2023

Characterization of blood and vaginal fluid using a microRNA expression profiling method

BOSTON UNIVERSITY

ARAM V. CHOBANIAN & EDWARD AVEDISIAN SCHOOL OF MEDICINE

Thesis

CHARACTERIZATION OF BLOOD AND VAGINAL FLUID USING A MICRORNA EXPRESSION PROFILING METHOD

by

ISABEL LUISA N. TAN PALANCA

B.S., University of the Philippines - Diliman, 2018

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science

2023

Approved by

First Reader

Robin W. Cotton, Ph.D.
Associate Professor of Anatomy & Neurobiology

Second Reader

Amy N. Brodeur, M.F.S., ABC-GKE
Assistant Professor of Anatomy & Neurobiology

Third Reader

Andrew J. Schweighardt, Ph.D., ABC-MB
Special Agent Forensic Scientist
Tennessee Bureau of Investigation

ACKNOWLEDGMENTS

I would like to express my most sincere gratitude to my family and friends for their love, unending patience, and unwavering support, to Dr. Robin Cotton for her wisdom and calming guidance throughout the project and writing process, to Prof. Amy Brodeur and Dr. Andrew Schweighardt for their valuable input and for taking the time to help me improve my work, to HTG Molecular Diagnostics, Inc. for providing us with the necessary training and equipment to conduct this study, to Simone for working with me on all laboratory and computer related experiments, and to the BMFS Class of 2023 for the friendship, fun experiences, and sweet memories.

CHARACTERIZATION OF BLOOD AND VAGINAL FLUID USING A MICRORNA EXPRESSION PROFILING METHOD

ISABEL LUISA N. TAN PALANCA

ABSTRACT

MicroRNAs (miRNAs) are short non-coding sequences that function in maintaining tissue and cell specificity. These characteristics have made them ideal candidiates for body fluid identification. This study aims to determine if miRNA expression profiles can be used to differentiate vaginal fluid samples collected at different time points within a single menstrual cycle, as well as miRNA markers that can be used to differentiate menstrual blood from peripheral blood. The HTG EdgeSeq® workflow was used to prepare and analyze the samples. The HTG Reveal software was used to analyze the expression profiles of these samples by conducting pairwise comparisons. The top differentially expressed probes based on "rawP" values were assessed for observable trends in their fold changes. Potential miRNA markers for vaginal fluid, peripheral blood, and menstrual blood in this study and those from existing literature were identified. Week 3 appears to be the most different among the vaginal fluid only samples within the menstrual cycle. Minimal differences were observed between the samples of Weeks 1 and 2, thus the identified differentially expressed probes could not be reliably used to differentiate all timepoints in the menstrual cycle. MiR-200b-3p is the only probe out of four possible menstrual blood markers with supporting evidence from the data collected in this study and in literature. Several other miRNAs (miR-106a, miR-144-3p, miR-16-5p, miR-25-3p, miR-451a, miR-486-5p, and miR-93-5p) show potential as blood markers without

distinguishing between peripheral and menstrual blood. Possible miRNA markers for vaginal fluid in general were also determined: miR-124-3p, miR-128-1-5, mir-147b, miR-193b-5p, miR-5585-3p, and miR-612.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	two dimensional
3D	three dimensional
bp	base pairs
CE	capillary electrophoresis
DE	differential expression
DNA	Deoxyribonucleic acid
HTG	HTG Molecular Diagnostics Inc.
miRNA	micro ribonucleic acid
mL	milliliters
MTC	multi-tissue control
mRNA	messenger RNA
mtDNA	mitochondrial DNA
NEB	New England Biolabs Inc.
NGS	Next-generation sequencing
NTC	No template control
PAS	periodic acid-Schiff
PCA	Principal component analysis
PCR	Polymerase chain reaction
pM	picomolar
QC	quality check
qPCR	Quantitative polymerase chain reaction

RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RSD	relative standard deviation
SAV	Sequencing Analysis Viewer
STR	short tandem repeat
Tris-HCl	Trizma hydrochloride
μL	microliters
WTA	Whole Transcriptome Assay

1. INTRODUCTION

1.1 Body Fluid Identification in Sexual Assault Cases

Body fluid identification is a driving force in forensics. This often determines whether evidence will be useful in creating linkages between people, places, or items involved in a case. Hence, it is considered the first critical test in the examination of biological evidence [1]. Most methods employed to test for the presence of body fluids are categorized as presumptive or confirmatory. Presumptive testing is similar to a triage in that these methods are used to determine the probative value of an unknown stain or fluid. This is typically the first step in body fluid identification as it dictates whether or not the unknown stain requires downstream analysis. Presumptive tests are designed to be speedy, sensitive, safe, and simple [2]. Examples of these include AP Spot test to detect the presence of semen and the Kastle-Meyer test to detect the presence of blood. Confirmatory tests, in comparison, positively identify that the suspected body fluid is present in the evidence. These tests are more tedious to perform but provide more definitive results [1]. Due to the inter- and intraperson variability of some body fluids, not all body fluids have confirmatory tests available [3].

1.1.1 Nucleic acid tesing

Nucleic acid testing, especially in the context of forensics, is almost synonymous to DNA analysis. Over the years, analyses have evolved to cover more nucleic acid types and different methods. Conducting nucleic acid testing on evidence typically results in data that could be compared to a reference sample. The most straightforward example of this

would be obtaining a DNA profile from an evidentiary sample after going through DNA analysis and comparing this profile to a DNA profile from a known individual. These tests yield incredibly valuable data, especially when combined with body fluid identification. The multi-step nature of the process increases the chance of error but this is mitigated by validated protocols, routine quality checks, and proper training of forensic scientists. Conducting DNA analysis is more expensive than perfoming a simple presumptive test, which is why presumptive tests are used first to determine if an item of evidence is worth further analysis by testing for DNA [3].

Nucleic acid tests have also shown to be advantageous when it comes analyzing and interpreting mixture samples. Results from a single analysis can show multiple donors present on the evidence. However, while these mixtures are from two different sources, they do not take into consideration the type of fluid or tissue from which these were taken.

1.1.2 Vaginal fluid as an evidentiary sample

Vaginal fluid is one of the most common evidentiary samples in sexual assault cases. Its composition varies widely; vaginal fluid is affected by the health and age of a person among other things [4, 5, 6, 7]. Vaginal fluid differences at the protein level could even be detected within a time frame as short as a month, such as over the course of a person's menstrual cycle [5, 7]. The microflora detected in vaginal fluid seem to also be influenced in part by a person's menstrual cycle [4]. Hormonal changes that occur, whether naturally during this period or introduced externally, have also shown to affect vaginal fluid [8].

Because vaginal fluid as a sample varies greatly, testing for its presence in a sample is not as straightforward as testing for other biofluids. The identification of glycogenated epithelial cells by staining with Lugol's iodide or periodic acid-Schiff (PAS) reagent may be useful in presumptively determining the presence of vaginal fluid. However, glycogenated epithelial cells are only present during the childbearing years of a female. Additionally, glycogenated cells are not isolated to the vagina; they can be found in lower concentrations on mouth, rectum, and penile swabs [1, 7]. Thus, the method can only be considered as a presumptive test for the presence of vaginal fluid.

There have been studies that make use of nucleic acid tests to distinguish vaginal fluid from other types of samples [7, 9, 10, 11, 12, 13, 14, 15]. A possible caveat of these observations is that the strength of the data is hinged on the uniqueness of the microbial composition of vaginal fluid. On the other hand, studies that have been able to identify some markers from vaginal fluid itself have not converged on a set of markers which might constitute a definitive test [3, 11]. While these efforts have made identifying vaginal fluid easier, interpretation becomes complicated when there is no health information available on the source of the fluid [4,7].

1.1.3 Menstrual blood as an evidentiary sample

Menstrual blood, despite being less common in crime scenes, is incredibly useful in providing or corroborating case context. There is almost no difference visually between menstrual and peripheral blood, hence initial testing would likely follow the protocol for peripheral blood without consideration of the details surrounding the case. D-dimer, a

degradation product of fibrin, is present in menstrual blood at high concentrations, thus assays have been created to test for this compound specifically [17]. One of the biggest limitations of D-dimer testing, however, is that false positives can still occur from people who have cardiovascular diseases or cancer [17].

Determining the presence of menstrual blood has also been done through mRNA and miRNA markers. Menstrual blood is inherently a mixture of blood, vaginal fluid, and some tissues so its markers may show some overlap with other biofluids [7]. Unfortunately, like vaginal fluid, there is no consensus on which markers are definitive for menstrual blood [3, 7, 11].

1.2 Next-Generation Sequencing in Forensics

Next-generation sequencing (NGS) is an incredible breakthrough for many scientific fields, especially with the completion of the Human Genome Project. Its application to forensic science has undergone substantial development to make the process feasible for routine experiments.

At present, DNA typing mainly makes use of capillary electrophoresis (CE). Short tandem repeats (STRs) that are present at highly discriminating loci in different regions of human DNA are amplified by the polymerase chain reaction (PCR) and tagged with fluorescent dyes. These STRs are then sorted by size and identified by color through CE, producing a DNA profile [18]. The data coming from CE, however, is limited to allele lengths; there is no sequence information apart from how many repeats are present [19].

Introducing NGS to the DNA analysis workflow allows for more data to be elucidated from the same amount of evidence. The same STRs that are already used for CE are sequenced such that reads of the same length can be differentiated by their sequences of bases [19]. Unlike the typical workflow, degraded DNA samples are not a major problem with NGS [18]. The overlapping reads from the same target region compensate for the gaps caused by DNA degradation. This is a big advantage considering that many non-reference forensic samples are never perfectly preserved prior to collection. Evidence found in crime scenes are often exposed to bacteria, harsh weather, and other contaminants that could interfere with DNA analysis. By the time they are processed, it is likely that some degradation would be observed. Multiplexing more STRs and running more samples in parallel are also some benefits of NGS over the conventional CE method [19, 20].

The integration of NGS into the forensic DNA analysis workflow shows great promise. Van Neste et al. (2013) have already created a pipeline that streamlines STR analysis using a dataset coming from an Illumina MiSeq [20]. NGS can further be used in identifying Y-chromosome related distinctions and epigenetic analysis of monozygotic twins. Non-human samples can also be processed with NGS, which is not possible with the current kits available for STR typing [19]. Body fluid identification has likewise been conducted through NGS by identifying unique markers for certain biofluids [3, 6, 7, 10, 11].

1.3 MicroRNA (miRNA)

MicroRNAs (miRNAs) are non-coding sequences, about 18-24 bp long, that work as post-transcriptional gene regulators [21, 22, 23]. Unlike DNA which has the exact same sequence within a person regardless of cell origin, there are about 2,000 different human miRNA sequences found to date [22]. The sequences and expression levels of miRNAs in each cell varies based on cell type, function, and location among others [21, 22]. Thus, the varying expression levels of miRNA capture the state of the tissue or fluid.

The transcription of miRNA begins with RNA polymerase II activity to form the primary miRNA transcript. The enzyme DROSHA then forms a complex with DGCR8 in the cell nucleus in order to cleave the primary miRNA transcript into a distinct loop structure [21, 22]. It is this particular structure, which is known as the miRNA precursor, that makes miRNA unique from other RNA with regulatory functions [23]. The final step before the double-stranded mature miRNA forms is the cleavage of the loop structure by Dicer, an endonuclease, in the cytoplasm [21, 22].

Although the mature miRNA has a duplex structure, typically only one strand goes onto the RNA-induced silencing complex (RISC) while the other is degraded [21, 22]. In cases where both strands proceed into the RISC, a distinction is made between the two by affixing "-3p" or "-5p" onto the name of the miRNA [22]. This indicates that the miRNA strand in the RISC comes from the 3' or 5' end respectively of the mature miRNA. It is the entire RISC complex that binds with messenger RNAs (mRNA) to work on gene expression regulation [21, 22, 23].

The presence of miRNAs is associated with many biological processes [21, 22, 23]. One example is the role of miRNAs in making sure that tissues are functioning correctly. Hammond (2015) writes about this by citing studies that have shown the development of a heart despite lacking cardiac-specific miRNA; leading to the suggestion that one possible purpose of miRNAs is to maintain differentiation of a tissue rather than assign its identity [22]. Knowing that there are some tissue or body fluid specific miRNAs has allowed them to serve as markers for tissue or fluid identification [3, 11, 12, 19]. The expression of some miRNAs has been also shown to be indicative or related to disease [22, 24, 25].

1.3.1 MiRNA in forensic science

MiRNAs have been making frequent appearances in forensic science research, especially associated with body fluid identification [3, 11, 12, 19]. Body fluid identification using miRNAs may be more tedious than chemical testing but it consumes less of the evidence and provides more specific information. This is supported by a study from Lewis, Layne, and Seashols-Williams (2019) who report that miRNA can still be detected after DNA extraction, thereby foregoing a separate RNA extraction to determine the kind of biofluid from which the DNA was sourced [26]. Co-extracting DNA and RNA increases the amount of information that can be obtained while reducing the amount of time and resources spent on the analysis. Furthermore, no additional costs are incurred for equipment as miRNA analyses use machines that are already present in most forensic laboratories to conduct DNA testing.

DNA testing cannot be used to determine the nature of a biological sample, hence analyses with nucleic acids that show some tissue or fluid specificity have been developed. Both mRNA and miRNA can be used but the length of miRNA is a significant advantage over mRNA; the shorter fragments make miRNA inherently more resistant to degradation compared to the longer mRNA sequences [14].

Silva et al. (2015) summarize some of the miRNA markers that have been identified for body fluids [19]. Although these markers can signify the kind of biofluid (e.g. whether a sample is saliva or vaginal fluid), some difficulty has been reported in finding miRNA markers that try to provide a more temporal or health-specific context [16, 14, 27]. The expression of miRNA markers is also influenced by age and disease, therefore, there is yet to be a true consensus of miRNA markers for all body fluids.

1.4 HTG EdgeSeq system

The HTG EdgeSeq® system (HTG Molecular Diagnostics Inc. (HTG); Tucson, AZ) semi-automates a nucleic acid extraction-free method to prepare mRNA and miRNA molecules for next-generation sequencing (NGS). Assays that have been developed to be used with the HTG EdgeSeq® system include the miRNA Whole Transcriptome Assay (WTA) and HTG EdgeSeq Oncology Biomarker Panel.

The biggest innovation of the system is its extraction-free chemistry which is done at room temperature. This is the target protection step of the protocol and occurs inside the HTG EdgeSeq® processor (liquid handler robot). HTG provides a set of reagents that contain a total of 2,102 probes protected by “wings” and unhybridized “wingmen”. When

the RNA sample and these reagents are combined, the targeted complementary sequence attaches onto the probe by DNA-RNA hybridization. These probes are protected from nuclease activity by wings. In order to similarly protect the RNA strand, wingmen hybridize to the wings, thus flanking the RNA sequence. After this, S1 nuclease digests any free floating single-stranded nucleic acids, e.g. the overhanging portions of the target RNA and unhybridized DNA probes. Heat is then applied to remove the DNA-RNA hybridization. The end products of the target protection step are single-stranded protected DNA probes that contain the complement of the target RNA sequence (F. Kero, personal communication, November 15, 2021) [28].

All succeeding steps in the HTG miRNA WTA occur outside of the HTG EdgeSeq® processor. Tags and adapters are added during the amplification step to prepare the samples for sequencing. Purification via magnetic beads is done to remove any unbound tags, thus having only tagged sequences in solution. An overview of the process is illustrated in Figure 1 [29]. While the total time to complete the process is similar to a typical extraction to sequencing workflow, HTG EdgeSeq® system limits hands-on time, hence its claim of semi-automation. This allows for better reproducibility and minimal human error [28, 29]. Further, because target protection is contained within the HTG EdgeSeq® processor, there is less sample manipulation and less exposure to contamination.

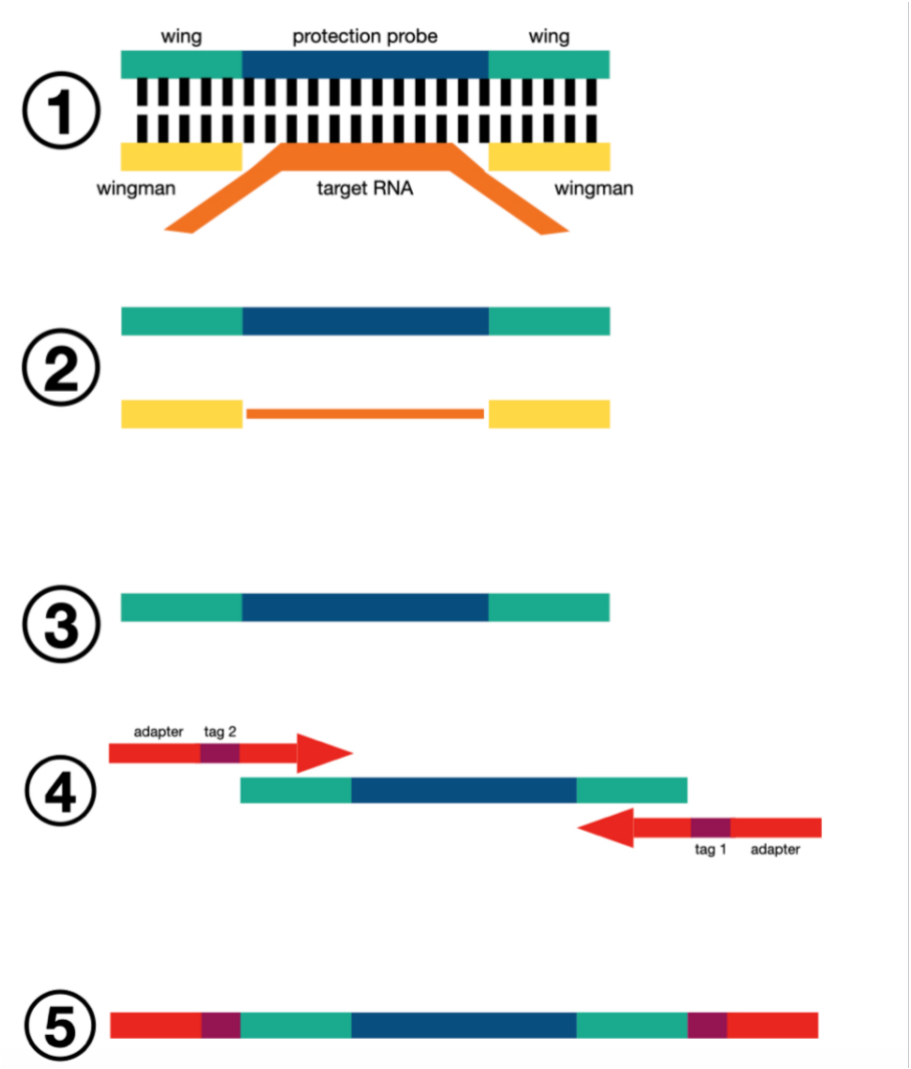

Figure 1. HTG EdgeSeq® chemistry. When the reagents interact with the sample lysate, the targeted RNA is flanked by wingmen to capture the necessary sequence (Step 1). Single-stranded RNAs are then formed in solution (Step 2). S1 nuclease and heat are applied, leaving behind only protected probes that contain the target sequences (Step 3). Steps 1-3 occur inside the HTG EdgeSeq® processor. Amplification is performed to add adapters and tags onto the probes (Step 4) and the solution is purified via magnetic beads so that only tagged probes remain (Step 5). Steps 4 and 5 occur outside of the HTG EdgeSeq® processor [29].

1.5 Objectives

Analyses that make use of nucleic acids are designed such that the data, such as evidentiary DNA profiles, could be used to make a comparison to a reference sample. Nucleic acid experiments are not routinely used to provide case context by body fluid identification.

There are three specific goals in this study. First, to produce miRNA expression profiles from vaginal fluid and blood samples using the HTG EdgeSeq® system following the miRNA Whole Transcriptome Assay (WTA). Using these profiles, the second objective is to determine if an miRNA signature can be identified based on differences in the miRNA expression levels obtained from vaginal samples that were collected from a single donor over the course of their menstrual cycle. Lastly, this study aims to determine if miRNA expression profiles can be used to differentiate peripheral blood from menstrual blood.

2. MATERIALS AND METHODS

All methods used are adapted from the HTG EdgeSeq® System User Manual (2021) and communication with HTG Molecular Diagnostics Inc. (HTG).

2.1 Validation of the use of dried body fluids with the HTG EdgeSeq® protocol

The HTG EdgeSeq® system has been used mostly with liquid samples. The use of dry swabs containing body fluids with the HTG EdgeSeq® was tested in advance by HTG to ensure that the method works. Preliminary testing was similarly done. The optimum amount of elution buffer to process each half swab was determined to be ~60µL. This ensures that there is enough liquid to be used inside the HTG EdgeSeq® machine after preparing the swab.

2.2 Sample collection and preparation

All samples were collected in compliance with procedures outlined in protocol H-26187, approved by the Boston University Institutional Review Board. Each swab contained varying amounts of biofluids; swabs that were prepared from liquid biofluids were spiked with 30µL of sample while swabs collected directly from donors were not quantified.

Vaginal swabs and menstrual swabs were collected from volunteers. All swabs used as menstrual cycle samples were received from a single volunteer over the course of 28 days. For the creation of blood swabs, approximately 30µL of liquid peripheral blood was placed onto sterile cotton swabs. To create the peripheral blood and vaginal fluid mixture swabs, 30µL of liquid peripheral blood was added onto a vaginal swab that was previously

collected. These were then left to dry at room temperature. A liquid multi-tissue control (MTC), which is a proprietary run-to-run consistency control, was provided by HTG.

Each sample type had three replicates. The samples corresponding to the menstrual cycle and the mixture were represented in technical replicates (i.e. from the same donor and same date of collection) while the samples for blood and vaginal fluid were biological replicates (i.e. from different donors). Table 1 lists the sample names used in the procedure and their corresponding descriptions.

Table 1. Legend of the sample names used all throughout the experiment

Group Name	Description
MTC	A control sample provided by HTG Molecular Diagnostics
V_Week1	Vaginal swabs collected during the first week of the volunteer's menstrual cycle
V_Week2	Vaginal swabs collected during the second week of the volunteer's menstrual cycle
V_Week3	Vaginal swabs collected during the third week of the volunteer's menstrual cycle
V_Week4	Vaginal swabs collected during the fourth week of the volunteer's menstrual cycle. This also serves as the menstrual blood sample.
Blood	30µL of liquid peripheral blood deposited onto clean cotton swabs
Vaginal	Vaginal swabs collected from volunteers without knowledge of their menstrual cycle
Blood_Vag	Vaginal swabs collected from volunteers without knowledge of their menstrual cycle spiked with 30µL of peripheral blood

To prepare the 24 samples for analysis on the HTG EdgeSeq processor, approximately half of each replicate swab was placed into a clean 2mL tube. To each 2mL tube, 60µL of 1X Plasma Lysis Buffer (HTG; Tucson, AZ) was added and left to incubate at room temperature for 10 minutes. The swabs were then removed from the liquid and placed into clean plastic spin basket. The tubes, which had both the spin basket and

remaining lysis buffer were centrifuged at 15,000 rpm (maximum speed) for 5 minutes. The spin basket was removed from the tube and 30µL of the lysis buffer was transferred to a clean 2mL tube. To each tube, 3µL of Proteinase K was added. The tubes were incubated at 50˚C in a shaking heat block for 3 hours.

2.3 Target protection

A Sample Plate file was created on Microsoft Excel on a separate computer prior to processing the samples on the HTG EdgeSeq®. This file is a template created by HTG that allows the HTG EdgeSeq® processor to identify the samples by their names and corresponding plate layout. The Sample Plate file also indicates the type of assay to be done. In this study, the miRNA Whole Transcriptome Assay (WTA) was used.

After the incubation with Proteinase K, 25µL of each sample was placed into a new Sample Plate (HTG; Tucson, AZ), following the plate layout indicated on the Sample Plate file. The Sample Plate was then placed inside the HTG EdgeSeq® processor along with the Stop Plate, Assay Reagent Pack, and plastic pipette tips supplied by HTG.

On the HTG EdgeSeq® computer that is connected to the HTG EdgeSeq® processor, the miRNA WTA was chosen and the Sample Plate file was uploaded. Once the information was uploaded and all plastics were in place, the target protection was initiated. This process takes approximately 20 hours and is completed at room temperature. Referring to Figure 1, target protection concludes by Step 3. The products from each sample are single-stranded protected DNA probes that contain the complement of the target RNA sequence which is ready for amplification and library preparation.

2.4 PCR amplification and tagging

Three master mixes (designated as F1, F2, and F3 after their respective forward primers) were created by combining 24µL Hemo KlenTaq® enzyme (New England Biolabs Inc. (NEB); Ipswich, MA), 60µL Hemo KlenTaq® buffer (NEB; Ipswich, MA), 6µL dNTP solution mix (NEB; Ipswich, MA), 120µL molecular grade water, and 30µL of the appropriate forward primer. The master mix was dispensed in 24µL aliquots into all the wells in a single column of a PCR plate. To each row, 3µL of the 8 different corresponding reverse primers was added. Figure 2 illustrates how the master mixes were distributed in the plate. Once each well contained the master mix and correct pair of primers, 3µL of the appropriate sample from the Stop Plate was added. Caution was exerted to avoid the oil on top of the sample wells in the Stop Plate. Amplification was done using the Applied Biosystems™ Veriti™ 96-Well Fast Thermal Cycler (ThermoFisher; Waltham, MA) following the PCR conditions listed on Table 2.

Table 2. PCR conditions for amplification and tagging

Cycles	Temperature	Time
1	95˚C	4 minutes
16	95˚C	15 seconds
	56˚C	45 seconds
	68˚C	45 seconds
1	68˚C	10 minutes
1	4˚C	∞

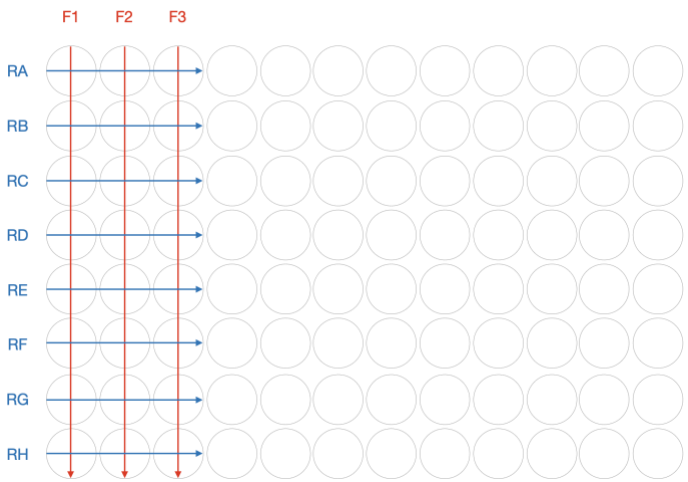

Figure 2. Distribution of master mixes F1, F2, and F3 into the 96-well PCR plate. The red arrows show how the master mixes were distributed on the 96-well plate while the blue arrows show how the reverse primers were distributed.

2.5 Post-amplification clean-up

A fresh 96-well plate was prepared by adding 37.5 µL of AMPure® magnetic beads (Beckman Coulter; Brea, CA) into each well and 15µL of the corresponding amplified samples. The contents of each well were thoroughly mixed until a uniform color was achieved. Incubation at room temperature was done for 5 minutes. The plate was placed onto a magnetic stand for 5 minutes, until the solutions appeared clear. The supernatant was carefully removed, with the plate still on the magnetic stand. A wash was done using 200µL 80% ethanol (Fisher Scientific; Waltham, MA) for each well and incubating the plate for 1 minute at room temperature before removing the ethanol. This step was repeated a second time before the beads were left for 10 minutes to air dry. The plate was then taken off of the magnetic stand and the beads were resuspended in 40µL 10mM Trizma hydrochloride pH8.0 (Tris-HCl) (Sigma-Aldrich; St. Loius, MO). Incubation at room

temperature was done for 5 minutes to release DNA from the beads before placing the plate again onto the magnetic stand. Once the solution in each well appeared clear (after 5 minutes), 30μL of the newly cleaned up product was transferred onto a new 96-well plate. The cleaned-up amplification products were stored at -20˚C.

2.6 Quantification

2.6.1 Sample dilution

The KAPA Library Quantification kit (Kapa Biosystems; Wilmington, MA) uses DNA standards with concentrations between 0.0002 – 20 pM, thus it is necessary to dilute the samples to concentrations within this range. Into the first 6 columns of a 96-well deep well plate, 297μL 10mM Tris-HCl pH8.0 (Sigma-Aldrich; St. Loius, MO) with 0.05% Tween 20 (Millipore Sigma; St. Louis, MO) was added. Using a multi-channel pipette, 3μL of the cleaned-up product was transferred to the corresponding column on the deep-well plate (i.e. samples in column 1 of the cleaned-up product plate were transferred into column 1 of the deep well plate). This step created 1:100 dilutions of the cleaned-up products. The samples were mixed thoroughly using a multi-channel pipette set to 200μL. To create 1:10,000 dilutions, 3μL of the 1:100 dilutions in column 1 were transferred into column 4. This step was repeated for column 2 into column 5 and column 3 into column 6. Figure 3 is a visual representation of this dilution process.

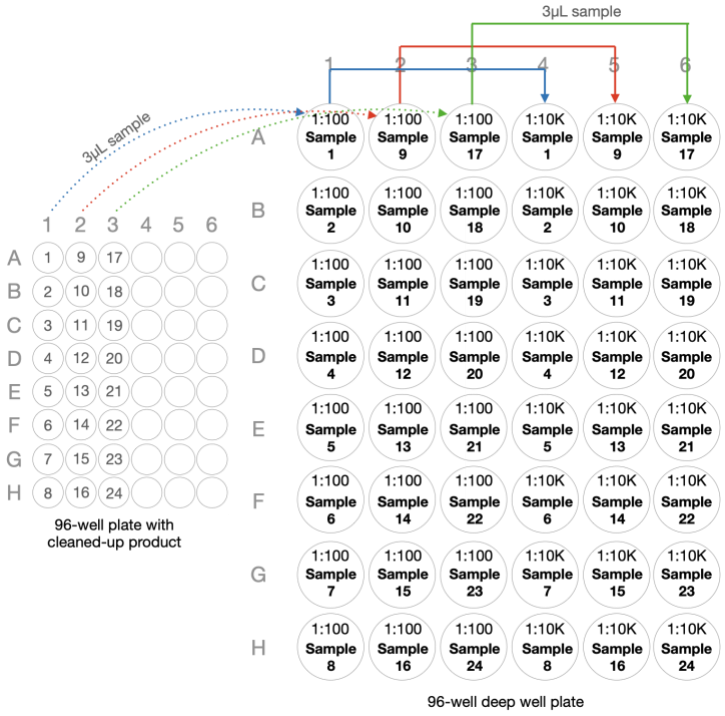

Figure 3. Schematic of the dilution process. Each well in columns 1-6 of the 96-well deep well plate contained 297uL of 10mM Tris-HCl pH8 with 0.05% Tween prior to adding the samples. The dotted arrows indicate that 3µL of the cleaned-up product was transferred into a clean 96-well deep well plate to create 1:100 dilutions of the samples while the solid arrows show the movement of the samples from the first to third columns into the fourth to sixth column of the same deep well plate to complete the 1:10000 dilution of the cleaned-up products. Blue arrows follow the path of Samples 1-8, red arrows follow the path of Samples 9-16, and green arrows follow the path of Samples 17-24.

2.6.2 Plate set-up and qPCR parameters

The KAPA Library Quantification Kit (ROX Low) (Kapa Biosystems; Wilmington, MA) master mix was prepared per manufacturer's instructions [30]. A fresh 96-well optical PCR plate was prepared by dispensing 16uL of the prepared master mix into each well (except wells G1, G2, and G3). Wells G1, G2, and G3 contained 20µL of water.

Following the plate layout illustrated in Figure 4, 4µL of the standards and 1:10,000 dilutions of the samples were added as triplicates into their corresponding wells. The no

template control (NTC) wells contained the master mix and 4µL of water. The library quantification was performed on the Applied Biosystems 7500 Real-Time PCR System (company) following the PCR conditions listed on Table 3.

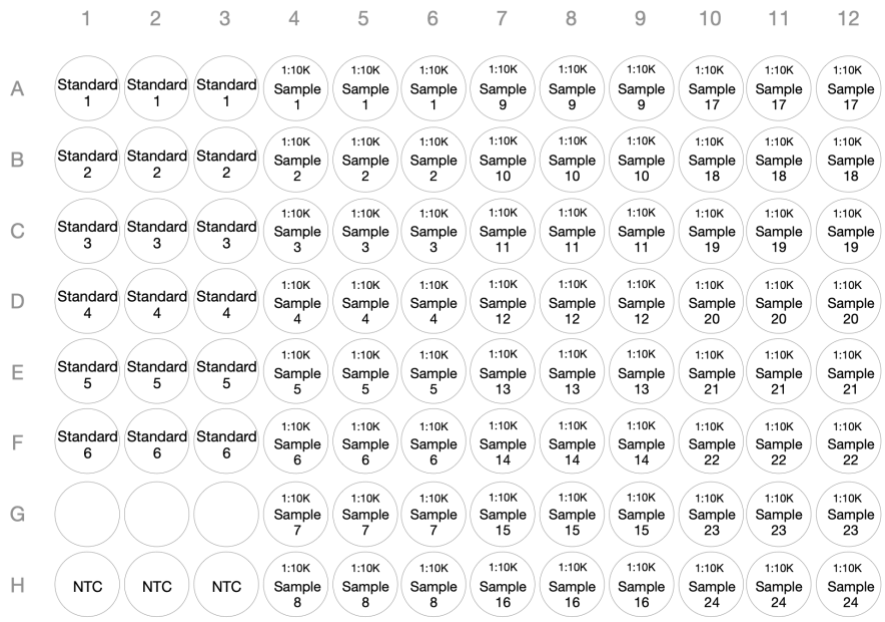

Figure 4. Plate map for quantification of individual samples. Each well (except for wells G1-G3) contains 16µL of the KAPA Library Quantification master mix and 4µL of the appropriate sample (wells A4-H12), standard (wells A1-F3), or water (wells H1-H3). Wells G1-G3 had 20µL of water. All samples, standards, and controls were present in triplicate.

Table 3. Parameters for qPCR

Cycles	Temperature	Time
1	95˚C	5 minutes
30	95˚C	30 seconds
	60˚C	45 seconds

2.7 Normalization and pooling

The data collected from the quantification step was transferred into the HTG Library Calculator to determine an appropriate normalization concentration. The cleaned-up products were transferred onto a clean 96-well plate in 5µL aliquots. These were then diluted with a calculated volume of Tris-HCl pH8.5 (Qiagen; Germantown, MD) that was determined through the HTG Library Calculator. The pooled library was then prepared by taking 8µL of each sample and placing them into a single 2mL tube. The pooled library was stored in -20˚C until sequencing.

2.8 Sequencing

The pooled library was denatured following the steps indicated by the HTG Library Calculator. This process created a 600µL pool with a final concentration of 20pM. The samples were sequenced with MiSeq Custom Primers (HTG; Tucson, AZ) on the MiSeqFGx Sequencing System (Verogen; San Diego, CA) using a 150-cycle sequencing cartridge (Illumina Inc.; San Diego, CA).

2.9 Data Analysis

The sequencing data was analyzed using the Illumina Sequencing Analysis Viewer (SAV). Cluster density, “Q30” score, cluster passing filter, and % alignment were the parameters that determined the success of the sequencing run. The sequencing sample sheet was processed through the HTG Parser to generate a new file that contained the total number of reads of each sample as well as the number of reads of each probe. The HTG

EdgeSeq Reveal software (reveal.htgmolecular.com) (HTG; Tucson, AZ) was used to show information about the sequencing quality, differential expression (DE) among the miRNA probes, and principal component analysis (PCA) plots. RStudio software (RStudio, PBC; Boston, MA) was used to create a 3D PCA plot for better visualization of the data.

3. RESULTS

3.1 Quality of the sequencing run

3.1.1 Illumina Sequencing Analysis Viewer (SAV)

The Illumina SAV is a software that allows the user to visualize their sequencing data and the different metrics that provide information about the sequencing run. Cluster density is a measure of how many clusters formed on the sequencing flow cell. For this sequencing run, the cluster density was 1820K/mm^2. "Q30" represents the probability of an incorrect base being called. Specifically, Q30 is the score where 1:1000 bp would be incorrectly identified. The Q30 score of this experiment was 92.9%, meaning that 92.9% of the bases are called correctly with 99.9% accuracy. Cluster passing filter is a metric that shows the purity of the signal being received from each cluster. The closer this value is to 100, the purer the signal. In an overloaded or underloaded flow cell, this number decreases. The sequencing run for the samples reported here had a score of 84.25%. The % alignment is based on the phiX control that was placed in the pooled library. This number should be equal to the concentration of the phiX in the library. While the expected % alignment was 4%, the observed value was 2.81%.

Illumina lists some target values for the sequencing metrics to determine whether sequencing is successful or not [31, 32]. An ideal cluster density would be 1000-1200K/mm^2 while anything over 90% is considered as a good Q30 score. The expected % aligned value for this sequencing run is 4%, which is higher than the 2.81% that was

observed. Although some metrics were not ideal, the sequencing data was still processed with the HTG Parser and subsequently, HTG Reveal.

3.1.2 HTG Reveal

The HTG Reveal Software has three levels of post-sequencing quality checks (QC). The definition and threshold of each metric were provided by HTG in an internal report (F. Kero, personal communication, November 15, 2021). The HTG metric, QC0, is a measure of RNA quality; if a sample fails at QC0, it is likely degraded. QC1 is a simple threshold for the number of reads from each sample. A sample that was able to generate 500,000 or more reads passes QC1. QC2 is the expression variability of each sample. This is measured in relative standard deviation (RSD) where anything less than 0.08% is considered a failure.

While the Illumina SAV showed that the metrics of the sequencing run were not ideal, all samples passed the QC standards set by HTG thus principal component analysis (PCA) and differential expression (DE) analysis was conducted.

3.2 Principal components analysis (PCA) of all samples

The HTG Reveal software is able to generate a two dimensional (2D) PCA plot (Figure 5) that shows how the samples are correlated to each other. The samples were normalized using counts per million (CPM). The annotations that indicate to the software which replicate samples belong to the same group for analysis, are uploaded as a separate file. Figure 5 shows that the MTC samples grouped together in a separate area from the other biological samples. The peripheral blood only samples (Blood group) showed clear

separation from the samples with vaginal fluid but did not cluster together. This group also appeared to be far from the samples that have venous blood and vaginal fluid combined (Blood_Vag group).

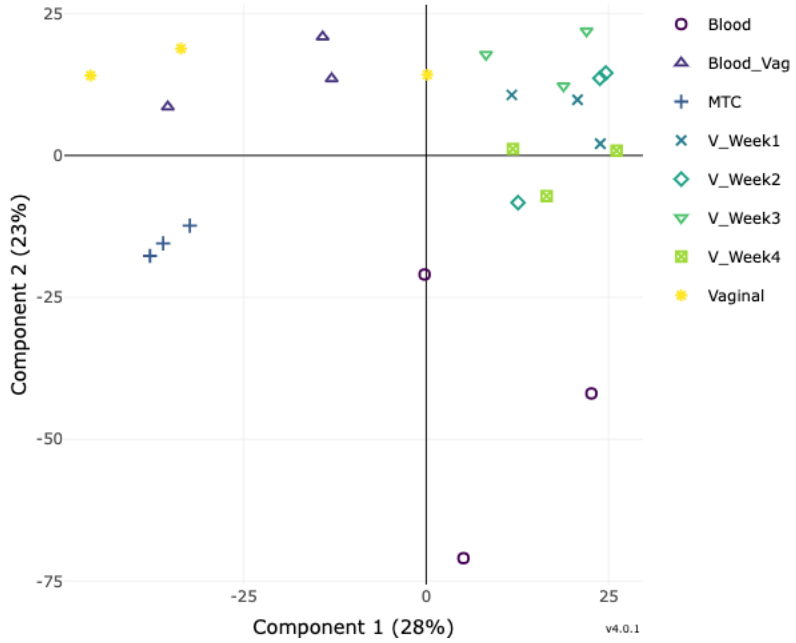

Figure 5. 2D PCA plot of all samples. Each shape on the graph represents a group of samples, with each group being represented in triplicate. Component 1 is the axis that represents the most differences among the samples.

RStudio was used to create a 3D PCA plot (Figure 6) that would show a different perspective of the clustering. The data was normalized before conducting the PCA. The MTC samples are still clearly separated from the rest of the samples, as expected. The venous blood only samples are very distinct from the other biological samples, although they did not closely group together. Notably, the samples that have a mixture of venous blood and vaginal fluid (Blood_Vag group) are clustered far from the venous blood samples (Blood group). A possible reason for this is the samples may have more vaginal

fluid than venous blood. While the amount of venous blood deposited was fixed, the amount of vaginal fluid on the swab was not standardized or quantified in any way.

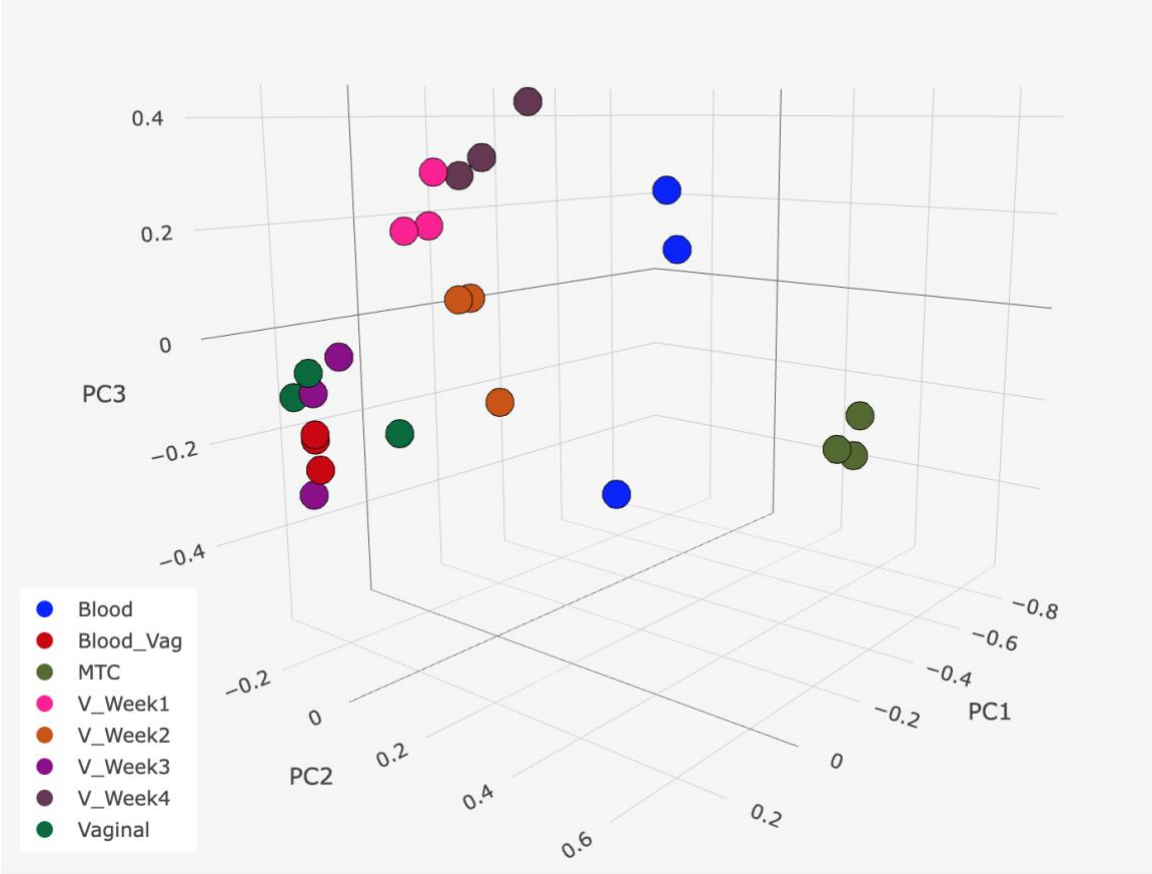

Figure 6. 3D PCA plot of all samples. Each color on the graph represents a group of samples, with each group being represented in triplicate. The most differences are observed along PC1, followed by PC2, and PC3.

3.3 Comparing vaginal fluid samples within one menstrual cycle

One of the objectives of this study is to determine if there are differences in the miRNA signatures obtained from vaginal samples collected from the same donor but at different points in the menstrual cycle. The samples that will be discussed in this section are V_Week1, V_Week2, V_Week3, and V_Week4. These will collectively be called the "menstrual cycle set" for ease of discussion.

3.3.1 PCA

Using the "menstrual cycle set", the HTG Reveal software was able to generate a 2D PCA plot (Figure 7). The separation of each sample type is clearly visible, though one of the V_Week2 replicates appears to be very far from the other two along Component 2.

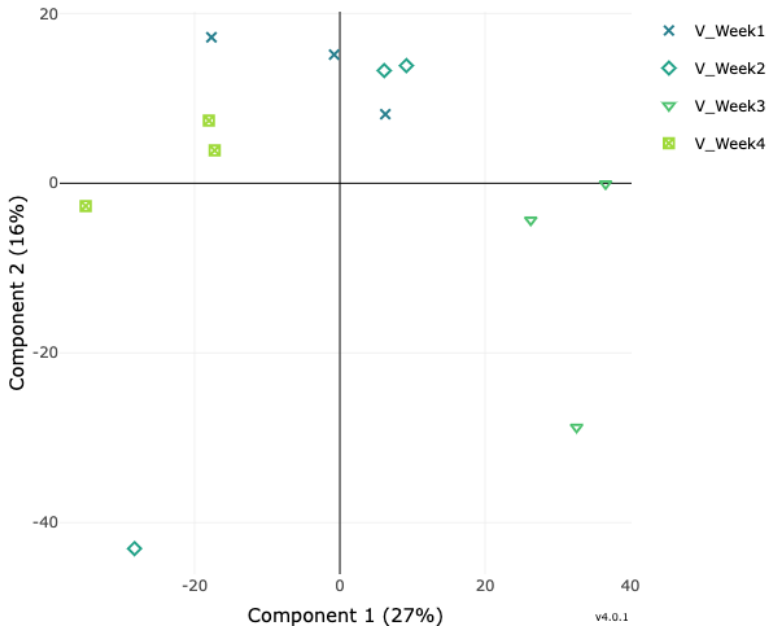

Figure 7. 2D PCA plot of the "menstrual cycle set" of samples. Each shape on the graph represents a group of samples, with each group being represented in triplicate. Component 1 is the axis that represents the most differences among the samples.

The 3D PCA plot of the same data (Figure 8) generated by the RStudio software shows the clustering of the samples more clearly. While one of the V_Week2 replicates is still far from the other two, it is now shown to be separated from any of the other samples. V_Week2 appears to be the most different of the samples without menstrual blood as it is the furthest from V_Week1 and V_Week3 across PC1. One V_Week4 replicate is distant from the other two points but it is clearly away from the other groups. Because both these

samples are separated from their respective groups along PC3, they are still more closely correlated to the other replicates in their group than the other samples.

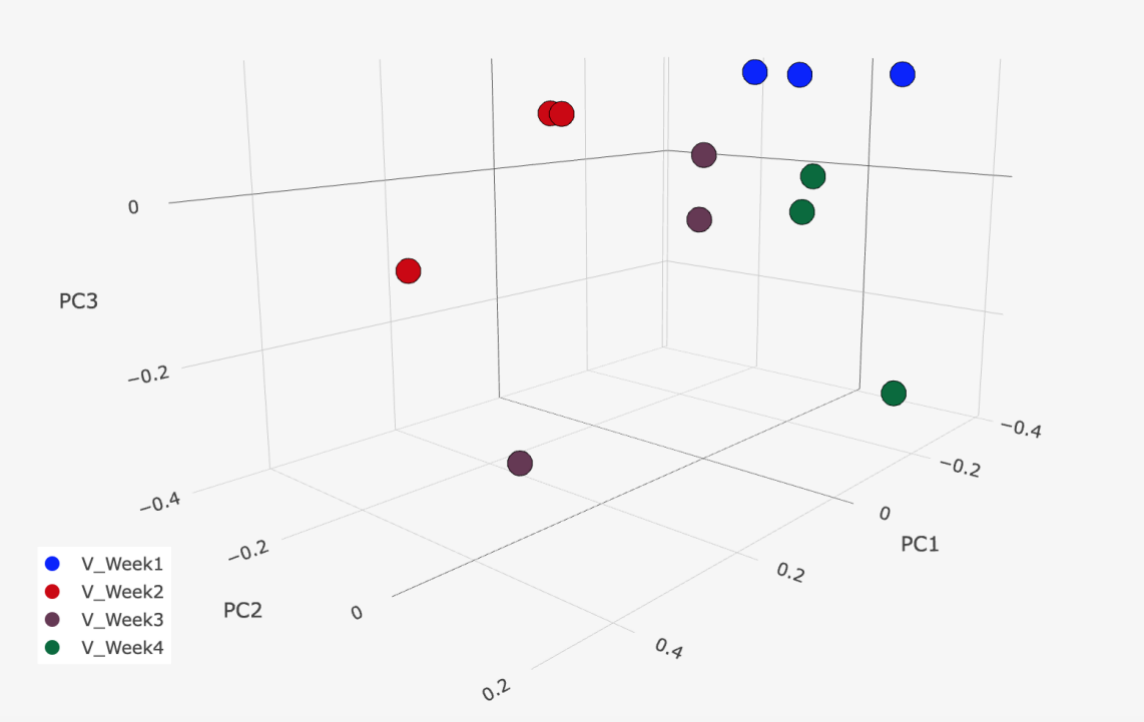

Figure 8. 3D PCA plot of the "menstrual cycle set" of samples. Each color on the graph represents a group of samples, with each group being represented in triplicate. The most differences are observed along PC1, followed by PC2, and PC3.

3.3.2 Differential Expression

The differential expression (DE) of miRNA probes were analyzed using HTG Reveal. A total of 6 pairwise comparisons were done between sample groups in the "menstrual cycle set" to calculate the fold change, "rawP" values, and adjusted values for every miRNA probe through the DESeq2 method. HTG Reveal automatically flags the top 10 upregulated and downregulated probes with the highest "rawP" values. The "rawP" values are unadjusted p-values calculated on the HTG Reveal software prior to the independent filtering done using DESeq2.

Tables 4a-c compile these flagged miRNA probes and show the corresponding fold changes under their respective pairwise comparison. A large magnitude of fold change represents a greater difference in the expression level of the probe between the groups. Positive fold change values indicate that the probe is more abundant in the comparison group while negative fold change values mean that the expression of the probe is higher in the reference group. The fold changes highlighted in yellow are the probes that HTG Reveal identified as part of the top 10 upregulated or downregulated probes in that particular pairwise comparison. Probes highlighted in green and orange have been found in published studies as vaginal and blood markers respectively (see references in the next paragraph). Probes highlighted in blue are potential miRNA markers for vaginal fluid or blood due to the trends observed across their expression levels.

In Tables 4a-c, only miR-124-3p was previously identified as a miRNA marker for vaginal fluid [3]. The miRNA probes that were flagged by HTG Reveal as up or downregulated and have already been shown to be markers for the presence of blood are: let-7b-5p, miR-106a-5p, miR-144-3p, miR-16-5p, miR-185-5p, miR-200b-3p, miR-451a, miR-486-5p, and miR-93-5p [3, 11, 12, 13, 14, 15, 16].

The probes highlighted in blue were chosen based on their observed fold change patterns, despite not being consistently flagged as a top upregulated or downregulated miRNA. These markers were identified as potential markers for vaginal fluid or blood. These can loosely be classified as relating to either Week 3 or Week 4 of the menstrual cycle. MiRNA probes that had a noticeably large fold change during Week 3 of the menstrual cycle were let-7a-5p, let-7b-5p, let-7c-5p, let-7f-5p, let-7i-5p, miR-200c-3p,

miR-22-3p, miR-3182, miR-34a-5p, miR-3607-3p, miR-362-5p, miR-664a-5p, and miR-7705. All of these miRNA seem to be downregulated compared to the other samples in the "menstrual cycle set". Meanwhile, miR-106b, in both -3p and -5p forms, appeared to be consistently upregulated during Week 4 of the menstrual cycle, with the largest fold change observed between Weeks 3 and 4.

Table 4a. **List of differentially expressed probes flagged by HTG Reveal as part of the top 10 upregulated or downregulated miRNAs across 6 pairwise comparisons**. The fold change observed for the designated miRNA is listed under each pairwise comparison. Positive fold change values show higher expression in the comparison group while negative values indicate higher expression in the reference group. Values highlighted in yellow indicate the probe was flagged by HTG Reveal as a top 10 upregulated or downregulated miRNA. Probes highlighted in green/orange are in published studies as miRNA markers for vaginal/ blood respectively. Probes highlighted in blue are potential markers for vaginal fluid or blood.

LEGEND	Flagged by HTG Reveal		Potential marker for vaginal fluid or blood	Literature Hit (vaginal fluid)		Literature Hit (blood/ menstrual blood)
Comparison Group	V_Week1	V_Week1	V_Week1	V_Week2	V_Week2	V_Week3
Reference Group	V_Week2	V_Week3	V_Week4	V_Week3	V_Week4	V_Week4
			Fold change			
HK_SNORA66	-3.25	-1.64	-1.89	2.08	1.73	-1.23
let-7a-5p	-5.24	77.35	-4.14	429.68	1.28	-328.85
let-7b-5p	-5.74	114.96	-2.81	698.22	2.06	-332.56
let-7c-5p	-5.74	138.5	-2.9	833.15	2	-416.77
let-7f-5p	-2.22	75.89	-6.69	179.95	-2.99	-486.68
let-7i-5p	-3.02	34.32	-8.76	107.23	-2.88	-320.75
miR-101-3p	10.18	9.94	-10.24	-1.01	-101.7	-105.87
miR-106a-5p	1.15	15.67	-24.99	15.28	-29.91	-404.04
miR-106b-3p	1.33	3.72	-15.53	6.7	-9.66	-58.59
miR-106b-5p	-1.65	6.1	-26.64	4.76	-35.98	-171.57
miR-124-3p	2.47	20.08	-1.36	8.45	-3.33	-28.9
miR-125b-5p	-2.77	10.16	-1.38	28.72	2.02	-14.88
miR-126-3p	2	1.61	-61.07	-1.23	-115.05	-105.6
miR-1287-5p	3.78	3.53	2.45	-1.02	-1.54	-1.54
miR-144-3p	0	-4.37	-2381.22	-3.55	-1976.07	-639.2
miR-147a	-2.1	-7.87	-1.07	-3.49	1.97	6.85
miR-147b	-2.11	-3.51	-1.5	-1.57	1.42	2.2
miR-16-5p	1.78	1.75	-29.84	1.02	-52.04	-54.45
miR-185-5p	-1.67	2.26	-11.94	4	-7.03	-28.54
miR-193a-5p	-2.55	-3.43	-1.27	-1.25	2.03	2.54
miR-193b-5p	-3.29	-3.64	-1.33	-1.03	2.51	2.58
miR-200b-3p	-1.35	121.04	-3.52	168.95	-2.6	-440.36
miR-200c-3p	-1.28	102.19	-2.77	134.81	-2.16	-291.85
miR-22-3p	-1.84	37.69	-2	71.98	-1.08	-79.18
miR-25-3p	-1.62	6.4	-13.1	10.96	-8.04	-85.97
miR-27a-5p	-2.72	3.89	-1.68	10.99	1.63	-7

Table 4b. List of differentially expressed probes flagged by HTG Reveal as part of the top 10 upregulated or downregulated miRNAs across 6 pairwise comparisons. The fold change observed for the designated miRNA is listed under each pairwise comparison. Positive fold change values show higher expression in the comparison group while negative values indicate higher expression in the reference group. Values highlighted in yellow indicate the probe was flagged by HTG Reveal as a top 10 upregulated or downregulated miRNA. Probes highlighted in green/orange are in published studies as miRNA markers for vaginal/ blood respectively. Probes highlighted in blue are potential markers for vaginal fluid or blood.

LEGEND	Flagged by HTG Reveal	Potential marker for vaginal fluid or blood	Literature Hit (vaginal fluid)	Literature Hit (blood/ menstrual blood)		
Comparison Group	V_Week1	V_Week1	V_Week1	V_Week2	V_Week2	V_Week3
Reference Group	V_Week2	V_Week3	V_Week4	V_Week3	V_Week4	V_Week4
			Fold change			
miR-2861	1.79	-1.22	1.19	-2.06	-1.5	1.37
miR-299-3p	7	48.24	4.16	7.37	-1.66	-12.4
miR-29c-5p	-2.78	-1.25	-1.72	2.34	1.62	-1.46
miR-3064-5p	-3.32	-1.88	-1.28	1.89	2.62	1.38
miR-3125	-1.92	-3.89	1.16	-1.9	2.24	4.23
miR-3175	1.16	12.12	2.8	10.91	2.42	-4.61
miR-3180	1.29	2.02	1.8	1.66	-1.41	-1.2
miR-3182	-1.02	119.91	1.82	126.84	1.86	-69.56
miR-3198	-1.36	-2.45	1.49	-1.69	2.04	3.43
miR-34a-5p	1.03	30.78	-2.04	30.89	-2.1	-65.15
miR-3607-3p	-3.09	25.04	-1.29	79.23	2.4	-34.22
miR-362-5p	-2.99	27.06	-3.05	84.69	-1.01	-83.54
miR-3674	1.6	-2.07	2.18	-3.11	1.37	4.23
miR-3907	-1.06	-3.03	2.14	-2.64	2.3	6.03
miR-3940-5p	1.26	1.06	1.63	-1.12	1.3	1.44
miR-425-3p	-2.43	-1.99	-4.27	1.27	-1.75	-2.26
miR-4257	-1.87	-2.38	-1.44	-1.19	1.31	1.56
miR-4299	-1.53	-1.93	1.61	-1.18	2.48	2.9
miR-4324	-1.76	14.54	1.8	25.98	3.2	-8.55
miR-4444	-2.15	-2.45	-1.5	-1.07	1.45	1.54
miR-4461	-1.18	-4.45	-1.27	-3.52	-1.07	3.29
miR-4483	-1.53	-10.79	1.06	-6.57	1.63	10.68
miR-4506	-2.32	-4.19	-1.04	-1.7	2.25	3.77
miR-451a	1.41	-3.26	-2303.8	-4.44	-3281.97	-736.42
miR-4525	1.31	1.63	2.16	1.32	1.66	1.24
miR-4534	1.34	-5.94	1.59	-7.42	1.19	8.88

Table 4c. List of differentially expressed probes flagged by HTG Reveal as part of the top 10 upregulated or downregulated miRNAs across 6 pairwise comparisons. The fold change observed for the designated miRNA is listed under each pairwise comparison. Positive fold change values show higher expression in the comparison group while negative values indicate higher expression in the reference group. Values highlighted in yellow indicate the probe was flagged by HTG Reveal as a top 10 upregulated or downregulated miRNA. Probes highlighted in green/orange are in published studies as miRNA markers for vaginal/ blood respectively. Probes highlighted in blue are potential markers for vaginal fluid or blood.

LEGEND	Flagged by HTG Reveal		Potential marker for vaginal fluid or blood	Literature Hit (vaginal fluid)		Literature Hit (blood/ menstrual blood)
Comparison Group	V_Week1	V_Week1	V_Week1	V_Week2	V_Week2	V_Week3
Reference Group	V_Week2	V_Week3	V_Week4	V_Week3	V_Week4	V_Week4
			Fold change			
miR-4653-3p	1.31	-1.68	3.41	-2.06	2.62	5.38
miR-4668-5p	-3.63	-7.39	1.48	-1.89	5.41	10.3
miR-4725-3p	-1.58	-1.65	1.35	1.03	2.16	2.08
miR-4736	3.75	4.74	1.2	1.32	-3.08	-4.18
miR-4747-5p	1.09	1.31	2.44	1.26	2.24	1.75
miR-4800-3p	1.02	1.42	2.13	1.48	2.1	1.41
miR-486-5p	-1.48	-1.80	-299.10	-1.15	-199.49	-176.23
miR-548d-5p	1.03	-2.45	1.32	-2.37	1.29	3.02
miR-5739	1.24	1.14	1.76	-1.02	1.42	1.45
miR-574-5p	-2.38	-1.31	1.35	1.95	3.24	1.66
miR-608	-1.71	-1.68	1.98	1.08	3.39	3.13
miR-612	-2.17	-1.6	-1.2	1.44	1.82	1.25
miR-6126	2.53	2.36	1.07	-1.02	-2.36	-2.35
miR-632	-2.49	-2.81	-1.26	-1.06	2	2.1
miR-664a-5p	-3.1	15.66	-1.94	50.77	1.61	-31.8
miR-6741-5p	1.9	12.08	1.14	6.62	-1.66	-11.2
miR-6780b-5p	2.07	1.64	1.61	-1.19	-1.28	-1.08
miR-6784-5p	1.12	-2.36	1.22	-2.48	1.09	2.69
miR-6802-5p	3.28	4.08	1.33	1.3	-2.45	-3.22
miR-6833-5p	2.51	1.58	2.2	-1.48	-1.13	1.3
miR-6869-5p	1.64	-1.82	1.41	-2.81	-1.15	2.41
miR-6877-5p	-1.07	-8.85	1.34	-7.55	1.46	11.08
miR-7151-3p	-1.67	-3.61	1.28	-2.01	2.16	4.31
miR-7154-3p	2.69	1.69	2.53	-1.48	-1.05	1.4
miR-7705	-1.89	33.94	-1.86	67.3	1.02	-66.95
miR-8071	-1.36	-2.87	1.25	-1.97	1.72	3.36
miR-93-5p	-1.21	10.23	-9.17	12.82	-7.59	-95.42

3.5 Comparing peripheral and menstrual blood

The second objective of this study is to determine if miRNA signatures can be used to differentiate peripheral blood from menstrual blood. The samples analyzed in this section are Blood, V_Week 4 which serves as the menstrual blood sample, and Blood_Vag that is the combination of vaginal fluid and peripheral blood. These samples will collectively be referred to as the "blood comparison set" for ease of discussion.

3.4.1 PCA

All samples that contained peripheral or menstrual blood were analyzed. HTG Reveal generated a 2D PCA plot (Figure 9). From this plot alone, there is a clear separation among all three groups. The replicates in the Blood group, although in a separate area from the other samples, did not cluster. This might reflect the fact that biological replicates were used (i.e. the replicates come from different donors) compared to the other two groups that are represented by technical replicates.

Interestingly, the peripheral blood only (Blood) group and peripheral blood with vaginal fluid (Blood_Vag) appear to have the most differences as they are on opposite sides of Component 1. Looking at the samples from a 3D perspective (Figure 10), this is not quite correct as it appears that the Blood_Vag group is closer than V_Week4 to the Blood group along PC1. The Blood_Vag group and the menstrual blood group (V_Week4) are still closer to each other along PC1. This clustering may arise from having more vaginal fluid present in these samples than blood.

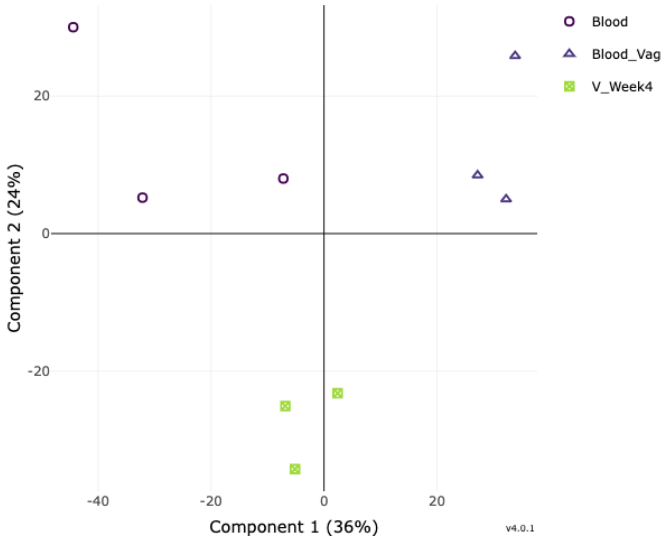

Figure 9. 2D PCA plot of samples in the "blood comparison set". Each shape on the graph represents a group of samples, with each group being represented in triplicate. Component 1 is the axis that represents the most differences among the samples.

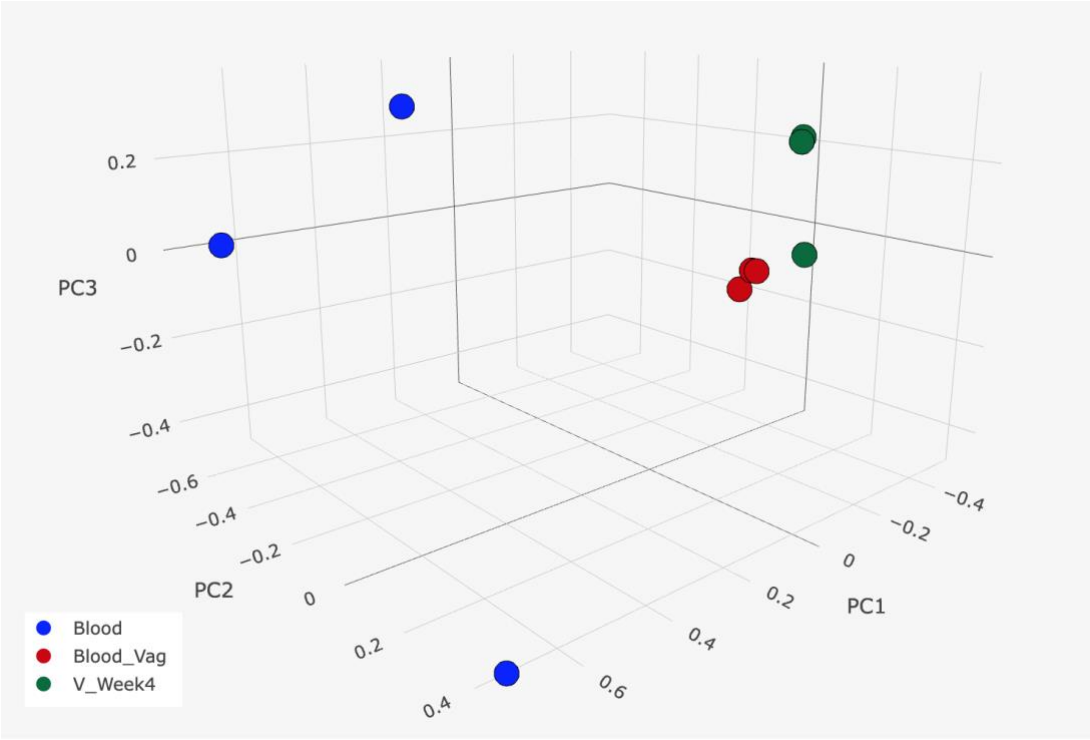

Figure 10. 3D PCA plot of samples in the "blood comparison set". Each color on the graph represents a group of samples, with each group being represented in triplicate. The most differences are observed along PC1, followed by PC2, and PC3.

3.4.2 Differential Expression

The differential expression (DE) of miRNA probes were analyzed using HTG Reveal. Three pairwise comparisons were done between the groups in the "blood comparison set" to calculate the fold change, "rawP" values, and adjusted values for every miRNA probe using the DESeq2 method. HTG Reveal automatically flags the top 10 upregulated and downregulated probes with the highest "rawP" values.

Tables 5a and 5b enumerate the miRNAs that HTG Reveal flagged in the pairwise comparisons. The fold changes observed for these probes are listed under their respective comparisons. Larger fold change values are indicative of bigger differences in the expression levels of the probes in the groups being compared. If the fold change is positive, the expression of the probe is higher in the comparison group than in the reference group while negative fold change values mean that the expression of the probe is higher in the reference group. The fold changes highlighted in yellow belong to the probes that HTG Reveal identified as part of the top 10 upregulated or downregulated probes in that particular pairwise comparison. Probes highlighted in orange have been found in published studies as blood markers while those highlighted in blue are potential miRNA markers for vaginal fluid or blood due to the trends observed across their expression levels.

Five of the probes in Tables 5a and 5b have appeared in existing literature: let-7b-5p, miR-185-5p, miR-200b-3p, miR-486-5p, and miR-93-5p [11, 12, 13, 16]. All of these probes also appear in the set of miRNA probes flagged in the pairwise comparisons between the groups in the "menstrual cycle set".

Of the probes that are of interest in this study, four appear to be upregulated in the menstrual blood samples and not in the samples with peripheral blood. These are miR-193b-3p, miR-200b-3p, miR-200c-3p, and miR-34a-5p. All four miRNA probes have low fold changes in the pairwise comparison between the Blood_Vag and Blood groups, at 1.26, -1.93, -5.75, and 2.42 respectively. This suggests that the DE observed does not come from vaginal fluid only but possibly from menstrual blood as a whole. Both miR-200c-3p and miR-34a-5p have been previously identified as markers of interest in Tables 4a and 4b because of their large fold change magnitudes observed in comparisons with samples from Week 3 of the menstrual cycle. Their expression is downregulated in the V_Week3 samples unlike in the V_Week4 samples which show these probes to be upregulated. Meanwhile, miR-200b-3p has been reported as a miRNA marker for peripheral blood [15].

MiR-25-3p is identified as a potential marker in Table 5a. The fold changes observed in all three pairwise comparisons were high, although the value in the Blood_Vag and Blood comparison is noticeably larger at -2078.42. This miRNA is also seen in Table 4a, with the biggest fold change at -85.97 under the pairwise comparison between V_Week3 and V_Week4. Although the fold changes observed between the other comparisons with menstrual blood in Table 4a do not have the same magnitude as those in Table 5a, miR-25-3p may be considered as an miRNA marker for blood.

The remaining miRNA probes of interest (miR-128-1-5p, miR-147b, miR-193b-5p, miR-25-3p, miR-5585-3p, miR-612) have large fold changes in the pairwise comparisons between V_Week4 and Blood groups as well as the Blood_Vag and Blood groups. Considering that these pairs involve both vaginal fluid and peripheral blood, the

miRNA probes of interest are thought to be possible markers for either body fluid based on Tables 5a and 5b alone. The probes in Tables 4a-c were assessed to see if they were already identified as vaginal fluid or blood markers. MiR-147b, miR-193b-5p, and miR-612 were previously identified by HTG Reveal as top DE probes in Tables 4a and 4c. Despite being flagged, the highest fold change observed for these probes across the 6 pairwise comparisons among the "menstrual cycle set" are only -3.51, -3.64, and -2.17 respectively. Meanwhile, both miR-128-1-5p and miR-5585-3p were not considered top DE probes at all in the "menstrual cycle set" pairwise comparisons. Considering the low fold changes and that all the samples in the "menstrual cycle set" have vaginal fluid, these probes of interest are likely to be more informative as miRNA markers for vaginal fluid than peripheral blood.

Table 5a. List of differentially expressed probes flagged by HTG Reveal as part of the top 10 upregulated or downregulated miRNAs across the pairwise comparisons in the "blood comparison set". The fold change observed for the designated miRNA is listed under each pairwise comparison. Positive fold change values show higher expression in the comparison group while negative values indicate higher expression in the reference group. Values highlighted in yellow indicate the probe was flagged by HTG Reveal as a top 10 upregulated or downregulated miRNA. Probes highlighted in orange are in published studies as miRNA markers for blood. Probes highlighted in blue are potential markers for vaginal fluid or blood.

LEGEND	Flagged by HTG Reveal	Potential marker for blood or vaginal fluid	Literature Hit (blood/ menstrual blood)

Comparison Group	V_Week4	Blood_Vag	Blood_Vag
Reference Group	Blood	Blood	V_Week4
		Fold change	
HK_SNORA66	1.13	-139.22	-198.83
let-7a-5p	-1.76	-905.52	-714.14
let-7b-5p	-3.38	-726.35	-299.02
let-7i-5p	-5.61	-1469.97	-372.71
miR-100-5p	12.76	-18.16	-305.79
miR-1255b-2-3p	5.45	56.2	8.08
miR-125b-5p	13.79	-8.72	-156.4
miR-128-1-5p	54.11	68.51	-1.01
miR-1285-5p	14.56	50.32	2.73
miR-1299	-1.58	4	5.14
miR-1306-5p	-5.47	4.29	18.79
miR-141-3p	33.19	-3.89	-166.31
miR-147a	7.04	34.46	3.85
miR-147b	23.47	36.98	1.23
miR-185-5p	-70	-240	-4.64
miR-193b-3p	77.31	1.26	-79.66
miR-193b-5p	38.29	52.82	1.1
miR-1976	-88.53	-22.49	3.01
miR-200b-3p	343.43	-1.93	-855.45
miR-200c-3p	107.17	-5.75	-796.26
miR-22-3p	-8.9	-566.46	-84.8
miR-222-3p	2.02	7.82	-62.62
miR-25-3p	-55.01	-2078.42	-50.52
miR-296-5p	-69.39	-28.41	1.95

Table 5b. List of differentially expressed probes flagged by HTG Reveal as part of the top 10 upregulated or downregulated miRNAs across the pairwise comparisons in the "blood comparison set". The fold change observed for the designated miRNA is listed under each pairwise comparison. Positive fold change values show higher expression in the comparison group while negative values indicate higher expression in the reference group. Values highlighted in yellow indicate the probe was flagged by HTG Reveal as a top 10 upregulated or downregulated miRNA. Probes highlighted in orange are in published studies as miRNA markers for blood. Probes highlighted in blue are potential markers for vaginal fluid or blood.

LEGEND	Flagged by HTG Reveal	Potential marker for blood or vaginal fluid	Literature Hit (blood/ menstrual blood)

Comparison Group	V_Week4	Blood_Vag	Blood_Vag
Reference Group	Blood	Blood	V_Week4
		Fold change	
miR-34a-5p	99.31	2.42	-54.41
miR-362-5p	-6.66	-260.85	-50.89
miR-3674	3.34	43	9.89
miR-375	63.51	3.34	-24.08
miR-4306	-50.42	-72.85	-1.94
miR-4534	2.98	37.85	10.05
miR-4732-3p	-115.44	-53.16	1.67
miR-486-5p	-28.12	-321.94	-14.97
miR-492	128.8	5.56	-29.52
miR-500b-5p	-10.1	-344.62	-43.84
miR-501-5p	-62.42	-81.13	-1.68
miR-548ay-5p	2.56	22.32	6.82
miR-548d-5p	3.83	36.94	7.56
miR-5585-3p	18.93	18.25	-1.33
miR-612	25.31	29.87	-1.06
miR-6741-5p	2.82	-12.76	-45.58
miR-6769b-3p	-10.56	-1.2	7.04
miR-6787-5p	2.96	14.45	3.91
miR-6892-3p	1.46	7.24	3.97
miR-7-5p	-7.63	-326.56	-55.01
miR-7976	-69.49	-36.18	1.47
miR-92a-3p	-104.41	-3394.21	-43.45
miR-92b-3p	-110.14	-2367.25	-29.91
miR-93-5p	-13.68	-415.73	-39.72
miR-99a-5p	17.18	-14.6	-338.13

4. DISCUSSION

MicroRNAs (miRNAs) are short non-coding sequences, about 18-24 bp long, that exhibit body tissue and fluid specificity [22]. The goal of this study is to identify differentially expressed (DE) miRNA markers that could indicate the presence of vaginal fluid and determine at which point in the menstrual cycle the sample was collected. The second objective is to find DE miRNA markers that can differentiate menstrual blood from peripheral blood.

4.1 Correlation among all samples by principal component analysis (PCA)

A total of 24 samples were analyzed in this study, 21 of which were experimental samples and 3 were controls. Figure 6 plots each sample in a single three-dimensional (3D) principal component analysis (PCA) graph to show any correlations to each other. The multi-tissue control (MTC) samples separate from the experimental samples while keeping the replicates clustered together. The MTC is a product proprietary to HTG Molecular Diagnostics Inc.; no information about its contents was provided. Based on information from HTG, the MTC performed as expected (HTG, personal communication, 2022). In comparison, the Blood group also shows separation from the remaining samples but the replicates are further apart from each other on the graph than the MTC replicates. This observation might reflect the use of replicates being from peripheral blood samples provided by different donors. However, the Vaginal group of samples did not show the same drastic separation between replicates despite coming from different donors as well. This may be influenced by the sheer number of samples with vaginal fluid in the PCA, i.e.

all samples except the MTC and Blood groups contain vaginal fluid. This may also explain the similar positioning among the replicates within the remaining groups on the graph. The Blood_Vag group, which represents a mix of peripheral blood and vaginal fluid, shows more similarity to the samples with vaginal fluid than peripheral blood. The swabs that were created for the Blood_Vag group had a fixed amount of liquid blood deposited onto swabs with an unknown amount of vaginal fluid. It may be possible that there was significantly more vaginal fluid on the swabs, which may have masked some of the characteristics of the peripheral blood.

The samples used to create Figure 8 all have vaginal fluid. Plotting the "menstrual cycle set" of samples eliminates the variance from the donor, so their differences as vaginal fluid samples become more apparent. Although each group of samples are separate from each other, the replicates are much more spread out than what was observed in Figure 6. Figure 8 also shows a shorter range for all three axes which may explain the larger separation observed among replicates.

Figure 10 is the 3D PCA plot for the blood comparison set of samples. Like the menstrual cycle set, there is a clear separation among the groups. However, the extreme differences among the blood replicates may have influenced the clustering of the other samples within their respective groups. Both Blood_Vag and V_Week4 show two replicates almost on top of each other. While this was also seen in Figure 6 for Blood_Vag, V_Week4 did not show the same overlap in Figures 6 and 8.

PCA is a method that organizes highly variable data in a reduced dimensionality format [33]. The information gathered from this method will always depend on what

samples are used in the analysis, thus making the comparisons relative rather than absolute. To identify the kind of body fluid in evidence by PCA alone would require standards that establish the differences among all possible types of body fluids. Considering how variable body fluids are, even within the same type and donor, this would not be practical. Thus, using PCA alone for body fluid identification would be difficult to standardize and validate for routine case work.

The information gleaned from the PCA plots, however may provide context to support the DE analysis that would lead to the identification of the body fluid. As previously mentioned, Figure 6 shows that the V_Week4 and Blood_Vag groups exhibit profiles that are more similar to vaginal fluid than peripheral blood. In the succeeding sections, this observation, along with the DE, is used to assess the probative values of miRNA markers for either peripheral blood or vaginal fluid.

4.2 Identification of possible miRNA markers for vaginal fluid at a specific time point in the menstrual cycle

This section analyzes four groups of vaginal fluid samples taken within one menstrual cycle of a single volunteer in order to identify possible miRNA markers. This is collectively referred as the “menstrual cycle set”. HTG Reveal was used to determine the top upregulated and downregulated genes according to their “rawP” values in 6 pairwise comparisons. The fold changes between the expression of miRNA probes were analyzed for possible trends. These can be seen in Tables 4a-c. MiRNA probes were then tagged as of interest in this study or have been previously reported in other studies.

MiR-124-3p is the only miRNA probe flagged by HTG Reveal that is also in existing literature; it was observed as a vaginal fluid marker [3, 16]. The fold changes with the greatest magnitudes are observed in the pairwise comparisons between the V_Week1 and V_Week3 groups and the V_Week3 and V_Week4 groups. Although its fold change in the remaining comparisons is relatively lower, there is no observable trend that could make miR-124-3p a reliable marker to differentiate the weeks within the menstrual cycle.

There are several probes that are of interest in this study based on the trends observed in their fold changes. The miRNAs belonging to the let family (let-7a-5p, let-7b-5p, let-7c-5p, let-7f-5p, and let-7i-5p) all show downregulation in the V_Week3 samples. Except for let-7i-5p, these probes were not flagged in the comparison between the samples from Week 2 and Week 3 of the menstrual cycle despite having a very large fold change. Upon checking the data generated by HTG Reveal, the analyses of these probes did not result in "rawP" values, hence they could not be called as top upregulated genes. The occurrence is explained in the HTG Reveal software. DESeq2, the chosen method of analysis, performs independent filtering that leads to the lack of "rawP" values for some of the miRNA probes.

Let-7 is a family of miRNAs that is present in many species, thus thought to be conserved [12, 22]. The differences among the members of the let-7 family lie in a few base pairs in their sequences [34]. In the context of body fluid identification, only let-7g and let let-7i have been reported as markers for normalization [15]. Rekker et al. (2013) lists let-7b-5p as one of the miRNA present in abundance in the plasma and blood of healthy women, not vaginal fluid [16].

A study by Scalici et al. (2016) reports that let-7b was found to be downregulated in women with polycystic ovary syndrome (PCOS) [35]. No health information about the donor is available and the expression of the let family miRNAs is not consistently downregulated across all the vaginal fluid samples. Thus, this study cannot confirm if the DE observed is consistent with the findings of Scalici et al. (2016).

The remaining probes that were highlighted as interesting in Tables 4a-c (miR-200c-3p, miR-22-3p, miR-3182, miR-34a-5p, miR-3607-3p, miR-362-5p, miR-664a-5p, and miR-7705) show significant downregulation the V_Week3 samples. These are all flagged as a top downregulated gene in the comparison between Weeks 2 and 3, with miR-200c-3p and miR-22-3p being flagged in all pairwise comparisons involving the V_Week3 samples. The large magnitude of the fold changes in these probes suggest that Week 3 of the menstrual cycle shows the most differences in miRNA expression of vaginal fluid. This is reinforced by the comparison between V_Week1 and V_Week2, where the highest fold change observed in the probes listed in Tables 4a-c is only 10.18 for miR-101-3p. Large fold changes were also observed in the V_Week4 comparisons, however the differences may arise from the menstrual blood in the samples and not from just the vaginal fluid.

The probes that are of interest in Tables 4a-c show possible probative value in relation to Week 3 of the menstrual cycle, although it is difficult to conclude whether the miRNA profiles can be used to determine the time in the menstrual cycle at which a vaginal sample was collected. Rekker et al. (2013) report that there were no significant differences in miRNA throughout the menstrual cycle, however their study made use of plasma and not vaginal fluid [16]. This study observes a similar pattern with Weeks 1 and 2 where the fold changes appear to only go as high as 10.18. It may be possible to differentiate a vaginal sample from Week 3 of the menstrual cycle from Weeks 1 and 2 using the miRNAs identified, however they may not be reliable when trying to analyze more than one donor. Further research with a larger sample set would be needed to confirm the reliability of these markers. Donor information, e.g. age and health, may strengthen the analysis as well considering that they affect miRNA expression.

4.3 Identification of possible miRNA markers for differentiating peripheral and menstrual blood

This section aims to identify miRNA markers that could be used to differentiate peripheral blood from menstrual blood by analyzing the "blood comparison set" of samples. The trends in the fold changes observed in Tables 4a-c, as well as Tables 5a and 5b are assessed in order to determine the probes that are of interest in this study. Probes that have been identified in existing literature are noted as well.

MiR-106b is consistently downregulated in the V_Week4 samples (Table 4a), suggesting some probative value in miR-106b as a menstrual blood marker. However, it is

not flagged in the pairwise comparisons within the “blood comparison set” so it may not be able to distinguish menstrual blood from peripheral blood. While miR-106a has been observed in body fluid identification literature as a marker for blood, miR-106b has been implicated as possibly being associated with breast cancer [36, 37].

Four miRNA probes that are of interest in Tables 5a and 5b are thought to be possible menstrual blood markers based on the large fold changes observed in the pairwise comparisons between the V_Week4 and Blood and the V_Week4 and Blood_Vag groups. This means that the DE observed could not come from the presence of vaginal fluid alone. These four probes are miR-193b-3p, miR-200b-3p, miR-200c-3p, and miR-34a-5p.

Seashols-Williams et al. (2016) report miR-200b-3p as a peripheral blood marker, however the data in this study suggests that miR-200b-3p might be useful as a marker for menstrual blood [15]. MiR-200b-3p appears to be upregulated in the menstrual blood samples in Tables 4a and 5a while having a small fold change of -1.93 in the comparison between the Blood_Vag and Blood groups.

In contrast, both miR-200c-3p and miR-34a-5p were previously identified as interesting miRNA probes relating to their downregulation in the V_Week3 group. The small magnitude of the fold changes in the other pairwise comparisons involving V_Week4 in Tables 4a and 4b undermine their potential as miRNA markers for menstrual blood. Similarly, miR-193b-3p appears to be a possible miRNA marker for menstrual blood, but

its absence in Tables 4a-c does not make it a reliable candidate. These probes have been seen in existing works for their roles in cancer [38, 39, 40].

There are two miRNAs that are of interest in this study that have been listed in other studies as candidates for menstrual blood markers [11, 12, 13]. Mir-185-5p was flagged by HTG Reveal as a top DE probe in the pairwise comparison between V_Week1 and V_Week4 (Table 4a). Although it is not flagged for the other two comparisons involving menstrual blood in the "menstrual cycle set", miR-185-5p shows consistent downregulation in the V_Week4 samples in Table 4a. However, in Table 5a, the biggest fold change is observed in the Blood_Vag and Blood comparison, which weakens its position as a marker for menstrual blood. This is further seen in the comparison between Blood_Vag and V_Week4 where the fold change is -4.64, showing not only that there is little difference in the expression of miR-185-5p, it is higher in the menstrual blood samples. Given these, perhaps miR-185-5p might have some probative value in showing that a mixture is present rather than having solely vaginal fluid or blood.

MiR-144-3p is shown to be upregulated in Table 4a. However, it is not considered as a top DE probe among the comparisons made with the "blood comparison set" of samples. This suggests that miR-144-3p can be used as a marker for blood without differentiating menstrual blood from peripheral blood.

While there are observable trends that indicate that the probes listed in this section could be used to identify menstrual blood, they do not seem to be reliable. MiR-106b and miR-144-3p appear to be able to at least confirm the presence of blood but are unable to differentiate peripheral from menstrual. Meanwhile miR-193b-3p, miR-200c-3p, and miR-

34a-5p show weak evidence as usable miRNA markers for menstrual blood. This study makes use of only one set of menstrual blood samples so a larger sample size may further the potential of these probes as markers for menstrual blood.

4.4 Identification of miRNA markers for blood and vaginal fluid in general

There are other miRNA markers in Tables 4a-c and Tables 5a and 5b that have been reported in literature for their ability to identify the presence of blood. MiR-451a was identified as a venous blood marker [3, 12]. This miRNA was only flagged as a top DE probe in Table 4b where its underexpression is observed in the menstrual blood samples. However, it does not appear in Tables 5a or 5b, suggesting that it may not be able to distinguish peripheral blood from menstrual blood. MiR-106a-5p and miR-185-5p have also been reported as miRNA markers for venous blood [11, 12]. This study reinforces their reliability by the observed upregulation of these probes in the V_Week4 group compared to the other samples in the "menstrual cycle set". Table 5a shows that miR-185-5p is differentially expressed in the Blood_Vag and Blood pairwise comparison. Considering Figure 6, which shows that the Blood_Vag group displays more characteristics from vaginal fluid than from peripheral blood, it is possible that the DE observed is due to the vaginal fluid present. Meanwhile, miR-106a-5p is not flagged by HTG Reveal as a top DE probe in the comparisons within the "blood comparison set", suggesting that it may indicate the presence of blood without discriminating between menstrual and peripheral blood.

Rekker et al. (2013) report miR-93-5p and miR-486-5p to be highly abundant in the plasma of healthy females [16]. The data in this study support their findings as Table 4c shows both probes to be upregulated in the V_Week4 samples. Further, Table 5b shows that miR-93-5p and miR-486-5p are both more abundant in peripheral blood alone than in the other two groups.

MiR-25-3p appears in both Table 4a and Table 5a as a top DE probe. Table 4a shows a large fold change value of –85.97 between the Week 3 and Week 4 groups, although the other pairwise comparisons that use V_Week4 are relatively smaller. In Table 5a, the fold changes for miR-25-3p all have a large magnitude, however the biggest difference is seen between the Blood_Vag and Blood groups. Considering that the Blood_Vag group mimics the characteristics of vaginal fluid more than peripheral blood, miR-25-3p may be a blood marker that is unable to differentiate between peripheral and menstrual blood. Literature search shows that upregulation of miR-25-3p is implicated in esophageal cancer [41].

There are some probes that were identified in Tables 5a and 5b as interesting because these probes have a large fold change in the Blood_Vag and Blood and V_Week4 and Blood pairwise comparisons. Probes that fit this criteria could either be markers for peripheral blood or vaginal fluid. Figure 6 and Tables 4a-c were assessed in order to determine whether these miRNA markers would correspond to blood or vaginal fluid. Figure 6 shows that the Blood_Vag and V_Week4 groups are more similar to vaginal fluid than peripheral blood. Further, all samples used in the comparisons between the “menstrual cycle set” of samples in Tables 4a-c have vaginal fluid. Taking these into account, it is

more likely that these probes may be miRNA markers for vaginal fluid in general. The miRNA probes that fit this criteria are miR-128-1-5, mir-147b, miR-193b-5p, miR-5585-3p, and miR-612.

MiR-147b, miR-193b-5p, and miR-612 also have entries in Tables 4a and 4c but the maximum fold change recorded for these probes are -3.51, -3.64, and -2.17 respectively. MiR-128-1-5p and miR-5585-3p do not appear in Tables 4a-c at all. These strengthen their use as vaginal fluid markers however they might not be useful if a temporal context is required.

5. CONCLUSION

5.1 Summary and Conclusion

Body fluid identification is a crucial process in forensic science as it determines if and how biological evidence would be analyzed. The most common approach makes use of chemical tests, however very few have been designed to confirm the presence of body fluids. The introduction of next generation sequencing (NGS) in forensic science has bridged traditional procedures to nucleic acid based methods for identification. Particularly, tissue and fluid specific miRNAs have been studied as a way of determining the nature of biological evidence [3, 11, 12, 13, 14, 15, 16].

In this study, the HTG EdgeSeq® workflow was employed to analyze vaginal fluid samples from a single menstrual cycle, peripheral blood samples, and a mixture of peripheral blood and vaginal fluid. Pairwise analysis was performed using the HTG Reveal software and the top differentially expressed (DE) probes according to the "rawP" values were identified. The top DE probes that had observable trends in their fold changes were tagged as of interest in this study. Probes that have been reported in existing literature were identified as well.

The data collected in this study suggest that Week 3 shows the most differences among the other vaginal fluid samples in the menstrual cycle. Although many of the DE probes show large fold changes in the comparisons involving Week 3 samples, minimal differences were observed when comparing Week 1 and Week 2 samples. Thus, these DE probes cannot reliably differentiate all weeks of the menstrual cycle. A total of four miRNA probes were identified as possible markers for menstrual blood but only miR-200b-3p has

strong evidence in this study to support its candidacy. MiR-144-3p and miR-185-5p have been reported as menstrual blood miRNA markers but the data in this study suggests that they are unable to distinguish between peripheral and menstrual blood. Other miRNA markers in this study that may be used as miRNA markers for blood in general include miR-106a-5p, miR-144-3p, miR-16-5p, miR-25-3p, miR-451a, miR-486-5p, and miR-93-5p. Possible miRNA markers for vaginal fluid in general were also determined: miR-124a, miR-128-1-5, mir-147b, miR-193b-5p, miR-5585-3p, and miR-612.

5.2 Future Directions

The information gathered from this study can be further explored by introducing some criteria when selecting donors. Additional health information on the donors, such as hormonal intake and age, may provide more context to the observations made in the study. Increasing the number of donors may also lead to more information on the reliability of the miRNA markers.

It would be interesting to see if the expression profile of a vaginal swab taken from an unknown time during the menstrual cycle shows consistency with the expression profiles obtained from this study. This would reinforce the trends observed and strengthen the potential of the miRNA marker candidates that were identified.

Hanson, Lubenow, and Ballentyne (2009) have suggested using a panel of markers for identification as they had difficulty finding a truly specific marker per biofluid. Given that there were very few DE probes seen in this study, creating an accurate panel may be difficult but is worth exploring.

BIBLIOGRAPHY

1. Virkler, K. & Lednev, I. (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

2. Cox, M. (1991). A Study of the Sensitivity and Specificity of Four Presumptive Tests for Blood. *Journal of Forensic Sciences, 36*(5), 1503-1511. https://projects.nfstc.org/workshops/resources/articles/A%20Study%20of%20the%20Sensitivity%20and%20Specificity%20of%20Four.pdf

3. Hanson, E., Lubenow, H., & Ballentyne, J. (2009). Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. *Analytical Biochemistry, 387(2),* 303-314. https://doi.org/10.1016/j.ab.2009.01.037

4. Eschenbach, D. et al. (2000). Influence of the Normal Menstrual Cycle on Vaginal Tissue, Discharge, and Microflora. *Clinical Infectious Diseases, 30 (6),* 901-907. https://doi.org/10.1086/313818

5. MacNeill, C. et al. (2012). Cyclic changes in the level of the innate immune molecule, surfactant protein-a, and cytokines in vaginal fluid. *American Journal of Reproductive Immunology, 68(3),* 244-250. https://doi.org/ 10.1111/j.1600-0897.2012.01155.x

6. Macgregor, K. et al. (2022). Changes in adipose tissue microRNA expression across the menstrual cycle in regularly menstruating females: a pilot study. *Physiological Genomics, 54(1),* 1-10. https://doi.org/10.1152/physiolgenomics.00088.2021

7. Jakubowska, J., Maciejewska, A., Pawłowski, R., & Bielawski, K. (2013). MRNA profiling for vaginal fluid and menstrual blood identification. *Forensic Science International, 7(2),* 272-278. https://doi.org/10.1016/j.fsigen.2012.11.005

8. Chappell, C. et al. (2014). The effects of reproductive hormones on the physical properties of cervicovaginal fluid. *American Journal of Obstetrics and Gynecology, 211(3),* 226.e1-226.e7. https://doi.org/10.1016/j.ajog.2014.03.041

9. Giampaoli, S., et al. (2014). Forensic interlaboratory evaluation of the ForFLUID kit for vaginal fluids identification. *Journal of Forensic and Legal Medicine, 21*, 60–63. https://doi.org/10.1016/j.jflm.2013.10.016

10. Benschop, C., Quaak, F., Boon, M., Sijen, T., & Kuiper, I. (2012). Vaginal microbial flora analysis by next generation sequencing and microarrays; can microbes indicate vaginal origin in a forensic context?. *International Journal of Legal Medicine, 126(2),*

303-310. https://link.springer.com/article/10.1007/s00414-011-0660-8

11. Zubakov, D. et al. (2010). MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation. *International Journal of Legal Medicine, 124(3),* 217-226. https://doi.org/10.1007/s00414-009-0402-3

12. Silva, S., Lopes, C., Teixeira, A., Sousa, M., & Medeiros, R. (2015). Forensic miRNA: Potential biomarker for body fluids?. *Forensic Science International: Genetics, 14,* 1-10. https://doi.org/10.1016/j.fsigen.2014.09.002

13. Sirker, M., Fimmers, R., Schneider, P., & Gomes, I. (2017). Evaluating the forensic application of 19 target microRNAs as biomarkers in body fluid and tissue identification. *Forensic Science International: Genetics, 27,* 41-49. https://doi.org/10.1016/j.fsigen.2016.11.012

14. Bexon, K. & Williams, G. (2015). Characterising the fluctuation of microRNA expression throughout a full menstrual cycle. *Forensic Science International: Genetics Supplementary Series, 5,* e264-e266. https://doi.org/10.1016/j.fsigss.2015.09.105

15. Seashols-Williams, S. et al. (2016). High-throughput miRNA sequencing and identification of biomarkers for forensically relevant biological fluids. *Electrophoresis, 37(21),* 2780-2788. https://doi.org/10.1002/elps.201600258

16. Rekker, K. Saare, M., Roost, A., Salumets, A., & Peters, M. (2013). Circulating microRNA profile throughout the menstrual cycle. *PLOS One, 8(11).* https://doi.org/10.1371/journal.pone.0081166

17. Baker, D., Grimes, E., & Hopwood, A. (2011). D-dimer assays for the identification of menstrual blood. *Forensic Science International, 212 (1-3),* 210-214. https://doi.org/10.1016/j.forsciint.2011.06.0

18. Butler, J. M. (2011). *Advanced topics in forensic DNA typing : Methodology*. Elsevier Science & Technology.

19. Yang, Y., Xie, B., & Yan, J. (2014). Application of next-generation sequencing technology in forensic science. *Genomics, Proteomics and Bioinformatics, 12(5),* 190-197. https://10.1016/j.gpb.2014.09.001

20. Van Neste, C., Vandewoestyne, M., Van Criekinge, W., Deforce, D., & Van Nieuwerburgh, F. (2013). My-Forensic-Loci-queries (MyFLq) framework for analysis of forensic STR data generated by massive parallel sequencing. *Forensic Science International: Genetics, 9,* 1-8. https://doi.org/10.1016/j.fsigen.2013.10.012

21. Desvignes, T. et al. (2015). MiRNA Nomenclature: A View Incorporating Genetic Origins, Biosynthetic Pathways, and Sequence Variants. *Trends in Genetics, 31(11),* 613-626. https://doi.org/10.1016/j.tig.2015.09.002

22. Hammond, S. (2015). An overview of microRNAs. *Advanced Drug Delivery Reviews, 87,* 3-14. https://doi.org/10.1016/j.addr.2015.05.001

23. Bartel, D. (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell, 136(2),* 215-233. https://doi.org/10.1016/j.cell.2009.01.002

24. Bao, C. et al. (2020). MiR-93 suppresses tumorigenesis and enhances chemosensitivity of breast cancer via dual targeting E2F1 and CCND1. *Cell Death & Disease, 11 (618).* https://doi.org/10.1038/s41419-020-02855-6

25. Ninawe, A. et al. (2021). miR-486-5p: A Prognostic Biomarker for Chronic Myeloid Leukemia. *American Chemical Society Omega, 2021(6),* 7711-7718. https://doi.org/10.1021/acsomega.1c00035

26. Lewis, C., Layne, T., & Seashols-Williams, S. (2019). Detection of microRNAs in DNA Extractions for Forensic Biological Source Identification. *Journal of Forensic Sciences, 64(6),* 1823-1830. https://doi.org/10.1111/1556-4029.14070

27. Corradini, B. et al. (2015). Estimation of the time of death through the analysis of clock miRNAs expression in blood and vitreous humor. *Forensic Science International: Genetics Supplementary Series, 5,* e204-e206. https://doi.org/10.1016/j.fsigss.2015.09.082

28. Brodeur, A., Kero, F., & Kerns, B. (2022). *Demonstration of a Semi-Automated, Extraction Free miRNA Gene Expression Profiling Method for the Simultaneous Detection of Biofluids* [Conference presentation]. American Academy of Forensic Sciences 2022, Seattle, WA, United States.

29. HTG Molecular Diagnostics. (n.d.). *HTG EdgeSeq System: Sample to Answer in 36 Hours*. Retrieved from https://www.htgmolecular.com/assets/htg/resources/C-003-05.22.R6_-_HTG-EdgeSeq_System_Brochure_.pdf

30. Kapa Biosystems (2020). KAPA Library Quantification Kit Illumina® Platforms: Technical Data Sheet. https://rochesequencingstore.com/wp-content/uploads/2017/10/KAPA-Library-Quantification-Kit.pdf

31. Illumina. (2021). *Cluster Optimization Overview.* https://support-docs.illumina.com/SHARE/ClusterOptimize/cluster-optimization-overview.pdf

32. Illumina. (2023). *Sequencing Analysis Viewer.* https://support-docs.illumina.com/SW/SAV/Content/SW/SAV/Overview.htm

33. Ringnér, M. (2008). What is principal component analysis?. *Nature biotechnology, 26,* 303-304. https://doi.org/10.1038/nbt0308-303

34. Wang, Y. et al. (2017). Quantification of distinct let-7 microRNA family members by a modified stem-loop RT-qPCR. *Molecular Medicine Reports, 17(3),* 3690-3696. https://doi.org/10.3892/mmr.2017.8297

35. Scalici, E. et al. (2016). Circulating microRNAs in follicular fluids, powerful tools to explore in vitro fertilization process. *Scientific Reports, 6(24976).* https://doi.org/10.1038/srep24976

36. Li, N. et al. (2017). MiR-106b and miR-93 regulate cell progression by suppression of PTEN via PI3K/Akt pathway in breast cancer. *Cell Death & Disease, 8(e2796).* https://doi.org/ 10.1038/cddis.2017.119

37. Zheng, R. et al. (2015). Prognostic value of miR-106b expression in breast cancer patients. *Journal of Surgical Research, 195(1),* 158-165. https://doi.org/10.1016/j.jss.2014.12.035

38. Antolín, S. et al. (2015). Circulating miR-200c and miR-141 and outcomes in patients with breast cancer. *BMC Cancer, 15(297),* https://doi.org/10.1186/s12885-015-1238-5

39. Chen, F. & Hu, S. (2011). Effect of microRNA-34a in cell cycle, differentiation, and apoptosis: A review. *Journal of Biochemical and Molecular Toxicology, 26(2),* 79-86. https://doi.org/10.1002/jbt.20412

40. Choi, K., Shin, C. Lee, W., Ji, H., & Kim, H. (2019). Dual-strand tumor suppressor miR-193b-3p and -5p inhibit malignant phenotypes of lung cancer by suppressing their common targets. *Bioscience Reports, 39(7).* *https://doi.org/10.1042/BSR20190634*

41. Zhang, L. et al. (2020). MiR-25-3p targets PTEN to regulate the migration, invasion, and apoptosis of esophageal cancer cells via the PI3K/AKT pathway. *Bioscience Reports, 40(10).* https://doi.org/10.1042/BSR20201901

CURRICULUM VITAE