewer false positive reactions than phenolphthalein-based tests, but some studies have shown it to interfere with subsequent serological analyses. The Luminol test is widely used in the field where the presence of blood is suspected but not visible. These presumptive tests are very sensitive but lack specificity to humans (Gaensslen, 1983).

If the presence of blood has been inferred by field-testing, serological confirmatory testing is utilized prior to subjecting the evidence for DNA analysis. The Takayama crystal test is one such confirmatory method. Pyridine ferroprotoporphyrin or hemochromogen crystals are produced by heating a mixture of heme and an alkaline solution of pyridine, in presence of a reducing sugar. These pink-colored, feather-shaped crystals (Figure 4) can be viewed under a microscope and their presence is inferred as a positive test for blood (James et al., 2005). The reaction relies upon the presence of hemoglobin only, resulting in the Takayama crystal test's non-specificity for human blood.

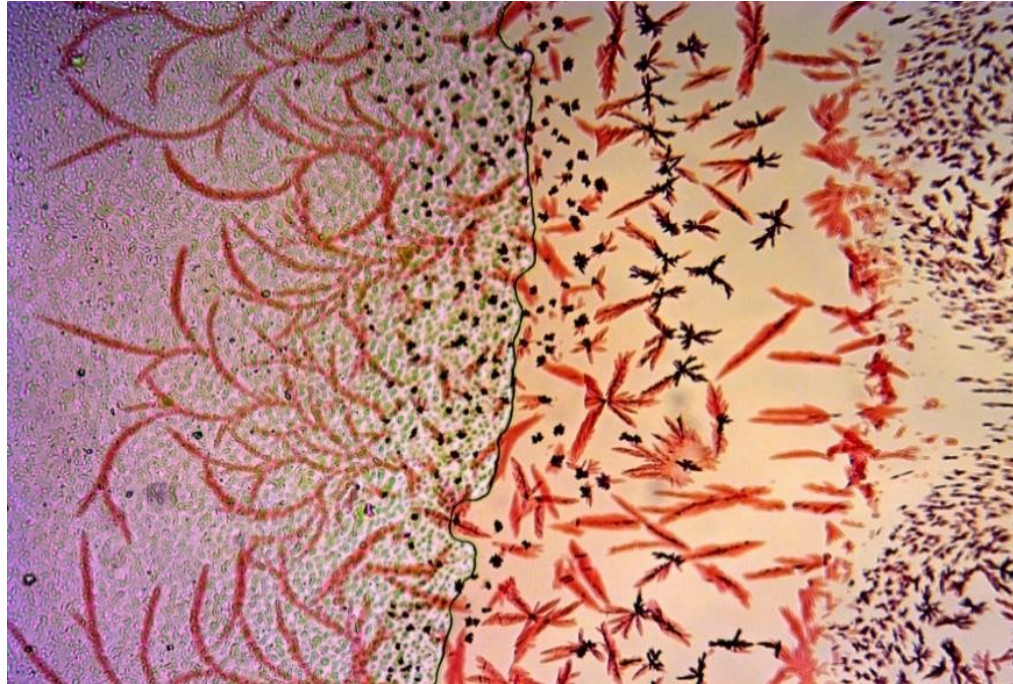


Figure 4. Pink colored, feather shaped Takayama hemochromogen crystals.

(Image courtesy-K. Balamurugan)

The ABACard Hematrace test strip is another confirmatory method that is considered mostly specific for humans but has been reported to give false positives for some animal blood (Ferret). A partially-diluted aqueous sample is placed in the sample window, where the strip has been treated with mobile antihuman hemoglobin antibodies that bind to any human hemoglobin antigens present in the sample. The antibody-antigen complex and attached dye molecule migrate to the testing window portion of the strip that has been treated with an immobilized antihuman hemoglobin antibody. The antigen-antibody complex binds in the testing window to the immobilized antibody. A single antibody-antigen-antibody complex is undetectable to the naked eye, but a pinkish line appears as more and more of these complexes aggregate in the test area. An internal control line appears separately to denote a functional test even in the absence of blood.



Though largely specific for humans, the ABACard has been shown to give false positives for the blood of ferrets and some primates (Reynolds M., 2004).

For the rapid identification of blood, the RSID-blood test identifies the presence of glycophorin-A, an antigen found in the cell membrane of red blood cells. The test also shows no inhibition by field presumptive test reagents such as Luminol (Turrina et. al. 2008).

### *Saliva*

The most common semi-quantitative approach for identifying the presence of saliva is the radial diffusion test. For radial diffusion, gel plates containing starch are prepared with multiple wells with a template attached to the back of the plate. Unknown evidentiary samples, known saliva standards, and non-saliva standards such as blood, urine, and semen are pipetted into the wells separately and the plate is incubated to allow the digestion of the starch within the gel to occur. After incubation, the plates are treated with iodine to stain remaining starch within the gel. Areas where a clear circle is observed indicate amylase activity, and the size of the clear area is directly proportional to the amount of amylase contained within the corresponding body fluid. Radial diffusion test is highly presumptive and non-species specific (Haltiner, 2008).

For rapid detection of saliva, the RSID-saliva test detects the presence of the  $\alpha$  - amylase enzyme present in salivary secretions. Currently-used tests detect the presence of amylase activity, such as the starch iodine radial diffusion test or the Phadebas test, both of which rely on  $\alpha$  -amylase digestion of starches in the tests and are interpreted by the absence or presence of a dark blue color (Myers & Adkins, 2008). The RSID-saliva

test detects the actual presence of  $\alpha$ -amylase antigens as opposed to their activity against starch (Casey & Price, 2010).

### *Semen*

Several tests exist for the presumptive identification of seminal fluid. One method to detect the presence of semen is the acid phosphatase (AP) test. Acid phosphatase is an enzyme that is secreted into semen in high quantities by the prostate gland. The test relies on Alpha-Naphthyl acid phosphate reacting with Brentamine Fast Blue dye in the presence of AP to rapidly produce a purple color. The AP test is no more than a presumptive test due to the presence of acid phosphatase in other bodily fluids, including vaginal fluid. Additionally, AP outside of the body degrades rapidly when exposed to heat and humidity (Noureddine, 2011). The standard confirmatory test for seminal fluid is the microscopic observation of spermatozoa using the Christmas tree staining technique. The stain consists of Kernechtrot (Nuclear Fast Red) and Picroindigocarmine (green) dyes that stain the sperm nuclei and acrosomal regions red, while staining the tail and neck portions greenish-blue. Though this confirmatory test is extremely effective for the identification of spermatozoa, it lacks efficacy for the identification of seminal fluid from males with a compromised sperm count (oligospermic) or ejaculatory fluid that may contain little or no spermatozoa (azoospermic) (Noureddine, 2011).

For the detection of semen, the RSID series includes a test for semenogelin, which is a human-specific antigen produced only in seminal vesicles and excreted in seminal fluid. Validation tests were performed to successfully rule out false positives in saliva, urine, breast milk, serum, and vaginal fluid, as well as species-specificity against

equine, canine, swine, and bovine semen samples (Sato et al., 2004). The Abacard Onestep PSA test kit (Abacus diagnostics, CA) is also available commercially to detect the semen-specific antigens. This test is based on the human-specific prostate antigens reacting with anti-human prostate antibodies, producing a colored band in the test kit.

### *Vaginal epithelial cells*

Though there are many methods for the identification of seminal fluid, saliva, and blood, there are no currently-implemented standard confirmatory methods for the identification of vaginal epithelial tissue or vaginal secretions. Many of the existing technologies, such as mRNA profiling, molecular identification of microbial signature, and protein markers are still in the validation stage.

### **mRNA profiling**

Several techniques involving mRNA profiling have surfaced, including one that combines mRNA with bacterial markers for a more comprehensive identification of vaginal fluid. Though the efficacy of this method has been shown, it requires additional steps not typically taken in a forensic laboratory and can consume portions of the sample that may have otherwise been used for DNA identification. Researchers also noted the variability of both individual vaginal microbiomes and expression of the targeted mRNA (Jakubowska et al., 2013), leading to some uncertainty regarding the reliability of this method. Giampaoli et al. added that mRNA in vaginal samples are highly unstable in the best of laboratory storage conditions and can degrade in as little as two weeks environmentally (Giampaoli et al., 2012), making mRNA profiling an unrealistic option for cases in which the evidence is not collected and processed rapidly.

## **Microbiome Differentiation**

Giampaoli et al. studied the microbial signature of vaginal flora to differentiate vaginal fluid from other body fluids. Though the results are promising for vaginal, oral, and anal sample differentiation based on the abundance of characteristic bacteria for each orifice, the technique becomes significantly less-applicable in the presence of mixtures or in the case of trace samples (Giampaoli et al., 2012). The microbiome analysis technique could be additionally problematic, as it does not specifically identify the presence of a bodily fluid or tissue, but infers its presence based on the bacteria associated with those areas of the body.

Protein-based assays, such as the RSID and PSA tests, are commonly used, but suffer many similar drawbacks as microbiome and mRNA analysis. Protein-based assays are neither highly specific nor sensitive and are unstable under environmental conditions (Harbison & Fleming, 2016). Furthermore, the use of these assays requires a large amount of sample be consumed with little probative return.

## **Methylation-Based Tissue Differentiation**

A study by Frumkin et al. was among the first with promising results for the differentiation of blood, saliva, semen, and skin tissues based on DNA methylation levels. The researchers used a methylation-specific restriction enzyme PCR technique to prepare samples for capillary electrophoresis. The ratios of percent methylation were calculated by comparing the low electropherogram signals for hypomethylation with the high signals associated with hypermethylation. This method allowed for successful discrimination between the tissue types (Frumkin et al., 2011).

A subsequent study by Lee et al. tested five tissue-specific differentially methylated regions (tDMRs) for differentiating blood, saliva, semen, menstrual blood, and vaginal fluid via bisulfite conversion and DNA sequencing through capillary electrophoresis. Though their results suggested two markers for the use of semen identification, they also noted that some of the tDMRs could be used in conjunction with each other to identify the presence of menstrual blood and/or vaginal fluid. Additionally, the study suggests that the methylation profiles of these tDMRs remain consistent in samples aged for 30 days at room temperature (Lee et al., 2012). In two similar studies, An et al. reported similar results regarding identification of vaginal fluid and menstrual blood (An et al., 2013). A review by Kader and Ghai suggests the necessity of further studies to identify more robust markers that can distinguish between vaginal fluid and menstrual blood (Kader & Ghai, 2015).

In a comprehensive study of global DNA methylation profiles, Park et al. utilized the Illumina HumanMethylation 450K bead array to assess the methylation values of 450,000 CpG sites. From the resulting methylation data, researchers identified two CpG sites per tissue for robust differentiation of blood, saliva, semen, and vaginal fluid. Additional potential CpG sites that can be used to differentiate tissues were also provided (Park et al., 2014).

Recently, Antunes et al. published a study specifically testing a small section of the PFN3-PFN3A marker that Lee et al. (2012) previously revealed for its use in discrimination of vaginal fluid. The study utilized bisulfite conversion and pyrosequencing to investigate the data from Lee et al. The researchers tested the assay for mixture and species specificity, with largely successful results, but they highlight the

necessity of identifying further locations with improved sensitivity for vaginal epithelial discrimination (Antunes et al., 2016).

Pursuant to the line of testing performed by Antunes et al., we seek to identify a set of markers for further investigation into novel CpG sites to distinguish vaginal epithelial cells from other relevant tissue types. CpG sites with notable differences in percent methylation of vaginal epithelial cells in relation to other tissues, as described by their beta values (Park et al., 2014) are considered candidates for investigation.

## CHAPTER III - MATERIALS AND METHODS

### **Sample Collection**

Blood, buccal cells, vaginal epithelial cells, and semen samples were collected under an IRB-approved protocol for a previous, related study and were available in the laboratory.

#### *Blood*

Samples were collected from volunteers by cleaning the volunteer's fingertip with absolute ethanol, then applying a lancet device to the fingertip of the volunteer. Blood was squeezed from the fingertip onto sterile cotton swabs. The swabs were air-dried, placed in labeled paper envelopes, and stored at -20°C until extraction.

#### *Buccal*

Saliva samples were collected from volunteers by providing the volunteers with sterile cotton swabs and instructing the volunteers to firmly swab the inside of the cheek for one minute. The swabs were air-dried, placed in labeled paper envelopes, and stored at -20°C until extraction.

#### *Semen*

Semen samples were collected from male volunteers by providing a specimen cup and instruction for deposit and return of sample. Volunteers were instructed to avoid sexual contact prior to sample collection. Sample cups were labeled and stored at -20°C until extraction.

#### *Vaginal Epithelial*

Vaginal epithelial samples were collected from female volunteers by providing sterile cotton swabs and instruction for collection and return of sample. Volunteers were

instructed to wait several days after sexual contact or menstrual activity before collecting the sample. Swabs were air-dried, placed in labeled paper envelopes, and stored at  $-20^{\circ}\text{C}$  until extraction.

### **DNA Extraction and Quantitation**

Standard organic DNA extraction procedures were performed on all samples (Budowle et al., 2000). The swab portion of buccal, blood, and vaginal epithelial samples were cut from the swab stick and placed in individually labeled 1.5mL tubes with  $400\mu\text{L}$  of stain extraction buffer and  $10\mu\text{L}$  of proteinase K (20mg/mL). The samples were incubated overnight at  $56^{\circ}\text{C}$ . Semen samples were extracted by combining  $25\mu\text{L}$  of sample in individually labeled 1.5mL tubes with  $150\mu\text{L}$  of TNE (Tris-HCL, NaCl, EDTA),  $50\mu\text{L}$  of 20% Sarkosyl,  $40\mu\text{L}$  of 0.39M DTT (dithiothreitol),  $150\mu\text{L}$  of  $\text{dH}_2\text{O}$ , and  $10\mu\text{L}$  of proteinase K. The samples were incubated at  $56^{\circ}\text{C}$  overnight. After incubation, the cotton swab material was placed in a Spin-X basket in each 1.5mL tube and centrifuged at 16,000 rcf for five minutes to force any liquid out of the material and into the tube. The Spin-X basket and swab material were removed and discarded.

$500\mu\text{L}$  of phenol/chloroform/isoamyl alcohol were added to each tube, vortexed to obtain a milky emulsion, and centrifuged for 10 minutes at 16,000 rcf to separate the aqueous and organic phases. The aqueous phases were pipetted into Amicon100 Ultra centrifugal filter devices (Millipore Corporation, Bedford, MA), and centrifuged for 10 minutes at 2,300 rcf. After centrifugation, approximately  $400\mu\text{l}$  of Tris-EDTA (TE) buffer were added to the column and centrifuged for 10 minutes at 2,300 rcf. This washing process was repeated three additional times before collection of the purified



DNA. The purified DNA was collected by inverting the filter unit in a clean 1.5ml tube and centrifuging at 600 rcf for 3 minutes. The DNA was stored frozen until further use.

Agarose gel (1%) was utilized to obtain both quantitative and qualitative information for the DNA samples (Budowle et al., 2000). 1 $\mu$ L of sample and 2 $\mu$ L of bromophenol blue (BPB) loading dye were loaded into wells of a 1% agarose gel containing ethidium bromide. The gel was electrophoresed in 1x TAE (Tris/Acetic Acid/EDTA) buffer at 120V for 20 minutes. Samples were visualized with a UV Transilluminator. The gel images were photographed and stored in a computer.

### **Bisulfite Conversion**

Bisulfite conversion was performed on genomic samples using the Qiagen EpiTect® Bisulfite Kit with manufacturer's recommendations (Qiagen Inc).

1. Bisulfite reaction components were added to 0.2mL PCR tubes based on the concentration of genomic DNA identified in each sample via gel electrophoresis. The components and volumes of individual reagents for bisulfite conversion are described in Table 1.

<b>Component</b>	<b>Volume per reaction (<math>\mu</math>L)</b>
DNA solution (1 ng-2 $\mu$ g)	Variable* (maximum 20 $\mu$ L)
RNase-free water	Variable*
Bisulfite Mix (dissolved)	85
DNA Protect Buffer	35
<b>Total volume</b>	<b>140</b>

Table 1. Bisulfite reaction components

\*The combined volume of DNA and RNase-free water is 20 $\mu$ L.

2. After a brief vortex and quick spin, the samples were placed in a thermal cycler.
3. The samples underwent a three-stage denaturation and incubation process:
  - a. The samples were denatured at 95°C for 5 minutes, then incubated at 60°C for 25 minutes.
  - b. The samples were denatured at 95°C for 5 minutes, then incubated at 60°C for 85 minutes.
  - c. The samples were denatured at 95°C for 5 minutes, then incubated at 60°C for 175 minutes.
4. The samples were held at 20°C until they were removed from the thermal cycler and subjected to a cleanup process to remove the bisulfite conversion reagents and purify the converted DNA.

#### **Bisulfite Conversion Cleanup Process**

1. The PCR tubes were removed from the thermal cycler and the bisulfite reactions were transferred to clean 1.5mL tubes.
2. 560 $\mu$ L of loading buffer (Buffer BL) containing 10 $\mu$ L of carrier RNA were added to each tube, and samples were vortexed and centrifuged briefly at 16,000 rcf.
3. Samples were transferred to *EpiTect*® DNA spin columns with collection tubes and centrifuged for 1 minute at 16,000 rcf.
4. The flow-through was discarded, and 500 $\mu$ L of wash buffer (Buffer BW) was added to the spin columns.
5. Samples were centrifuged at 16,000 rcf for 1 minute, and the resulting flow-through was discarded.

6. 500 $\mu$ L desulphonation buffer (Buffer BD) was added, followed by incubation at room temperature for 15 minutes.
7. Samples were centrifuged for 1 minute at 16,000 rcf and the resulting flow-through was discarded.
8. 500 $\mu$ L of wash buffer (BW) was added, followed by 1 minute of centrifugation at 16,000 rcf. The resulting flow through was discarded.
9. The washing step using buffer BW (step 8) was repeated once.
10. The columns were transferred to new 2mL tubes and centrifuged for another minute, then incubated for 5 minutes at 60°C with open lids to encourage ethanol to dissipate.
11. To elute the converted DNA from the spin columns, the columns were placed in clean 1.5mL tubes, and 20 $\mu$ L of eluting buffer was pipetted directly onto the column membrane.
12. The tubes were allowed to incubate for 1 minute at room temperature, then eluted via centrifugation at 16,000 rcf.
13. Tubes containing the bisulfite converted DNA were stored at -20°C until use.

### **CpG Marker Selection and Primer Design**

The prospective markers were chosen from data provided in Park et al., 2014, which detailed a list of potential CpG sites that may be used to differentiate vaginal epithelial cells from other tissues. CpG sites with notable differences in hyper- or hypomethylation of vaginal epithelial cells in relation to other tissues, denoted by beta values, were considered candidates for investigation. Selected CpG site information, such as the location of the marker in a chromosome and its position in the genomic DNA,

was entered into the University of California Santa Cruz (UCSC) Genome Browser using the human genome assembly build 37 (GRCh37/hg19). After locating the CpG site within the human genome, a sequence of approximately 200 bases on both the 5' and, 3' side was downloaded. This downloaded sequence was used to design primers for PCR amplification and sequencing using the Pyromark assay design software (Qiagen). Three CpG markers were selected after screening a pool of candidate markers. The primer specifications are listed in Appendix A.

### **Polymerase Chain Reaction (PCR)**

Bisulfite-converted DNA from at least seventeen samples of each tissue type, blood, buccal cells, sperm, and vaginal epithelial tissue were chosen for PCR amplification using the designed PCR primers. PCR components per sample are detailed in Table 2.

Bisulfite-converted DNA	2 $\mu$ L
10x PCR Primer set	2 $\mu$ L
Coral load solution	2 $\mu$ L
Q-solution	4 $\mu$ L
2x PCR Master mix	10 $\mu$ L
Reaction Volume:	20 $\mu$ L

Table 2. PCR components per sample for sample amplification

Negative controls were included to check for contamination in reagents. Annealing temperatures were determined by subtracting 5°C from the melting temperatures ( $T_m$ ) of the forward and reverse primers. The PCR cycling conditions for the amplification are detailed in Table 3.

Initial Incubation Step	Denaturation	Annealing	Extension	Final Extension	Final Soak
<b>HOLD</b>	<b>CYCLE (45 cycles)</b>			<b>HOLD</b>	<b>HOLD</b>
95°C 15 min	94°C 30 sec	T <sub>m</sub> -5°C 30 sec	72°C 30 sec	72°C 10 min	4°C ∞

Table 3. PCR cycling conditions for sample amplification.

To check the robustness of amplification, a 2% agarose gel quantitation was utilized. 2 $\mu$ L of amplified products was pipetted into each well of the gel, electrophoresed at 120V for 20min, and visualized using a UV Transilluminator.

### Pyrosequencing

Pyrosequencing assays were created using the PyroMarkQ24 software. Twenty-four bisulfite-converted samples that had undergone amplification and quantitation were selected for each pyrosequencing run.

Amplified samples were arranged in three rows of eight per the assay design, then 18 $\mu$ L of PCR product was aliquoted into each well of a multi-well tray containing 62 $\mu$ L of pyrosequencing cocktail (2 $\mu$ L streptavidin beads, 40 $\mu$ L binding buffer, 20 $\mu$ L diH<sub>2</sub>O). The multi-well tray was covered with strip caps and shaken on a microplate shaker at 1000 rpm for 10 minutes. After removing the plate from the shaker and removing the strip caps, the PCR products attached to the streptavidin beads were lifted using a vacuum pump and processed using the PyroMark Q24 workstation. The streptavidin beads and PCR products were released onto a pyrosequencing plate containing 25 $\mu$ L of

1x sequencing primer in each well. The plate was incubated on a hot plate at 80°C for 2 minutes, then allowed to rest at room temperature for 10 minutes.

The pyrosequencing cartridge was prepared with enzyme, substrate, and dNTPs in amounts determined by the PyroMark Q24 software. The samples underwent pyrosequencing in the PyroMark Q24 pyrosequencer as per manufacturer's protocol. Once sequencing was completed, the Pyromark software was used to analyze the percent methylation of each CpG site.

### **Data Analysis**

After data collection was complete, the methylation data of individual CpG sites for all samples for a marker was entered in an Excel spread sheet, grouping each tissue separately. Mean percent methylation and standard deviation values were calculated by averaging the methylation values at each CpG site for each group of tissues tested. Mean methylation values of the different tissues were compared using a one-way ANOVA and Tukey's Post Hoc pair-wise comparisons to determine if there were statistically-significant differences between the mean methylation values of the tissues studied. Methylation differences were considered statistically significant when p- values were less than 0.05 ( $p < 0.05$ ). Statistical analysis for the significant difference in percent methylation level between tissues was determined by SPSS software package version 22 (IBM) (Balamurugan et al., 2014).

### **Bisulfite Controls**

Bisulfite conversion controls were present in each assay to ensure appropriate levels of bisulfite conversion within samples. The bisulfite conversion process changes all unmethylated cytosines to thymine; a control is implemented by selecting a cytosine

that is not part of a CpG site, so the unmethylated cytosine will convert fully to thymine. When the sample is pyrosequenced, the results should indicate complete conversion of the control site 'C' to 'T'. A control site that retains any percentage of unconverted cytosines is indicative of incomplete bisulfite modification. Samples that failed the bisulfite control were rejected due to incomplete conversion.

## CHAPTER IV – RESULTS

DNA quantity estimates were made by comparing the intensity of DNA standards and the samples. Figure 5 shows a gel image of the extracted DNA samples with a 1KB ladder for comparison.

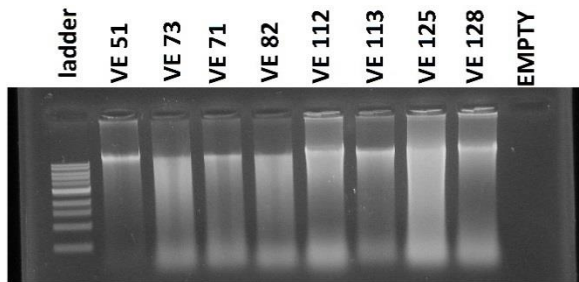


Figure 5. 1% agarose gel quantitation of genomic DNA.

The gel image was used to roughly estimate the amount of genomic DNA necessary for bisulfite conversion. Seventeen bisulfite-converted DNA samples for each tissue type per marker were used to amplify the target segment containing multiple CpG sites. Figure 6 shows the agarose gel quantitation of PCR product from one of the unused markers.

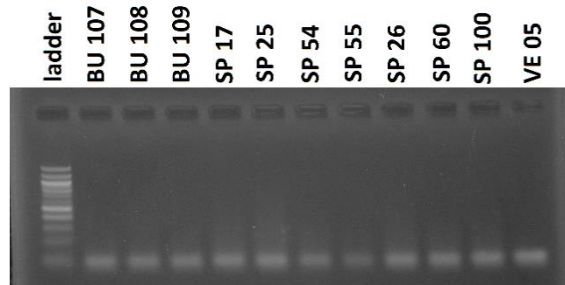


Figure 6. 2% agarose gel quantitation of PCR products

Vaginal epithelial marker 5167251.

Most of the samples amplified produced the expected size of amplicons. In some of the agarose gel images, primer dimers were identified by the observation of fluorescent



bands smaller than the target amplicon size. These are largely either non-specific amplification or primer dimers. The presence of primer dimers did not affect the quality of the pyrosequencing results.

The pyrosequencing results were analyzed using the Pyromark Q24 software and the resulting data were displayed as pyrograms. Figures 7 and 8 show pyrograms for samples that are hypermethylated and hypomethylated, respectively.

Well: C6  
Assay: VEpi\_9323727  
Sample ID: VE 112  
Note:  
Analysis version: 2.0.6

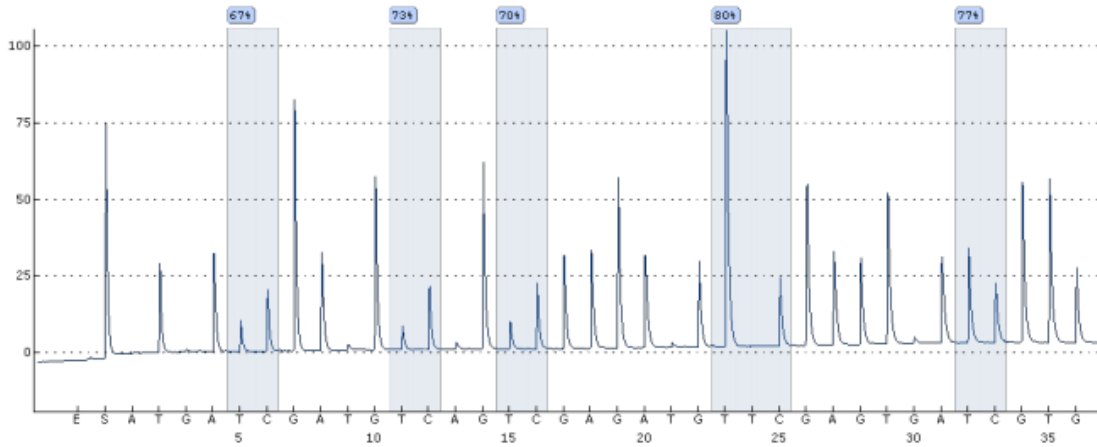


Figure 7. Hypermethylated Pyrogram

Vaginal epithelial marker cg-9323727, vaginal epithelial sample 112.

The CpG sites that were analyzed for this marker are indicated by the shaded area with percent methylation displayed above the shaded box. At the first CpG site, the percent methylation recorded by the pyrosequencer was 67%. This means that 67% of the cytosines at this site were methylated, with the remaining unmethylated.

Well: A3  
Assay: VEpi\_9323727  
Sample ID: BL\_36  
Note:  
Analysis version: 2.0.6

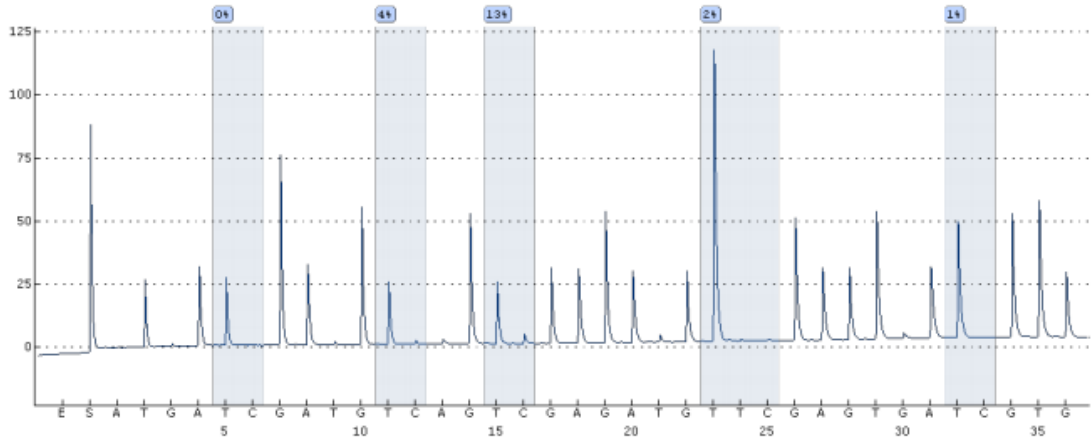


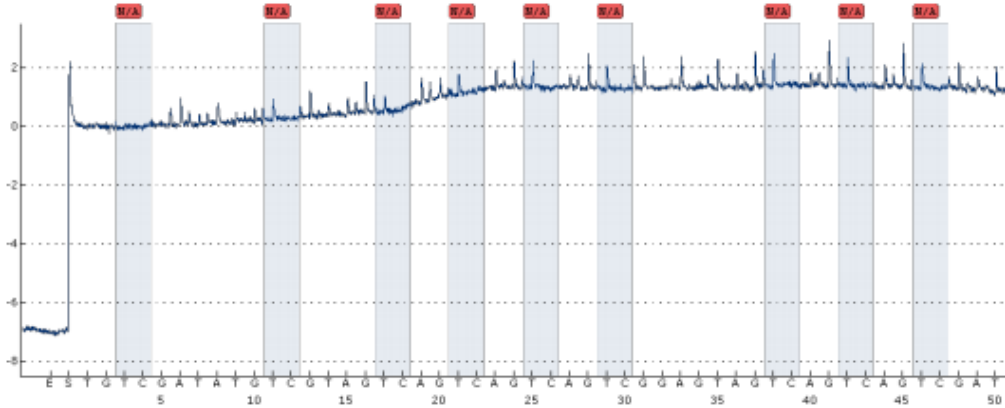
Figure 8. Hypomethylated Pyrogram

Vaginal epithelial marker cg-9323727, blood sample 36.

In Figure 8, at CpG site 1, the percent methylation recorded by the pyrosequencer was zero. This means that this site was completely unmethylated.

The reagent blanks used during DNA extraction were also used for PCR to check for any possible reagent contamination. Figure 9 shows the pyrogram of a reagent blank for vaginal epithelial marker cg-4739647 with no detectable peaks.

Well: C8  
 Assay: VEpi\_4739647  
 Sample ID: Reagent Blank  
 Note:  
 Analysis version: 2.0.6



Sequence to analyze:

GYGAATTAGTGGTGGTGGGGYGTGGYGGGGAGTGTGGYGGYGGGATAGTGAATGTTGGGTGGGAGT

Position	1	2	3	4	5	6	7	8	9
Quality	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Meth (%)	.	.	.	.	.	.	.	.	.

Figure 9. Reagent Blank Pyrogram

Vaginal epithelial marker cg-4739647, sample reagent blank.

If no amplifiable DNA is present in the sample, the results are displayed as a flat line as opposed to individual peaks.

If a band was identified in the reagent blank well of agarose quantitation, the reagent blank was pyrosequenced to determine whether the band was a primer dimer or contamination. If no band was identified in the agarose quantitation, one reagent blank was pyrosequenced per marker to ensure a lack of contamination in pyrosequencing reagents.

### Vaginal Epithelial Marker 4739647

Of the samples pyrosequenced for the epigenetic marker 4739647, 11 blood, 18 buccal, 14 vaginal epithelial, and 13 sperm samples provided viable data for analysis. Percent methylation data per CpG site for all samples was recorded in an excel spread

sheet and average standard deviations were calculated with the appropriate formula within the spreadsheet. Table 4 shows the percent methylation averages and values per tissue type of each CpG site for marker cg4739647.

		<b>CpG</b>	<b>CpG</b>	<b>CpG</b>	<b>CpG</b>	<b>CpG</b>	<b>CpG</b>	<b>CpG</b>	<b>CpG</b>	<b>CpG</b>
		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
<b>Blood</b>	Avg.	13	6	9	8	9	8	10	8	11
	Std.	15.7	6.7	13.6	12.7	13.7	11.7	13.4	12.3	16.1
	Dev.									
<b>Buccal</b>	Avg.	20	21	15	14	19	16	17	14	14
	Std.	10.7	14	12.2	10.6	14.9	11.9	13.2	11.4	11.2
	Dev.									
<b>Vaginal Epithelial</b>	Avg.	42	44	41	37	42	39	44	37	34
	Std.	18.4	23.2	17.5	17.6	21.5	16.9	23	16.8	14.6
	Dev.									
<b>Sperm</b>	Avg.	10	9	10	8	9	8	10	8	8
	Std.	13.2	11.6	13.4	10.9	11.4	11.3	12.1	11.1	9.95
	Dev.									

Table 4. Percent Methylation Averages and Standard Deviations-Marker 4739647

Percent methylation averages and standard deviations arranged by tissue type and CpG site.

Figure 10 shows the average methylation data and corresponding standard deviations for individual CpG sites for all tissues for the v.epi. marker 4739647.

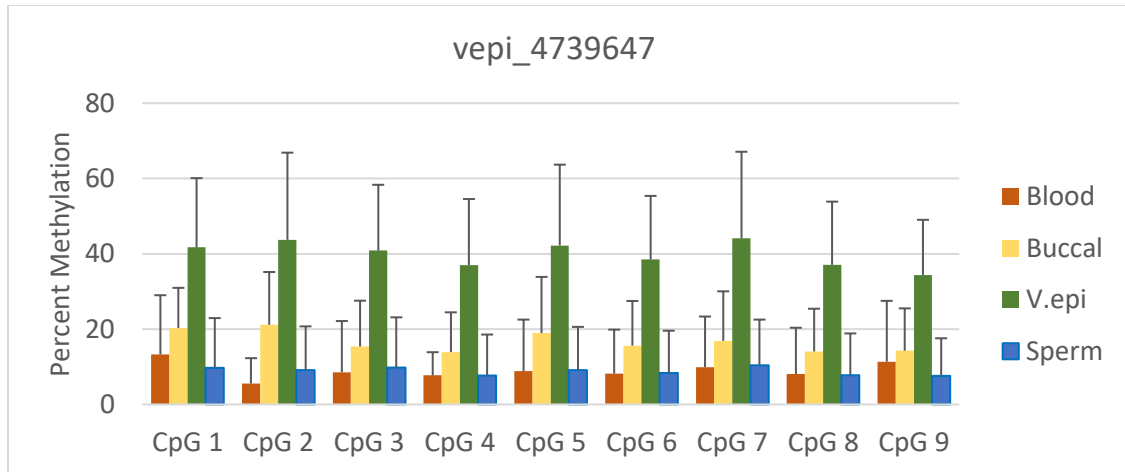


Figure 10. Histogram for 4739647

Average methylation data for all four tissues studied for the vaginal epithelial marker 4739647. Bars above each data indicate standard deviation.

Statistical analysis of the mean methylation values for vaginal epithelial tissue in comparison with blood, buccal, and sperm tissues was performed using a one-way ANOVA and Tukey's HSD test. The significance values for each tissue as compared to vaginal epithelial for marker 4739647 are shown in Table 5.

<b>v.epi. 4739647</b>									
	<b>CpG1</b>	<b>CpG2</b>	<b>CpG3</b>	<b>CpG4</b>	<b>CpG5</b>	<b>CpG6</b>	<b>CpG7</b>	<b>CpG8</b>	<b>CpG9</b>
<b>Blood</b>	0	0	0	0	0	0	0	0	0
<b>Buccal</b>	0.002	0.002	0	0	0.002	0	0	0	0
<b>Sperm</b>	0	0	0	0	0	0	0	0	0

Table 5. Levels of significance (p) for 4739647

For all CpG sites of v. epi marker compared to other tissues.

The p values of all CpG sites of v. epi. marker 4739647 were found to be below 0.05, indicating significant difference in percent methylation values. The epigenetic

marker 4739647 was found to differentiate vaginal epithelial cells from blood, buccal, and sperm cells ( $p < 0.05$ ).

### Vaginal Epithelial Marker 6266993

Of the samples pyrosequenced for the epigenetic marker 6266993, 14 blood, 11 buccal, 12 vaginal epithelial, and 12 sperm samples provided viable data for analysis. Percent methylation data for all CpG sites were recorded and averaged by tissue type. Table 6 shows the percent methylation average and standard deviation values for all CpG sites for the marker 6266993.

		<b>CpG 1</b>	<b>CpG 2</b>	<b>CpG 3</b>	<b>CpG 4</b>
<b>Blood</b>	Avg.	9	5	8	5
	Std. Dev.	13.3	8.8	10.9	6.6
<b>Buccal</b>	Avg.	25	20	28	19
	Std. Dev.	27.2	28.4	24	26.5
<b>Vaginal Epithelial</b>	Avg.	48	43	40	51
	Std. Dev.	24.7	24	20	25.7
<b>Sperm</b>	Avg.	2	2	6	3
	Std. Dev.	3.1	1	3.8	3

Table 6. Percent Methylation Averages and Standard Deviations-Marker 6266993

In Figure 11, the percent methylation averages and standard deviation values were organized into a histogram to visualize the correlations.

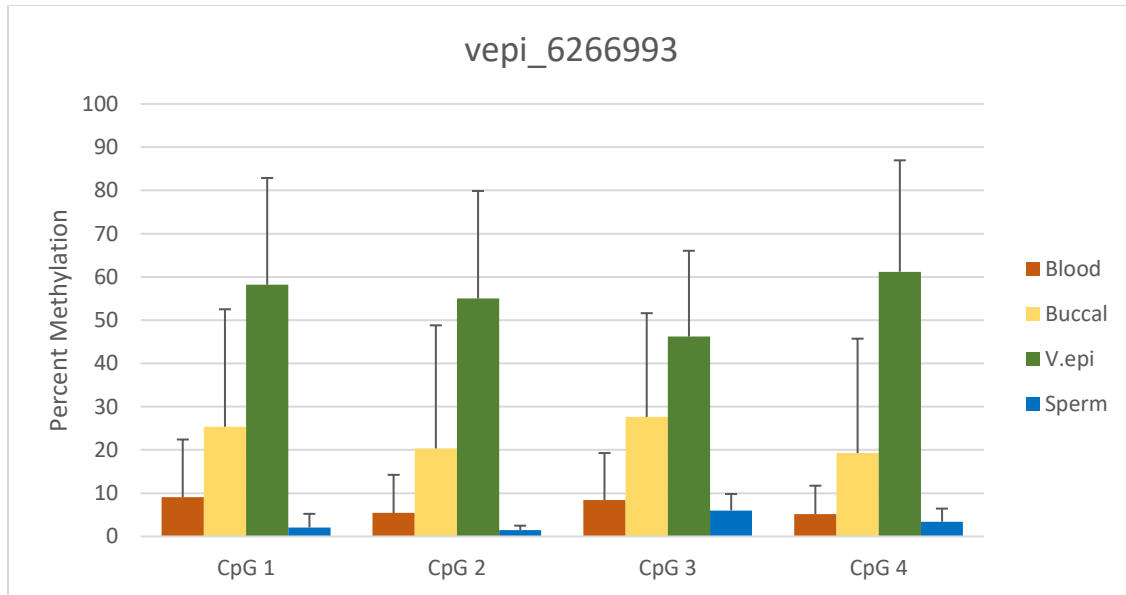


Figure 11. Histogram for 6266993

Average methylation data for all four tissues studied for the vaginal epithelial marker 6266993. Bars above each data indicate standard deviation.

Statistical analysis of the mean methylation values for vaginal epithelial tissue in comparison with blood, buccal, and sperm tissues was performed using a one-way ANOVA and Tukey's HSD test. The level of significance for each tissue as compared to vaginal epithelial are shown in Table 7.

<b>v. epi. 6266993</b>				
	<b>CpG1</b>	<b>CpG2</b>	<b>CpG3</b>	<b>CpG4</b>
<b>Blood</b>	0	0	0	0
<b>Buccal</b>	0.038	0.03	0.316	0
<b>Sperm</b>	0	0	0	0

Table 7. Levels of significance (p) for 6266993

For all CpG sites of v. epi marker compared to other tissues.

The p values of all CpG sites of v. epi. marker 6266993 were found to be below 0.05 except CpG 3, indicating significant difference in percent methylation values. The epigenetic marker 6266993 was found to differentiate vaginal epithelial cells from blood, buccal, and sperm cells ( $p < 0.05$ ) for the samples tested except for CpG 3 ( $p > 0.05$ ).

### **Vaginal Epithelial Marker 9323727**

Of the samples pyrosequenced for the epigenetic marker 9323727, 9 blood, 12 buccal, 9 vaginal epithelial, and 11 sperm samples provided viable data for analysis.

Percent methylation for all CpG sites was recorded and averaged for all samples.

Standard deviations for each CpG site for all tissues were calculated. Table 8 shows the percent methylation averages and standard deviation values for all CpG sites tested for marker 9323727.

		<b>CpG 1</b>	<b>CpG 2</b>	<b>CpG 3</b>	<b>CpG 4</b>	<b>CpG 5</b>
<b>Blood</b>	Avg.	2	6	8	8	14
	Std. Dev.	4	6	8.9	9.9	30.7
<b>Buccal</b>	Avg.	6	10	11	10	13
	Std. Dev.	4.9	7.3	7.1	5.7	8.8
<b>Vaginal Epithelial</b>	Avg.	40	49	47	38	37
	Std. Dev.	27.4	30.7	30.8	26.6	26.24
<b>Sperm</b>	Avg.	3	4	4	3	5
	Std. Dev.	2.5	4.4	3.7	3.1	4.6

Table 8. Percent Methylation Averages and Standard Deviations-Marker 9323727

Figure 12 shows the relative percent methylation averages and standard deviation values for all sites tested for the marker 9323727.



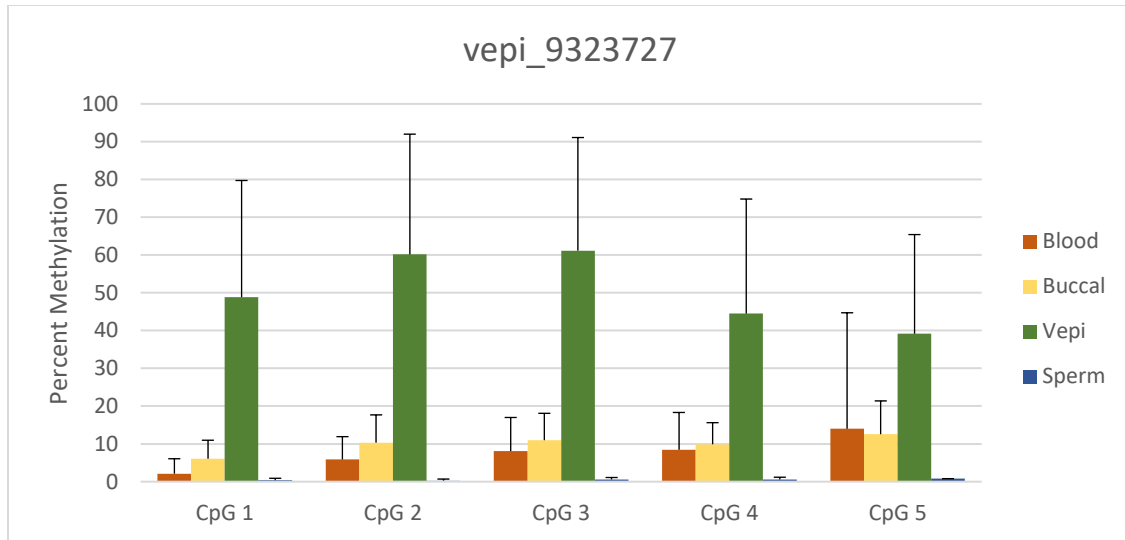


Figure 12. Histogram for 9323727

Average methylation data for all four tissues studied for the vaginal epithelial marker 9323727. Bars above each data indicate standard deviation.

Table 9 shows the level of significance of the methylation data of v.epi. tissue compared to all other tissues tested.

<b>v. epi. 9323727</b>					
	<b>CpG1</b>	<b>CpG2</b>	<b>CpG3</b>	<b>CpG4</b>	<b>CpG5</b>
<b>Blood</b>	0	0	0	0.001	0.069
<b>Buccal</b>	0	0	0	0	0.031
<b>Sperm</b>	0	0	0	0	0.006

Table 9. Levels of significance (p) for 9323727

For all CpG sites of v. epi. marker compared to other tissues

The p values of all CpG sites of v.epi. marker 9323727 were found to be below 0.05 except CpG 5, indicating significant difference in percent methylation values. The epigenetic marker 9323727 was found to differentiate vaginal epithelial cells from blood, buccal, and sperm cells in four of the five sites studied ( $p < 0.05$ ).

## CHAPTER V – DISCUSSION

The importance of tissue identification cannot be overemphasized in the serological process in a criminal case. Human identification through serological analysis and DNA profiling plays a critical role by supporting or rejecting an individual's involvement in a crime based on the presence or absence of a suspect's or victim's DNA on probative evidentiary items. Analysis of biological material in a criminal case serves to clarify the events that took place by distinguishing between casual and criminal contact between individuals. For example, the presence or transfer of skin tissues or saliva may be taken as casual contact during a handshake or kiss, but the presence of blood or semen may suggest more intimate or violent contact. Thus, the presence of a specific body fluid indicates the type of assault that may have occurred in criminal cases. If the biological origin of an evidence item cannot be readily inferred, a test for discrimination between forensically-relevant tissues could significantly enhance the item's probative value. Forensic identification of saliva, blood, and semen using methylation markers has been reported (Park et al 2014, Balamurugan et al 2014) but such markers for the identification of vaginal fluid are scarce.

In this study we sought to analyze certain methylation markers that have the potential to distinguish vaginal fluid from other tissues such as blood, buccal cells, and sperm. Several potential vaginal epithelial cell markers were reported by Park, et al., (2014) using Illumina human methylome 450K bead chip study, and we used this preliminary information to explore the extent to which those markers were useful in distinguishing vaginal fluid from other tissues.

Among several markers screened, a set of three epigenetic markers was identified that are promising for distinguishing vaginal epithelial cells from blood, buccal cells, and spermatozoa. The three markers that were identified for differentiating v. epi. cells from other tissues were cg-6266993, cg-9323727, and cg-4739647. These markers were identified by comparing the relative methylation levels of the four tissues studied using pyrosequencing.

The methylation data for marker cg-6266993 shows that the v. epi. samples are hypermethylated in general while all other tissues were hypomethylated. The data for this marker displayed a high standard deviation from the average when analyzed. These high standard deviations are due to certain outliers in the sample population. The methylation data for marker cg-9323727 also shows that the v. epi. samples are hypermethylated compared to other tissues that are hypomethylated. The data also showed a relatively high standard deviation from the average, due to a couple of outlier samples in the sample pool. The third marker cg-4739647 also shows that the v. epi. samples are hypermethylated in relation to the methylation levels of other tissues. The differences between methylation levels of vaginal epithelial tissue and other tissues were statistically significant ( $p < 0.05$ ), suggesting that markers cg-6266993, cg-9323727, and cg-4739647 may be used to identify vaginal epithelial tissue, and the methylation differences are tissue-specific and not by chance.

The study of epigenetic modifications of DNA and tissue specific methylation markers has received widespread attention among the forensic community in recent years (Frumkin, et al., 2011; Balamurugan, et al., 2014; Park, et al., 2014; Jenkins, et al., 2018). Epigenetics is a chemical process by which the gene expression is altered without

changing the genetic code of the gene involved. Some areas of DNA undergo an epigenetic change known as methylation, which has been demonstrated to be conserved over time (Vidaki, et al., 2016; Lee, et al., 2012). Appearing most often as a regulatory mechanism in the promoter region of genes, methylation is the addition of a methyl group (CH<sub>3</sub>) at the 5-carbon of a Cytosine (C) ring (followed by a Guanine-G), resulting in 5-methylcytosine (5mC) (Kader & Ghai, 2015). Because methylation is an epigenetic change, the methylation information is lost during the PCR process; performing bisulfite conversion of the DNA sample converts the unmethylated C to T, while the methylated C are not altered. This ratio of converted/unconverted 'C' is used during the analysis to estimate the methylation status of a marker.

In recent years, it has become evident that the momentum on the study of methylation markers for forensic purposes has been rapid and progressive. Reliable methods have become available for the analysis of methylation and the level of methylation of a CpG site has been used to accurately predict the tissue source of a DNA sample; for example, blood and semen identification as reported by Frumkin, et al., (2011), Lee, et al., (2012), and Balamurugan, et.al., (2014). These studies have successfully differentiated tissue types through differential methylation patterns and have served as forerunners for this current study.

The advantages of these methylation markers for tissue identification are several-fold. First, DNA-based tissue differentiation circumvents the challenges of evidence that has been improperly stored or degraded since DNA is more stable than proteins and degrades at a slower rate. Second, this method allows for the conservation of limited probative sample. Once DNA is extracted from a sample in routine HID analysis, a small

amount of the extract can be used for tissue identification without consuming additional evidentiary material. Third, bisulfite conversion and pyrosequencing can be performed with the existing lab equipment without any additional cost to the laboratory. Fourth, this technology has the potential for automation and multiplexing, thereby reducing the time and the amount of sample expended. The development of a universal methylation multiplex kit for forensic tissue and age identification is not too far away. Since the pyrosequencing technology provides quantitative methylation data, it becomes easy to compare multiple samples and multiple tissues for their similarities and differences. Fifth, besides tissue identification, this pyrosequencing technology can also be used in the determination of the age of a sample donor, studies related to obesity, alcoholism, smoking, twin studies, and cancer research.

Recently, the Illumina human methylome 450K bead chip has been used to study the overall methylation pattern of the entire human genome (Jenkins, et al. 2014, Lee, et al., 2015). However, this study is very expensive and specific methylation quantitation values are not derived from the data. Alternatively, the pyrosequencing technology provides the exact methylation quantitation values for each CpG site studied thereby making this technology more cost-efficient and a better forensic application.

Evidence items can be subjected to harsh environmental conditions before collection and during storage. Serological analysis of body fluid is carried out routinely using enzyme and protein markers, but exposure to heat, moisture, and the passage of time expedites the degradation of these proteins. The detection of acid phosphatase in semen and amylase in saliva are two common examples involving those tests. Studies involving mRNA profiling have also been reported, but again the tendency of mRNA to

degrade over time becomes problematic (Hass, et al., 2009). To overcome these shortcomings, other reliable methods should be explored; one such method to be explored is the use of DNA methylation for tissue identification. DNA degrades at a slower rate than enzymes, and partially-degraded DNA is still suitable for human identification (Edwards et al., 1991), thus making DNA a better candidate for tissue source attribution of biological evidence.

In serological and DNA analysis, reducing the amount of probative sample expended for each test becomes priority. Evidence conservation is also considered prudent as technology is ever-changing, and future innovations may be retroactively applied to unused portions of evidence. This is evident by the fact that the previous DNA technology, Restriction Fragment Length Polymorphism (RFLP) required 100 ng or more DNA for analysis, while the current Polymerase Chain Reaction (PCR) technology requires only one nanogram of DNA samples to produce a complete DNA profile. What was not possible two decades ago is possible today, thanks to the advancement in DNA technology as well as a rapid progress of computer applications.

With an eye toward future research, the higher standard deviations caused by the outliers need to be explored to see if any other parameters such as age, smoking, obesity, or alcoholism may contribute to the differences in individual methylation values. Rando and Verstrepen (2007) have reported that epigenetic changes can occur due to environmental factors such as diet and smoking. Several medical conditions and environmental factors such as obesity (de Mello et al., 2014), smoking (Besingi and Johansson, 2017), cancers, and age (Klutstein et. al. 2016; Jenkins et. al. 2018) are also

known to affect methylation levels. Additional research may be needed to explore the ways in which these external factors affect methylation patterns in different tissues.

In conclusion, this study shows how DNA methylation can be used for differentiating one tissue from a variety of other bodily fluids. All three markers detailed in this report were found to be hypermethylated in vaginal fluid, while the relative methylation levels of other tissues were hypomethylated. Thus, the three markers reported herein have the potential to differentiate vaginal fluid from blood, saliva, and sperm samples. The full potential of these methylation markers needs to be explored further as they are not only a good indicator of the tissue source of DNA samples but can also be used to determine the age of the donor of the samples when the perpetrator is unknown.

APPENDIX A – Primer Specifications

Marker	Primer Sequence	PCR Product Size	CPG Site in Park Paper
4739647	<p>Forward</p> <p>5' ATGGTGGAGGTTGTAGTT</p> <p>Reverse</p> <p>5' TCCCACCCAACATTACACTA</p> <p>Sequencing</p> <p>5' GTGGAGGTTGTAGTTT</p>	89bp	4th
6266993	<p>Forward</p> <p>5' GGTGATTTTGGAGGGTTTGAT</p> <p>Reverse</p> <p>5' AACCCCCCCCCCATATTTAA</p> <p>Sequencing</p> <p>5' GGGTTTGATTTGTGGA</p>	137bp	4th
9323727	<p>Forward</p> <p>5' TGGGGAGATTGTAGTTTTTAAGT</p> <p>Reverse</p> <p>5' ACCCCATTCCTTCCTCCT</p> <p>Sequencing</p> <p>5' GGGAATAAGTATTTTTAGGTTGG</p>	194bp	3rd



## REFERENCES

- An, J., Choi, A., Shin, K., Yang, W., & Lee, H. (2013). DNA methylation-specific multiplex assays for body fluid identification. *International Journal of Legal Medicine*, 127, 35-43.
- Antunes, J., Silva, D., Balamurugan, K., Duncan, G., Alho, C. S., & McCord, B. (2016). Forensic discrimination of vaginal epithelia by DNA methylation analysis through pyrosequencing. *Electrophoresis*, 37, 2751-2758.
- Besingi W., Johansson A. (2014). Smoke-related DNA methylation changes in the etiology of human disease. *Human Molecular Genetics*, 23(9), 2290-2297.
- Balamurugan, K., Bombardi, R., Duncan, G., and McCord, B. (2014). Identification of spermatozoa by tissue-specific differential DNA methylation using bisulfite modification and pyrosequencing. *Electrophoresis*, 35, 3079–3086.
- Botstein, D., White, R., Skolnick, M., & Davis, R. (1980). Construction of a Genetic Linkage Map in Man Using Restriction. *American Journal of Human Genetics*, 32, 314-331.
- Brief History of Forensic DNA Typing. (2010, August 24). Retrieved March 9, 2017, from <http://www.cstl.nist.gov/strbase/ppt/intro.pdf>.
- Budowle, B. (2000). Capillary Electrophoresis for Short Tandem Repeat Loci Typing and Real Time Fluorescent Detection. In *DNA Typing Protocols: Molecular Biology and Forensic Analysis* (pp. 157-188).
- Budowle, B., Smith, J., Moretti, T., & DiZinno, J. (2000). *DNA Typing Protocols: Molecular Biology and Forensic Analysis*. Natick, MA: Eaton Publishing.
- Butler, J. M. 2010. *Fundamentals of Forensic DNA Typing*. London: Academic Press.

- Casey, D. G., & Price, J. (2010). The sensitivity and specificity of the RSID-saliva kit for the detection of human salivary amylase in the Forensic Science Laboratory, Dublin, Ireland. *Forensic Science International*, 194(1-3), 64-71.
- Chakarov, S., Petkova, R., Russev, G. C., & Zhelev, N. (2014). DNA damage and mutation. Types of DNA damage. *BioDiscovery*, 11(1). doi:10.7750/BioDiscovery.2014.11.1
- DNA methylation: a guide. (2017, April 2). Retrieved from <http://www.abcam.com/epigenetics/dna-methylation-a-guide>.
- de Mello VD, Pulkkinen L, Lalli M, Kolehmainen M, Pihlajamaki J, Uusitupa M.. (2014). DNA methylation in obesity and type 2 diabetes. *Annals of Medicine*, 46(3), 103-113.
- Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V.K., Attwood, J., Burger, M., Burton, J., Cox, T.V., Davies, R., Down, T.A., Haefliger, C., Horton, R., Howe, K., Jackson, D.K., Kunde, J., Koenig, C., Liddle, J., Niblett, D. Otto, T., Pettett, R., Seemann, S., Thompson, C., West, T., Rogers, J., Olek, A., Berlin, K., & Beck, S. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. *National Genetics*, 38(12), 1378-1385.
- Edwards, A., Civitello, A., Hammond, H., & Caskey, C. (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics*, 49(4), 746-756.
- Frumkin, D., Wasserstrom, A., Budowle, B., & Davidson, A. (2011). DNA methylation-based forensic tissue identification. *Forensic Science International: Genetics*, 5, 517-524.
- Gardiner-Garden, M., Frommer, M., *Journal of Molecular Biology*. 1987, 196, 261–282.
- Gaensslen, R. (1983). *Sourcebook in forensic serology, immunology, and biochemistry*. National Institute of Justice, US Department of Justice.

- Giampaoli, S., Berti, A., Valeriani, F., Gianfranceschi, G., Piccolella, A., Buggiotti, L., . . .  
Spica, V. R. (2012). Molecular identification of vaginal fluid by microbial signature. *Forensic Science International: Genetics*, 6(5), 559-564.
- Gill, P., Jeffreys, A., & Werrett, D. (1985). Forensic application of DNA 'fingerprints'. *Nature*, 318(6046), 577-579.
- Haas, C., Klessner, B., Maake, C., Bar, W., & Kratzer, A. (2009). mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR. *Forensic Science International: Genetics* (3), 80–88.
- Haltiner, C. C. (2008). Comparison of Radial Diffusion, Rapid Stain Identification Series Test for Saliva (RSID), and DNA Test Positivity in Saliva Specimens Over Time. Masters Thesis, University of California Davis.
- Harbison, S., & Fleming, R. (2016). Forensic body fluid identification: state of the art. *Research and Reports in Forensic Medical Science*, 2016(6), 11-23.
- Jakubowska, J., Maciejewska, A., Pawlowski, R., & Bielawski, K. P. (2013). mRNA profiling for vaginal fluid and menstrual blood identification. *Forensic Science International: Genetics*, 7(2), 272-278.
- James, S. H., Nordby, J. J., Bell, S., & Jon J. Nordby, P. D. (2005). *Forensic Science: An Introduction to Scientific and Investigative Techniques* (Second Edition ed.). CRC Press.
- Jenkins, T.G., Aston, K.I., Pflueger, C., Cairns, B.R., & Carrell, D.T. (2014). Age-Associated Sperm DNA Methylation Alterations: Possible Implications in Offspring Disease Susceptibility. *PLoS Genetics*, 10(7). e1004458. doi:10.1371/journal.pgen.1004458.
- Jenkins, T., Aston, K., Cairns, B., Smith, A., & Carrell, D. (2018). Paternal germ line aging: DNA methylation age prediction from human sperm. *BMC Genomics*, 19 (763).

- Kader, F., & Ghai, M. (2015). DNA methylation and application in forensic sciences. *Forensic Science International*, 249, 255-265.
- Klutstein M., Nejman D., Greenfield R., Cedar H. (2016). DNA Methylation in Cancer and Aging. *Cancer Research*, 76(12), 3446-3450.
- Landsteiner, K. (1901). Agglutination phenomena in normal human blood. *Wiener klinische Wochenschrift*, 14, 1132-1134.
- Landsteiner, K. & Wiener, A.S., (1940) An agglutinable factor of human blood recognizable by immune sera for rhesus blood. *Proceedings of the Society for Experimental Biology and Medicine*, 43, 223.
- Ledgerwood, S. (2015). High Resolution Melt Analysis of Samples with Differential DNA Methylation to Identify Tissue Source of Origin. Masters Thesis, University of Southern Mississippi.
- Lee, H.Y., An, J.H., Jung, S., Oh, Y.N., Lee, E.Y., Choi, A., Yang, W.I., Shin, K. (2015). Genome-wide methylation profiling and a multiplex construction for the identification of body fluids using epigenetic markers. *Forensic Science International: Genetics*, 17, 17–24.
- Lee, H. Y., Park, M. J., Choi, A., An, J. H., Yang, W. I., & Shin, K.-J. (2012). Potential forensic application of DNA methylation profiling to body fluid identification. *International Journal of Legal Medicine*, 126, 55-62.
- Lister, R., Pelizzola, M., Dowen, R. H., Hawkings, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B., Ecker, J. R. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462(7271), 315-322.

- 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, 1736-1745.
- Magalhaes, T., Dinis-Oliveira, R. J., Silva, B., Corte-Real, F., & Vieira, D. N. (2015). Biological Evidence Management for DNA Analysis in Cases of Sexual Assault. *Scientific World Journal*(2015), 365674.
- McGrew, M., & Rosenthal, N. (1993). Quantitation of genomic methylation using ligation-mediated PCR. *Biotechniques*, 15(4), 722-729.
- Myers, J. R., & Adkins, W. K. (2008). Comparison of Modern Techniques for Saliva Screening. *Journal of Forensic Sciences*, 53(4), 862-867.
- Ng, R. K., & Gurdon, J. B. (2008). Epigenetic inheritance of cell differentiation status. *Cell Cycle*, 7(9), 1173-1177.
- Noureddine, M. (2011, October 19). Forensic Tests for Semen: What you should know. Retrieved March 14, 2017, from Forensic Resources: <https://ncforensics.wordpress.com/2011/10/19/forensic-tests-for-semen-what-you-should-know/>.
- Park, J.-L., Kwon, O.-H., Kim, J. H., Yoo, H.-S., Lee, H.-C., Woo, K.-M., . . . Kim, Y. S. (2014). Identification of body fluid-specific DNA methylation markers for use in forensic science. *Forensic Science International: Genetics*(13), 147-153.
- Patterson, K., Molloy, L., Qu, W., & Clark, S. (2011). DNA Methylation: Bisulphite Modification and Analysis. *Journal of Visualized Experiments*, 56, 3170.

- Patzelt, D. (2004). History of forensic serology and molecular genetics in the sphere of activity of the German Society for Forensic Medicine. *Forensic Science International*, 144(2-3), 185-191.
- Preston, A. (n.d.). Forensic Biology Screening Workshop. Retrieved from SlidePlayer: <http://slideplayer.com/slide/4462911/14/images/82/Takayama+Crystals+Red+feathery+appearance.jpg>.
- Rando, O. & Verstrepen, K. (2007). Timescales of genetic and epigenetic inheritance. *Cell*, 128, 4 (655–668).
- Reynolds, M. (2004, June). The ABACard HemaTrace- A Confirmatory Identification of Human Blood Located at Crime Scenes. *International Association of Bloodstain Pattern Analysts Newsletter*, pp. 4-10.
- Reynolds, R. M., Jacobsen, G. H., & Drake, A. J. (2013). What is the evidence in humans that DNA methylation changes link events in utero and later life disease? *Clinical Endocrinology*, 78, 814-822. doi:10.1111/cen.12164.
- Saiki RK, et al. (1985). Enzymatic Amplification of  $\beta$ -globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia. *Science*, 230, 1350-1354.
- Sato, I., Kojima, K., Yamasaki, T., Yoshida, K., Yoshiike, M., Takano, S., . . . Iwamoto, T. (2004). Rapid detection of semenogelin by one-step immunochromatographic assay for semen identification. *Journal of Immunological Methods*, 287(1-2), 137-145.
- Sodium Bisulfite Conversion of Unmethylated Cytosines in DNA. (2009). In Qiagen, EpiTect® Bisulfite Handbook (pp. 16-20).

- Stach, D., Schmitz, O. J., Stilgenbauer, S., Benner, A., Dohner, H., Wiessler, M., & Lyko, F. (2003). Capillary electrophoretic analysis of genomic DNA methylation levels. *Nucleic Acids Research*, 31(2), e2.
- Tost, J., & Gut, I. G. (2007). DNA methylation analysis by pyrosequencing. *Nature Protocols*, 2, 2265-2275.
- Turrina, S., Filippini, G., Atzei, R., Zaglia, E., & Leo, D. D. (2008). Validation studies for rapid stain identification-blood (RSID-blood) kit in forensic caseworks. *Forensic Science International: Genetics Supplement Series*, 1(1), 74-75.
- Vidaki, A., Giangasparo, F., & Court, D. S. (2016). Discovery of potential DNA methylation markers for forensic tissue identification using bisulphite pyrosequencing. *Electrophoresis*, 37, 2767-2779.
- Wasserstrom, A., Frumkin, D., Davidson, A., Shpitzen, M., Herman, Y., & Gafny, R. (2013). Demonstration of DSI-semen--A novel DNA methylation-based forensic semen identification assay. *Forensic Science International. Genetics.*, 7(1), 136-142.