

2023

# Biofluid identification in mock sexual assault samples using a semi-automated, extraction-free microRNA gene expression profiling method

---

<https://hdl.handle.net/2144/49507>

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

BOSTON UNIVERSITY

ARAM V. CHOBANIAN & EDWARD AVEDISIAN SCHOOL OF MEDICINE

Thesis

**BIOFLUID IDENTIFICATION IN MOCK SEXUAL ASSAULT SAMPLES  
USING A SEMI-AUTOMATED, EXTRACTION-FREE MICRORNA GENE  
EXPRESSION PROFILING METHOD**

by

**XIAOMENG YANG**

B.S., University of Washington, 2021

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

2023



Approved by

First Reader

---

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

Second Reader

---

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

Third Reader

---

Lynn Schneeweis, M.S.  
Deputy Chief Science Officer/Assistant Lab Director  
Massachusetts State Police Crime Laboratory

## **ACKNOWLEDGMENTS**

I would like to express my deepest gratitude to several individuals and organizations who have made this thesis possible. To my family and friends for the constant support and encouragement during my master's program. To Boston University School of Medicine and the Biomedical Forensic Sciences program for providing the academic and research opportunities that have significantly contributed to my learning and growth in this field. To my thesis advisor, Prof. Amy Brodeur, for her unwavering guidance, patience, and insight throughout this journey. Her expertise has been an invaluable asset to my research. My heartfelt appreciation also extends to Dr. Robin Cotton for the opportunity to work with her on this project. Her knowledge, collaboration, and motivational words have been an unending source of inspiration. My sincere thanks to my readers and the program faculty for their invaluable suggestions and constructive critiques that greatly enhanced the quality of this work. I would like to acknowledge HTG Molecular Diagnostics, Inc. for their cooperation and assistance during the research process. The opportunity to work in a real-world setting has been instrumental in deepening my understanding and enriching my experience in the field of forensic sciences. This thesis would not have been possible without your support, and I am eternally grateful.

**BIOFLUID IDENTIFICATION IN MOCK SEXUAL ASSAULT SAMPLES  
USING A SEMI-AUTOMATED, EXTRACTION-FREE MICRORNA GENE  
EXPRESSION PROFILING METHOD**

**XIAOMENG YANG**

**ABSTRACT**

Biofluid identification plays a pivotal role in forensic case investigation, as it assists in crime scene reconstruction and provides guidance for subsequent analysis. Current forensic biofluid detection techniques, such as alternate light sources, catalytic color tests, and lateral flow immunochromatographic assays, exhibit limitations in terms of sensitivity and specificity, and most can only detect one biofluid at a time. Developing a multiplex and confirmatory assay will significantly benefit forensic biofluid identification, with nucleic acid-based assays emerging as the most promising candidates. Among various nucleic acids, microRNAs boast biofluid specificity and exceptional stability, rendering them an optimal choice for biofluid identification assay and the focal point of this study.

In this study, we investigated a semi-automated, extraction-free microRNA gene expression profiling method developed by HTG Molecular Diagnostics, Inc., namely HTG EdgeSeq microRNA Whole Transcriptome Assay, utilizing seven biofluid types frequently encountered in sexual assault cases, including peripheral blood, menstrual blood, saliva, vaginal fluid, and semen with varying sperm counts. The research focused on three aspects: the assay's compatibility with whole blood samples and dried biofluid swab

samples, its ability to identify and classify different biofluid types based on their microRNA expression patterns, and its potential to differentiate biofluids within mixtures.

Our results demonstrate that HTG EdgeSeq microRNA Whole Transcriptome Assay effectively identified several differentially expressed microRNAs in each biofluid. Examples that align with the existing literature include miR-451a and 486-5p for all kinds of blood, miR-185-5p for peripheral blood, and miR-888-5p and 891a-5p for semen. Previously unreported microRNAs with biofluid-specific expressions were also discovered, such as miR-4306 for peripheral blood and miR-184 for saliva. The scatter plots and heat maps generated from experimental data exhibit distinctive microRNA expression profiles for different biofluid types, potentially enabling their identification and classification. Furthermore, the analysis of mixtures reveals that differentially expressed microRNAs in mixtures corresponded to those detected in the single-source biofluids, signifying the assay's capacity to resolve biofluid mixtures. The assay also yields optimal results for whole blood and dried biofluid swab samples, displaying its compatibility with sample types frequently found in forensic casework.

The outcomes from this study suggest the potential of HTG EdgeSeq microRNA Whole Transcriptome Assay to be employed in forensic biofluid identification, as it demonstrates high reproducibility, multiplex capacities, and compatibility with common forensic casework samples. Future studies encompassing a broader spectrum of biofluid types, and a greater number of individuals are required to validate the findings in this study and optimize the workflow for applying the assay in forensic laboratories.

## TABLE OF CONTENTS

ACKNOWLEDGMENTS .....	iv
ABSTRACT.....	v
TABLE OF CONTENTS.....	vii
LIST OF TABLES.....	x
LIST OF FIGURES .....	xi
LIST OF ABBREVIATIONS.....	xiii
1. INTRODUCTION .....	1
1.1 Current Methods for Biofluid Identification.....	1
1.1.1 Alternate Light Source.....	2
1.1.2 Catalytic Color Tests.....	2
1.1.3 Lateral Flow Immunochromatographic Assays.....	3
1.1.4 Microscopic Examination for Spermatozoa.....	4
1.2 Nucleic Acid-Based Biofluid Identification Assays.....	4
1.2.1 DNA Methylation Markers .....	5
1.2.2 Messenger RNAs .....	5
1.2.3 Microbial Nucleic Acids .....	6
1.3 MicroRNA in Biofluid Identification .....	6
1.3.1 Co-Analysis of DNA and miRNA Through Capillary Electrophoresis.....	8
1.3.2 Microarray and Reverse Transcription-Quantitative PCR.....	8
1.3.3 Next-Generation Sequencing Assays.....	9
1.4 HTG EdgeSeq miRNA Whole Transcriptome Assay.....	10
1.4.1 Purposes of the Study.....	12

2. MATERIALS AND METHODS.....	14
2.1 Sample Collection.....	14
2.2 Mixture Preparation .....	14
2.3 Sample Preparation and Quantitative Nuclease Protection Assay.....	15
2.4 PCR Amplification and Tag Addition .....	17
2.5 Library Cleanup .....	19
2.6 Library Quantification .....	20
2.7 Normalization and Pooling of Libraries .....	21
2.8 Next-Generation Sequencing of Libraries .....	23
2.9 Data Analysis .....	23
3. RESULTS AND DISCUSSION.....	25
3.1 MicroRNA Expression Profiling for Liquid and Dried Biofluid Samples .....	25
3.1.1 Data Quality of Samples from Three Runs.....	25
3.1.2 Comparison of miRNA Expression Profiles.....	27
3.2 Single-source Biofluid Identification.....	31
3.2.1 Principal Component Analysis .....	32
3.2.2 Differential Expression Analysis .....	34
3.2.2.1 Previously Reported miRNAs in Peripheral and Menstrual Blood .....	38
3.2.2.2 Previously Reported miRNAs in Saliva and Vaginal Fluid.....	41
3.2.2.3 Previously Reported miRNAs in Normal, Azoospermic, and Oligospermic Semen.....	42
3.2.2.4 Discrepancies With the Literature .....	44
3.2.2.5 Functions of Identified Biofluid-Specific miRNAs.....	46
3.2.3 Heat Map of Selected miRNAs.....	48

3.3 Analysis of Mixture Samples.....	50
3.3.1 Mixtures of Peripheral Blood and Vaginal Fluid.....	51
3.3.2 Mixtures of Menstrual Blood and Semen .....	55
3.3.3 Other Mixtures .....	57
3.4 Comparison to Preliminary Study.....	60
3.4.1 Principal Component Analysis .....	62
3.4.2 Common Differentially Expressed miRNAs .....	66
3.5 Method Optimization and Troubleshooting.....	70
3.5.1 Sample Preparation Protocol for Dried Biofluid Samples.....	70
3.5.2 Optimization of Workflow.....	71
4. CONCLUSIONS .....	74
4.1 Findings of the Study .....	74
4.2 Future Directions .....	77
BIBLIOGRAPHY.....	79
CURRICULUM VITAE.....	90

## LIST OF TABLES

Table 1. Top up-regulated and down-regulated miRNAs in peripheral blood (PB), menstrual blood (MB), normal semen (Sem), saliva (Sal), vaginal fluid (VF), azoospermic semen (AS), and oligospermic semen (OS) .....	35
Table 2. Up-regulated miRNAs shared between mixture and single-source biofluids, including vaginal fluid, saliva, normal semen, and oligospermic semen ...	58
Table 3. Common differentially expressed miRNAs in samples from preliminary and current studies, including saliva, and peripheral and menstrual blood, and normal and azoospermic semen .....	66

## LIST OF FIGURES

Figure 1. Molecular level illustration of HTG proprietary quantitative nuclease protection assay (qNPA) .....	11
Figure 2. Overview of adapted HTG EdgeSeq miRNA Whole Transcriptome Assay (WTA) workflow .....	15
Figure 3. Three-dimensional principal component analysis (3D PCA) scatter plot for 20 samples from the first run .....	28
Figure 4. Heat map of 18 samples from the first run .....	29
Figure 5. Three-dimensional principal component analysis (3D PCA) scatter plot for 47 single-source biofluid samples and positive controls .....	32
Figure 6. Heat map of 19 most differentially expressed miRNAs in 40 single-source biofluid samples .....	48
Figure 7. Expression profiles of top five up-regulated miRNAs in the mixture of peripheral blood and vaginal fluid that are common to a) peripheral blood and b) vaginal fluid .....	51
Figure 8. Three-dimensional principal component analysis (3D PCA) scatter plot for blood-vaginal mixtures, vaginal fluid, and peripheral and menstrual blood .....	54
Figure 9. Expression profiles of top five up-regulated miRNAs in the mixture of menstrual blood and normal semen that are common to a) menstrual blood and b) normal semen .....	56

Figure 10. Three-dimensional principal component analysis (3D PCA) scatter plot for samples from preliminary and current studies..... 62

Figure 11. Heat map of azoospermic, oligospermic, and normal semen samples from the preliminary and current studies..... 68

## LIST OF ABBREVIATIONS

ALS	Alternate light source
AS	Azoospermic semen
CE	Capillary electrophoresis
DI	Deionized
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
HCl	Hydrochloric acid
M	Molar
MB	Menstrual blood
miR, miRNA, microRNA	Micro ribonucleic acid
mL	Milliliter
mM	Millimolar
mRNA, messenger RNA	Messenger ribonucleic acid
MTC	Multi-tissue control
N	Normal
NaCl	Sodium chloride
NaOH	Sodium hydroxide

NGS	Next-generation sequencing
nM	Nanomolar
NPPs	Nuclease protection probes
OS	Oligospermic semen
PB	Peripheral blood
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pM	Picomolar
p-value	Probability value
QC	Quality control
qNPA	Quantitative nuclease protection assay
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
RT-PCR	Reverse transcription-polymerase chain reaction
RT-qPCR	Reverse transcription-quantitative polymerase chain reaction
RUO	Research use only
Sal	Saliva
SAV	Sequencing analysis viewer
Sem	Normal semen
tDMRs	Tissue-specific differentially methylated regions

Tris-HCl	Trizma hydrochloride
UHR	Universal human reference
VF	Vaginal fluid
WTA	Whole transcriptome assay
$\mu\text{L}$	Microliters
$^{\circ}\text{C}$	Degrees Celsius
3D	Three-dimensional

## **1. INTRODUCTION**

Biofluid identification constitutes a crucial component of crime scene investigation as it aids in reconstructing the scene and determining the order of events. It may serve as a prerequisite to forensic deoxyribonucleic acid (DNA) profiling, assisting forensic scientists in evaluating the probative value of collected samples and selecting those for further analysis. The mere existence of an individual's DNA at a crime scene does not definitively implicate them in criminal activities, as factors such as shared living spaces and DNA transfer exist. The presence of specific biofluids may occasionally be indicative of the crime. Moreover, a linkage between victims, suspects, evidence, and the scene can be established by associating the biofluid type with the DNA profile obtained from the biofluid samples.

### **1.1 Current Methods for Biofluid Identification**

In current forensic casework, biofluid identification typically begins with the visual examination of stains exhibiting the characteristic appearance or odor of specific biofluids. Chemical enhancement methods or alternate light sources facilitate the visualization and collection of inconspicuous biofluid samples. Subsequently, secondary screening or confirmatory tests are performed on the collected biofluid samples to detect the presence of various biofluid types and to guide further testing. Common approaches include colorimetric tests, which provide presumptive indications for the biofluid types, and the lateral flow immunochromatographic assay, which can be presumptive or confirmatory for

different biofluid types. For presumptive semen samples, a microscopic examination is conducted to verify the presence of spermatozoa.

#### 1.1.1 Alternate Light Source

An alternate light source (ALS) is a specialized light source with single or variable wavelengths across ultraviolet, visible, and infrared regions (1). ALS assists in visualizing biofluid stains by utilizing specific wavelengths of light to cause certain biofluids, such as semen and saliva stains, to emit fluorescent light that is observable using appropriate light filters or goggles (1,2). The technique is expeditious and user-friendly, with portable models designed for on-site application (2). Although ALS is generally non-destructive to biological samples, DNA in bloodstains may suffer damage if exposed to shortwave ultraviolet light for extended durations (3). Nonetheless, ALS lacks specificity to the biofluid types, and its sensitivity varies depending on the fluorescent properties of biofluids and substrate type (1,2).

#### 1.1.2 Catalytic Color Tests

Catalytic color tests employ chemical reagents to detect specific enzymatic activities of proteins in biofluids. Upon reagent addition to a biofluid sample, a particular protein catalyzes the chemical reaction, leading to a color appearance or alteration (4,5). For example, the phenolphthalein or Kastle-Meyer test produces a pink color in the presence of hemoglobin, a blood protein exhibiting peroxidase-like activity (4,6). Catalytic color tests are convenient, with certain tests demonstrating high sensitivity and specificity

for the detected biofluids (6,7). However, catalytic color tests lack human-specificity because they detect all non-human substances capable of initiating catalytic reactions (4). Common substances like plants and detergents may induce false positive or negative testing results (5). Furthermore, catalytic color tests can be destructive to biological samples, and variations in enzymatic activity levels in biofluids can lead to significant discrepancies in outcomes (6,8). Thus, catalytic color tests provide presumptive biofluid identity indications requiring confirmation through other analyses.

### 1.1.3 Lateral Flow Immunochromatographic Assays

Lateral flow immunochromatographic assays hinge on the interaction between antibodies and specific protein antigens in biofluids. When target antigens are detected in a sample, visible colored lines emerge in both test and control areas on the lateral flow immunoassay card, indicating the presence of specific biofluids (9,10). Lateral flow immunoassays exhibit high sensitivity and human specificity if the antibodies are designed to bind human antigens. These immunoassays demonstrate robust specificity for identified biofluid types, including blood and semen (11–13). However, lateral flow immunoassays for human saliva display lower specificity, as the target antigens are present in biofluids other than saliva (14). Certain lateral flow immunoassays are susceptible to the high-dose hook effect, wherein an excessive amount of antigen may yield a false negative result (15,16).

#### 1.1.4 Microscopic Examination for Spermatozoa

Forensic laboratories apply microscopic examination to confirm the presence of spermatozoa in samples, which serves as indisputable evidence of semen. Nevertheless, the manual search for spermatozoa is labor-intensive and time-consuming for laboratory technicians, especially when handling semen samples with low sperm counts and a profusion of distracting elements such as vaginal epithelial cells, bacteria, and yeasts in the background. In sexual assault cases involving penetration, spermatozoa detection via microscopy is primarily influenced by the timing of sample collection post-ejaculation and the persistence of spermatozoa in oral, vaginal, and anal environments (17,18).

### **1.2 Nucleic Acid-Based Biofluid Identification Assays**

Present methods for forensic biofluid identification exhibit various limitations, such as inconsistent sensitivity, lack of human-specificity, laborious processes, and a tendency for false positive results. Moreover, most current approaches can only identify a single biofluid at a time, necessitating the consumption of a substantial amount of sample when sequentially employing techniques to detect multiple biofluids. Recent research has been focused on nucleic acid-based assays for developing a confirmatory multiplex biofluid identification method. Nucleic acid-based assays, including DNA, messenger ribonucleic acid (mRNA), and micro ribonucleic acid (miRNA) assays, possess exceptional sensitivity and human specificity. These assays can incorporate multiple markers in a single panel to identify different biofluids simultaneously. Minimal sample quantities are required for

nucleic acid-based biofluid identification assays, as polymerase chain reaction (PCR) can amplify the number of nucleic acid copies within a sample (19).

### 1.2.1 DNA Methylation Markers

Epigenetic modification regulates gene expression levels without altering the DNA sequence (20). DNA methylation, a type of epigenetic modification, mainly occurs at the C5 position of cytosine within a CpG dinucleotide sequence (20,21). DNA methylation profiles have been shown to be unique in different human tissues, and numerous tissue-specific differentially methylated regions (tDMRs) have been identified in biofluids and tissues such as blood, sperm, lung, liver, and uterus (22). Forensic researchers have investigated DNA methylation profiles using microarray assays to select tDMRs in forensic-relevant biofluids, including peripheral blood, menstrual blood, saliva, semen, and vaginal fluid (23). Additionally, they have established multiplex systems to differentiate various biofluids by quantifying the methylation levels of candidate tDMRs (23,24). However, certain tDMRs can vary in methylation levels due to environmental and aging factors (24,25).

### 1.2.2 Messenger RNAs

Messenger RNA (mRNA) is a transcript from genomic DNA that carries protein synthesis instructions (26). Multiple mRNA markers with specific or differential expressions have been discovered in biofluids such as peripheral blood, menstrual blood, saliva, semen, sweat, urine, and vaginal fluid (27–31). Messenger RNA-based biofluid

identification assays generally involve the reverse transcription-polymerase chain reaction (RT-PCR), a technique employed to synthesize DNA from extracted RNA (32). Real-time analyses like quantitative PCR (30,31) or end-point analyses like capillary electrophoresis (27–29,31) are applied following RT-PCR to examine mRNA expression profiles in different biofluids and identify biofluid-specific mRNA markers. Multiplex assays incorporating several biofluid-specific mRNA markers have been developed to differentiate biofluids frequently encountered in forensic casework, including peripheral blood, semen, and saliva (30,31,33).

### 1.2.3 Microbial Nucleic Acids

Beyond human DNA and RNA, microbial DNA and RNA markers are exploited to detect forensic-relevant biofluids. Microbes, such as bacteria and fungi, are remarkably abundant and diverse in human organs such as the skin, gut, and mouth (34). Researchers have utilized approaches like next-generation sequencing, microarray, and quantitative PCR analysis to investigate the gene expression of microbes in various parts of the human body and identify specific microbial markers present in vaginal fluid (35,36), feces (37), and saliva (38).

### 1.3 MicroRNA in Biofluid Identification

MicroRNA (miRNA) is a single-stranded RNA of 18 to 25 nucleotides in length originating from a hairpin-shaped transcript (39–41). It acts as a gene expression regulator by controlling mRNA stability or translational efficiency (42–44). Several miRNAs have

been shown to display specific or differential expression in various tissues and biofluids, including the brain, heart, liver, and peripheral blood (45–48). Clinical studies have extensively explored tissue-specific miRNA expression patterns for cancer prognosis and monitoring cancer progression in tissues such as the prostate, esophagus, liver, and breast (49–54).

Although both miRNA and mRNA exhibit tissue specificity, miRNA boasts greater stability and environmental degradation resistance due to its small size (47,55,56). Extracellular miRNAs in human plasma are further stabilized by RNA-binding proteins that protect miRNAs from ribonuclease degradation (57,58). MicroRNAs have demonstrated resilience under severe environmental and chemical conditions, including ultraviolet radiation, heat, repeated freeze-thaw cycles, boiling, extreme pH, detergents, and bleach (56,59,60). In a study comparing miRNA and mRNA stability in environmentally compromised samples, miRNAs were detected throughout a 180-day study, whereas mRNAs became undetectable after 30 days (56). MicroRNAs have also maintained high stability in aged samples stored at frozen, ambient, and elevated temperatures (61–63).

Owing to its tissue specificity and stability, miRNA has emerged as a popular research subject for forensic biofluid identification. Various methods have been employed, including capillary electrophoresis, microarray, reverse transcription-quantitative PCR, and next-generation sequencing, to identify specific miRNAs in forensically relevant biofluids and develop a multiplex forensic biofluid identification panel.

### 1.3.1 Co-Analysis of DNA and miRNA Through Capillary Electrophoresis

Capillary electrophoresis (CE) is applied in forensic DNA profiling to separate DNA fragments based on the fragment size and fluorescent labeling (64). The primary objective of developing a CE-based miRNA biofluid identification method is to co-extract and co-analyze DNA and miRNA by making minor adjustments to existing DNA profiling protocols. Researchers have employed DNA extraction kits to co-isolate DNA and RNA (65,66) or used commercially available DNA and RNA co-extraction kits (67,68). Specially designed miRNA primers and an additional RT-PCR step are incorporated to convert RNA into DNA and amplify DNA products. The resulting products are analyzed by CE, generating either a single electropherogram displaying all DNA and miRNA markers (65,66) or separate electropherograms for each (67,68). This approach permits forensic scientists to simultaneously determine the identities of individuals and biofluids, thereby enhancing the accuracy of association between the DNA profile and the biofluid type.

### 1.3.2 Microarray and Reverse Transcription-Quantitative PCR

Microarray assays involve the hybridization of miRNAs to a microarray slide and the generation of a miRNA expression profile based on the hybridization intensity (69). These assays enable the concurrent detection and analysis of hundreds of miRNA markers. Researchers have predominantly applied microarray assays to screen for differentially expressed miRNA markers in peripheral blood, menstrual blood, semen, vaginal fluid, and

saliva (63,70–73). Only a few differentially expressed miRNAs are chosen for each biofluid type, requiring validation through quantitative PCR analysis.

Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) is typically a two-step process combining RT-PCR and quantitative PCR (74). RT-qPCR assays for miRNA exhibit exceptional sensitivity, detecting miRNAs in samples containing as little as two picograms of total RNA (63). RT-qPCR is often performed after microarray screening to confirm the biofluid specificity of candidate miRNAs by measuring their expression levels in different biofluids (63,70–73). In addition, RT-qPCR can be conducted independently to examine known miRNAs selected from databases or previous research (47,75–78). Reference miRNA markers stably and robustly expressed in all biofluid types can be identified through RT-qPCR and used to normalize miRNA expression data (73,77,79,80). Although no genuinely biofluid-specific miRNAs have been discovered, researchers have devised biofluid identification strategies using combinations of differentially expressed miRNAs and their unique expression patterns in biofluids (47,76,81).

### 1.3.3 Next-Generation Sequencing Assays

In recent years, next-generation sequencing (NGS), also known as high-throughput or massively parallel sequencing, has gained widespread recognition as a valuable tool for biofluid identification. Sequencing reveals the genetic details of a DNA or RNA fragment, including the composition and order of the nucleotides. Compared to Sanger sequencing, NGS offers improved speed, sensitivity, throughput, and versatility (82). While both NGS

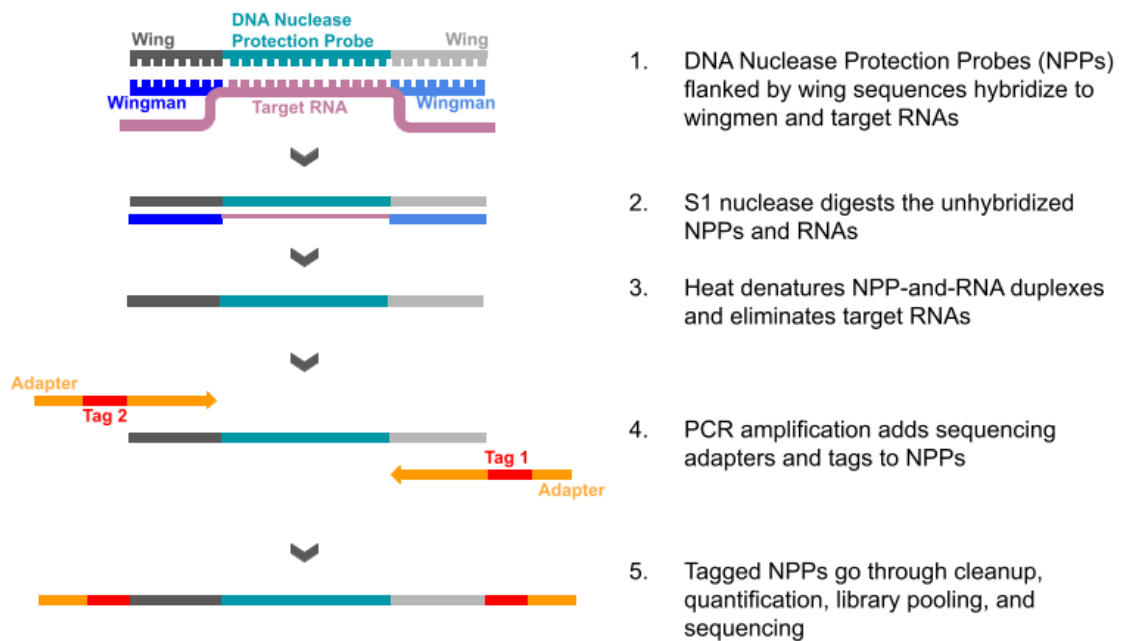
and microarray are multiplex techniques, NGS provides greater sensitivity, accuracy, and detection range than microarray assays. Moreover, NGS can detect new RNA transcripts, whereas microarrays are limited to identifying previously known RNAs (83). Forensic scientists have harnessed NGS's high-throughput capabilities to evaluate the expression levels of thousands of miRNAs and isolate differentially expressed miRNAs in various forensic-relevant biofluids, such as blood, saliva, semen, and vagina fluid (84,85). By coupling NGS with advanced bioinformatics analysis, novel miRNAs have also been discovered in forensically relevant biofluids and tissues (86,87). NGS has empowered researchers to discover novel genetic markers and accurately identify specific targets among numerous biological markers, consequently gaining popularity in clinical and forensic research.

#### **1.4 HTG EdgeSeq miRNA Whole Transcriptome Assay**

The majority of previously discussed miRNA assays, including CE, microarray, RT-qPCR, and NGS, require RNA extraction and DNA synthesis via RT-PCR before further analysis using different platforms. These sample-handling processes result in increased labor and time demands. In addition, the RNA extraction procedure may introduce contamination and inter-operator variation (88). A favorable improvement for biofluid identification using miRNA markers would be an approach that circumvents the need for additional extraction and RT-PCR steps.

In this study, we employed the HTG EdgeSeq miRNA Whole Transcriptome Assay (WTA), a semi-automated, extraction-free miRNA expression profiling method developed

by HTG Molecular Diagnostics, Inc. (hereafter referred to as HTG), to identify biofluid types that are commonly encountered in forensic casework. The HTG EdgeSeq miRNA WTA combines the proprietary quantitative nuclease protection assay (qNPA) with NGS-mediated quantification for expression profiling of 2,083 human miRNAs from the miRBase v20 database, as well as one miRNA internal positive control, five plant genes as internal negative controls, and 13 housekeeping genes (89). The qNPA is RNA extraction-free and automated to mitigate potential error from manual handling and enhance the reproducibility of the process.



**Figure 1. Molecular level illustration of HTG proprietary quantitative nuclease protection assay (qNPA).** Adapted from HTG EdgeSeq System User Manual (RUO) (90).

As illustrated in Figure 1, the qNPA commences with the hybridization of nuclease protection probes (NPPs) to target RNAs. Universal DNA wingman probes are hybridized to the flanking universal DNA wing sequences of NPPs to prevent S1 nuclease digestion.

Subsequently, S1 nuclease is introduced to remove excess non-hybridized NPPs and RNAs, resulting in a one-to-one ratio of NPPs to target RNAs. This process is followed by heat denaturation of NPP-and-RNA duplexes to eliminate target RNAs, leaving only NPPs for sequencing adapter and tag addition via PCR amplification. Tagged NPPs are then cleaned, pooled, and quantified using the Illumina MiSeq sequencing system.

A minimal sample volume, as small as 12.5 microliters ( $\mu\text{L}$ ) of liquid biofluid, is required for HTG EdgeSeq miRNA WTA. Quality control studies have demonstrated the assay's high fidelity and reproducibility (89,91). The assay has proven compatible with several sample types, including formalin-fixed paraffin-embedded tissue specimens (92,93), isolated RNAs (94), and plasma and serum samples (95,96). HTG EdgeSeq miRNA WTA has been practiced in clinical studies to investigate the miRNA differential expression profiles associated with sex (94) and diseases like non-small cell lung cancer (92), skin cancer (93), breast cancer (95), and type 1 diabetes (96).

A preliminary study exploring the application of HTG EdgeSeq miRNA WTA to forensic biofluid identification has been conducted by HTG researchers using liquid biofluid samples provided by our laboratory, encompassing peripheral blood, menstrual blood, saliva, normal semen, azoospermic semen, breast milk, and urine. The findings of this preliminary study (97) will be discussed in detail later in this paper.

#### 1.4.1 Purposes of the Study

To evaluate the HTG EdgeSeq miRNA WTA for forensic biofluid identification in our laboratory, we prepared samples that closely resembled those encountered in forensic

casework, including dried biofluid stains on cotton swabs and biofluid mixtures composed of two distinct biofluid types. Seven biofluid types found in sexual assault cases - saliva, vaginal fluid, menstrual blood, peripheral blood, and normal, oligospermic, and azoospermic semen - were examined to determine if HTG EdgeSeq miRNA WTA could accurately identify and classify these biofluids based on their miRNA expression profiles.

Liquid and dried biofluid samples were analyzed to ascertain if HTG EdgeSeq miRNA WTA could provide similar miRNA expression profiles for the same biofluid, regardless of its form. Furthermore, the comparison of miRNA expression profiles between liquid and dried biofluid samples aimed to comprehend the impact of the sample preparation protocol for dried biofluid samples on miRNA expression profiling.

Various biofluid mixtures were prepared using two distinct biofluid types. The miRNA profiles derived from single-source and mixed biofluid samples were scrutinized to assess HTG EdgeSeq miRNA WTA's ability to distinguish biofluids within the mixtures. The mixtures of peripheral blood and vaginal fluid were included in determining if the assay could differentiate between menstrual blood and a combination of peripheral blood and vaginal fluid, as observed in cases of bleeding resulting from vaginal injuries. Whole blood samples were examined to evaluate the assay's compatibility with such samples, given that previous clinical studies exclusively utilized plasma and serum samples.

## **2. MATERIALS AND METHODS**

### **2.1 Sample Collection**

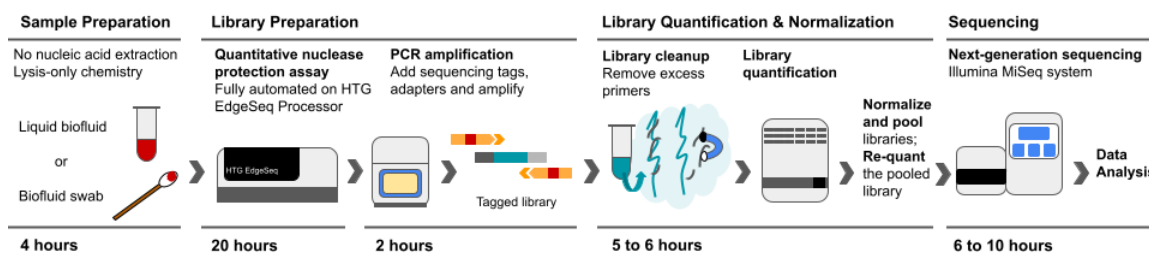
Peripheral blood, menstrual blood, saliva, normal semen, oligospermic semen, and vaginal fluid were collected from nine donors in compliance with a protocol approved by the Boston University School of Medicine Institutional Review Board. In addition, one milliliter (mL) of azoospermic/vasectomy semen was purchased from Lee BioSolutions, Inc. (Maryland Heights, MO). All liquid biofluid samples were stored at -20 degrees Celsius (°C) prior to use. All biofluid swabs were stored at room temperature. Peripheral blood was collected from three donors. For each donor, 30  $\mu$ L of peripheral blood was collected via puncture of the fingers and transferred to a cotton swab to prepare a peripheral blood swab. Liquid menstrual blood was provided by one donor using a menstrual cup and a 15-mL conical tube. Menstrual blood was also provided by another donor on cotton swabs. Normal semen was provided by one donor in a specimen container. Normal semen swabs were prepared by adding 30  $\mu$ L of liquid semen onto a cotton swab. Oligospermic semen was provided by one donor in a specimen container. Liquid saliva was provided by one donor in a 1.5-mL microcentrifuge tube. The saliva swab was prepared by adding 30  $\mu$ L of liquid saliva onto a cotton swab. Vaginal fluid was provided by three donors on cotton swabs.

### **2.2 Mixture Preparation**

Mixtures of saliva and normal semen, and saliva and oligospermic semen were prepared by adding 20  $\mu$ L of each liquid biofluid to a microcentrifuge tube, pipetting to

mix well, and depositing 30  $\mu$ L of the mixed biofluid onto a cotton swab. The mixtures of vaginal fluid and peripheral blood, vaginal fluid and normal semen, vaginal fluid and oligospermic semen, vaginal fluid and saliva, and menstrual blood and normal semen were prepared by adding 15  $\mu$ L of liquid biofluid (peripheral blood, saliva, normal semen, or oligospermic semen) onto the collected vaginal fluid or menstrual blood swabs.

### 2.3 Sample Preparation and Quantitative Nuclease Protection Assay



**Figure 2. Overview of adapted HTG EdgeSeq miRNA Whole Transcriptome Assay (WTA) workflow.** Adapted from HTG EdgeSeq miRNA WTA Verification Report (89).

HTG EdgeSeq miRNA WTA kit containing four plates and 24 samples per plate was purchased from HTG Molecular Diagnostics, Inc. (Tucson, AZ). An overview of our laboratory’s adapted HTG EdgeSeq miRNA WTA workflow is illustrated in Figure 2. The original procedures of HTG EdgeSeq miRNA WTA can be found in the HTG EdgeSeq System User Manual, Research Use Only (RUO), Code number 10383239, Revision B (HTG Molecular Diagnostics, Inc.) (90).

A total of three runs with 24 samples in each run was performed. For the first run, universal human reference (UHR) RNA was purchased from Thermo Fisher Scientific (Waltham, MA) and used as the positive control for the entire process. For the second and

third runs, multi-tissue control (MTC) lysates were provided by HTG and used as positive controls. All samples and controls were prepared in triplicates.

For each liquid biofluid sample, 15  $\mu$ L of liquid biofluid was combined with 15  $\mu$ L of Plasma Lysis Buffer and 3  $\mu$ L of Proteinase K from the HTG EdgeSeq miRNA WTA Sample Prep Pack (HTG Molecular Diagnostics, Inc.) in a microcentrifuge tube. The mixture was incubated on a heat block at 50 °C for 3 hours with agitation.

A sample preparation protocol for dried biofluids on cotton swabs was not provided in the HTG EdgeSeq miRNA WTA user manual. The following protocol was adapted from the HTG's sample preparation procedures for swab samples in other studies and modified by preliminary experiments in our laboratory. For each biofluid swab, half of the swab was cut and soaked into 60  $\mu$ L of 1x Plasma Lysis Buffer, which was prepared by diluting Plasma Lysis Buffer with deionized (DI) water in a one-to-one ratio. After 10 minutes of incubation, the swab was transferred to a spin basket and placed back in the original tube. The tube was centrifuged for 5 minutes at 14,000 revolutions per minute. After centrifugation, the spin basket with the swab was removed. Thirty microliters of the eluted product were transferred to a new microcentrifuge tube, combined with 3  $\mu$ L of Proteinase K, and incubated on a heat block at 50 °C for 3 hours with agitation.

The following procedures were the same for liquid biofluid and biofluid swab samples. After 3 hours of incubation, 25  $\mu$ L of the incubated sample was transferred into a sample plate from the EdgeSeq Plate Pack (HTG Molecular Diagnostics, Inc.). UHR RNA positive control was prepared and loaded in the same way as liquid biofluid samples. For each MTC positive control, 25  $\mu$ L of MTC lysate was directly loaded into the sample

plate without any preparation. The HTG sample plate template was prepared using Microsoft Excel software (Microsoft, Redmond, WA) and imported into the HTG Host software on a computer connecting to the HTG EdgeSeq Processor (HTG Molecular Diagnostics, Inc.). The sample plate, stop plate, assay reagent tray, and pipette tip box were loaded to their corresponding positions on the HTG EdgeSeq Processor, per the HTG EdgeSeq System User Manual (RUO) (90). The quantitative nuclease protection assay (qNPA) of miRNA targets was run automatically in the HTG EdgeSeq Processor for approximately 20 hours. After the run, the stop plate was removed from HTG EdgeSeq Processor and used immediately for the following procedures. Subsequently, the stop plate was stored at -30 °C.

#### **2.4 PCR Amplification and Tag Addition**

For the first run, PCR amplification and tagging were performed using OneTaq® Hot Start 2x Master Mix with Standard Buffer (New England Biolabs, Ipswich, MA) and forward primers and reverse primers provided in the HTG EdgeSeq miRNA WTA Sequencing Tag Pack for Illumina NGS systems (HTG Molecular Diagnostics, Inc.). Column master mix was prepared for the number of columns required, each with a different forward primer. For this study, 24 samples were tested per run; therefore, three column master mixes were prepared with three different forward primers (F1, F2, F3). For each PCR amplification reaction, the column master mix was prepared by combining 30 µL of OneTaq® Hot Start 2x Master Mix, 14 µL of DI water, and 6 µL of the forward primer. Fifty microliters of column master mix were added to each well on a 96-well microplate in

the corresponding columns. Eight different reverse primers were used (RA-RH). Six microliters of reverse primer were added to each well in the corresponding rows. Four microliters of samples from the stop plate were added to the appropriate wells for a total amplification reaction volume of 60  $\mu$ L.

For the second and third runs, PCR amplification and tagging were performed using the Hemo KlenTaq® kit (New England Biolabs) and forward primers and reverse primers provided in the HTG EdgeSeq miRNA WTA Sequencing Tag Pack for Illumina NGS systems. Three column master mixes were prepared with three different forward primers (F1, F2, F3). For each PCR amplification reaction, the column master mix was prepared by combining 2.4  $\mu$ L of Hemo KlenTaq® Enzyme, 6  $\mu$ L of 5x Hemo KlenTaq® Buffer, 0.6  $\mu$ L of Deoxynucleotide Solution Mix (New England Biolabs), 12  $\mu$ L of DI water, and 3  $\mu$ L of the forward primer. Twenty-four microliters of column master mix were added to each well on a 96-well microplate in the corresponding columns. Eight different reverse primers were used (RA-RH). Three microliters of reverse primer were added to each well in the corresponding rows. Three microliters of samples from the stop plate were added to the appropriate wells for a total amplification reaction volume of 30  $\mu$ L.

The microplate was sealed with flat strip caps and placed on the thermocycler. The thermocycler setting was listed as follows: initial denaturation at 95 °C for 4 minutes, followed by 16 cycles of 95 °C for 15 seconds, 56 °C for 45 seconds, and 68 °C for 45 seconds; the reaction was then run at 68 °C for 10 minutes and held at 4 °C.

## 2.5 Library Cleanup

For the first run, 37.5  $\mu\text{L}$  of AMPure® XP magnetic beads (Beckman Coulter, Inc., Brea, CA) and 100  $\mu\text{L}$  of Cleanup Buffer were added to each well on a new 96-well microplate. For each sample, 100  $\mu\text{L}$  of cleanup Buffer was prepared by combining 39  $\mu\text{L}$  of five-molar sodium chloride (5 M NaCl) (Thermo Fisher Scientific), 31.25  $\mu\text{L}$  of 40% polyethylene glycol (PEG) 8000 (Sigma-Aldrich, St. Louis, MO), and 29.75  $\mu\text{L}$  of DI water. Fifty-five microliters of the amplified product were then transferred to the appropriate well on the microplate and pipetted to mix well.

For the second and third runs, 37.5  $\mu\text{L}$  of AMPure® XP magnetic beads and 15  $\mu\text{L}$  of the amplified product were combined in each well on a new 96-well microplate and pipetted to mix well.

The following procedures were the same for all runs. For each sample, 400  $\mu\text{L}$  of 80% ethanol was prepared by combining 320  $\mu\text{L}$  of 200 proof ethanol (Thermo Fisher Scientific) and 80  $\mu\text{L}$  of DI water. The microplate was incubated at room temperature for 5 minutes and then placed on the magnetic stand for another 5 minutes or until all wells looked clear. With the microplate on the magnetic stand, the supernatant in each well was removed, and 200  $\mu\text{L}$  of 80% ethanol was added to each well and discarded after 1 minute of incubation at room temperature. The ethanol wash step was repeated one more time. A second aspiration might be needed to remove as much ethanol as possible. The bead pellets were then air-dried at room temperature for 10 minutes with the microplate on the magnetic stand. After air-drying, the microplate was taken off the magnetic stand, and the bead pellet in each well was resuspended in 40  $\mu\text{L}$  of ten-millimolar Trizma hydrochloride (10

mM Tris-HCl), pH 8.0. Ten-millimolar Tris-HCl, pH 8.0, was prepared by diluting 1 M Trizma® hydrochloride solution, pH 8.0 (Sigma-Aldrich), with DI water. The microplate was incubated at room temperature for 5 minutes and on the magnetic stand for another 5 minutes. Thirty microliters of the cleaned-up library were transferred from each well to appropriate wells on a new 96-well microplate. The cleaned-up libraries were used immediately for the library quantification or stored at -20 °C.

## **2.6 Library Quantification**

Library quantification of individual and pooled libraries was performed using Applied Biosystems™ 7500 Real-Time PCR Systems (Thermo Fisher Scientific) and KAPA Library Quantification ROX Low Kit for Illumina® platforms (Kapa Biosystems, Inc., Wilmington, MA). All samples and controls were quantified in triplicate. For a new KAPA library quantification kit, the KAPA master mix was prepared by adding 1 mL of Primer Mix to 5 mL of KAPA SYBR® FAST ROX Low quantitative PCR (qPCR) Master Mix. For each qPCR reaction, the qPCR master mix was prepared by combining 12.4 µL of the KAPA master mix and 3.6 µL of DI water. Sixteen microliters of the qPCR master mix were added to each well on a new 96-well microplate.

The cleaned-up libraries were serially diluted to a concentration of 1:10,000-fold using 10 mM Tris/0.05% Tween buffer, which was prepared by combining proper proportions of 10 mM Trizma hydrochloride, pH 8.0, and Tween 20 (Millipore Sigma, Burlington, MA). Three microliters of cleaned-up libraries were transferred to a new 96-well deep well plate, and 297 µL of 10 mM Tris/0.05% Tween were added to corresponding

wells to prepare 1:100-fold diluted cleaned-up libraries. This step was repeated one more time to prepare 1:10,000-fold diluted cleaned-up libraries.

On the qPCR reaction microplate, 4  $\mu$ L of DI water were added to appropriate wells as negative controls. Four microliters of 1:10,000-fold diluted cleaned-up libraries were added to appropriate wells. DNA standards 1-6 were provided by the KAPA library quantification kit, which covered a DNA detection range from 20 to 0.0002 picomolar (pM). Four microliters of each DNA standard were added to appropriate wells, working from least concentrated (standard 6) to most concentrated (standard 1). The qPCR reaction microplate was sealed with flat strip caps or optical adhesive film and loaded into the qPCR instrument.

In the HID Real-Time PCR Analysis Software (Thermo Fisher Scientific), the following experiment properties were used: 7500 (96 Wells 0.2-mL), Quantitation – Standard Curve, SYBR Green Reagents, Standard (approximately 2 hours to complete a run). The run method was set up as follows: 1 cycle of 95 °C for 5 minutes, followed by 30 cycles of 95 °C for 30 seconds and 60 °C for 45 seconds.

## **2.7 Normalization and Pooling of Libraries**

The qPCR results were exported from the HID Real-Time PCR Analysis Software and imported into the HTG EdgeSeq RUO library calculator (HTG Molecular Diagnostics, Inc.). The cleaned-up libraries were diluted using 10 mM Tris, pH 8.5, or Buffer EB (Qiagen, Hilden, Germany) and pooled together following the instructions generated from the HTG EdgeSeq RUO library calculator. The pooled library had an arbitrary

concentration estimated by the HTG EdgeSeq RUO library calculator. The concentration of the pooled library was verified before sequencing through library quantification. The pooled library was first serially diluted to concentrations that fitted into the detection range of the KAPA library quantification kit, which was 20 to 0.0002 pM. Then the serial dilutions of the pooled library were run in triplicate, along with the qPCR controls and DNA standards, using the same protocol mentioned in the Library Quantification section.

Depending on the qPCR results of the pooled library, the pooled library was denatured and diluted using two-normal (2 N) or 0.2 N sodium hydroxide (NaOH) (Thermo Fisher Scientific), 2 N hydrochloric acid (HCl) (Thermo Fisher Scientific) or 200 mM Tris, pH 7.4 (Sigma-Aldrich), and cold HT1 buffer (Illumina, San Diego, CA) per the instructions generated from HTG EdgeSeq RUO library calculator. The 12.5 pM PhiX control was prepared by denaturing and diluting ten-nanomolar (10 nM) PhiX control library V3 (Illumina), referring to Illumina MiSeq System Denature and Dilute Libraries Guide, Document number 15039740 v10. Ten-nanomolar PhiX control was first diluted to 4 nM using 10 mM Tris-Cl, pH 8.5 with 0.1% Tween 20, and denatured using 0.2 N NaOH. The denatured 4 nM PhiX control was serially diluted to 12.5 pM using the cold HT1 buffer. The denatured 12.5 pM PhiX control was used immediately or stored at -20 °C for up to two weeks. The denatured pooled library and 12.5 pM PhiX control, and cold HT1 buffer were combined in a 2-mL tube, resulting in a final 600 µL volume loaded onto the sequencing reagent cartridge.

## **2.8 Next-Generation Sequencing of Libraries**

Next-generation sequencing was performed using the MiSeq FGx Sequencing System (Verogen, San Diego, CA) and Illumina MiSeq V3 Sequencing Kit, 150-cycle (Illumina). The reagent cartridge and flow cell in the MiSeq sequencing kit were prepared for sequencing according to the HTG EdgeSeq System User Manual (90). On the reagent cartridge, 600  $\mu$ L of the denatured library was added into the well labeled “Load Sample,” and 600  $\mu$ L of HTG EdgeSeq MiSeq Primer Mix (HTG Molecular Diagnostics, Inc.) was added into well 18, which was the well for Read 1 custom primers. The reagent cartridge, flow cell, and incorporation buffer were placed in the appropriate positions according to the instructions on the MiSeq sequencer. The sample sheets for MiSeq sequencing were obtained from the HTG Host software and uploaded to the MiSeq sequencer. The following sequencer run parameters were shown in the sample sheet and used for the sequencing runs: Cycles per read – 50, Reads - 1 (Single Read), Index reads - 2 (Dual Index), Index length - 6.

## **2.9 Data Analysis**

The sequencing data were exported from the MiSeq sequencer as FASTQ files and imported into the HTG Host Parser software (HTG Molecular Diagnostics, Inc.) for the alignment of sequencing data files to miRNA reference sequences from the miRBase v20 database. The annotation files were generated using Microsoft Excel software and the parsed sequencing data files were uploaded to HTG EdgeSeq Reveal software Version v4.0.1 (HTG Molecular Diagnostics, Inc.) for quality control, correlation study, and

normalization of the parsed data. A correlation study for samples originating from the same donor or positive controls within the same run was performed to show the intra-plate reproducibility. If samples from different runs were analyzed together, a correlation study for samples originating from the same donor or positive controls used in different runs was performed to show the inter-plate reproducibility of the assay. The Pearson and Spearman correlation acceptance criteria were equal to or higher than 0.85.

HTG EdgeSeq Reveal software was also used for miRNA differential expression analysis and generation of heat maps. On HTG EdgeSeq Reveal software, the parsed data was normalized using the counts per million method. The differential expression analysis was performed using the DESeq2 method (98). Differential expression analyses on HTG EdgeSeq Reveal software were limited to pairwise comparison. A reference group was selected if more than two groups were analyzed. Each additional group was compared to the reference group, and an overall variation test was performed across all groups. The top up- and down-regulated genes were provided by the software based on the fold change of gene expression and adjusted probability value (p-value). A smaller adjusted p-value indicates a more statistically significant difference. The samples on the heat map were clustered using the Complete or Average method (99).

The parsed data was also normalized and analyzed using RStudio software (Posit, Boston, MA) for creating three-dimensional PCA scatter plots. The sequencing run parameters, including the cluster density, percentage of clusters passing filter, percentage of aligned PhiX control, and Q30 score, were verified on the Illumina Sequencing Analysis Viewer (SAV) (Illumina).

### **3. RESULTS AND DISCUSSION**

#### **3.1 MicroRNA Expression Profiling for Liquid and Dried Biofluid Samples**

##### **3.1.1 Data Quality of Samples from Three Runs**

In the first run, 24 samples underwent analysis using HTG EdgeSeq miRNA WTA, including normal semen, saliva, menstrual blood, azoospermic semen, vaginal fluid samples, and UHR RNA positive controls. The samples were prepared in triplicate because the HTG EdgeSeq Reveal software requires a minimum of three replicates for differential expression analysis. Menstrual blood and azoospermic semen were provided in liquid form, while vaginal fluid was collected as dried biofluid stains on the swabs. Normal semen and saliva were examined in both liquid and dried forms. Dark clots were observed in two menstrual blood replicates during the sample preparation. Library quantification results revealed that the PCR amplification for these two clotted menstrual blood samples did not yield adequate amplified products for library pooling and sequencing. Consequently, two unsuccessful menstrual blood samples were excluded from sequencing. The remaining 22 samples were combined into a library for next-generation sequencing.

In the second run, 24 samples were analyzed, including vaginal fluid, menstrual blood, peripheral blood swabs, swab mixtures of vaginal fluid and peripheral blood, and MTC positive controls. Vaginal fluid and peripheral blood swabs were supplied by three donors, while menstrual blood swabs originated from a single donor. The mixture was prepared using the vaginal fluid swab and peripheral blood from an identical donor, and this donor's single-source peripheral blood sample was included in the second run for comparative purposes. Clotting was not observed in the peripheral blood samples during

this run. Quantification results demonstrated that all samples produced ample amplified products for library pooling. Thus, 24 samples were pooled and sequenced.

In the third run, 24 samples were examined, including triplicates of six different biofluid mixtures, a liquid oligospermic semen sample, and the MTC positive control. Quantification results indicated that all samples were suitable for pooling and sequencing.

The initial step in statistical analysis using the HTG EdgeSeq Reveal software entails evaluating the quality of parsed sequencing data. Three quality control (QC) metrics were employed, designated as QC0, QC1, and QC2. The QC0 metric discerns samples with degraded RNA or suboptimal quality and quantity RNA by examining the percentage of sequencing reads distributed to the internal positive controls. Samples exhibiting inferior quality and quantity RNA have a high percentage of reads directed at these controls. Conversely, samples with high-quality RNA should have fewer than 14% of the reads for internal positive controls. The QC1 metric detects samples with insufficient sequencing reads, specifically below 500,000. The QC1 cut-off was established at 500,000 reads since samples exceeding this threshold demonstrated improved sample repeatability for low-expressed genes. The QC2 metric identifies samples with low expression variability across probes by calculating a relative standard deviation using the number of reads from the probes. A small relative standard deviation corresponds to a sample with minimal variation across the probes, indicating S1 nuclease inhibition or sample issues.

Optimal-quality samples should satisfy all three QC metrics. However, the HTG EdgeSeq miRNA WTA Verification Report (89) states that data from samples failing the QC1 metric may still be usable, albeit with the caveat that the repeatability of low-

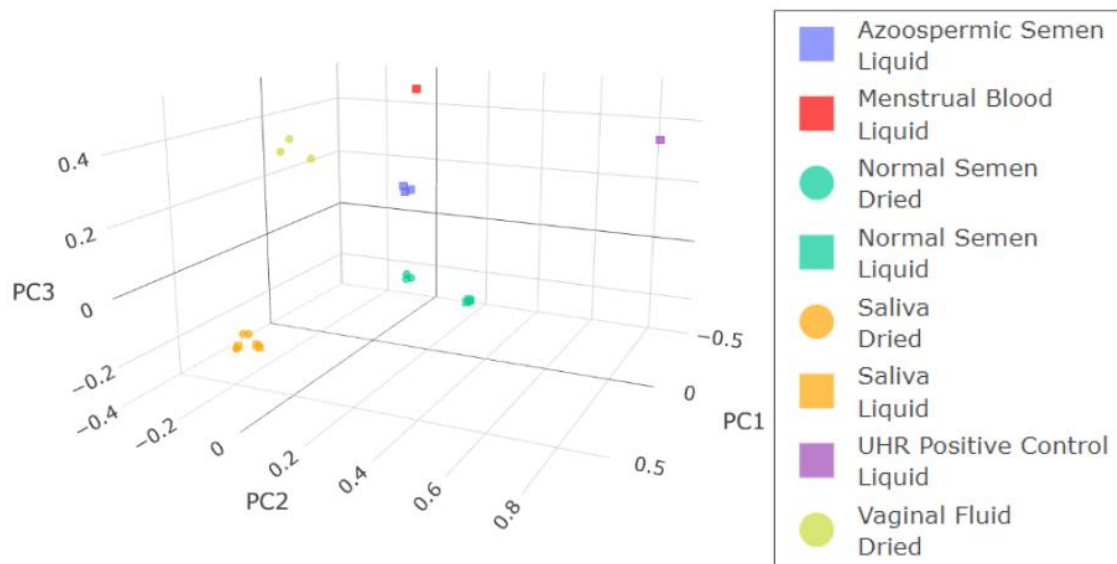
expressed genes could be impacted. In this study, samples with fewer than 500,000 but more than 200,000 reads were considered acceptable and included in the statistical analysis. The QC1-failed samples we included had at least two replicates passing all three QC metrics. The sequencing runs with failed samples demonstrated adequate sequencing quality parameters when assessed on the Illumina SAV. We recommend obtaining optimal-quality data in future studies to avoid complications with sample repeatability.

For the first run, three samples passed QC0 and QC2 but failed QC1. Failed samples encompassed two UHR RNA positive control replicates, with approximately 43,000 and 174,000 reads, respectively. Another failed sample was one replicate of the azoospermic semen sample, with approximately 280,000 reads. For the second run, 24 samples passed all three QC metrics and were included in the data analysis. For the third run, two samples failed QC1, including one replicate of vaginal fluid and oligospermic semen mixture with approximately 420,000 reads and one replicate of saliva and semen mixture with approximately 380,000 reads. Other samples passed all three QC metrics. For reasons mentioned earlier, we included the failed azoospermic semen sample from the first run and two failed mixtures samples from the third run in the statistical analysis.

### 3.1.2 Comparison of miRNA Expression Profiles

The principal component analysis (PCA) is employed to visualize sample differences within a reduced-dimensional space. PCA condenses the expressions of numerous probes into a compact set of principal components that account for the totality of variation from the original data (100). Each principal component constitutes a linear

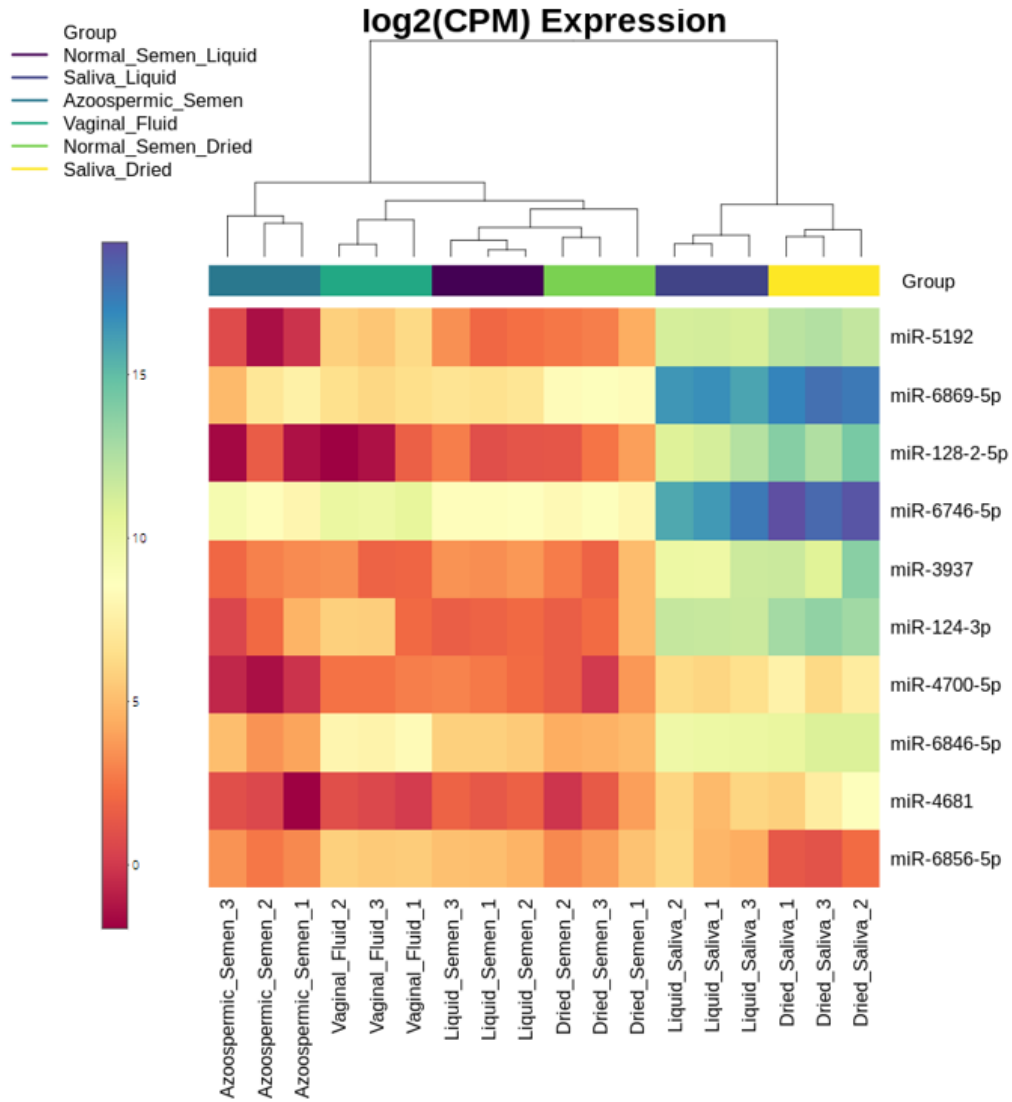
combination of probe expressions, with the first component exhibiting the most significant variation and each subsequent component representing progressively less variation. Each sample is assigned a principal component score based on these linear combinations, enabling sample comparisons across one or more components. Samples with similar component scores cluster within a multi-dimensional space.



**Figure 3. Three-dimensional principal component analysis (3D PCA) scatter plot for 20 samples from the first run.** Created with RStudio software.

A three-dimensional (3D) PCA scatter plot was generated for 20 samples from the first run, with each point in the plot representing a replicate from the experiment (Figure 3). Figure 3 features two shapes corresponding to dried and liquid samples, while different colors denote the control and various biofluid types. Replicates of samples clustered in Figure 3, signifying that the HTG EdgeSeq miRNA WTA analysis was correctly conducted, and a substantial degree of consistency was observed between replicates. Normal semen, saliva, azoospermic semen, menstrual blood, and vaginal fluid samples were separately grouped, demonstrating differential expression of miRNAs among biofluid

types. For normal semen and saliva, their dried and liquid forms occupied adjacent spaces, indicating that HTG EdgeSeq miRNA WTA yielded similar miRNA expression profiles for the same biofluid in both its dried and liquid forms.



**Figure 4. Heat map of 18 samples from the first run.** Samples included liquid normal semen, liquid saliva, azospermic semen, vaginal fluid, dried normal semen, and dried saliva samples. Created with HTG EdgeSeq Reveal software.

The differential expression analysis for data from the first run was performed using HTG EdgeSeq Reveal software, with dried saliva samples serving as the reference group. The top ten differentially expressed miRNAs among vaginal fluid, normal semen, saliva, and azoospermic semen were selected by the software to create a heat map of 18 samples from the first run (Figure 4). The UHR RNA positive control and menstrual blood samples were excluded because the differential expression analysis required at least three replicates, and these samples only had one successfully sequenced replicate. The heat map portrays normalized gene expression data through color hue and intensity, facilitating the simultaneous visualization of gene expression patterns across samples and probes. As depicted in Figure 4, samples were organized in alignment with their respective biofluid types. Notably, for normal semen and saliva samples, both their dried and liquid forms displayed similar miRNA expression patterns, resulting in their close proximity within Figure 4.

Since all dried biofluid samples from three runs produced analyzable sequencing data, the HTG EdgeSeq miRNA WTA protocol demonstrated compatibility with dried biofluid stains on cotton swabs. The PCA scatter plot and heat map of samples from the first run demonstrated consistent miRNA differential expression profiles between the liquid and dried forms of the same biofluid, indicating that the sample preparation protocol for dried biofluid samples did not noticeably affect the miRNA expression profiling.

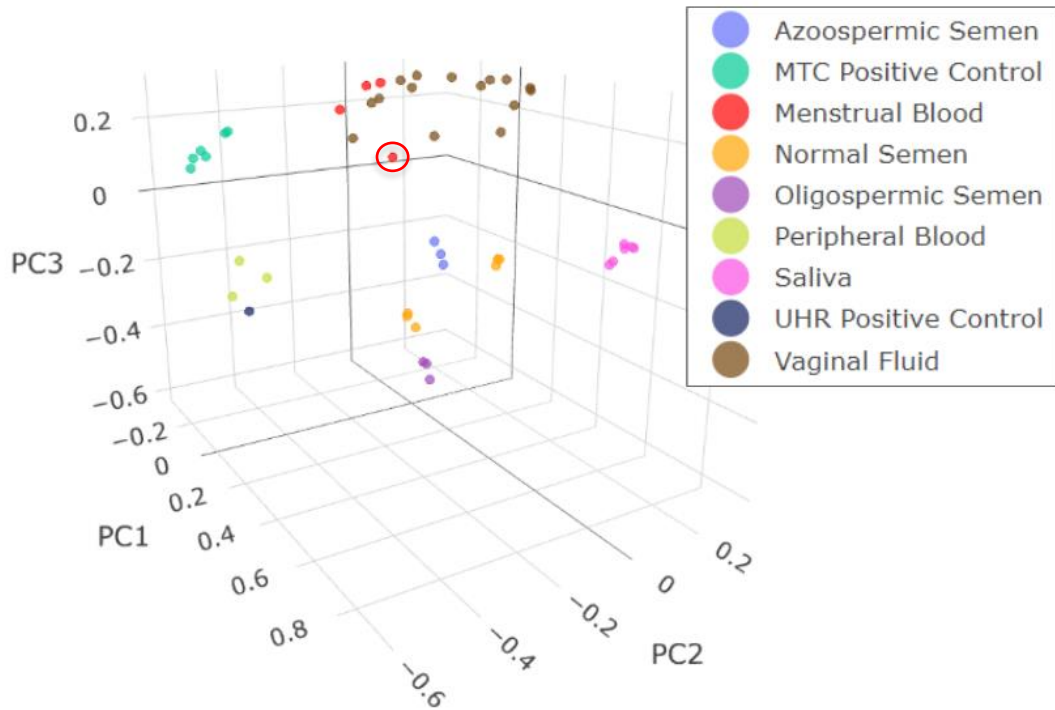
The ability to effectively analyze dried biofluids on cotton swabs constitutes a critical characteristic for forensic biofluid identification assays, considering that dried biofluid swab samples represent the predominant form of biofluid evidence encountered at

crime scenes. In this study, the HTG EdgeSeq miRNA WTA utilized half of a swab, corresponding to approximately 15  $\mu$ L of liquid biofluid, for generating a miRNA expression profile of over 2,000 miRNAs that potentially confirms the identity of a biofluid.

### **3.2 Single-source Biofluid Identification**

All single-source biofluid samples and controls from three runs were incorporated into the data analysis for this section. Seven biofluid types were examined: normal semen, azoospermic semen, oligospermic semen, menstrual blood, peripheral blood, saliva, and vaginal fluid. Samples were annotated solely based on biofluid types. Therefore, menstrual blood, saliva, and normal semen samples, regardless of liquid or dried forms, were annotated as menstrual blood, saliva, and normal semen, respectively. Correlation studies were conducted on the HTG EdgeSeq Reveal software to demonstrate sufficient intra- and inter-plate reproducibility across the three runs for their combined analysis.

### 3.2.1 Principal Component Analysis



**Figure 5. Three-dimensional principal component analysis (3D PCA) scatter plot for 47 single-source biofluid samples and positive controls.** The point circled in red is the menstrual blood sample collected using the menstrual cup, while other menstrual blood samples were collected using swabs. Created with RStudio software.

A 3D PCA scatter plot for 47 single-source samples and positive controls was generated, with each point corresponding to a replicate (Figure 5). Samples were color-coded based on their respective biofluid types. The clustering of MTC positive controls in Figure 5 underscored the assay's high reproducibility within and across different runs.

As displayed in Figure 5, vaginal fluid samples appeared more dispersed but remained confined to a separate space from other biofluid types. Vaginal fluid samples were collected from three donors at different times. Therefore, small variations of miRNA expression profiles between vaginal fluid samples may be attributed to the differences in

individuals and collection times. A study has shown that miRNA expression in vaginal fluid significantly varied throughout the menstrual cycle and between individuals (101).

Swab-collected menstrual blood samples were closely grouped with vaginal fluid samples. Conversely, the liquid menstrual blood sample collected using a menstrual cup (circled in red in Figure 5) was located between vaginal fluid and peripheral blood samples. This discrepancy may be ascribed to the collection techniques. Swabbing the vagina for menstrual blood collection inevitably amasses more vaginal epithelial cells and vaginal fluid components compared to collecting the liquid menstrual blood via a menstrual cup. Menstrual blood directly collected from victims of sexual assaults mirrors the swab-collected menstrual blood. In contrast, menstrual blood found at the crime scene may more closely resemble the menstrual blood collected using menstrual cups as it is deposited in liquid form and collected without directly swabbing the vaginal area.

Azoospermic, oligospermic, and normal semen samples occupied a similar space. The primary variation between these semen samples is the number of spermatozoa present. Azoospermic semen has no spermatozoa, while oligospermic semen has a low sperm count. Nonetheless, these semen samples all contain biofluids secreted by male reproductive organs, such as seminal vesicles and prostates, which constitute a significant fraction of semen (102). The similarity of the miRNA expression patterns between different semen types could be attributed to the shared miRNAs in the biofluid portion of semen. Identifying differentially expressed miRNAs for spermatozoa and biofluids in semen might aid in distinguishing semen samples with varying sperm counts.

Peripheral blood samples from different donors were successfully amplified, sequenced, and clustered together using PCA, demonstrating that the HTG EdgeSeq miRNA WTA protocol is compatible with whole blood samples collected with cotton swabs. This finding implies that HTG EdgeSeq miRNA can be utilized on bloodstain swabs, circumventing the necessity for plasma or serum extraction, thereby streamlining the process if implemented in forensic laboratories.

Figure 5 exhibits clustering of most biofluid types at unique locations, thereby highlighting the potential of HTG EdgeSeq miRNA WTA to identify and classify different biofluid types based on their distinctive miRNA expression profiles. However, additional data points are crucial to differentiate semen with varied sperm counts and to determine whether swab-collected menstrual blood can be separated from vaginal fluid.

### 3.2.2 Differential Expression Analysis

The differential expression analysis was performed using HTG EdgeSeq Reveal software. The pairwise comparison revealed that, on average, approximately 786 miRNAs displayed differential expression between two biofluid types. The most significant difference was observed between saliva and vaginal fluid samples, with 1668 differentially expressed miRNAs. The comparison between menstrual blood and vaginal fluid samples offered the smallest difference, with 146 miRNAs exhibiting differential expression.

**Table 1. Top up-regulated and down-regulated miRNAs observed in peripheral blood (PB), menstrual blood (MB), normal semen (Sem), saliva (Sal), vaginal fluid (VF), azoospermic semen (AS), and oligospermic semen (OS).** The miRNAs in the table are significantly up-regulated or down-regulated compared to the biofluid types listed in the parentheses; the common biofluid-specific miRNAs between our study and the existing literature are indicated in bold.

Biofluid Type	Up-regulated miRNAs	Down-regulated miRNAs
Peripheral Blood	4306, <b>185-5p</b> (MB, Sem, Sal, VF, AS, OS) <b>451a</b> , <b>486-5p</b> (Sem, Sal, VF, AS, OS) 4732-3p (MB, Sem, AS, OS) 1180-3p (MB, Sem, VF, AS) 92a-3p (MB, Sal, VF) 25-3p (MB, Sal, OS) 320a, 320b, 320c, 320e (Sem, Sal, OS) <b>144-3p</b> (VF, AS) <b>126-3p</b> (VF) <b>223-3p</b> (Sem)	200a-3p, <b>200b-3p</b> , 200c-3p (MB, Sem, AS) 141-3p, 339-3p, 671-5p (Sem, AS, OS) 1285-5p (MB, Sal, VF) 375, 128-1-5p (MB, VF) 3714 (AS, OS)
Menstrual Blood	<b>451a</b> , <b>486-5p</b> (Sem, Sal, VF, AS, OS) 147b (PB, Sem, Sal, AS) <b>144-3p</b> (Sem, VF, AS, OS) 612, 492 (Sem, AS, OS) 124-3p (PB, Sem, AS) <b>106a-5p</b> , <b>126-3p</b> , <b>16-5p</b> , <b>185-5p</b> , <b>20a-5p</b> (VF) 375, 34a-5p (PB)	4306, 185-5p (PB) 648, 4522, 6080 (AS, OS) 4697-5p (Sal, OS) 147a (VF)

Saliva	184, <b>583</b> , 4736, 128-2-5p, 3605-5p (MB, PB, Sem, VF, AS, OS) 6869-5p (MB, PB, Sem, VF, AS) 6787-5p (MB, PB, Sem, VF) 6802-5p (MB, Sem, VF, AS) 6511a-5p (PB, VF, AS, OS) 5192 (PB, AS, OS)	3197 (MB, Sem, VF, AS) 125b-5p (MB, Sem, VF) 451a, 486-5p (MB, PB) 99a-5p (MB, Sem) 612 (MB, VF) 648, 193a-3p (AS, OS) let-7c-5p, 10b-5p, 29a-3p (Sem)
Vaginal Fluid	147a (MB, PB, Sem, Sal, AS, OS) 612, <b>193b-5p</b> (PB, Sem, Sal, AS, OS) 147b (MB, PB, Sem, Sal, AS) 4496 (Sal, AS, OS) 320a, 320c (Sem, Sal, OS) 4653-3p (MB, AS, OS) <b>124-3p</b> (Sem)	126-3p, <b>144-3p</b> , 451a, 486-5p (MB, PB) 339-3p (Sem, AS, OS) 363-3p (Sem, AS) 648 (AS, OS)
Normal Semen	6824-5p (MB, PB, Sal, VF) 6782-5p, <b>888-5p</b> , <b>891a-5p</b> , <b>891b</b> (MB, PB, VF, AS) 363-3p, <b>26b-5p</b> , <b>29b-3p</b> (MB, Sal, VF) <b>671-5p</b> (PB, Sal) <b>890</b> , <b>891a-3p</b> , <b>892a</b> , <b>892b</b> (AS) <b>10b-5p</b> , <b>3197</b> (Sal)	124-3p (MB, Sal, VF) 451a, 486-5p (MB, PB) 612, 147b (MB, VF) 4429, 320a, 320b, 320c (PB, VF) 367-5p, 3681-5p (AS, OS) <b>144-3p</b> (MB)

Azoospermic Semen	648, 4522, 6080, 339-3p (MB, PB, Sal, VF) 4793-5p, 619-3p (MB, Sal, VF) 3169, 4664-3p (MB, PB, VF) 497-3p, 619-3p (MB, VF) <b>671-5p</b> (PB, Sal) <b>3197</b> (Sal)	144-3p, 451a, 486-5p (MB, PB) 612, 147a, 147b (MB, VF) 890, 888-5p, 891a-2p, 891a-5p, 891b, 892a, 892b (Sem)
Oligospermic Semen	648, 4519, 4522, 619-3p (MB, PB, Sal, VF) 6080 (MB, Sal, VF) 3169, 6770-3p (PB, Sal, VF) 3681-5p (MB, Sem, VF) 497-3p (MB, VF) <b>671-5p</b> (PB, Sal) <b>943</b> (AS)	451a, 486-5p (MB, PB) 320a, 320c (PB, VF) 612, 4496, 193b-5p (MB, VF) 3674, 1255b-2-3p, 1285-5p, 548d-5p (Sem, AS)

Table 1 enumerates the miRNAs that were consistently highly expressed (up-regulated), and those that were consistently lowly expressed (down-regulated) in one biofluid compared to other types of biofluids. The differentially expressed miRNAs identified in our study were compared to the biofluid-specific miRNAs reported in the existing literature to discern similarities and discrepancies.

### *3.2.2.1 Previously Reported miRNAs in Peripheral and Menstrual Blood*

For peripheral blood, several highly expressed miRNAs have been previously identified as peripheral blood markers in the literature, including miR-185-5p (63,72), miR-451a (47,70,72,73,75,85), 486-5p (71,72,85), 144-3p (63,73,85), 126-3p (70,72), and 223-3p (103). Multiple highly expressed miRNAs were shared between peripheral and menstrual blood, including miR-185-5p, 451a, 486-5p, 144-3p, and 126-3p.

Our study revealed a significant up-regulation of miR-185-5p in peripheral blood compared to all other types of biofluids. This observation aligns with the research conducted by Zubakov et al. (63), wherein they detected elevated expression levels of miR-185 in peripheral blood relative to all other evaluated biofluids, including menstrual blood. In our study, peripheral blood exhibited a 90-fold increase in miR-185 expression compared to menstrual blood (adjusted p-value =  $1.10 \times 10^{-8}$ ). As a result, miR-185-5p holds the potential for differentiating between peripheral and menstrual blood.

On the contrary, miR-451a, 486-5p, and 144-3p demonstrated high expression in both peripheral and menstrual blood, with no significant difference in expression levels detected between these two kinds of blood in our research. Various studies have suggested that miR-451a (47,77), 486-5p (71), and 144-3p (73) are valuable markers for all types of blood samples, rendering these miRNAs unsuitable for differentiating between peripheral and menstrual blood. Landgraf et al. (103) proposed that miR-144 is highly specific to hematopoietic cells, the precursor cells with the potential to differentiate into various blood cell types, such as white and red blood cells. This cellular specificity could explain the increased expression level of miR-144 in both peripheral and menstrual blood.

In certain research, miR-144-3p has been recognized as a menstrual blood marker (81,87). Nevertheless, studies by Zubakov et al. (63) and Sauer et al. (73) revealed that menstrual blood displayed a slightly decreased expression of miR-144 than peripheral blood. Our study is consistent with these findings, demonstrating that menstrual blood exhibited a 1.3-fold reduced expression of miR-144-3p relative to peripheral blood (adjusted p-value = 0.852). Thus, miR-144-3p can distinguish menstrual blood from other biofluid types but not from peripheral blood.

Seashols-Williams et al. (84) discovered the ubiquitous presence of miR-451a in a variety of biofluid types they tested, including blood, menstrual blood, saliva, semen, vaginal fluid, urine, feces, and sweat. Similarly, Liu et al. (87) identified miR-451a as a reference gene expressed in multiple biofluids. Our study verified the presence of miR-451a in all biofluid samples examined, although its expression level was substantially elevated in blood samples in comparison to non-blood samples. We observed, on average, a 92,000-fold increase in miR-451a expression in peripheral blood and a 10,600-fold increase in miR-451a expression in menstrual blood compared to other biofluid types. As evidenced by our research, miR-451a appears to be a robust marker closely affiliated with blood rather than a mere reference gene.

MicroRNA-126-3p was previously identified as a peripheral blood marker by Courts and Madea (70) and Park et al. (72). Our results indicated that miR-126-3p displayed significantly elevated expression in both peripheral and menstrual blood compared to vaginal fluid. It is worth noting that Courts and Madea (70) and Park et al. (72) did not include menstrual blood samples in their investigations. Based on the insights

from our research, miR-126-3p could serve as a marker for all blood samples, including peripheral and menstrual blood, a possibility that has not been previously reported.

In addition to up-regulated miRNAs in peripheral blood, Seashols-Williams et al. (84) pinpointed miR-200b-3p as a potential blood marker due to its pronounced down-regulation in peripheral blood relative to other biofluids. Our study similarly demonstrated that miR-200b-3p expression in peripheral blood was significantly lower than in menstrual blood, semen, and azoospermic semen. Consequently, miR-200b-3p may function as a peripheral blood marker characterized by its consistent down-regulation.

Compared to vaginal fluid, several microRNAs substantially up-regulated in menstrual blood have been identified as peripheral blood markers, including miR-16-5p (47,72,85), 20a-5p (63,72), and 106a-5p (63). Wang et al. (71) observed an elevated expression of miR-16 in both peripheral and menstrual blood. Conversely, Sirker et al. (77) suggested miR-16 as a menstrual blood marker, as their research revealed heightened miR-16 expression in menstrual blood compared to peripheral blood. Contrary to Sirker et al.'s findings, our results demonstrated a 9-fold increase in miR-16-5p expression in peripheral blood relative to menstrual blood, albeit not statistically significant. MicroRNA-20a-5p exhibited increased expression levels in peripheral blood compared to menstrual blood in a study conducted by Zubakov et al. (63). This observation aligns with our research, which revealed a 3-fold increase in miR-20a-5p expression in peripheral blood relative to menstrual blood (adjusted p-value = 0.398). Despite the slightly elevated expression levels of miR-16-5p and 20a-5p in peripheral blood, they did not rank among

the most up-regulated miRNAs of peripheral blood in our study. Hence, they were not included in the peripheral blood section of Table 1.

While Hanson et al. (81) described miR-124 as a saliva marker, Sirker et al. (77) suggested that miR-124 exhibited a strong association with both vaginal fluid and menstrual blood. Our findings detected markedly increased expression levels of miR-124-3p in menstrual blood relative to normal and azoospermic semen and in vaginal fluid compared to normal semen. Therefore, our observations align with Sirker et al.'s conclusions, indicating miR-124-3p as a marker for vaginal fluid and menstrual blood.

#### *3.2.2.2 Previously Reported miRNAs in Saliva and Vaginal Fluid*

Although our study identified several miRNAs that consistently demonstrated increased expression levels in saliva, only miR-583 has been previously mentioned as a potential saliva marker (63). In their study, Zubakov et al. initially selected miR-583 as a saliva marker through microarray screening, but its specificity to saliva was later refuted by RT-qPCR analysis. In our analysis, miR-583 exhibited a significantly elevated expression level, with an average 323-fold increase in expression, compared to all other biofluids. Consequently, we claim that miR-583 represents a robust marker linked to saliva.

In the case of vaginal fluid, miR-124-3p, as noted earlier, was highlighted as a marker for both vaginal fluid and menstrual blood in a study conducted by Sirker et al. (77). Likewise, Hanson et al. (81) pinpointed miR-124a as a vaginal fluid marker. Furthermore, Sauer et al. (73) incorporated miR-124-3p into their decision tree algorithm

to differentiate vaginal fluid from other biofluid types, although it was not categorized as an exclusive marker for vaginal fluid. In concordance with these findings, our study revealed substantially increased expression levels of miR-124-3p in vaginal fluid compared to normal semen. Another previously identified vaginal fluid marker is miR-193b-5p (87). Our study indicated that miR-193b-5p exhibited significantly heightened expression levels in vaginal fluid than other biofluid types, except for menstrual blood.

In addition to up-regulated miRNAs, Zubakov et al. (63) illustrated that miR-144-3p displayed low expression in semen, saliva, and vaginal fluid. This finding is consistent with our study, which showed that miR-144-3p was down-regulated in vaginal fluid and semen compared to menstrual blood.

### *3.2.2.3 Previously Reported miRNAs in Normal, Azoospermic, and Oligospermic Semen*

In previous studies, miR-29b-3p (48), 26b-5p (84), and 10b-5p (47,73,77) have been identified as semen markers. Our study corroborated these findings, demonstrating elevated expressions of these miRNAs in normal semen relative to other biofluid types.

Park et al. (72) identified miR-3197 as a semen marker, and our study revealed high expression levels of miR-3197 in normal and azoospermic semen compared to saliva. Nevertheless, we observed that miR-3197 expression in saliva was substantially lower than in semen, azoospermic semen, menstrual blood, and vaginal fluid. As a result, our findings imply that miR-3197 is down-regulated in saliva compared to other biofluids rather than being exclusively up-regulated in semen.

Weber et al. (48) demonstrated that miR-671-5p exhibited high expression levels in semen. In line with Weber et al.'s findings, our analysis indicated that miR-671-5p was up-regulated in normal, azoospermic, and oligospermic semen compared to peripheral blood and saliva. As discussed earlier, semen comprises spermatozoa and biofluids secreted by male reproductive organs. Our observation suggests that miR-671-5p may serve as a marker for the biofluid component of semen rather than for spermatozoa, given its up-regulation in semen samples with varying sperm counts.

Zubakov et al. (63) identified miR-943 as a semen marker. Our study exhibited an expression level of miR-943 in oligospermic semen that was 6 times higher than in azoospermic semen (adjusted p-value =  $2.12 \times 10^{-8}$ ). In contrast, its expression in normal semen was 1.37 times lower compared to azoospermic semen (adjusted p-value = 0.567). We deduce that the expression level of miR-943 lacks correlation with the sperm count in the sample. The difference in expression levels between oligospermic and azoospermic semen is not substantial and might be attributed to individual variations.

Various studies have reported miR-888-5p (71,87,103), 890 (87), 891a-3p (63,71–73,84), 891a-5p (63,71–73,84,87), 891b (103), 892a (48,81), and 892b (103) as semen markers. Our research discovered that miRNA-888-5p, 891a-5p, and 891b displayed elevated expression levels in normal semen compared to vaginal fluid, peripheral blood, and menstrual blood. Notably, all these semen markers demonstrated significantly increased expression in normal semen compared to azoospermic semen. Landgraf et al. (103) suggested that miR-888, 891a, 891b, and 892b were specifically expressed in the epididymis, the duct connected to the testicle for spermatozoa maturation and storage

(104). Hence, we hypothesize that miR-888-5p, 891a, 891b, 892a, and 892b are overexpressed in spermatozoa, leading to their high expression levels in normal semen compared to azoospermic semen.

While the azoospermic semen was purchased commercially, the normal and oligospermic semen samples were collected in our laboratory. Microscopic examination was conducted on the oligospermic semen samples to confirm a low sperm count, but the precise number of spermatozoa in the samples remained undetermined. Further research integrating sperm counting experiments alongside miRNA expression profiling of semen samples with varying sperm counts would be beneficial for identifying the miRNAs specific to spermatozoa and biofluids in semen.

#### *3.2.2.4 Discrepancies With the Literature*

Some differentially expressed miRNAs identified in our research exhibit discrepancies with existing literature. Saliva, in particular, displayed the most inconsistent findings when compared to previous studies using liquid and dried saliva samples. Wang et al. (85) identified several miRNAs as saliva markers, such as miR-141-3p, 375, 34a-5p, let-7c-5p, 125b-5p, 99a-5p, and 29a-3p. However, our study indicated that none of these miRNAs were significantly up-regulated in saliva; some even demonstrated down-regulation relative to other biofluids. For instance, in our study, semen exhibited a 131-fold increased expression of miR-let-7c-5p compared to saliva (adjusted p-value =  $1.27 \times 10^{-12}$ ). Similarly, menstrual blood showed a 139-fold increased expression (adjusted p-value =  $2.11 \times 10^{-16}$ ), and normal semen displayed a 141-fold increased expression of miR-

99a-5p compared to saliva (adjusted p-value =  $4.88 \times 10^{-17}$ ). Furthermore, saliva revealed a reduced expression of miR-125b-5p than menstrual blood, normal semen, and vaginal fluid, with an average of a 256-fold decrease.

Another inconsistency between our study and existing literature involves miR-320c. Seashols-Williams et al. (84) illustrated that miR-320c was abundantly expressed in feces while exhibiting relatively low expression levels in peripheral blood. However, our study demonstrated that peripheral blood and vaginal fluid have high expression levels of miR-320c compared to normal semen, saliva, and oligospermic semen. A possible explanation for this discrepancy could be the absence of feces samples in our study. Encompassing a broader range of biofluid types in future research could delve deeper into the biofluid specificity of miR-320c.

Discrepancies among biofluid-specific miRNAs discovered in different studies, including ours, may be attributed to the diverse biofluid types included in the research. For instance, Wang et al. (85) only incorporated peripheral blood and saliva samples in their analysis. As a result, the specificity of saliva markers they observed was in relation to peripheral blood and not any other biofluid types. It is possible that other untested biofluid types exhibit more substantially different expression levels of these miRNAs. Moreover, identifying specific markers for saliva and vaginal fluid has historically proven challenging (63,71–73), potentially due to significant individual variations and heterogeneous collection conditions. Differentiating saliva from vaginal fluid is further complicated by the similar cell types present in oral and vaginal cavities, such as epithelial cells (72). Thus, the biofluid-specific miRNAs identified in our study may be restricted by the biofluid types

we examined. Forensic scientists continue to broaden the scope of biofluid types and increase sample sizes to seek unambiguous miRNAs for biofluid identification.

#### *3.2.2.5 Functions of Identified Biofluid-Specific miRNAs*

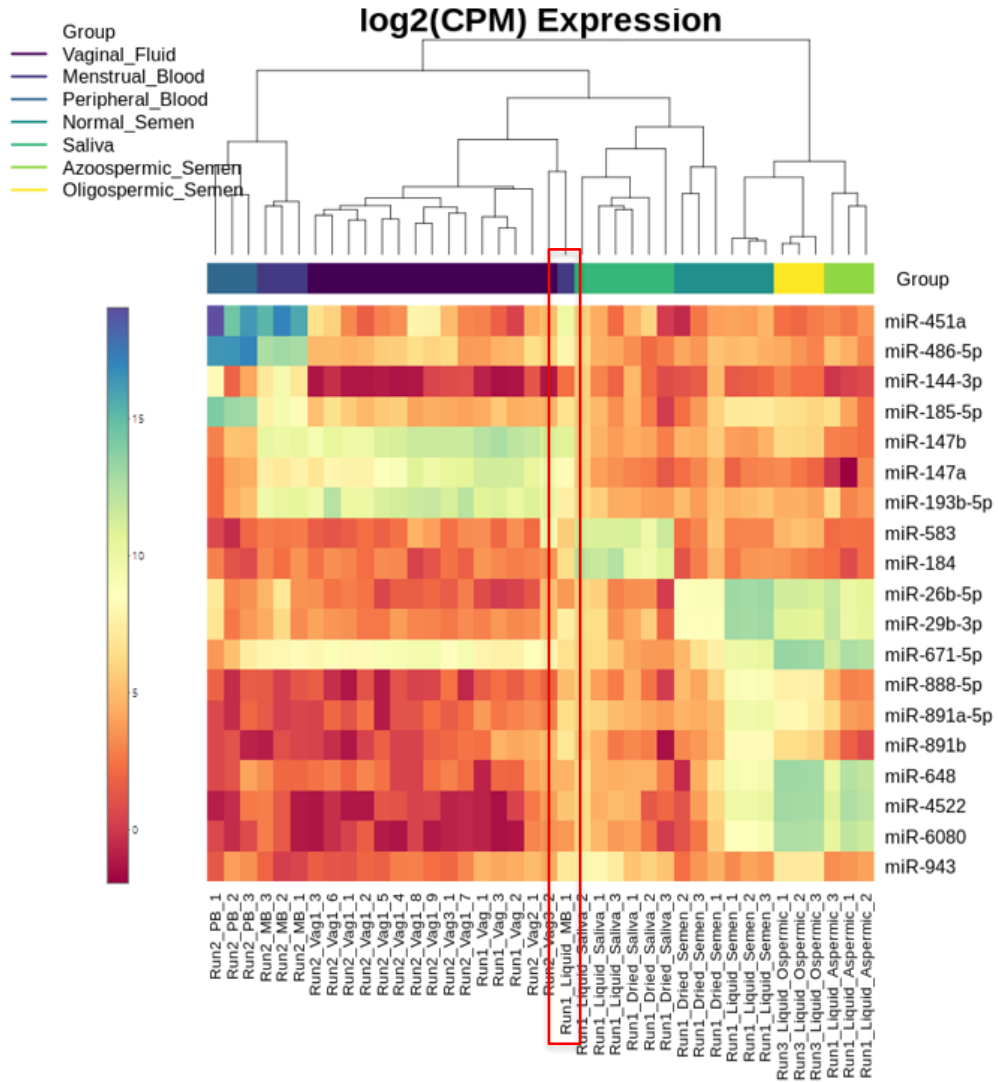
Several miRNAs with biofluid-specific expression patterns have not been documented in existing biofluid identification studies we have reviewed. To attempt to explain their biofluid specificity, we conducted a literature search to understand the functions of these miRNAs.

For peripheral blood, miR-4306 exhibited up-regulation compared to all other biofluid types, while miR-1180-3p demonstrated increased expression in peripheral blood relative to four other biofluid types. These miRNAs have been associated with lung cancer cell lines (105,106). MicroRNA-147b, found to be up-regulated in menstrual blood and vaginal fluid in our study, has been observed to display elevated expression levels in ovarian cancer tissues and cell lines (107). The specificity of miR-147b to ovarian tissues may be correlated with its unique expression pattern for menstrual blood and vaginal fluid. Regarding up-regulated miRNAs in saliva, miR-184 and 4736 have been linked to pancreatic cancer cell lines (108,109). Azoospermic and oligospermic semen shared several up-regulated miRNAs, including miR-648 and miR-497-3p. MicroRNA-648 has been identified as a regulator for vasoconstrictor endothelin-1 (110), while miR-497-3p has been associated with ovarian cancer cell lines (111).

No apparent relationship exists between the roles these miRNAs play in cancer prognosis or regulation, as evidenced by clinical studies, and their specificity in different

biofluids. It is crucial to note that our study was conducted with a limited number of samples from a few individuals. To validate the biofluid specificity of the miRNAs we have identified through our differential expression analysis, it would be necessary to undertake a comprehensive study encompassing a wider array of biofluid types and a larger sample size.

### 3.2.3 Heat Map of Selected miRNAs



**Figure 6. Heat map of 19 most differentially expressed miRNAs in 40 single-source biofluid samples.** Samples include peripheral blood, menstrual blood, normal semen, saliva, vaginal fluid, azoospermic semen, and oligospermic semen samples. The sample marked by the red square is the outlier menstrual blood sample. Created with HTG EdgeSeq Reveal software.

A heat map of 40 single-source biofluid samples was generated utilizing nineteen miRNAs that offered the most explicit differentiation of seven biofluid types (Figure 6). MicroRNA-451a, 486-5p, and 144-3p were chosen to distinguish blood samples from non-blood samples. MicroRNA-185-5p was employed to segregate peripheral blood from all

other biofluid types, including menstrual blood. MicroRNA-147b was utilized to differentiate menstrual blood and vaginal fluid from other biofluids, while miR-147a and 193b-5p were selected to differentiate vaginal fluid from other biofluid types. MicroRNA-583 and 184 exhibited high expression levels in saliva compared to other biofluids. MicroRNA-26b-5p, 29b-3p, 671-5p, 888-5p, 891a-5p, and 891b were incorporated to differentiate semen samples from non-semen samples. MicroRNA-888-5p, 891a-5p, and 891b also functioned as markers to distinguish normal semen from azoospermic semen. MicroRNA-648, 4522, and 6080 were significantly up-regulated in both azoospermic semen and oligospermic semen. MicroRNA-943, with elevated expression in oligospermic semen, was employed to differentiate oligospermic semen from azoospermic semen, as they shared similar miRNA expression patterns.

As illustrated in Figure 6, the majority of biofluid samples were correctly categorized according to their biofluid types, except for one menstrual blood sample (marked by the red square in Figure 6), which was grouped with vaginal fluid instead of the other menstrual blood samples. Given the presence of vaginal mucus and vaginal epithelial cells in the menstrual blood, it is explainable that menstrual blood exhibits a similar miRNA expression pattern to that of vaginal fluid. The outlier menstrual blood sample was the liquid menstrual blood analyzed in the first run, where two other replicates were not sequenced due to PCR amplification failure. The protocol was modified for the subsequent two runs to optimize the HTG EdgeSeq miRNA WTA workflow for the whole blood samples. The adapted protocol yielded high-quality data for blood samples in the second and third runs. This technical issue may account for the inconsistent miRNA

expression pattern and misclassification of the liquid menstrual blood sample from the first run. Otherwise, HTG EdgeSeq miRNA WTA has proven its efficacy in identifying and classifying biofluids using the biofluid-specific miRNAs discovered in the differential expression analysis.

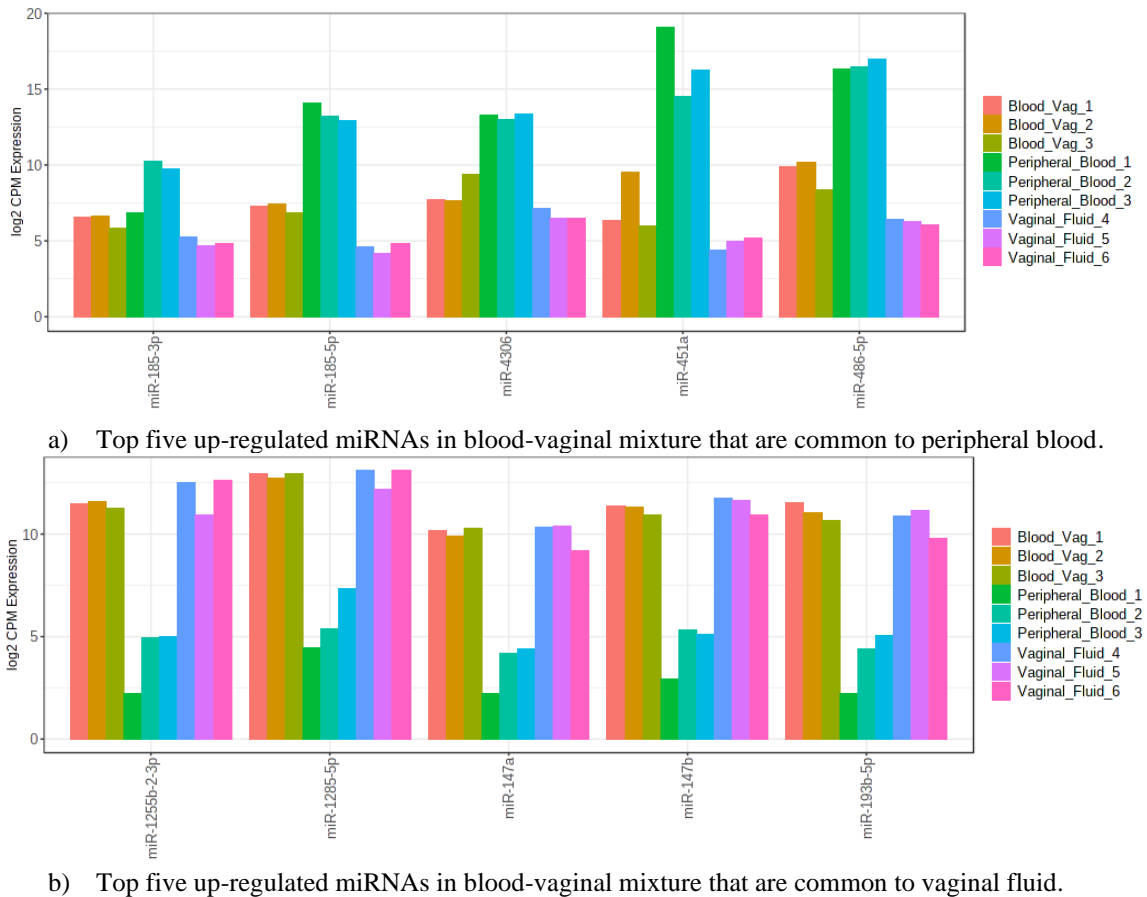
The consistency of biofluid-specific miRNA expression patterns across replicates and various donors demonstrates the high reproducibility of the HTG EdgeSeq miRNA WTA between samples and plates. Notably, peripheral blood and vaginal fluid samples exhibited remarkably similar miRNA expression patterns across samples from three distinct donors. Future research utilizing HTG EdgeSeq miRNA WTA could benefit from incorporating a larger number of donors to substantiate the biofluid-specific miRNA expression patterns across individuals.

### **3.3 Analysis of Mixture Samples**

Seven distinct biofluid mixtures were prepared in the second and third runs, each comprising two of the seven biofluid types. The combinations included vaginal fluid and peripheral blood, vaginal fluid and normal semen, vaginal fluid and oligospermic semen, vaginal fluid and saliva, menstrual blood and normal semen, saliva and normal semen, and saliva and oligospermic semen. We examined the miRNA expression profiles of biofluid mixtures and compared them to the single-source biofluid samples. The objective was to determine whether the mixtures exhibited differentially expressed miRNAs similar to single-source biofluids.

The differential expression analysis was conducted individually for each type of mixture and their corresponding single-source biofluid samples. To establish potential correlations, the up-regulated miRNAs in the mixture were identified and compared with the highly expressed miRNAs in the single-source samples. The expression profiles of up-regulated miRNAs were generated using the HTG EdgeSeq Reveal software to visualize the miRNA expression levels in mixed and single-source biofluids.

### 3.3.1 Mixtures of Peripheral Blood and Vaginal Fluid



**Figure 7. Expression profiles of top five up-regulated miRNAs in the mixture of peripheral blood and vaginal fluid that are common to a) peripheral blood and b) vaginal fluid.** Created with HTG EdgeSeq Reveal software.

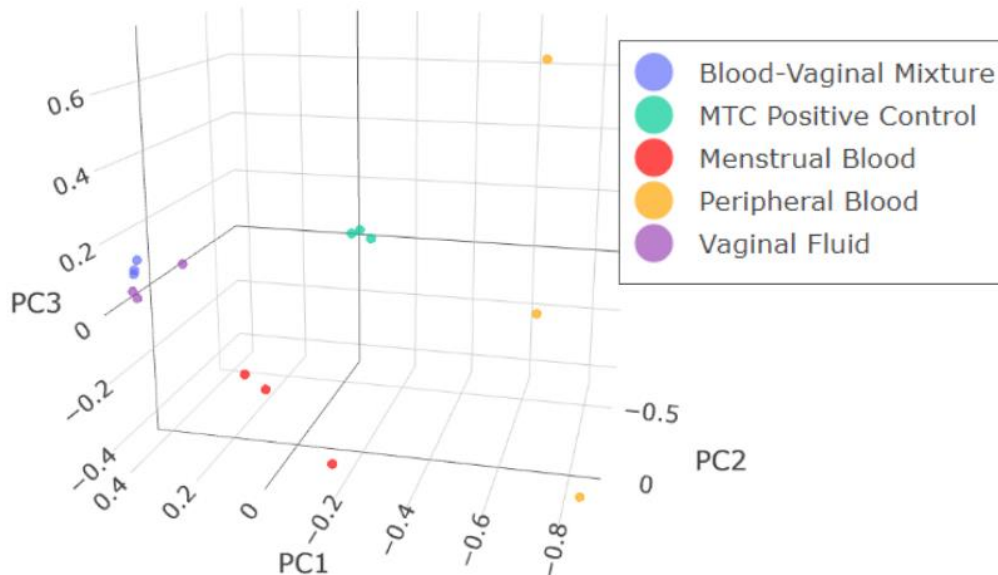
The differential expression analysis and visualized expression profiles (Figure 7) illustrated that several significantly up-regulated miRNAs in the mixture of peripheral blood and vaginal fluid (referred to as the blood-vaginal mixture) correspond to the most differentially expressed miRNAs in vaginal fluid and peripheral blood. These miRNAs include miR-185-5p, 451a, 486-5p, 4306, 1255b-2-3p, 1285-5p, 147a, 147b, and 193b-5p. This finding suggests that the biofluid mixture contains differentially expressed miRNAs derived from both sources.

As depicted in Figure 7, the up-regulated miRNAs shared between the blood-vaginal mixture and vaginal fluid displayed similarly high expression levels. However, the up-regulated miRNAs shared between the blood-vaginal mixture and peripheral blood did not reach the same elevated expression levels in the mixture as they did in peripheral blood. This observation indicates that the blood-vaginal mixture bears a closer resemblance to vaginal fluid than peripheral blood in terms of miRNA expression. However, this disparity may arise from the way we prepared the blood-vaginal mixture. We applied a known quantity of peripheral blood to a collected vaginal swab, but the exact volume of vaginal fluid on the swab was unknown. Additionally, the vaginal fluid may contain a greater number of cells than peripheral blood, which may contribute to a more substantial portion of miRNA expression. Despite this potential bias, HTG EdgeSeq miRNA WTA retains the capacity to identify differentially expressed miRNAs in the blood-vaginal mixture that align with the highly expressed miRNAs found in single-source vaginal fluid and peripheral blood.

The blood-vaginal mixture piques particular interest, as its composition resembles menstrual blood, which inherently consists of peripheral blood and vaginal fluid. Therefore, we were intrigued to determine if HTG EdgeSeq miRNA WTA could differentiate menstrual blood from a blood-vaginal mixture. At crime scenes, the co-occurrence of vaginal fluid and peripheral blood may result from bleeding caused by injury to the vaginal area. We identified five up-regulated miRNAs in menstrual blood compared to the blood-vaginal mixture: miR-200b-3p, 144-3p, 16-5p, 125b-5p, and 99a-5p.

MicroRNA-16-5p and 144-3p were previously recognized as menstrual blood markers in existing literature (47,71,77,87). However, our study showed that menstrual blood exhibited slightly lower expression levels of these miRNAs than peripheral blood. The significantly low expression of miR-16-5p and 144-3p in a mixture containing peripheral blood might be due to the biases in preparing blood-vaginal mixtures, as mentioned earlier. MicroRNA-200b-3p, formerly identified as down-regulated in peripheral blood in our study and by Seashols-Williams et al. (84), also displayed decreased expression in the blood-vaginal mixture. MicroRNA-125b-5p and 99a-5p exhibited considerable up-regulation in menstrual blood compared to saliva, but the difference was insignificant compared to other biofluids, as shown by our study.

These results demonstrate the potential to differentiate menstrual blood from blood-vaginal mixtures using specific miRNAs. In future studies, differentially expressed miRNAs could be identified in other components of menstrual blood that are not shared with vaginal fluid or peripheral blood to facilitate the distinction.



**Figure 8. Three-dimensional principal component analysis (3D PCA) scatter plot for blood-vaginal mixtures, vaginal fluid, and peripheral and menstrual blood.** Created with RStudio software.

A 3D PCA plot was generated utilizing blood-vaginal mixtures, vaginal fluid, menstrual blood, and peripheral blood to visually represent miRNA expression patterns in mixed and single-source biofluids (Figure 8).

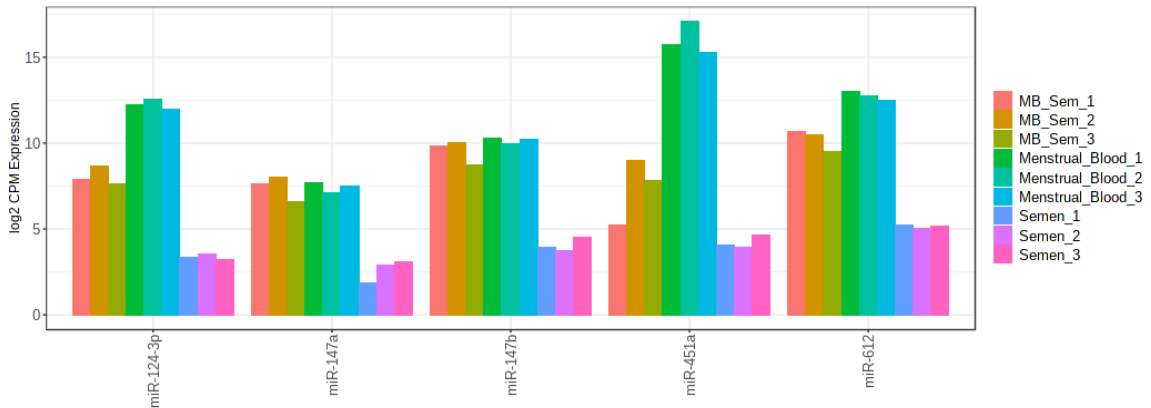
As illustrated in Figure 8, blood-vaginal mixtures demonstrated a greater affinity with vaginal fluid than peripheral blood. This observation reinforces the notion that the miRNA expression patterns of blood-vaginal mixtures more closely resembled those of vaginal fluid, possibly due to biases in sample preparation.

Intriguingly, menstrual blood samples occupied an intermediate position between vaginal fluid and peripheral blood samples. This observation is congruent with the concept that menstrual blood is a natural mixture of vaginal fluid and peripheral blood, containing one set of differentially expressed miRNAs shared with vaginal fluid and another set shared with peripheral blood.

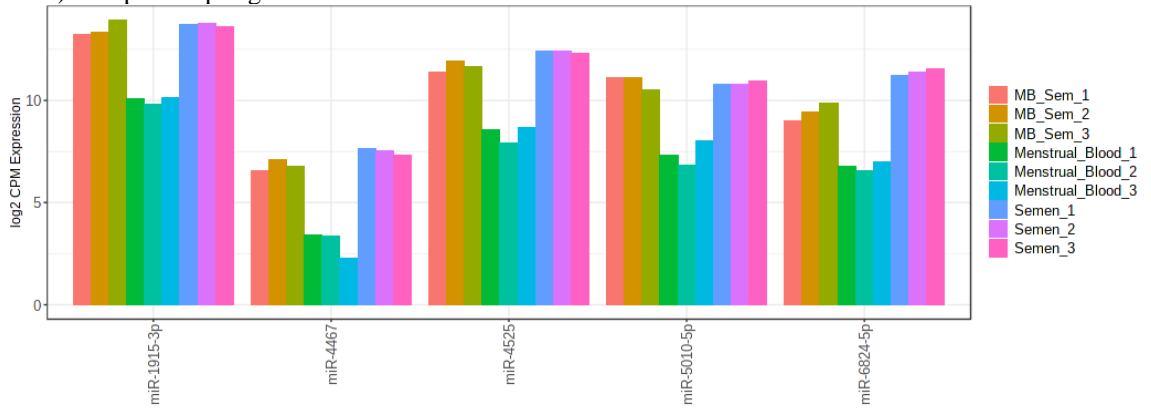
Peripheral blood samples exhibited varied miRNA expression patterns between samples. As previously explained, the first component of PCA analysis reveals the most substantial variation, with the following components displaying successively less variation. Thus, principal component 1 (PC1) exhibited the most variation, while PC3 had the least among the three principal components. As seen in Figure 8, peripheral blood samples displayed variations only along the Z-axis, corresponding to PC3, which had less variation. The subtle variation between peripheral blood samples could be ascribed to individual differences since they were collected from three donors.

### 3.3.2 Mixtures of Menstrual Blood and Semen

The mixture comprising menstrual blood and normal semen (referred to as the menstrual-semen mixture) is intriguing because menstrual blood is a natural mixture of peripheral blood and vaginal fluid. Consequently, the menstrual-semen mixture embodies a confluence of three discrete biofluid types.



a) Top five up-regulated miRNAs in menstrual-semen mixture that are common to menstrual blood.



b) Top five up-regulated miRNAs in menstrual-semen mixture that are common to semen.

**Figure 9. Expression profiles of top five up-regulated miRNAs in the mixture of menstrual blood and normal semen that are common to a) menstrual blood and b) normal semen.** Created with HTG EdgeSeq Reveal software.

Our differential analysis and expression profiles (Figure 9) revealed that several highly expressed miRNAs in the menstrual-semen mixture were common to up-regulated miRNAs in menstrual blood and vaginal fluid, including miR-124-3p, 147a, 147b, and 612. These miRNAs exhibited markedly elevated expression levels in menstrual blood and menstrual-semen mixtures compared to normal semen, as shown in Figure 9.

MicroRNA-451a, a prominently expressed miRNA in peripheral and menstrual blood, also demonstrated high expression in menstrual-semen mixtures compared to semen. However, Figure 9 illustrates that the expression level of miR-451a was relatively

lower in the menstrual-semen mixture than in single-source menstrual blood. Other highly expressed miRNAs identified in blood samples did not exhibit substantially differential expression patterns in menstrual-semen mixtures.

The menstrual-semen mixture's decreased expression of blood-related miRNAs compared to menstrual blood could be attributed to the fact that, despite concurrently being collected from the same donor, different menstrual blood swabs were utilized for mixture preparation and single-source biofluid analysis. Therefore, the quantity and composition of menstrual blood deposited on swabs may differ, leading to variations in the expression levels of blood-related miRNAs.

Differentiating a mixture containing multiple biofluid types presents a challenge for biofluid identification assays. The complexity of mixture resolution escalates when different biofluids possess vastly disparate proportions in a mixture sample, which may result in disproportionate miRNA expression. Incorporating additional biofluid-specific miRNAs or establishing a decision tree strategy with multiple miRNAs may be necessary to recognize different biofluids in a mixture using miRNA profiling methods.

### 3.3.3 Other Mixtures

The remaining mixtures were prepared using vaginal fluid, saliva, normal semen, and oligospermic semen. A summary of the up-regulated miRNAs shared between mixtures and single-source biofluids is provided in Table 2.

**Table 2. Up-regulated miRNAs shared between mixture and single-source biofluids, including vaginal fluid, saliva, normal semen, and oligospermic semen.**

Types of Mixtures	Up-regulated miRNAs Shared between Mixture and Single-source Biofluids			
	Vaginal Fluid	Saliva	Normal Semen	Oligospermic Semen
Vaginal-Saliva Mixture	320c, 320e, 612, 3197, 4429	184, 583, 4736, 3605-5p, 6802-5p	/	/
Vaginal-Semen Mixture	147a, 147b, 193b-5p, 612, 4257	/	21-5p, 26b-5p, 29c-3p, 141-3p, 200c-3p	/
Vaginal-Oligospermic Mixture	147a, 193b-5p, 320c, 4701-3p, 4496	/	/	885-3p, 4665-5p, 5010-5p, 6131, 6870-3p
Saliva-Semen Mixture	/	184, 583, 4736, 6746-5p, 6802-5p	3197, 6131, 6821-5p, 6824-5p, 6875-5p	/
Saliva-Oligospermic Mixture	/	184, 583, 4701-3p, 4736, 4784	/	885-3p, 3197, 6780b-5p, 6821-5p, 6875-5p

Previous analysis of single-source biofluids indicated that vaginal fluid and saliva displayed remarkable consistency in miRNA expression across replicates and plates. In the case of vaginal fluid, differentially expressed miRNAs were also consistent across individuals. MicroRNAs such as miR-147a, 147b, and 612 manifested elevated expression in single-source vaginal fluid and mixture samples containing vaginal fluid. Similarly,

miR-184, 583, and 4736 were highly expressed in single-source saliva and mixture samples incorporating saliva. Furthermore, miR-3197 persistently demonstrated down-regulation in saliva, whether presented individually or within a mixture. It is critical to note that only the top five up-regulated miRNAs were selected. Therefore, minor variations in common up-regulated miRNAs were seen in Table 2.

For normal and oligospermic semen, the differentially expressed miRNAs appeared contingent on the biofluid type employed for comparison. Up-regulated miRNAs in normal and oligospermic semen did not exhibit significantly increased expression levels in mixture samples. Future studies should involve additional samples collected from a greater number of individuals to identify miRNAs with expression patterns specific to semen.

An interesting observation was made with mixtures comprising saliva and normal semen (referred to as saliva-semen mixtures). A miRNA commonly associated with blood, miR-451a, was up-regulated in the saliva-semen mixture compared to single-source saliva and normal semen. The expression level of miR-451a in the saliva-semen mixture increased 55-fold compared to saliva (adjusted p-value =  $4.0 \times 10^{-5}$ ) and 37-fold compared to normal semen (adjusted p-value =  $1.9 \times 10^{-4}$ ). A possible explanation is that the saliva or normal semen utilized to prepare the saliva-semen mixture contained a minuscule amount of blood, which was detected through miRNA expression profiling. As a supplementary experiment, a catalytic color test or lateral-flow immunoassay could be performed on the saliva and normal semen samples used for mixture preparation to confirm the presence of blood. However, the blood quantity is likely so minimal that it falls below

the detection limit of commonly employed methods for blood detection and might only be identified by a nucleic acid-based assay with exceptional sensitivity.

The HTG EdgeSeq miRNA WTA demonstrated the ability to identify differentially expressed miRNAs shared between mixtures and single-source biofluids. Hence, this assay could serve as a tool to resolve biofluid mixtures if appropriate miRNAs are incorporated into the analysis. Comprehensive investigations involving more biofluid types and individuals could be undertaken to select and validate biofluid-specific miRNAs for developing a mixture differentiation assay.

### **3.4 Comparison to Preliminary Study**

We submitted seven types of biofluids, including peripheral blood, menstrual blood, saliva, normal semen, azoospermic semen, breast milk, and urine, to HTG for a preliminary study of applying HTG EdgeSeq miRNA WTA to forensic biofluid identification. The entire process, encompassing experimentation and data analysis, was conducted by the researchers at HTG. The reports generated by HTG researchers demonstrated that HTG EdgeSeq miRNA WTA identified approximately 1,600 miRNAs with expression differentiated among seven different liquid biofluid types. As shown by PCA scatter plots and heat maps generated using the experimental data, samples were categorized into their corresponding biofluid types based on miRNA differential expression (97).

A total of 29 liquid biofluid samples, including two repeated samples, were submitted to HTG for two separate runs conducted in August and November 2021. In the

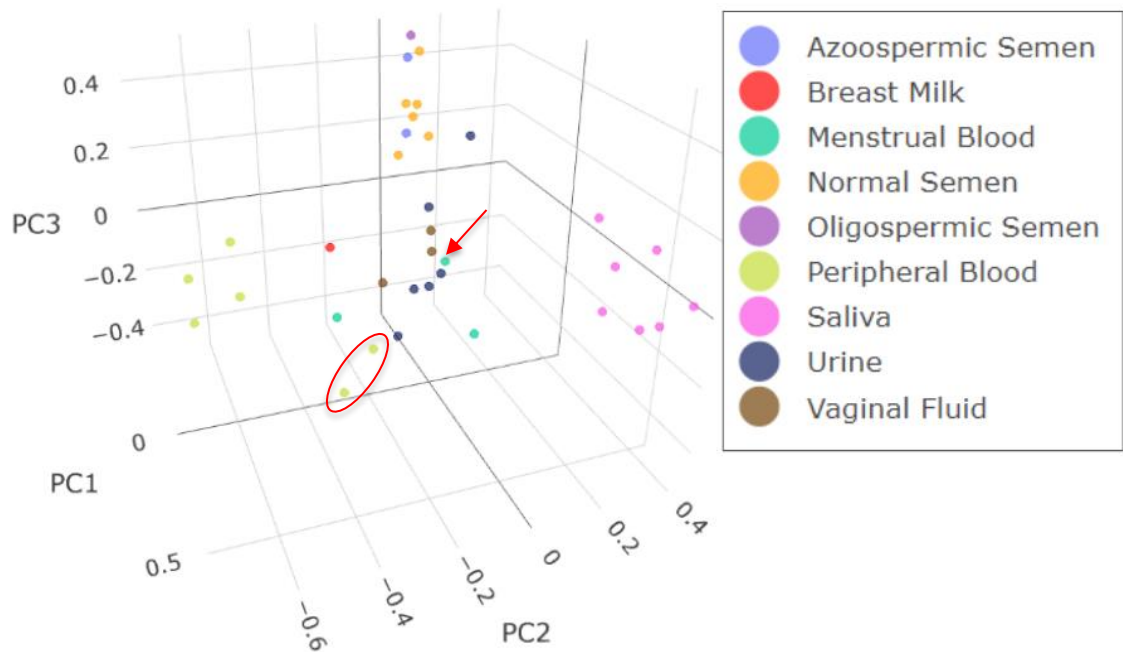
August run, one of three peripheral blood samples failed to produce sufficient amplified products for library pooling and was therefore excluded from the sequencing. It was confirmed that the peripheral blood samples sent to HTG in August did not contain ethylenediaminetetraacetic acid (EDTA), a common blood preservative that prevents clotting. HTG researchers assumed that the absence of blood preservatives caused the PCR amplification failure of the liquid peripheral blood sample.

The unsuccessful peripheral blood sample was reacquired from the same donor, stored in a container with EDTA, and submitted to HTG for the November run, accompanied by three other peripheral blood samples preserved in EDTA. The previously failed peripheral blood sample was successfully amplified and sequenced in the November run. However, all three additional peripheral blood samples experienced PCR amplification failure and were omitted from the sequencing. The detailed workflow performed by HTG researchers on the peripheral blood samples remains undisclosed, but at that time, the HTG EdgeSeq miRNA WTA appeared to encounter difficulties with whole blood samples.

In addition to the three unsuccessful peripheral blood samples, one semen sample in the August run failed QC2 when assessed on HTG EdgeSeq Reveal software. A QC2 failure indicates minimal expression variability, potentially stemming from S1 nuclease inhibition or an issue with the sample. The identical semen sample was resubmitted to HTG for the November run but persistently exhibited QC2 failure, implying the problem originated from the sample itself.

### 3.4.1 Principal Component Analysis

Twenty-three successfully sequenced samples with acceptable data quality were included in the analysis. The biofluid samples were collected from different donors, and no replicates were prepared for the preliminary study. We obtained the raw experimental data from HTG and incorporated it into the current data to examine and compare the miRNA expression patterns observed in the preliminary and current studies. For the samples from the current study, the miRNA expression levels were averaged among replicates to include only one data point for each donor's sample, better synchronizing with the preliminary study. Twelve data points from the current study were combined with the 23 data points from the preliminary study to generate a 3D PCA scatter plot (Figure 10).



**Figure 10. Three-dimensional principal component analysis (3D PCA) scatter plot for samples from preliminary and current studies.** Samples included azoospermic semen, breast milk, menstrual blood, oligospermic semen, peripheral blood, saliva, normal semen, urine, and vaginal fluid. Points circled in red are two peripheral blood samples in the August and November runs conducted by HTG researchers. The point indicated by the red arrow is the liquid menstrual blood sample submitted to HTG for the August run. Created with RStudio software.

As depicted in Figure 10, most peripheral blood samples congregated in a similar region, except for two samples from the August and November runs conducted by HTG researchers (circled in red in Figure 10). We remain uncertain regarding their workflow for whole blood samples; however, their protocol appeared ill-suited for such specimens, as three liquid peripheral blood samples from different donors failed to amplify in the runs they performed. In contrast, all three peripheral blood samples examined in our study yielded sufficient amplified products and high-quality sequencing results, subsequently clustering together in Figure 10.

Beyond the disparities in workflow, another notable distinction lies in the sample forms: the peripheral blood samples sent to HTG were in liquid form, while those examined in our study were obtained using swabs. The blood samples submitted to HTG for the August run lacked preservatives, which may have contributed to the amplification failure due to clotting or degradation. A similar issue occurred with the liquid menstrual blood samples in our first run, as the liquid menstrual blood formed dense clots and failed to amplify. However, the peripheral blood samples submitted to HTG for the November run were preserved with EDTA to mitigate the blood clotting issue. Yet, three EDTA-preserved samples experienced PCR amplification failure. The blood volume in both liquid and swab samples should be the same, as we gauged the amount of blood to deposit on the cotton swabs. Our adapted sample preparation protocol may optimize HTG EdgeSeq miRNA WTA for handling whole blood samples. Alternatively, the HTG EdgeSeq miRNA WTA may have yielded superior results with swab-collected whole blood samples rather than liquid blood samples due to factors such as clotting.

Menstrual blood samples exhibited a dispersed distribution between peripheral blood and vaginal fluid samples. As previously discussed, menstrual blood constitutes a natural mixture of peripheral blood and vaginal fluid, containing differentially expressed miRNAs from both biofluids. Variability in collection methods and individual discrepancies could contribute to the diverse miRNA expression profiles observed in menstrual blood samples. Interestingly, the liquid menstrual blood sample collected using a menstrual cup in the August run (indicated by the red arrow in Figure 10) grouped most closely with the vaginal fluid samples. In contrast, our study revealed that menstrual blood collected with a menstrual cup displayed a more pronounced demarcation from vaginal fluid compared to swab-collected samples. Based on the observation from the current study, we postulate that swabbing the vaginal area may accumulate more vaginal fluid components and vaginal epithelial cells. However, this hypothesis is not supported by the outcome of the preliminary study. Additional data points are required to ascertain whether this result is more related to differences in collection techniques or variations among individuals.

Urine samples displayed a scattered arrangement, potentially owing to the composition of urine, which predominantly consists of water and a negligible quantity of cells. The urine composition also significantly varies among individuals, as it is subject to diet, hydration, activity level, and environmental conditions. As a result, identifying a specific marker or unique miRNA expression pattern for urine could prove challenging.

In contrast, saliva, another biofluid with a highly varied composition, demonstrated evident clustering separate from all other biofluid types in Figure 10. This observation

indicates that miRNA expression patterns in saliva are markedly specific compared to other biofluids examined. Implementing miRNA expression profiling for forensic biofluid identification to discern saliva could prove advantageous, as saliva is among the most frequently encountered biofluids at crime scenes. Moreover, most current detection methods for saliva are non-confirmatory and may exhibit cross-reactivity with multiple biofluids, rendering miRNA expression profiling a more reliable option.

Semen samples tended to cluster and were not distinctly separated based on sperm counts. Although discerning semen samples from non-semen samples might be relatively simple using a select few differentially expressed miRNAs in semen, distinguishing between normal, azoospermic, and oligospermic semen may necessitate employing additional miRNAs specific to spermatozoa and biofluids in semen.

A single breast milk sample was included in the analysis, demonstrating unique miRNA expression patterns from other biofluids. Additional breast milk samples are required to identify miRNAs specifically or differentially expressed in breast milk.

Vaginal fluid was not examined in the preliminary study conducted by HTG but was included when we performed the HTG EdgeSeq miRNA WTA in our laboratory. Vaginal fluid samples exhibited a relatively high degree of similarity among themselves, and a distinct separation compared to semen, saliva, and peripheral blood samples.

### 3.4.2 Common Differentially Expressed miRNAs

**Table 3. Common differentially expressed miRNAs in samples from preliminary and current studies, including saliva, and peripheral and menstrual blood, and normal and azoospermic semen.**

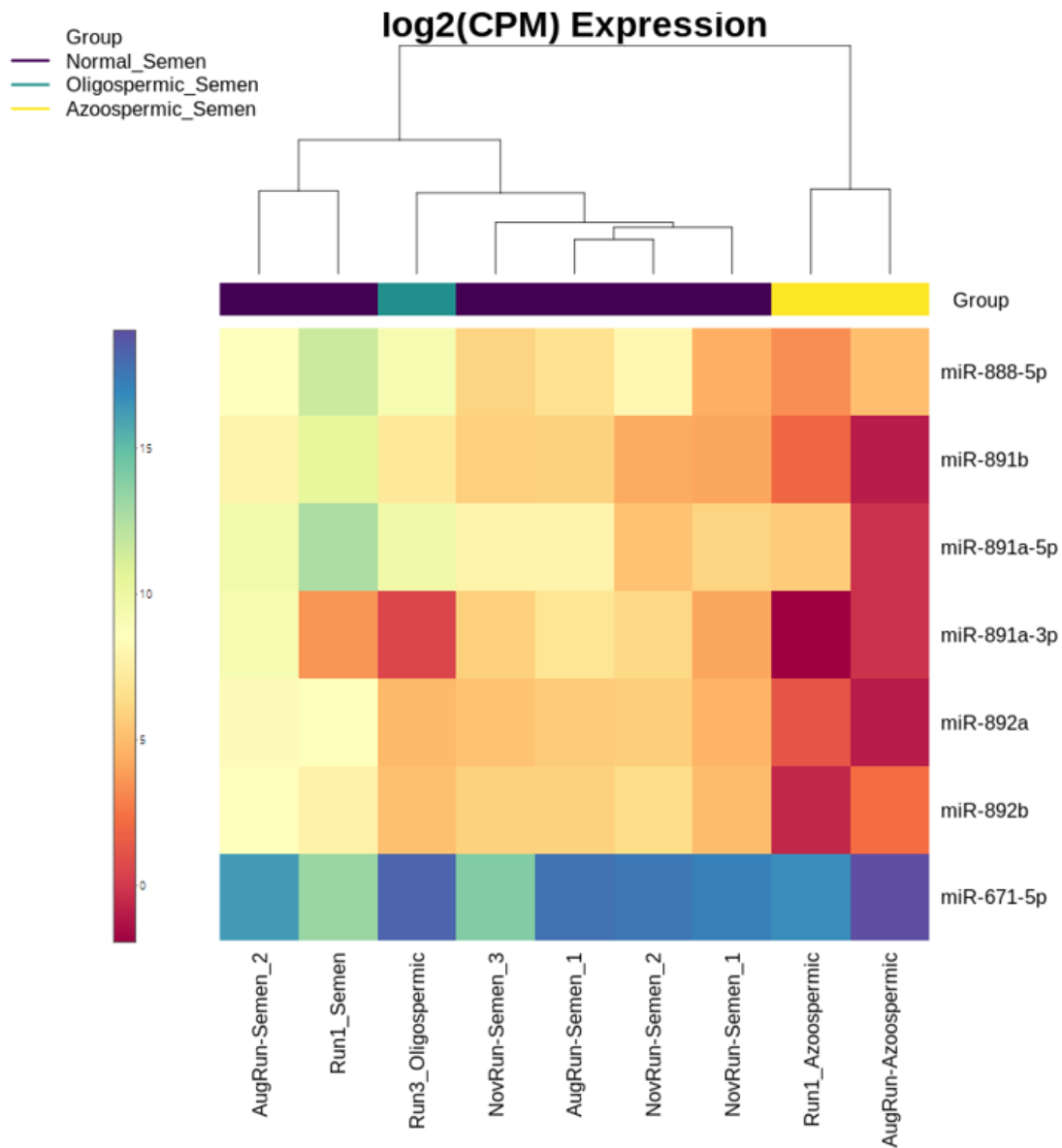
	<b>Common Differentially Expressed miRNAs</b>
<b>Peripheral Blood</b>	Up-regulated: 451a, 486-5p, 185-3p, 185-5p, 144-3p, 144-5p, 4306  Down-regulated: 200a-3p, 200b-3p, 200c-3p
<b>Menstrual Blood</b>	Up-regulated: 451a, 486-5p, 144-3p, 124-3p, 612, 492, 147b, 99a-5p
<b>Saliva</b>	Up-regulated: 184, 583, 128-2-5p, 4736, 3605-5p, 6802-5p, 6511a-5p
<b>Normal Semen</b>	Up-regulated: 10b-5p, 26b-5p, 29b-3p, 888-5p, 891a-3p, 891a-5p, 891b, 892b
<b>Azoospermic Semen</b>	Up-regulated: 648, 671-5p, 3169, 4664-3p, 6080  Down-regulated: 891a-3p (compared to semen)

A comprehensive differential expression analysis was conducted on the biofluid types examined in both preliminary and current studies, including peripheral blood, menstrual blood, saliva, normal semen, and azoospermic semen. The analysis identified several differentially expressed miRNAs in each biofluid type that are consistent with findings from the current study. Table 3 lists common differentially expressed miRNAs between preliminary and current studies.

MicroRNA-451a, 486-5p, and 144-3p appeared to serve as markers for both peripheral and menstrual blood, while miR-185-5p and 185-3p demonstrated more

predominant expression in peripheral blood. Saliva exhibited a remarkably consistent group of up-regulated miRNAs, including miR-184, 583, 128-2-5p, 4736, 3605-5p, 6802-5p, and 6511a-5p. The high consistency of differentially expressed saliva markers is also evident in Figure 10, wherein saliva samples from diverse individuals clustered tightly at a unique location.

We observed that previously identified semen markers, including 888-5p, 891a-3p, 891a-5p, 891b, and 892b, displayed elevated expression levels in semen samples from the preliminary study. Moreover, these semen markers were not highly expressed in the azoospermic semen sample from the preliminary study, potentially suggesting that these miRNAs are more strongly associated with spermatozoa. Conversely, microRNA-671-5p, detected as up-regulated across all semen types in the current study, exhibited a substantially increased expression in azoospermic semen samples from both preliminary and current studies. This observation implies that miRNA-671-5p functions as a marker for biofluids in semen rather than spermatozoa, given its elevated expression in semen samples regardless of sperm counts. The significant up-regulation of miR-671-5p in azoospermic semen could be attributed to individual variation or a higher proportion of biofluids. This supposition stems from the lack of spermatozoa in azoospermic semen, allowing for a larger fraction of biofluids within a fixed volume of semen.



**Figure 11. Heat map of azoospermic, oligospermic, and normal semen samples from the preliminary and current studies.** Generated using common semen markers observed in preliminary and current studies. Created with HTG EdgeSeq Reveal software.

A heat map incorporating the semen markers mentioned above was generated to visually represent miRNA expression patterns in semen samples from preliminary and current studies (Figure 11). The oligospermic semen sample from the current study was included, although oligospermic semen had not been examined in the preliminary study.

As illustrated in Figure 11, azoospermic semen samples were separated from other semen samples due to their distinctive expression patterns of the selected miRNAs. However, the oligospermic semen sample exhibited similar expression patterns to normal semen samples, resulting in its indistinguishability from normal semen samples. This resemblance between oligospermic and normal semen is understandable, considering that oligospermic semen contains spermatozoa, albeit at a relatively low level.

As shown in Figure 11, miR-671-5p was highly expressed across all semen types, especially in azoospermic semen samples. On the contrary, the semen markers presumably associated with spermatozoa, including miR-888-5p, 891a-3p, 891a-5p, 891b, and 892b, exhibited reduced expression in azoospermic semen samples, with miR-891a-3p demonstrating the most significant down-regulation. Likewise, miR-891a-3p was down-regulated in oligospermic semen samples. These findings suggest that miR-891-3p holds the greatest potential as a marker for spermatozoa. Nonetheless, the normal semen sample from the current study also displayed a relatively low miR-891a-3p expression, possibly due to individual variation. It is crucial to mention that the precise sperm counts in the sample remained unknown. Additional studies are needed to explore the correlation between expression levels of these semen-specific miRNAs and the number of spermatozoa.

The biofluid specificity of the miRNAs identified in our study is substantiated by analyzing supplementary samples from the preliminary study. Evaluating additional samples in future investigations is essential to bolster the validity of these findings.

### **3.5 Method Optimization and Troubleshooting**

#### **3.5.1 Sample Preparation Protocol for Dried Biofluid Samples**

Predominantly, biofluid samples retrieved from crime scenes present as dried stains collected on cotton swabs. The HTG EdgeSeq miRNA WTA was designed for liquid biofluids and other biological sample types routinely encountered in clinical settings. Therefore, a sample preparation protocol specifically tailored for dried biofluid stains on cotton swabs was not readily available in the user manual. In this study, we adapted the sample preparation protocol from another study conducted by HTG, which utilized HTG EdgeSeq miRNA WTA on biofluid swabs. Our laboratory conducted supplementary experiments to refine the sample preparation protocol for biofluid swabs. We determined that 60  $\mu\text{L}$  of Plasma Lysis Buffer for half of a cotton swab represented an appropriate volume, ensuring sufficient eluted product for subsequent procedures without over-diluting the biofluid within the swab.

In addition to the dried biofluid samples collected in the swab form, we prepared biofluid swabs for this study by adding 30  $\mu\text{L}$  of liquid biofluid to an entire cotton swab. Half of the swab should contain approximately 15  $\mu\text{L}$  of the biofluids, corresponding to the amount of liquid biofluid sample used for HTG EdgeSeq miRNA WTA. However, quantifying the biofluid in the swab poses a challenge for dried biofluid samples utilized in this study and in actual forensic casework. Our research demonstrated that the biofluids contained in half of a swab yielded a complete miRNA profile for over 2,000 miRNAs, verifying that such a sample size is adequate for the assay.

### 3.5.2 Optimization of Workflow

Two replicates of UHR RNA positive controls failed QC1 on the HTG EdgeSeq Reveal software in the first run. Our results suggested that UHR RNA might be unsuitable as a positive control for intra- and inter-plate correlations in HTG EdgeSeq miRNA WTA. Therefore, MTC lysates were acquired from HTG and employed as positive controls for the second and third runs. MTC positive controls exhibited highly reproducible miRNA expression patterns both within and between runs.

In the first run, two replicates of liquid menstrual blood formed dark clots in the sample preparation step, leading to PCR amplification failure. Dried menstrual blood collected using swabs was tested in the second run, and clotting was not observed during sample preparation. Following HTG researchers' advice, we purchased the Hemo KlenTaq® enzyme from New England Biolabs for PCR amplification in the second and third runs. Hemo KlenTaq® enzyme demonstrated enhanced resistance to blood inhibitors, yielding amplified products of superior quality and quantity for peripheral and menstrual blood swab samples, as shown in our study.

Quantification is a critical part of the entire protocol, as it directly impacts the quality of sequencing data. We discovered that the freshness of the KAPA library quantification kit was vital for successful quantification runs with reliable data. Empirically, utilizing the KAPA kit within a week of combining the primer mix with the master mix yielded optimal quantification run parameters, such as anticipated cycle threshold values for DNA standards and consistent outcomes among replicates.

Notwithstanding the exceptional quality parameters of our quantification runs, discrepancies arose between quantification results and sequencing cluster density. Consequently, we performed an additional quantification on the combined library before sequencing. This additional library quantification validated our pooling process and provided a more accurate concentration for the combined library, resulting in improved sequencing run performance. Ideally, quantifying the combined library should yield a concentration aligning with our expectations. If the quantification result suggests otherwise, downstream normalization of the library is adjusted accordingly to prevent under- or over-clustering the sequencing flow cell.

The process of diluting and pooling individual libraries into a combined library is susceptible to pipetting errors. For instance, normalizing a highly concentrated sample requires a considerable quantity of diluent. Generally, a one-step dilution using a large volume of diluent should be avoided, as it may result in an uncertain amount of sample incorporated into the combined library. HTG updated their library calculator and supplied it to us for the second and third runs. This latest version of the library calculator contains the instruction for a serial dilution of the individual and combined libraries, thereby reducing the risk of over-dilution when utilizing a large amount of diluent.

Upon examining the sequencing data file via Illumina SAV, we observed an uneven distribution of reads across different samples. We surmised that the biases originated from the library pooling process. Although the HTG RUO library calculator instructed users to dilute individual libraries by introducing diluents followed by samples, we found a more balanced distribution of reads when the sample was added first, succeeded by the diluent.

Our theory for this observation is that when the diluent was added subsequently, the mixing of the sample and diluent was completed using a pipette tip of a larger volume, promoting a more effective mixing of these two constituents. As a result, a more consistent concentration was achieved within each individual sample, facilitating a more even distribution of samples when they were pooled into a combined library.

Our optimization and troubleshooting efforts for the HTG EdgeSeq miRNA WTA workflow have yielded positive outcomes, with each successive run demonstrating improved sequencing run and data quality. Whole blood samples have greatly benefited from refined protocols, producing optimal quantification results and sequencing data. However, the inconsistency between quantification results and sequencing performance remains a challenge yet to be resolved. A correlation study could be conducted to understand the relationship between quantification results and sequencing cluster density for the miRNA expression profiling method.

## **4. CONCLUSIONS**

### **4.1 Findings of the Study**

Biofluid identification, as an integral aspect of forensic case investigations, has captured the attention of forensic scientists. MicroRNA has emerged as a particularly compelling area of research, given its specific expression patterns across various forensic-relevant biofluids and tissues, as well as its remarkable resistance to degradation. Numerous studies have been conducted on miRNA-based assays to identify biofluid-specific miRNA markers, with the aim of establishing a multiplex and confirmatory forensic biofluid identification panel.

In this study, we explored the potential application of HTG EdgeSeq miRNA WTA, a semi-automated, extraction-free miRNA expression profiling method developed by HTG Molecular Diagnostics, Inc., for forensic biofluid identification. Our findings demonstrate that HTG EdgeSeq miRNA WTA provided distinctive miRNA expression profiles across a diverse range of biofluid types, including saliva, normal semen, oligospermic semen, azoospermic semen, vaginal fluid, menstrual blood, and peripheral blood.

Several differentially expressed miRNAs identified herein align with existing literature, while others were novel findings that correspond with our preliminary results. Notably, miR-451a and 486-5p exhibited significant up-regulation in all blood samples, including peripheral and menstrual blood, while miR-185-5p displayed substantially increased expression in peripheral blood relative to other biofluids. Saliva samples from various individuals consistently presented up-regulated miRNAs such as miR-184 and 583. Vaginal fluid, untested in the preliminary study, revealed a unique miRNA expression

pattern consistent between individuals, as shown by the current study. Semen samples were readily distinguishable from non-semen samples, though differentiating among semen samples with variable sperm counts proved challenging. We hypothesize that miR-888-5p, 891a-3p, 891a-5p, and 891b are associated with spermatozoa, as their expression levels were significantly elevated in normal semen compared to azoospermic semen. Conversely, miR-671-5p could serve as a marker for biofluids in the semen, as it was up-regulated in all semen types regardless of sperm counts.

Furthermore, HTG EdgeSeq miRNA WTA demonstrated high reproducibility within and between plates, as demonstrated by clustering replicates and positive controls. MicroRNA expression patterns for biofluid samples collected from diverse donors were substantially consistent. A comparison of miRNA expression profiles between liquid biofluid and dried biofluid swab samples exhibited striking similarities in miRNA expression between the liquid and dried forms of the same biofluid.

Analysis of mixtures comprising two biofluid types revealed highly expressed miRNAs from both single-source biofluids, suggesting that HTG EdgeSeq miRNA WTA could indicate the biofluids present in the mixtures. However, as the number of biofluid types in a mixture rises, differentiating every biofluid within the mixture becomes increasingly complex due to common differentially expressed miRNAs and disproportionate miRNA expressions. Blood-vaginal mixtures, resembling menstrual blood in composition, are of particular interest. Our study illustrated that specific miRNAs were highly expressed in menstrual blood but not in blood-vaginal mixtures, potentially

differentiating menstrual blood from blood-vaginal mixtures encountered at crime scenes due to bleeding resulting from vaginal injuries.

The comparison between the preliminary and current studies demonstrates that the HTG EdgeSeq miRNA WTA provided miRNA expression profiles specific to various biofluid types, potentially facilitating biofluid identification and classification. Saliva, in particular, is effortlessly distinguishable from other tested biofluids. Semen samples can also be readily segregated from non-semen samples, although more specific miRNAs are needed to differentiate between semen samples with different sperm counts. Improving sample preparation and PCR amplification techniques for whole blood samples may yield more consistent expression profiles, allowing peripheral blood to display a distinctive expression pattern from other biofluids. Differentiating urine samples from other biofluids may prove challenging, given that their components are highly diverse and subject to individual variation. Additional vaginal fluid, menstrual blood, and breast milk samples are required to further illustrate biofluid-specific expression patterns in these biofluids.

HTG EdgeSeq miRNA WTA displayed remarkable compatibility with dried biofluid swab samples, remaining unaffected by the sample preparation protocol. The assay efficiently analyzed whole blood samples, eliminating the need for plasma extraction from bloodstain specimens. Heat maps and PCA scatter plots underscored the assay's proficiency in distinguishing between commonly encountered biofluid types in sexual assault cases based on their distinctive miRNA expression patterns. Furthermore, the assay's commendable inter- and intra-plate reproducibility facilitated the comparison of results across different runs. Finally, the assay revealed the potential for resolving biofluid

mixtures by identifying differentially expressed miRNAs for various biofluids within a mixture.

Based on the results from this study, HTG EdgeSeq miRNA WTA emerges as a promising forensic biofluid identification assay, boasting advantages such as exceptional sensitivity and specificity, multiplex abilities, minimal sample input requirements, and compatibility with common forensic casework samples like whole blood and dried biofluid stains collected using swabs. Future research involving a larger sample size is required to validate the findings presented in this study and further explore the capabilities of HTG EdgeSeq miRNA WTA in identifying a more comprehensive range of biofluid types.

#### **4.2 Future Directions**

For semen samples with diverse sperm counts, sperm counting techniques could be employed in conjunction with miRNA expression profiling to pinpoint miRNAs specific to spermatozoa. Expanding future investigations to encompass a more diverse range of biofluids and tissue types, including sweat, skin, urine, breast milk, and feces, could facilitate the detection of specifically expressed miRNAs in these biofluid types.

The current study used half of a swab to generate a complete miRNA expression profile from a dried biofluid swab sample. An examination into whether HTG EdgeSeq miRNA WTA can yield a comprehensive miRNA expression profile with a reduced sample size—perhaps a quarter or even an eighth of a swab—could prove advantageous, particularly in forensic casework when only one or two swabs are available for analysis.

Devising a panel or decision tree strategy comprising biofluid-specific miRNAs could prove beneficial for implementing HTG EdgeSeq miRNA WTA in forensic casework. Upon the establishment of a panel consisting of selected biofluid-specific miRNAs, unknown samples can be examined utilizing the panel to refine the chosen miRNAs and evaluate its ability for differentiating biofluid types. Moreover, a comparative study assessing labor, cost, and time between the HTG EdgeSeq miRNA WTA and current biofluid identification techniques employed in forensic laboratories is crucial to examine the assay's efficiency. From a practical standpoint, it is also worth exploring the possibility of decreasing sample preparation time. Other potential examinations include assessing the differences in miRNA profiling between samples used immediately and those kept frozen for an extended period, as well as investigating the variability in the quantity of eluted cellular materials from samples incubated in lysis buffer for varying durations.

Varied miRNA expression patterns were observed in menstrual blood samples collected using menstrual cups versus swabs. Further investigations are required to determine whether the disparate expression patterns stem from individual variations or discrepancies in collection techniques. Furthermore, studies should be conducted to address potential issues concerning liquid blood samples and inconsistencies between library quantification and sequencing outcomes.

## BIBLIOGRAPHY

1. Tay JW, Joudo J, Tran T, Ta H, Botting JL, Liew YC, et al. Comparison of Crime-lite® 82S, Polilight® PL400 and Polilight® PL500 for the detection of semen and saliva stains. *Australian Journal of Forensic Sciences* 2021;53(4):483–93. doi: 10.1080/00450618.2020.1718208.
2. Vandenberg N, Oorschot RAH. The use of Polilight® in the detection of seminal fluid, saliva, and bloodstains and comparison with conventional chemical-based screening tests. *Journal of Forensic Sciences* 2006;51(2):361–70. doi: 10.1111/j.1556-4029.2006.00065.x.
3. Andersen J, Bramble S. The effects of fingermark enhancement light sources on subsequent PCR-STR DNA analysis of fresh bloodstains. *Journal of Forensic Sciences* 1997;42(2):14116J. doi: 10.1520/JFS14116J.
4. Watson N. The analysis of body fluids. In: White PC, editor. *Crime scene to court*. Cambridge: Royal Society of Chemistry, 2004;377–413.
5. Vennemann M, Scott G, Curran L, Bittner F, Tobe SS. Sensitivity and specificity of presumptive tests for blood, saliva and semen. *Forensic Science, Medicine, and Pathology* 2014;10(1):69–75. doi: 10.1007/s12024-013-9515-6.
6. Tobe SS, Watson N, Daéid NN. Evaluation of six presumptive tests for blood, their specificity, sensitivity, and effect on high molecular-weight DNA. *Journal of Forensic Sciences* 2007;52(1):102–9. doi: 10.1111/j.1556-4029.2006.00324.x.
7. Johnston E, Ames CE, Dagnall KE, Foster J, Daniel BE. Comparison of presumptive blood test kits including Hexagon OBTI. *Journal of Forensic Sciences* 2008;53(3):687–9. doi: 10.1111/j.1556-4029.2008.00727.x.
8. Virkler K, Lednev IK. Analysis of body fluids for forensic purposes: From laboratory testing to non-destructive rapid confirmatory identification at a crime scene. *Forensic Science International* 2009;188(1–3):1–17. doi: 10.1016/j.forsciint.2009.02.013.
9. Posthuma-Trumpie GA, Korf J, van Amerongen A. Lateral flow (immuno)assay: Its strengths, weaknesses, opportunities and threats. A literature survey. *Analytical and Bioanalytical Chemistry* 2009;393(2):569–82. doi: 10.1007/s00216-008-2287-2.
10. Foley MM, Brown CO, Westring CG, Danielson PB, McKiernan HE. Effects of organic acids and common household products on the occurrence of false positive test results using immunochromatographic assays. *Forensic Science International* 2020;308:110165. doi: 10.1016/j.forsciint.2020.110165.

11. Schweers BA, Old J, Boonlayangoor PW, Reich KA. Developmental validation of a novel lateral flow strip test for rapid identification of human blood (Rapid Stain Identification™-Blood). *Forensic Science International: Genetics* 2008;2(3):243–7. doi: 10.1016/j.fsigen.2007.12.006.
12. Old J, Schweers BA, Boonlayangoor PW, Fischer B, Miller KWP, Reich K. Developmental validation of RSID™-Semen: A lateral flow immunochromatographic strip test for the forensic detection of human semen. *Journal of Forensic Sciences* 2012;57(2):489–99. doi: 10.1111/j.1556-4029.2011.01968.x.
13. Boward ES, Wilson SL. A comparison of ABACard® p30 and RSID™-Semen test kits for forensic semen identification. *Journal of Forensic and Legal Medicine* 2013;20(8):1126–30. doi: 10.1016/j.jflm.2013.09.007.
14. Park H-Y, Son B-N, Seo Y-I, Lim S-K. Comparison of four saliva detection methods to identify expectorated blood spatter. *Journal of Forensic Sciences* 2015;60(6):1571–6. doi: 10.1111/1556-4029.12864.
15. Miles LEM, Lipschitz DA, Bieber CP, Cook JD. Measurement of serum ferritin by a 2-site immunoradiometric assay. *Analytical Biochemistry* 1974;61(1):209–24. doi: 10.1016/0003-2697(74)90347-9.
16. Ross GMS, Filippini D, Nielen MWF, Salentijn GIJ. Unraveling the hook effect: A comprehensive study of high antigen concentration effects in sandwich lateral flow immunoassays. *Analytical Chemistry* 2020;92(23):15587–95. doi: 10.1021/acs.analchem.0c03740.
17. Allard JE. The collection of data from findings in cases of sexual assault and the significance of spermatozoa on vaginal, anal and oral swabs. *Science & Justice* 1997;37(2):99–108. doi: 10.1016/S1355-0306(97)72154-3.
18. Allery J-P, Telmon N, Mieusset R, Blanc A, Rougé D. Cytological detection of spermatozoa: Comparison of three staining methods. *Journal of Forensic Sciences* 2001;46(2):14970J. doi: 10.1520/JFS14970J.
19. Sijen T. Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers. *Forensic Science International: Genetics* 2015;18:21–32. doi: 10.1016/j.fsigen.2014.11.015.
20. Handy DE, Castro R, Loscalzo J. Epigenetic modifications: Basic mechanisms and role in cardiovascular disease. *Circulation* 2011;123(19):2145–56. doi: 10.1161/CIRCULATIONAHA.110.956839.

21. Bird AP. CpG-rich islands and the function of DNA methylation. *Nature* 1986;321(6067):209–13. doi: 10.1038/321209a0.
22. Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Gräf S, et al. An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Research* 2008;18(9):1518–29. doi: 10.1101/gr.077479.108.
23. Lee HY, An JH, Jung S-E, Oh YN, Lee EY, Choi A, et al. Genome-wide methylation profiling and a multiplex construction for the identification of body fluids using epigenetic markers. *Forensic Science International: Genetics* 2015;17:17–24. doi: 10.1016/j.fsigen.2015.03.002.
24. An JH, Choi A, Shin K-J, Yang WI, Lee HY. DNA methylation-specific multiplex assays for body fluid identification. *International Journal of Legal Medicine* 2013;127(1):35–43. doi: 10.1007/s00414-012-0719-1.
25. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genetics* 2009;5(8):e1000602. doi: 10.1371/journal.pgen.1000602.
26. Feher J. DNA and protein synthesis. In: Feher J, editor. *Quantitative human physiology*. Academic Press, 2017;120–9.
27. Haas C, Hanson E, Anjos MJ, Bär W, Banemann R, Berti A, et al. RNA/DNA co-analysis from blood stains—Results of a second collaborative EDNAP exercise. *Forensic Science International: Genetics* 2012;6(1):70–80. doi: 10.1016/j.fsigen.2011.02.004.
28. Haas C, Hanson E, Anjos MJ, Banemann R, Berti A, Borges E, et al. RNA/DNA co-analysis from human saliva and semen stains—Results of a third collaborative EDNAP exercise. *Forensic Science International: Genetics* 2013;7(2):230–9. doi: 10.1016/j.fsigen.2012.10.011.
29. Haas C, Hanson E, Anjos MJ, Ballantyne KN, Banemann R, Bhoelai B, et al. RNA/DNA co-analysis from human menstrual blood and vaginal secretion stains: Results of a fourth and fifth collaborative EDNAP exercise. *Forensic Science International: Genetics* 2014;8(1):203–12. doi: 10.1016/j.fsigen.2013.09.009.
30. Park S-M, Park S-Y, Kim J-H, Kang T-W, Park J-L, Woo K-M, et al. Genome-wide mRNA profiling and multiplex quantitative RT-PCR for forensic body fluid identification. *Forensic Science International: Genetics* 2013;7(1):143–50. doi: 10.1016/j.fsigen.2012.09.001.

31. Xu Y, Xie J, Cao Y, Zhou H, Ping Y, Chen L, et al. Development of highly sensitive and specific mRNA multiplex system (XCYR1) for forensic human body fluids and tissues identification. *PLoS ONE* 2014;9(7):e100123. doi: 10.1371/journal.pone.0100123.
32. Farkas DH, Holland CA. Overview of molecular diagnostic techniques and instrumentation. In: Tubbs RR, Stoler MH, editors. *Cell and tissue based molecular pathology*. Churchill Livingstone, 2009;19–32.
33. Lynch C, Fleming R. One-step endpoint RT-PCR assays for confirmatory body fluid identification. *Forensic Science International: Genetics* 2023;64:102856. doi: 10.1016/j.fsigen.2023.102856.
34. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. *Science* 2009;326(5960):1694–7. doi: 10.1126/science.1177486.
35. Akutsu T, Motani H, Watanabe K, Iwase H, Sakurada K. Detection of bacterial 16S ribosomal RNA genes for forensic identification of vaginal fluid. *Legal Medicine* 2012;14(3):160–2. doi: 10.1016/j.legalmed.2012.01.005.
36. Benschop CCG, Quak FCA, Boon ME, Sijen T, Kuiper I. Vaginal microbial flora analysis by next generation sequencing and microarrays; can microbes indicate vaginal origin in a forensic context? *International Journal of Legal Medicine* 2012;126(2):303–10. doi: 10.1007/s00414-011-0660-8.
37. Quak FCA, de Graaf M-LM, Weterings R, Kuiper I. Microbial population analysis improves the evidential value of faecal traces in forensic investigations. *International Journal of Legal Medicine* 2017;131(1):45–51. doi: 10.1007/s00414-016-1390-8.
38. Jung JY, Yoon HK, An S, Lee JW, Ahn E-R, Kim Y-J, et al. Rapid oral bacteria detection based on real-time PCR for the forensic identification of saliva. *Scientific Reports* 2018;8(1):10852. doi: 10.1038/s41598-018-29264-2.
39. Bartel DP. MicroRNAs. *Cell* 2004;116(2):281–97. doi: 10.1016/S0092-8674(04)00045-5.
40. He L, Hannon GJ. MicroRNAs: Small RNAs with a big role in gene regulation. *Nature Reviews Genetics* 2004;5(7):522–31. doi: 10.1038/nrg1379.
41. Kim VN. Small RNAs: Classification, biogenesis, and function. *Molecules and Cells* 2005;19(1):1–15.

42. Mendell JT. MicroRNAs: Critical regulators of development, cellular physiology and malignancy. *Cell Cycle* 2005;4(9):1179–84. doi: 10.4161/cc.4.9.2032.
43. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nature Reviews Genetics* 2008;9(2):102–14. doi: 10.1038/nrg2290.
44. Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. *Nature Reviews Molecular Cell Biology* 2009;10(2):126–39. doi: 10.1038/nrm2632.
45. Liang Y, Ridzon D, Wong L, Chen C. Characterization of microRNA expression profiles in normal human tissues. *BMC Genomics* 2007;8(1):166. doi: 10.1186/1471-2164-8-166.
46. Cortez MA, Calin GA. MicroRNA identification in plasma and serum: A new tool to diagnose and monitor diseases. *Expert Opinion on Biological Therapy* 2009;9(6):703–11. doi: 10.1517/14712590902932889.
47. Hanson EK, Lubenow H, Ballantyne J. Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. *Analytical Biochemistry* 2009;387(2):303–14. doi: 10.1016/j.ab.2009.01.037.
48. Weber JA, Baxter DH, Zhang S, Huang DY, How Huang K, Jen Lee M, et al. The microRNA spectrum in 12 body fluids. *Clinical Chemistry* 2010;56(11):1733–41. doi: 10.1373/clinchem.2010.147405.
49. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nature Reviews Cancer* 2006;6(11):857–66. doi: 10.1038/nrc1997.
50. Feber A, Xi L, Luketich JD, Pennathur A, Landreneau RJ, Wu M, et al. MicroRNA expression profiles of esophageal cancer. *The Journal of Thoracic and Cardiovascular Surgery* 2008;135(2):255–60. doi: 10.1016/j.jtcvs.2007.08.055.
51. Li T, Li D, Sha J, Sun P, Huang Y. MicroRNA-21 directly targets MARCKS and promotes apoptosis resistance and invasion in prostate cancer cells. *Biochemical and Biophysical Research Communications* 2009;383(3):280–5. doi: 10.1016/j.bbrc.2009.03.077.
52. Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, et al. Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. *Oncogene* 2006;25(17):2537–45. doi: 10.1038/sj.onc.1209283.

53. Iorio MV, Ferracin M, Liu C-G, Veronese A, Spizzo R, Sabbioni S, et al. MicroRNA gene expression deregulation in human breast cancer. *Cancer Research* 2005;65(16):7065–70. doi: 10.1158/0008-5472.CAN-05-1783.
54. Chen X-M. MicroRNA signatures in liver diseases. *World Journal of Gastroenterology* 2009;15(14):1665. doi: 10.3748/wjg.15.1665.
55. Alshehhi S, McCallum NA, Haddrill PR. Quantification of RNA degradation of blood-specific markers to indicate the age of bloodstains. *Forensic Science International: Genetics Supplement Series* 2017;6:e453–5. doi: 10.1016/j.fsigss.2017.09.175.
56. Mayes C, Houston R, Seashols-Williams S, LaRue B, Hughes-Stamm S. The stability and persistence of blood and semen mRNA and miRNA targets for body fluid identification in environmentally challenged and laundered samples. *Legal Medicine* 2019;38:45–50. doi: 10.1016/j.legalmed.2019.03.007.
57. Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF, et al. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proceedings of the National Academy of Sciences* 2011;108(12):5003–8. doi: 10.1073/pnas.1019055108.
58. Wang K, Zhang S, Weber J, Baxter D, Galas DJ. Export of microRNAs and microRNA-protective protein by mammalian cells. *Nucleic Acids Research* 2010;38(20):7248–59. doi: 10.1093/nar/gkq601.
59. Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, et al. Characterization of microRNAs in serum: A novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Research* 2008;18(10):997–1006. doi: 10.1038/cr.2008.282.
60. Layne TR, Green RA, Lewis CA, Nogales F, Dawson Cruz TC, Zehner ZE, et al. MicroRNA detection in blood, urine, semen, and saliva stains after compromising treatments. *Journal of Forensic Sciences* 2019;64(6):1831–7. doi: 10.1111/1556-4029.14113.
61. Grasedieck S, Schöler N, Bommer M, Niess JH, Tumani H, Rouhi A, et al. Impact of serum storage conditions on microRNA stability. *Leukemia* 2012;26(11):2414–6. doi: 10.1038/leu.2012.106.
62. Fang C, Zhao J, Li J, Qian J, Liu X, Sun Q, et al. Massively parallel sequencing of microRNA in bloodstains and evaluation of environmental influences on miRNA candidates using realtime polymerase chain reaction. *Forensic Science International: Genetics* 2019;38:32–8. doi: 10.1016/j.fsigen.2018.10.001.

63. Zubakov D, Boersma AWM, Choi Y, van Kuijk PF, Wiemer EAC, Kayser M. MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation. *International Journal of Legal Medicine* 2010;124(3):217–26. doi: 10.1007/s00414-009-0402-3.
64. Smith A, Nelson RJ. Capillary electrophoresis of DNA. *Current Protocols in Nucleic Acid Chemistry* 2003;13(1). doi: 10.1002/0471142700.nc1009s13.
65. van der Meer D, Uchimoto ML, Williams G. Simultaneous analysis of micro-RNA and DNA for determining the body fluid origin of DNA profiles. *Journal of Forensic Sciences* 2013;58(4):967–71. doi: 10.1111/1556-4029.12160.
66. Williams G, Uchimoto ML, Coult N, World D, Beasley E, Avenell P. Characterisation of body fluid specific microRNA markers by capillary electrophoresis. *Forensic Science International: Genetics Supplement Series* 2013;4(1):e274–5. doi: 10.1016/j.fsigss.2013.10.140.
67. Li Y, Zhang J, Wei W, Wang Z, Prinz M, Hou Y. A strategy for co-analysis of microRNAs and DNA. *Forensic Science International: Genetics* 2014;12:24–9. doi: 10.1016/j.fsigen.2014.04.011.
68. Mayes C, Seashols-Williams S, Hughes-Stamm S. A capillary electrophoresis method for identifying forensically relevant body fluids using miRNAs. *Legal Medicine* 2018;30:1–4. doi: 10.1016/j.legalmed.2017.10.013.
69. Liu C-G, Calin GA, Volinia S, Croce CM. MicroRNA expression profiling using microarrays. *Nature Protocols* 2008;3(4):563–78. doi: 10.1038/nprot.2008.14.
70. Courts C, Madea B. Specific micro-RNA signatures for the detection of saliva and blood in forensic body-fluid identification. *Journal of Forensic Sciences* 2011;56(6):1464–70. doi: 10.1111/j.1556-4029.2011.01894.x.
71. Wang Z, Zhang J, Luo H, Ye Y, Yan J, Hou Y. Screening and confirmation of microRNA markers for forensic body fluid identification. *Forensic Science International: Genetics* 2013;7(1):116–23. doi: 10.1016/j.fsigen.2012.07.006.
72. Park J-L, Park S-M, Kwon O-H, Lee H, Kim J, Seok HH, et al. Microarray screening and qRT-PCR evaluation of microRNA markers for forensic body fluid identification. *Electrophoresis* 2014;35(21–22):3062–8. doi: 10.1002/elps.201400075.
73. Sauer E, Reinke A-K, Courts C. Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR. *Forensic Science International: Genetics* 2016;22:89–99. doi: 10.1016/j.fsigen.2016.01.018.

74. Pandita P, Bhalla R, Saini A, Mani I. Emerging tools for studying receptor endocytosis and signaling. In: Singh V, Mani I, editors. Progress in molecular biology and translational science. Elsevier, 2023;19–48.
75. Omelia EJ, Uchimoto ML, Williams G. Quantitative PCR analysis of blood- and saliva-specific microRNA markers following solid-phase DNA extraction. *Analytical Biochemistry* 2013;435(2):120–2. doi: 10.1016/j.ab.2012.12.024.
76. Wang Z, Zhang J, Wei W, Zhou D, Luo H, Chen X, et al. Identification of saliva using microRNA biomarkers for forensic purpose. *Journal of Forensic Sciences* 2015;60(3):702–6. doi: 10.1111/1556-4029.12730.
77. Sirker M, Fimmers R, Schneider PM, Gomes I. Evaluating the forensic application of 19 target microRNAs as biomarkers in body fluid and tissue identification. *Forensic Science International: Genetics* 2017;27:41–9. doi: 10.1016/j.fsigen.2016.11.012.
78. He H, Han N, Ji C, Zhao Y, Hu S, Kong Q, et al. Identification of five types of forensic body fluids based on stepwise discriminant analysis. *Forensic Science International: Genetics* 2020;48:102337. doi: 10.1016/j.fsigen.2020.102337.
79. Sirker M, Liang W, Zhang L, Fimmers R, Rothschild MA, Gomes I, et al. Impact of using validated or standard reference genes for miRNA qPCR data normalization in cell type identification. *Forensic Science International: Genetics Supplement Series* 2015;5:e199–201. doi: 10.1016/j.fsigss.2015.09.080.
80. Fujimoto S, Manabe S, Morimoto C, Ozeki M, Hamano Y, Tamaki K. Optimal small-molecular reference RNA for RT-qPCR-based body fluid identification. *Forensic Science International: Genetics* 2018;37:135–42. doi: 10.1016/j.fsigen.2018.08.010.
81. Hanson EK, Rekab K, Ballantyne J. Binary logistic regression models enable miRNA profiling to provide accurate identification of forensically relevant body fluids and tissues. *Forensic Science International: Genetics Supplement Series* 2013;4(1):e127–8. doi: 10.1016/j.fsigss.2013.10.065.
82. Behjati S, Tarpey PS. What is next generation sequencing? *Archives of Disease in Childhood - Education & Practice Edition* 2013;98(6):236–8. doi: 10.1136/archdischild-2013-304340.
83. Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics* 2009;10(1):57–63. doi: 10.1038/nrg2484.
84. Seashols-Williams S, Lewis C, Calloway C, Peace N, Harrison A, Hayes-Nash C, et al. High-throughput miRNA sequencing and identification of biomarkers for

- forensically relevant biological fluids. *Electrophoresis* 2016;37(21):2780–8. doi: 10.1002/elps.201600258.
85. Wang Z, Zhou D, Cao Y, Hu Z, Zhang S, Bian Y, et al. Characterization of microRNA expression profiles in blood and saliva using the Ion Personal Genome Machine ® System (Ion PGM™ System). *Forensic Science International: Genetics* 2016;20:140–6. doi: 10.1016/j.fsigen.2015.10.008.
  86. Friedländer MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, et al. Discovering microRNAs from deep sequencing data using miRDeep. *Nature Biotechnology* 2008;26(4):407–15. doi: 10.1038/nbt1394.
  87. Liu Z, Wang Q, Wang N, Zang Y, Wu R, Sun H. A comprehensive characterization of small RNA profiles by massively parallel sequencing in six forensic body fluids/tissue. *Genes* 2022;13(9):1530. doi: 10.3390/genes13091530.
  88. Rice J, Roberts H, Burton J, Pan J, States V, Rai SN, et al. Assay reproducibility in clinical studies of plasma miRNA. *PLoS ONE* 2015;10(4):e0121948. doi: 10.1371/journal.pone.0121948.
  89. HTG Molecular Diagnostics, Inc.. HTG EdgeSeq miRNA whole transcriptome assay verification report. Tucson, AZ: 2022.
  90. HTG Molecular Diagnostics, Inc.. HTG EdgeSeq system user manual (research use only), code 10383239, revision B. Tucson, AZ: 2021.
  91. Thompson DM, Botros IW, Harrison H, Modur V. Automated high fidelity miRNA expression profiling using nuclease protection coupled with next generation sequencing. Merck Technology Symposium; 2014, Long Branch, NJ.
  92. Girard L, Rodriguez-Canales J, Behrens C, Thompson DM, Botros IW, Tang H, et al. An expression signature as an aid to the histologic classification of non–small cell lung cancer. *Clinical Cancer Research* 2016;22(19):4880–9. doi: 10.1158/1078-0432.CCR-15-2900.
  93. Bell A, Bell D, Chakravarti N, Ma J, Henton N, Prieto VG. Detection of a microRNA molecular signature of ultraviolet radiation in the superficial regions of melanocytic nevi on sun-exposed skin. *Modern Pathology* 2018;31(11):1744–55. doi: 10.1038/s41379-018-0088-5.
  94. Lizarraga D, Huen K, Combs M, Escudero-Fung M, Eskenazi B, Holland N. MiRNAs differentially expressed by next-generation sequencing in cord blood buffy coat samples of boys and girls. *Epigenomics* 2016;8(12):1619–35. doi: 10.2217/epi-2016-0031.

95. Loke SY, Munusamy P, Koh GL, Chan CHT, Madhukumar P, Thung JL, et al. A circulating miRNA signature for stratification of breast lesions among women with abnormal screening mammograms. *Cancers* 2019;11(12):1872. doi: 10.3390/cancers11121872.
96. Satake E, Pezzolesi MG, Md Dom ZI, Smiles AM, Niewczas MA, Krolewski AS. Circulating miRNA profiles associated with hyperglycemia in patients with type 1 diabetes. *Diabetes* 2018;67(5):1013–23. doi: 10.2337/db17-1207.
97. Brodeur A, Kero F, Kerns BJ. Demonstration of a semi-automated, extraction-free microRNA (miRNA) gene expression profiling method for the simultaneous detection of biofluids. *Proceedings of the American Academy of Forensic Sciences; 2022, Seattle, WA.*
98. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 2014;15(12):550. doi: 10.1186/s13059-014-0550-8.
99. Bar-Joseph Z, Gifford DK, Jaakkola TS. Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics* 2001;17(suppl\_1):S22–9. doi: 10.1093/bioinformatics/17.suppl\_1.S22.
100. Jolliffe IT, Cadima J. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 2016;374(2065):20150202. doi: 10.1098/rsta.2015.0202.
101. Bexon K, Williams G. Characterising the fluctuation of microRNA expression throughout a full menstrual cycle. *Forensic Science International: Genetics Supplement Series* 2015;5:e264–6. doi: 10.1016/j.fsigss.2015.09.105.
102. Owen DH. A review of the physical and chemical properties of human semen and the formulation of a semen simulant. *Journal of Andrology* 2005;26(4):459–69. doi: 10.2164/jandrol.04104.
103. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, et al. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 2007;129(7):1401–14. doi: 10.1016/j.cell.2007.04.040.
104. James ER, Carrell DT, Aston KI, Jenkins TG, Yeste M, Salas-Huetos A. The role of the epididymis and the contribution of epididymosomes to mammalian reproduction. *International Journal of Molecular Sciences* 2020;21(15):5377. doi: 10.3390/ijms21155377.

105. Liu L, He L, Li W, Zhao T, Li G, Xiu X, et al. The miR-4306/IGF2R axis modulates the lung adenocarcinoma response to irradiation in vitro and in vivo. *Translational Lung Cancer Research* 2021;10(12):4494–510. doi: 10.21037/tlcr-21-890.
106. Zhou Z, Zhou X, Jiang Y, Qiu M, Liang X, Lin Q, et al. Clinical significance of miR-1180-3p in hepatocellular carcinoma: A study based on bioinformatics analysis and RT-qPCR validation. *Scientific Reports* 2020;10(1):11573. doi: 10.1038/s41598-020-68450-z.
107. Kleemann M, Bereuther J, Fischer S, Marquart K, Hänle S, Unger K, et al. Investigation on tissue specific effects of pro-apoptotic micro RNAs revealed miR-147b as a potential biomarker in ovarian cancer prognosis. *Oncotarget* 2017;8(12):18773–91. doi: 10.18632/oncotarget.13095.
108. Li S, Li H, Ge W, Song K, Yuan C, Yin R. Effect of miR-184 on proliferation and apoptosis of pancreatic ductal adenocarcinoma and its mechanism. *Technology in Cancer Research & Treatment* 2020;19:1–8. doi: 10.1177/1533033820943237.
109. Geng X, Li L, Luo Y, Yang W, Hu J, Zhao Z, et al. Tumor cell derived Lnc-FSD2-31:1 contributes to cancer-associated fibroblasts activation in pancreatic ductal adenocarcinoma progression through extracellular vesicles cargo miR-4736. *Advanced Science* 2023;10:2203324. doi: 10.1002/advs.202203324.
110. Li C, Gonsalves CS, Eiyimo Mwa Mpollo M-S, Malik P, Tahara SM, Kalra VK. MicroRNA 648 targets ET-1 mRNA and is cotranscriptionally regulated with MICAL3 by PAX5. *Molecular and Cellular Biology* 2015;35(3):514–28. doi: 10.1128/MCB.01199-14.
111. Jie Y, Ye L, Chen H, Yu X, Cai L, He W, et al. ELFN1-AS1 accelerates cell proliferation, invasion and migration via regulating miR-497-3p/CLDN4 axis in ovarian cancer. *Bioengineered* 2020;11(1):872–82. doi: 10.1080/21655979.2020.1797281.

## CURRICULUM VITAE

