4-2019

Small RNA Sequencing and RT-qPCR Validation of Forensically Relevant Body Fluids

Karly Johannsen
University of New Haven

THE UNIVERSITY OF NEW HAVEN

GRADUATE SCHOOL

SMALL RNA SEQUENCING AND RT-qPCR VALIDATION OF FORENSICALLY RELEVANT BODY FLUIDS

A THESIS

Submitted in partial fulfillment

Of the requirements for the degree of

MASTERS OF SCIENCE IN FORENSIC SCIENCE

BY

Karly Johannsen
Advisor: Dr. Claire L. Glynn

University of New Haven
West Haven, CT
April 2018

ACKNOWLEDGEMENTS

I would like to acknowledge the assistance of Dr. Claire L. Glynn, Dr. David San Pietro and Mr. Jason Kramer. I would also like to thank the Forensic Science Department of University of New Haven for the funding for this study. Lastly, I would like to thank the participants of this study for allowing it to happen.

ABSTRACT

Body fluid identification is an important aspect of forensic investigations because it assists with the reconstruction of crime scenes and can refute and/or support witness statements. Currently, there is no universal method for body fluid identification. Each body fluid has several tests for its identification, both presumptive and confirmatory. A universal method for body fluid identification that is sensitive, specific, efficient, and minimally destructive is necessary.

In recent years miRNAs have been heralded as novel biomarkers for the identification of body fluids. Several research groups across the US and internationally have been formed to determine which miRNAs are suitable for body fluid identification, however there has been little agreement between them in their conclusions. With the advent of Next Generation Sequencing (NGS), it is now possible to sequence all forensically relevant body fluids for both known and novel miRNAs, with the expressed interest to identify panels of miRNAs for each body fluid.

The goal of this study was to determine if it was feasible to use NGS in the form of the MiSeq FGx platform (Illumina) to determine a panel of miRNAs that are specific to forensically relevant body fluids including venous blood, semen, saliva, vaginal fluid, and menstrual blood. Subsequently, a panel of miRNAs was identified from previously published research that were reported to have potential applications to forensic body fluid identification. Ultimately, it was the goal of this study to determine if the chosen miRNAs were suitable for body fluid identification. This was done using RT-qPCR, which is the preferred method of miRNA expression analysis.

It was determined that the MiSeq FGx was not an optimal NGS platform for this purpose, as no miRNAs were identified. The RT-qPCR validation of the chosen miRNAs resulted in both agreement and disagreement with previously published results, illustrating the need for more extensive research to determine an appropriate panel of miRNAs for body fluid identification.

Table of Contents

List of Figures

List of Tables

List of Equations

CHAPTER 1: INTRODUCTION

1 CHAPTER 1: INTRODUCTION

1.1 Forensic Body Fluid Identification

1.1.1 Introduction

The identification of forensically relevant body fluids including venous blood, menstrual blood, semen, vaginal material, and saliva is an extremely important aspect of forensic investigations. Some ask whether this identification is necessary when obtaining a DNA profile from body fluids is possible and seems more useful. The identification of the body fluid is just as important as obtaining a DNA profile because the correct identification of a body fluid can help reconstruct the events that took place and can potentially corroborate or refute witness statements. Body fluid identification is very important in all forensic investigations but can be especially important in sexual assault investigations. For example, if a condom was worn by the perpetrator during an assault, the correct identification of the body fluids on the outside of the condom can be very helpful to the case. If the victim claims there was forced vaginal intercourse, while the suspect claims that there was only voluntary oral intercourse, the identification of the body fluid on the outside of the condom could be helpful in determining which statement is more truthful. The repercussions of the two scenarios are very different, with one leading to the potential arrest of the male in the scenario. In this scenario and others like it, body fluid identification is just important as the information obtained from a DNA profile. If both a DNA profile can be obtained and a body fluid identification can be made, then it could be possible to link a suspect to the scene of the crime and aid the reconstruction of events, both of which are important within a forensic investigation.

1.1.2 Current Methods

Currently, body fluid identification occurs in two steps. First, presumptive tests are performed to determine whether an unknown stain is more than likely a specific body fluid (Virkler et al, 2009). While these tests are highly sensitive, they are not highly specific. Therefore, each presumptive test can react positively with materials other than the targeted body fluid. That is why a second, confirmatory test is required for a confirmed identification of the questioned stain. These tests are considered highly specific to human body fluids and are extremely sensitive (Virkler et al, 2009). Generally, all these methods are destructive to the questioned stain, which can be problematic when only a small amount of sample is available. Because each body fluid has its own presumptive and confirmatory tests, there can be a potential for large amounts of a questioned stain to be consumed if it is suspected to be one of several body fluids.

In the current body fluid identification system, venous blood, semen, and saliva have established and reliable tests used for their identification (Virkler et al, 2009). The most common presumptive test used for the identification of venous blood is the use of the Kastle-Meyer reagent, used in conjunction with hydrogen peroxide to create a distinct color change (Virkler et al, 2009). This presumptive test relies on the heme within hemoglobin acting as a catalyst for the oxidation-reduction reaction between phenolphthalein and hydrogen peroxide. This reaction can occur in reaction to other chemicals, which is why this is considered a presumptive test. A common confirmatory test for venous blood is the ABA Hematrace Test, which is a an immunochromatographic assay, meaning it contains antigens and antibodies. This test is considered confirmatory because it uses the specific interaction between antigens and antibodies specific to human hemoglobin to garner a positive result.

Similarly, saliva identification relies on a presumptive test with a distinct color change and an immunochromatographic assay for a confirmatory test. The Phadebas Reagent is the most commonly used presumptive test for saliva identification. A positive result from this test occurs when amylase, an enzyme found in saliva, cleaves starch from a dye molecule. This is what causes the distinct color change (James et al, 2014). Amylase, while found in saliva, is found in other substances and even other body fluids in smaller amounts. This is why the Phadebas Reagent can only be used as a presumptive test. Recently, a commercial immunochromatographic assay, RSID-saliva, was developed for the confirmatory testing of saliva. It specifically tests for the alpha-amylase found in human saliva, which is considered specific to humans. It is currently one of the only validated confirmatory tests for saliva (Old et al, 2009).

Semen identification is slightly different than venous blood and saliva identification. The most common presumptive testing method is the Acid Phosphatase test. Like the other two presumptive tests mentioned, a positive result is indicated by a color change. This color change is due to the reaction between alpha-napththyl phosphate and seminal acid phosphatase. In this case, Fast Blue B is the color developer (James et al, 2014). Again, like other presumptive tests, other materials can induce a positive reaction. The most commonly used confirmatory testing method is microscopy identification of sperm cells with the aid of a staining agent like the Christmas Tree Stain. Because human sperm cells are dissimilar from other animals in size and shape, it is a trusted confirmatory method used within the majority of forensic laboratories (James et al, 2014).

While venous blood, saliva, and semen all have reliable testing methods, other body fluids including vaginal fluid and menstrual blood do not have reliable testing methods. Several

presumptive tests exist for these body fluids, but there are some issues with them (James et al, 2014). The major issues are that they lack specificity and can be subjective. Additionally, there are no confirmatory tests that can be used reliably. This is problematic because these body fluids are very important in forensic investigations, especially in sexual assault cases. While reliable methods do not exist currently, there are several emerging methods that have taken this, and the other issues mentioned before, into consideration in their development.

1.1.3 Emerging Methods

Currently, the issues with body fluid identification include the tests being destructive, no universal method for their identification existing, a lack of specificity for the presumptive tests, and the complicated nature of some of the confirmatory tests. The emerging techniques for body fluid identification have been attempting to reduce the effect of some of these points. Therefore, the ideal body fluid identification method should be universal, non-destructive, sensitive, specific, and relatively simple. A representation of this need can be seen in Figure 1.

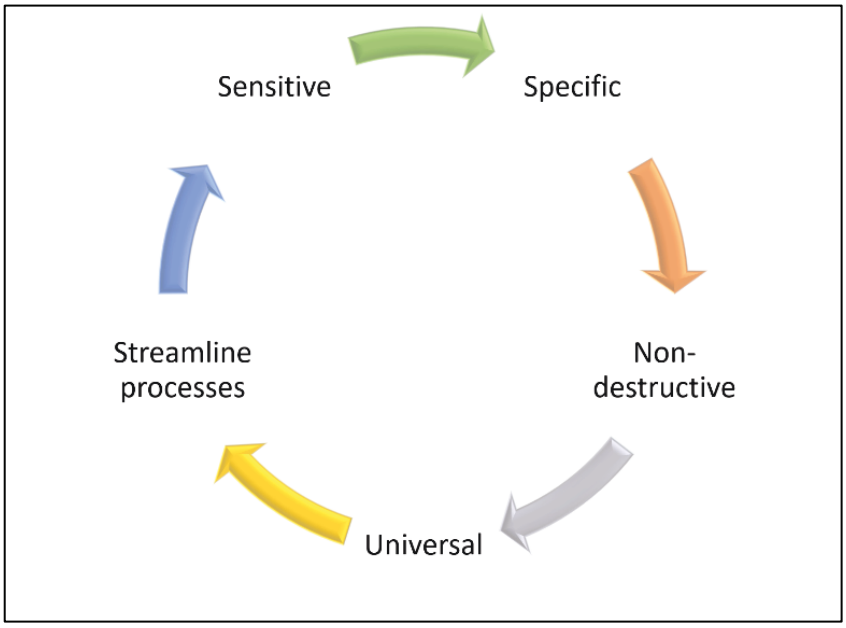

Figure 1: Characteristics of Ideal Body Fluid Identification Test

Over the last few years, there has been substantial research into the development of new methods for body fluid identification. In a general sense, there are two types of emerging identification techniques: molecular techniques and spectroscopic techniques. Molecular techniques generally focus on unique aspects of genomic material, while spectroscopic techniques generally focus on the outward characteristics of the body fluid.

Spectroscopic techniques for body fluid identification have become popular recently due to the many benefits that it provides, including minimal sample preparation and no destruction to the sample. Within this area, the use of Raman spectroscopy has been gaining momentum for the identification of body fluids. While researchers have been able to use Raman spectroscopy for the differentiation of semen, vaginal, fluid, sweat, saliva and blood, there still exists some issues with this method (Virkler & Lednev, 2008). One of the main issues with this methodology is that the material that the body fluid is on can interfere with the identification of the body fluid. While statistical models can be used to correct this, it can be complicated and time-consuming (McLaughlin et al, 2015). Raman spectroscopy could be a good candidate for the identification of body fluids, but the method needs more fine-tuning.

Another emerging spectroscopic technique is the use of Attenuated Total Reflectance Fourier Transformed Infrared (ATR-FTIR) Spectroscopy. The premise behind this technique is that each body fluid type should absorb infrared light at unique wavelengths (Quinn et al, 2017). While some researchers have been able to differentiate between the forensically relevant body fluids using ATR-FTIR, there has been little research on the effect that the substrate background has on the differentiation of body fluids (Quinn et al, 2017). Because body fluids can be found on any number of substrates, the unknown effect that it has on ATR-FTIR is a major downfall to this technique.

Unlike spectroscopic techniques, molecular techniques do require part of the sample to be consumed during testing. This amount, however, is generally very small and does not necessarily require the entire consumption of a sample. Additionally, molecular techniques analyze the genetic content of body fluids, which can be more complicated that spectroscopic techniques. In recent years, extensive research has been done on the specificity of messenger RNAs (mRNAs) to forensically relevant body fluids. Since the conception of this idea, a panel of mRNAs that have been proven to be specific to forensically relevant body fluids (Nussbaumer et al, 2006). Despite this, the utilization of mRNAs for body fluid identification is not often used in the field of forensic science, mainly because mRNAs are prone to degradation. This can be an issue when body fluids are aged or exposed to unfavorable weather conditions (Zubocov et al, 2010).

Another molecular technique that has been gaining momentum is the use of DNA methylation sites for body fluid identification. It was proposed that different body fluids could contain DNA methylation sites specific to each body fluid (Park et al, 2014). There are, however, issues with this method. Currently, there does not appear to be any studies published that involved menstrual blood, which is an important body fluid within forensic investigations. Additionally, the sample preparation can be considered extensive and complicated, which can be problematic in a laboratory setting.

Another molecular technique for body fluid identification for body fluid identification that shows immense promise is the use of microRNAs (miRNAs). There has been great success in the last few years in identifying miRNAs that can be used for body fluid identification. While many miRNAs have been found to be helpful in the differentiation of body fluids, there has been little agreement on which miRNAs are the most appropriate for this purpose. It is, however, the

purpose of this study is to determine which miRNAs throughout literature are the most appropriate for the identification of forensically relevant body fluids.

1.2 MicroRNAs

1.2.1 Introduction

miRNAs are short, non-coding strands of RNA that are approximately 18-22 base pairs in length. Up until 1993, miRNAs were thought to have no genetic purpose and were deemed unimportant. It was at this time, however, that researchers studying nematode genetics discovered that two small RNA transcripts from the lin-4 were partly responsible for the regulation of the translation of the lin-14 gene (Lee et al, 1993). These small transcripts were later termed small RNAs, or microRNAs. It wasn't until 2000, however, that more miRNAs were discovered (Bartel, 2014). In the following years, more miRNAs were discovered. These miRNAs were found to be specific to gene regulations within mammals, fish, worms, flies, and humans (Bartel, 2004).

1.2.2 MicroRNA Biogenesis

Since the discovery of miRNAs, much has been discovered about their specificity and biogenesis. miRNA biogenesis begins in the nucleus of the cell. The biogenesis pathway of a representative miRNA can be seen in Figure 2.

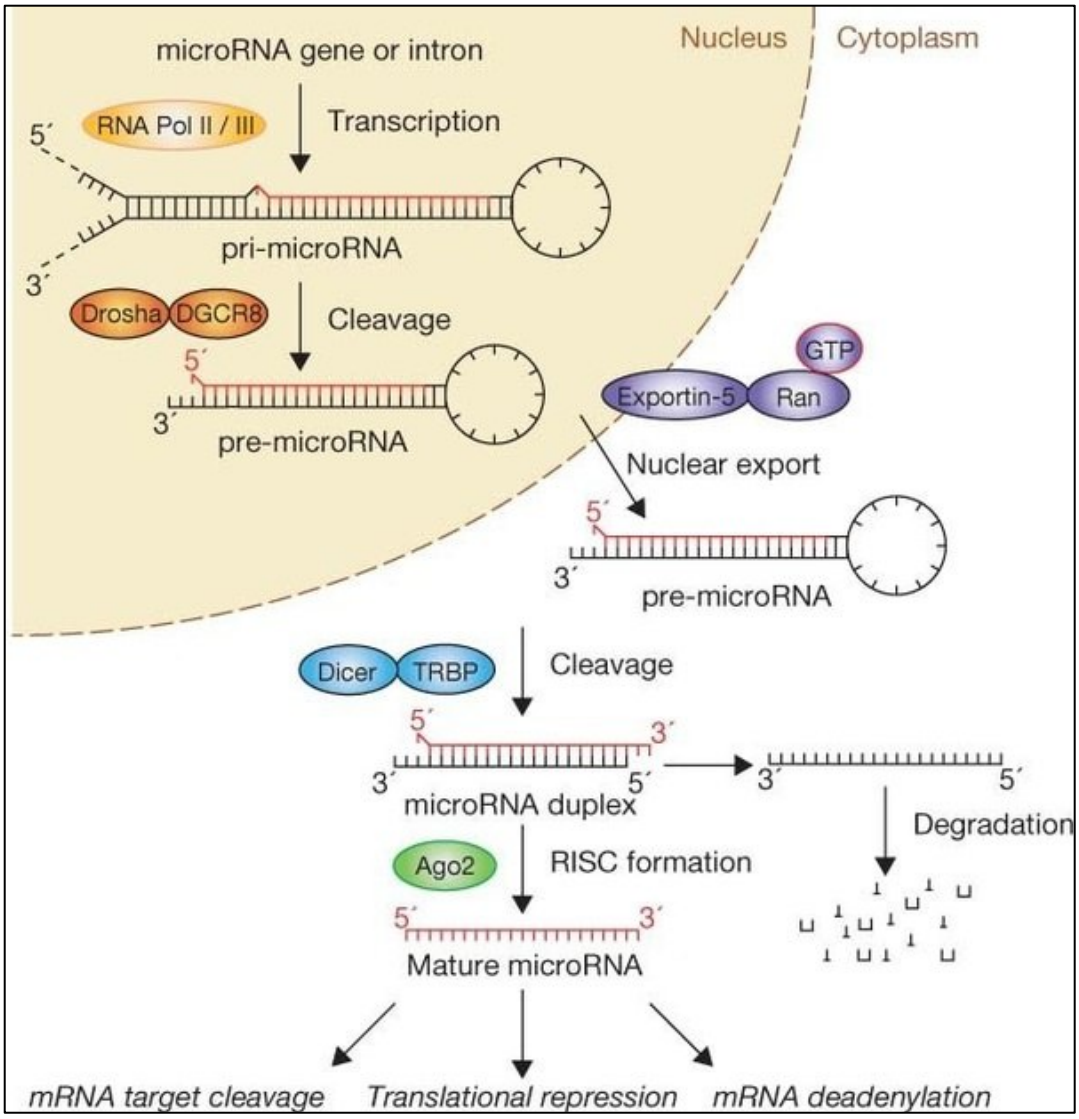

Figure 2: miRNA Biogenesis Pathway

Adapted from Winter, J., Jung, S., Keller, S., Gregory, R. I., & Diederichs, S. (2009). Many roads to maturity: MicroRNA biogenesis pathways and their regulation. *Nature Cell Biology, 11*(3), 228-234. doi:10.1038/ncb0309-228

First, miRNAs genes or introns are transcribed within the nucleus by RNA polymerases II and III. This transcription event results in primary-miRNAs (pri-miRNA), which are precursors to mature miRNAs (MacFarlane & Murphy, 2010). The pri-miRNAs take a stem-loop structure as depicted in Figure 2. While still in the nucleus, the pri-miRNAs are cleaved by a protein complex commonly called the microprocessor that contains the Drosha enzyme and an RNA-binding protein DGCR8 (MacFarlane & Murphy, 2010). After this cleavage event, the pri-miRNAs are transformed into pre-miRNAs, which is one step closer to mature miRNA. The pre-miRNAs are then transported across the membrane and into the cytoplasm. It is at this point that the pre-miRNAs are cleaved again by an RNase II Dicer enzyme, Dicer-1. This cleavage results in two miRNA products, the mature miRNA strand and another strand that gets degraded. The mature miRNA then goes on to serve its function within the cell, while still retaining the stem-loop structure (MacFarlane & Murphy, 2010).

1.2.3 MicroRNA Analysis

Currently, the gold standard method for analyzing miRNAs is an expression assay using Real Time Quantitative Polymerase Chain Reaction (RT-qPCR). Generally, analysis is done to determine whether a miRNA or panel of miRNAs are expressed differently amongst two or more tissue types. This comparison is done by using the cycle thresholds at which each of the miRNAs are recognized during the analysis process. Each individual cycle threshold can be correlated to the original concentration of the miRNA in a specific sample or tissue. By comparing these cycle thresholds, a comparison can be made on the quantity of miRNAs in each tissue type. One of two conclusions can be made from the results from the RT-qPCR: either a miRNA is found to only

be expressed in one tissue or is found to be expressed in a higher or lower abundance compared to other tissues.

1.2.4 Biomedical Applications of microRNAs

miRNAs have also been extensively researched within humans, especially within cancer research. Abnormal miRNA expression in human cancer was first described by Calin et al in 2002 (Visone & Croce, 2009). It was found that miR-15 and miR-16 were expressed less in B-cell chronic lymphocytic leukemia cells compared to control cells (Visone & Croce, 2009). The differential expression of these miRNAs provided a new potential avenue for the diagnosis of B-cell chronic lymphocytic leukemia. Because of the potential seen with miRNAs, this research study sparked more research of miRNAs in cancer types. As research progressed, miRNAs were discovered that showed true specificity to certain cancer cell types (Visone & Croce, 2009). True specificity means that that miRNA was only found in a specific cancer type compared to normal cells and similar cancer types. Consequently, it was shown that miRNAs can be extremely tissue specific. At the very least, miRNAs have been shown to be differentially expressed between body tissues.

1.2.5 Forensic Applications of microRNAs

Through extensive research, miRNAs have been shown to be highly specific. For one, it has been shown that miRNAs are species specific (Bartel et al, 2010). Additionally, researchers have been able to extensively show that miRNAs are both tissue specific and even specific to certain cancers. Because body fluids are a type of body tissue, it is reasonable to assume that it is possible that miRNAs exist that are specific to forensically relevant body fluids including venous blood, semen, menstrual blood, vaginal material, and saliva. As mentioned before, miRNAs have

been shown to be resistant to degradation. This characteristic of miRNAs suits forensic investigations because body fluids are exposed to the aging process and poor environmental environments. If miRNAs can withstand these conditions, they would be beneficial for the identification of body fluids.

Because miRNAs have been gaining much attention as a potential biomarker for body fluids, different techniques have been used for the discovery of miRNAs unique to body fluids. Some of these include microarray assays and capillary electrophoresis. More recently, however, next generation sequencing has been used by select researchers because it has the ability sequence the entire miRNA content of body fluids in a short amount of time. The potential data that could be garnered from this technique may provide many opportunities to discover new miRNAs that are more specific to body fluids.

1.3 Next Generation Sequencing

1.3.1 Early Development

Next generation sequencing (NGS) has the capacity for the efficient and complete sequencing of genetic material by sequencing millions, if not billions, of DNA molecules at the same time (Yang et al, 2014). A full genome can be sequenced by an NGS platform in less than a day, which took the Human Genome Project 12 years and billions of dollars using the traditional Sanger sequencing methods (Muzzey et al, 2015). The first iteration of this technology, the 454 Genome Sequencing System, was introduced in 2005. It was considered a high-throughput sequencing system and was capable of generating 200,000 reads of genomic material 110 base pairs in lengths (Yang et al, 2014). In 2007, Applied Biosystems introduced a new platform that relied on an oligonucleotide ligation method and a two-base encoding system. While this

platform was able to perform 30 million reads, its base pair length capacity was only 35 base pairs long (Yang et al, 2014). Over the years, these methods have been refined and the current methods are capable of performing more reads, with longer read lengths.

1.3.2 Current platforms

NGS technology is a collection of different instruments, each with different sequencing methods and different uses. Some of these sequencing methods include ion semiconductor sequencing, single-molecule, sequencing by ligation, and sequencing by synthesis (Muzzey et al, 2015). With ion semiconductor sequencing, the nucleotides are sequenced based on the protons that are released during processing and the intensity of the resulting signal (Muzzey et al, 2015). Sequencing by ligation involves using the innate matching abilities of DNA ligase to determine the sequence of the nucleotides in a DNA sequence (Wehling, 2015). While these two methods are valid and provide valuable information, it is the sequencing by synthesis technique that dominates the current field of NGS technology. The main reason this technique dominates the field is because it is used by the most popular company that provides NGS platforms, Illumina (Muzzey et al, 2015). Illumina has even released a NGS platform, the MiSeq FGx that is designed specifically for forensic genomics research.

Because sequencing by synthesis (SBS) is one of the more popular next generation sequencing techniques currently used, it is important to understand how it works. Figure 3

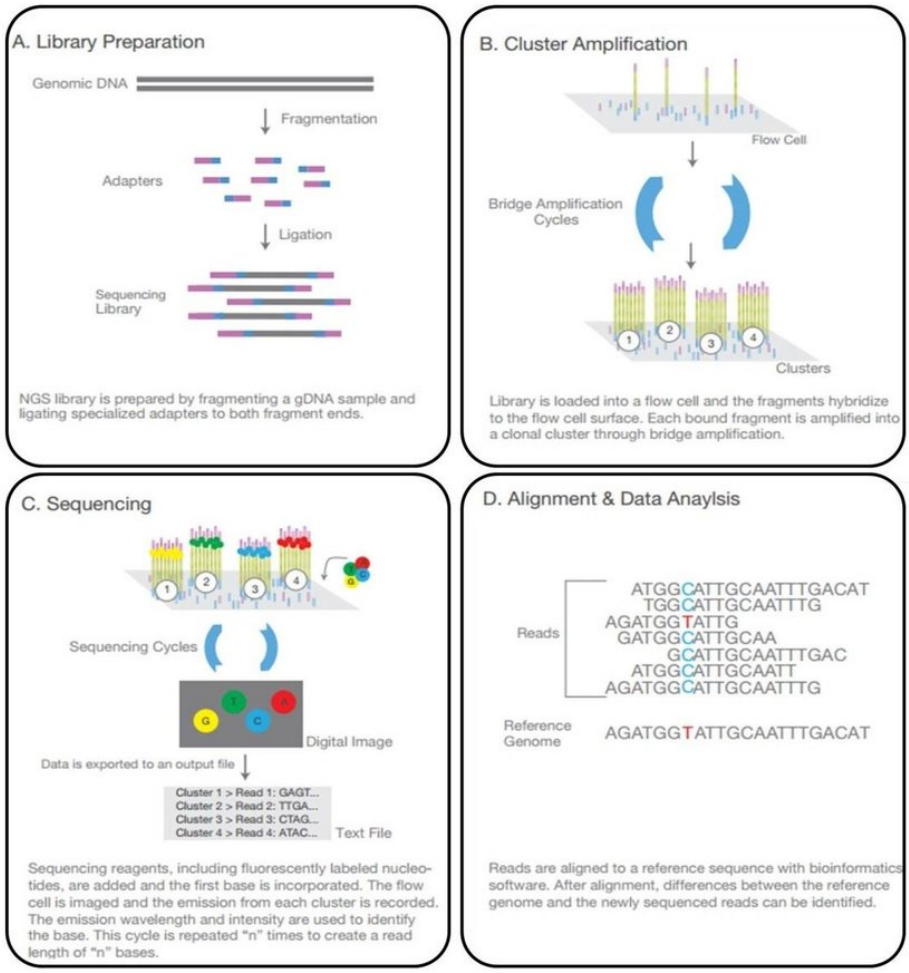

Figure 3: SBS Workflow

Adapted from Illumina. (2015). *An introduction to next generation sequencing technology*. San Diego, CA: Illumina

depicts the general workflow and biochemistry behind the sequencing by synthesis methodology used by Illumina. This technique takes advantage of the natural process of DNA replication. Essentially, DNA replication requires a template DNA strand, free bases to add to the new DNA strand, and polymerases to link the bases together. SBS utilizes this idea on a much grander scale and the bases that are being added are fluorescently tagged (Muzzey et al, 2015). Each different

base type (Adenine, Tyrosine, Guanine, and Cytosine) has a specific color assigned to it which allows for their identification during the detection phase of the sequencing. If a certain color is detected, then that specific base will be identified at a specific position in the genome. There are several general steps to SBS. Firstly, the RNA samples are prepared for sequencing, which includes fragmentation and reverse transcription into cDNA. Secondly, hundreds of thousands of template DNA strands are bound to the surface where the sequencing will take place, also called a flow cell (Muzzey et al, 2015). Next, these template strands are copied hundreds of times until a cluster of strands are formed. Once a certain amount of template strands has been formed in each cluster, the template strands will be replicated at the same time using the fluorescently labeled nucleotides. At specified times, the replicate pauses and a digital image is formed of all the strands and the nucleotides are identified based on the color conformation of the digital images. This stepwise process repeats until all the genetic material is sequenced and reported by the instrument. The instrument automatically assembles a readable sequence of the genetic material and statistical manipulation can be done on the information if needed.

It is important to note that each fragment within each sample receives more than one read from the instrument. Ultimately, a sample could receive anywhere between thousands to millions of reads. The success of the sequencing depends partly on the number of reads a sample receives. Generally, the more reads a sample receives, the more successful the sequencing will be. Additionally, there is a minimum number of reads that each sample needs for successful sequencing to occur. This minimum number is usually dependent on what is being sequenced. For example, miRNAs need a minimum of 1-2 million read for successful sequencing. More reads, however, is preferred.

1.3.3 Forensic Application

Next Generation Sequencing technology allows for many more avenues for the forensic science community. For example, it has the potential to be used with forensic laboratories within casework. In this sense, NGS can be used to garner more information about the genomic content of body fluids found at the scene. At this stage in its development, it would generally only be used to assist forensic scientists in determining the identity of an unknown individual.

NGS can also be used within forensic genomic research. For example, it can be used in research assessing the level of degradation in DNA and how efficient different DNA restoration techniques are (Gorden et al, 2018). Other potential research uses for NGS include species identification, identical twin studies, and body fluid identification. Within body fluid identification, NGS technology can even be used to identify known and discover novel miRNAs that are unique to individual body fluids.

1.4 **Thesis Aims**

The overall goal of this research study is to evaluate different miRNAs for their potential use as specific biomarkers for the identification of body fluids (venous blood, semen, vaginal fluid, menstrual blood, and saliva). The miRNAs chosen for evaluation will be garnered from either NGS data or a literature review and then validated using RT-qPCR. For example, if one miRNA seems to be consistently expressed within venous blood, that miRNA would be chosen for validation as a potential biomarker for venous blood. If no usable data can be obtained from the NGS, then a literature search will be performed and a panel of miRNAs will be chosen for validation. Regardless of which avenue is chosen, validation of the chosen miRNAs will be achieved through reverse transcription quantitative PCR (RT-qPCR), a gold standard method for

miRNA validation. Statistical analysis will be performed on this validation and it will be inferred from there whether any of the evaluated miRNAs are truly specific biomarkers for any of the body fluids. A miRNA could be considered a specific biomarker for a body fluid if it is expressed in higher or lower abundance compared to other body fluids that also express the same miRNA or if it is only expressed in one body fluid. The expression levels need to be statistically different for differentiation between body fluids to be possible. Thus, this ability to differentiate between body fluids using a miRNA has the potential to become a universal method for body fluid identification.

In summary, the four main goals for this thesis research are:

1. Performing Next Generation Sequencing to identify known or discover novel miRNAs that can be utilized for the identification of five forensically relevant body fluids including venous blood, menstrual blood, semen, vaginal material, and saliva.
2. Performing a literature review of existing studies to determine whether there is any agreement amongst the known miRNAs that have been shown to be potential biomarkers for the five body fluids being studied.
3. Utilizing real-time quantitative PCR (RT-qPCR) to validate the miRNAs chosen for evaluation from the information garnered from the next generation sequencing and the literature review.
4. Performing statistical analysis on the results from the RT-qPCR to determine whether the miRNAs expression levels are statistically different and will allow for body fluid differentiation and identification. Therefore, this step will allow us to

infer whether a universal identification method for body fluids is possible with the set of miRNAs chosen for evaluation.

CHAPTER 2: LITERATURE REVIEW

2 CHAPTER 2: LITERATURE REVIEW

2.1 **Methods of body fluid collection for research purposes**

Throughout literature, there was some disagreement on the collection technique for forensically relevant body fluids including venous blood, semen, menstrual blood, vaginal fluid, and saliva. The current study chose the most popular collection technique within the literature.

Within research studies using human venous blood samples, there were generally two methods used for collection. One of those methods involve blood being collected from the vein by venipuncture into dry vacutainer tubes that contain some form of EDTA to prevent coagulation of the blood (Sauer et al., 2014). From there, the blood was either immediately deposited on a sterile cotton swab or refrigerated until needed. The second method involved pricking the finger with lancets and immediately depositing it on a sterile cotton swab (Shi et al., 2017, Sirker et al., 2017). The sample could then be stored at -20˚C until its use. The current research study utilized the venipuncture method and stored the collected blood at 4˚C. Afterwards, the whole blood itself was extracted based on the extraction protocol of the extraction kit used.

Regardless of the research study, the collection of semen was the same throughout. The male participants were asked to ejaculate into a container, either a conical tube or a specimen cup (Sauer et al., 2014, Sauer et al., 2016). After the sample was delivered to the appropriate area, the sample was then either stored at -20˚C or immediately prepared for extraction. The preparation included either depositing a portion of the sample on a sterile cotton swab or extracting directly from the semen (Sauer et al., 2014, Sauer et al., 2016).

Within research that utilizes human saliva samples, two collection methods were used. One included spitting into a container, usually a conical tube, until enough, (1-5 mL), of saliva

was produced (Sauer et al., 2016, Schweighardt et al., 2015). The other method involved taking buccal swabs of the inside of the check (Courts & Madea, 2011, Sauer et al., 2014, Sauer et al., 2016). It is unclear from the literature which technique is superior, but the current study utilized the spitting method.

Unlike the other body fluids, the collection of vaginal fluid was consistent throughout the literature. The female participants were asked to collect their own samples by inserting a sterile cotton swab into the vagina and swabbing the vaginal wall (Sauer et al., 2014, Sauer et al., 2016). Participants were generally asked to refrain from sexual intercourse for at least 4 days before sample collection prevent contamination from semen and to collect the samples at least two weeks after the end of their last menstrual cycle to ensure that there was no contamination of menstrual blood.

Within research studies using human menstrual blood samples, there were generally two methods used for sample collection. One method included utilizing used tampons of the female participants and the other method involved having female participants collecting their own samples by inserting a sterile cotton swab into the vagina (Sirker et al., 2017, Sauer et al, 2016). The second method using the sterile cotton swabs was utilized in the current research study to avoid any contamination from material found in a tampon.

2.2 **Early studies for the forensic application of microRNAs**

The first mention that miRNAs could be used for body fluid identification was within the groundbreaking study performed in 2009 by Hanson et al (Hansen et al., 2009). In this study, Hanson et al was able to show that for each body fluid examined, several different miRNAs were differentially expressed, thereby allowing the differentiation of several body fluids using

miRNAs. Due to this development, other researchers have been striving to determine which miRNAs can be utilized as specific biomarkers for body fluids (Courts & Madea, 2011, Li et al., 2017, Sauer et al., 2016, Sirker et al., 2017). Despite the research that has been performed in this area, a consistent panel of miRNAs for body fluid differentiation has yet to be identified.

2.3 **Recent studies for the forensic application of microRNAs**

Since Hanson's study in 2009, several research studies have focused on linking miRNAs to specific body fluids, but there has been a lack of a consistent panel of miRNAs that could be used for body fluid identification (Courts & Madea, 2011, Li et al., 2017, Sauer et al., 2016, Sirker et al., 2017). For instance, in 2011, Courts and Madea examined six separate miRNAs within the context of distinguishing between blood and saliva. Three were chosen to be specific for saliva and three were chosen to be specific for blood. The researchers were unable to isolate a miRNA that was exclusive to blood or saliva amongst the chosen miRNAs (Courts & Madea, 2011). Researchers did, however, hypothesize that it could be possible to identify blood or saliva by using a combination or panel of the miRNAs chosen for each. However, results from this study can only be applied to saliva and blood. Therefore, whether the chosen miRNAs are expressed in other relevant body fluids, such as semen or vaginal fluid, has yet to be determined. The current study looks to investigate five different, forensically relevant, body fluids including saliva, semen, vaginal fluid, menstrual blood, and venous blood.

Several research studies choose to focus on the identification of only one body fluid compared to other body fluids. For example, Li et al evaluated several miRNAs for specificity between menstrual and peripheral blood. The miRNAs were chosen for investigation based on an extensive literature review of previous research (Li et al., 2017). After validation via RT-qPCR,

results showed that miR-141-3p was highly expressed in menstrual blood and expressed at much lower levels in peripheral blood, which allowed for the differentiation between menstrual and venous blood (Li et al., 2017) After this was discovered, the researchers developed a multi-step protocol for blood identification. Several miRNAs that were seen to be generalized for blood was used for the identification of both types of blood. After this, menstrual blood could be identified by the newly discovered miRNA. While this multi-step process does allow for the identification of menstrual and peripheral blood, it requires extensive sample preparation and the total testing time would require several hours. Additionally, more validation on the menstrual blood-specific miRNA, miR-141-3p, would need to be performed before it could be trusted as menstrual blood specific.

Another study that focused on the identification of one type of body fluid was performed by Wang et al in 2015, focused only on saliva (Wang et al., 2015). Eight miRNAs were chosen for further evaluation due to their apparent specificity to saliva in past research articles and past microarray studies done by the researchers. The purpose of the study was to determine whether the miRNAs chosen for validation could be used for the identification of saliva. After validation, none of the evaluated miRNAs were found to be truly specific to saliva (Wang et al., 2015). Researchers were, however, able to develop a stepwise strategy using several of the miRNAs to identify saliva (Wang et al, 2015). While the identification of saliva was eventually possible, the process that it took to achieve this was highly involved and somewhat complicated. As with the previously mentioned study involving menstrual blood identification, a multi-step process may not be conducive to a laboratory testing.

Rather than just focus on the identification of one body fluid at a time, some researchers attempt to identify as many body fluids as possible by using a panel of miRNAs. Recently, Sirker

et al. examined 19 miRNAs as potential biomarkers for body fluid identification (Sirker et al. 2017). Of the 19 miRNAs examined, only six showed any level of specificity towards the body fluids they were thought to be specific for. The researchers found that of those six, miR-451 showed clear differentiation capabilities for peripheral blood and miR-10b showed clear differentiation capabilities for semen, which enabled identification of both body fluids (Sirker et al. 2017). Showing clear differentiation between body fluids through the use of miRNA demonstrates that the aim of the current research to discover miRNAs that can be used to identify all five body fluids (saliva, semen, vaginal fluid, menstrual blood, and venous blood) has immense potential. It is important to note however, that the identifications of the body fluids were based on the differential expression of the miRNAs in multiple body fluids (Sirker et al. 2017). For example, miR-451 was expressed in a higher abundance in peripheral blood as compared to the other body fluids studied. This trend of differentiation, and not true specificity, has occurred in the majority of research investigating the link between miRNAs and body fluid identification. Very few research studies have been able to discover a miRNA that is truly specific to a body fluid (i.e. isn't expressed in any other body fluid at all). While identification of body fluids can be done with differential expression of miRNAs, discovering miRNAs that are truly specific to an individual body fluid would be the best possible result achievable.

Although miRNAs with true specificity to one body fluid have been difficult to isolate, one study performed by Sauer et al in 2016 was successful in finding a miRNA that was truly specific to semen (Sauer et al., 2016). This miRNA, miR-891-5p, was shown to be specific to semen via RT-qPCR (Sauer et al., 2016). Semen was the only body fluid that showed expression of miR-891-5p, leading to the conclusion that this miRNA seems to be truly specific to only semen. This study is very important because it shows that true specificity is possible when trying

to link a miRNA to a body fluid. It is important to note, however, that further validation should be done to ensure that miR-891-5p is actually semen specific and is not unique to this study. It is hopeful, however, that this study did show the potential of miRNAs to be truly specific to forensically relevant body fluids.

2.4 **Next Generation Sequencing for the forensic application of microRNAs**

Next Generation Sequencing has seldomly been used very often in relation to identifying forensically relevant body fluids and even fewer have incorporated miRNAs. One research study, performed by Giampaoli *et al* in 2017, investigated the ability of the MiSeq platform (Illumina) to sequence and identify vaginal secretions among other samples (oral, fecal, and yogurt) and compared this ability to that of the more traditional method of RT-qPCR analysis (Giampaoli et al., 2017). Rather than use miRNAs for this purpose, this study simply used extracted DNA from the samples in question. Results showed that the RT-qPCR method is usually sufficient for vaginal fluid identification. The MiSeq system, however, was shown to be effective in determining if a sample was vaginal material when RT-qPCR analysis was inconclusive (Giampaoli et al., 2017). Because this NGS system showed some success with vaginal fluid identification, more research could be done with other forensically important body fluids. Because Next Generation Sequencing platforms have the ability to sequence RNA and even miRNAs, there is the potential that it can be used to identify miRNAs that could be used for body fluid identification.

One study did, in fact, use Next Generation Sequencing technology to identify a panel of miRNAs that appeared to be specific to forensically relevant body fluids. In 2016, Seashols-Williams et al used the HiSeq platform supplied by Illumina® to sequence the small RNAs

found in venous blood, semen, vaginal fluid, menstrual blood, saliva, urine, feces, and perspiration (Seashols-Williams et al, 2016). Results from the NGS analysis showed several miRNAs that appeared to be specific to individual body fluids based on their expression. The panel of miRNAs developed from the NGS data was validated using RT-qPCR, which is the standard method. It was found that none of the miRNAs identified as potentially specific to body fluids by the HiSeq platform actually showed true discriminatory capabilities for the body fluids investigated (Seashols-Williams, 2016). The identification of the body fluids could only performed by the traditional differential expression analysis. Therefore, there seems to be a distinct lack of predictive capabilities of the HiSeq platform to discover miRNAs that are body fluid specific. However, there were distinct limitations with this study, including the fact that the cohort of donors were of a similar demographic and small sample sizes were used for the sequencing step, which could have limited the extent of miRNAs that were sequenced. With a larger sample size and a more diverse cohort of donors, there is the potential that NGS could successfully predict miRNAs that are specific to body fluids.

CHAPTER 3: MATERIALS AND METHODS

3 CHAPTER 3: MATERIALS AND METHODS

3.1 **Scope of Research**

This research study investigated the use of miRNAs as a biomarker for identifying forensically relevant body fluids including venous blood, semen, saliva, menstrual blood, and vaginal fluid. Real time quantitative polymerase chain reaction (RT-qPCR) was used to validate the miRNAs chosen from an extensive literature review.

3.2 **Body Fluid Collection**

Following University of New Haven Institutional Review Board (IRB) approval, venous blood (n=9), semen (n=7), saliva (n=10), menstrual blood (n=10), and vaginal fluid (n=10) was collected from volunteers with informed written consent. Each participant was assigned a random code for each of their donated samples. Any individualizing information connected to a participant remained only in the hands of the co-principal investigator, Dr. Claire Glynn.

For semen, salvia, menstrual blood, and vaginal fluid donation, participants were permitted to take home a collection kit and collect their sample personally. Once the samples were collected, participants were asked to return the samples to the designated freezer in room 415 on the fourth floor of Dodd's Hall for storage at -20˚C until further use. For venous blood donation, sample collection was performed by medical personnel at the University of New Haven's Health Services clinic via venipuncture. Samples were returned to the designated refrigerator in room 415 on the fourth floor of Dodd's Hall at the University of New Haven for storage at 4˚C until further use.

For semen donation, participants were asked to ejaculateonce into a sterile 50 mL Falcon Polypropylene Conical Tube after proper hygiene practices were performed.

For saliva donations, participants were asked to collect fluid saliva by spitting into a 15 mL Falcon Polypropylene Conical Tube to at least the 2.0 mL mark. Participants were also asked to refrain from eating or drinking anything for at least thirty minutes prior to both collections.

For menstrual blood donation, participants were asked to collect the samples on the heaviest day of the menstrual cycle, which was done to ensure that there was ample menstrual blood to collect from. Participants were given three sterile swabs to swab the internal vaginal wall, making sure to swab a different area each time.

For vaginal fluid donation, participants were asked to abstain from sexual intercourse for at least four days prior to sample collection to prevent contamination from semen in the vaginal cavity. Additionally, participants were asked to wait until at least two weeks after the end of their menstrual cycle to collect the samples to ensure that there was no contamination from menstrual blood. Sample collection was carried out using the same protocol as menstrual blood collection.

For the sample collections that were performed by the participants themselves, the participants were asked to ensure that the samples were kept as cold as possible up until they were placed in the -20 ˚C freezer in Dodds Hall. This will ensure that the sample will be in its best condition before extraction.

3.3 Methods

3.3.1 RNA Extraction Procedure: miRNeasy Mini Kit

Before extraction occurred, donated samples were taken out of storage and allowed to thaw to room temperature. Liquid fluids (venous blood, semen, and saliva) were extracted using 500 mL of the fluid. Vaginal material and menstrual blood were collected using sterile swabs and

extraction, therefore, was performed directly on those swabs. The swab head was removed from the swab using sterile techniques and the extraction was performed on the swab head. Figure 4

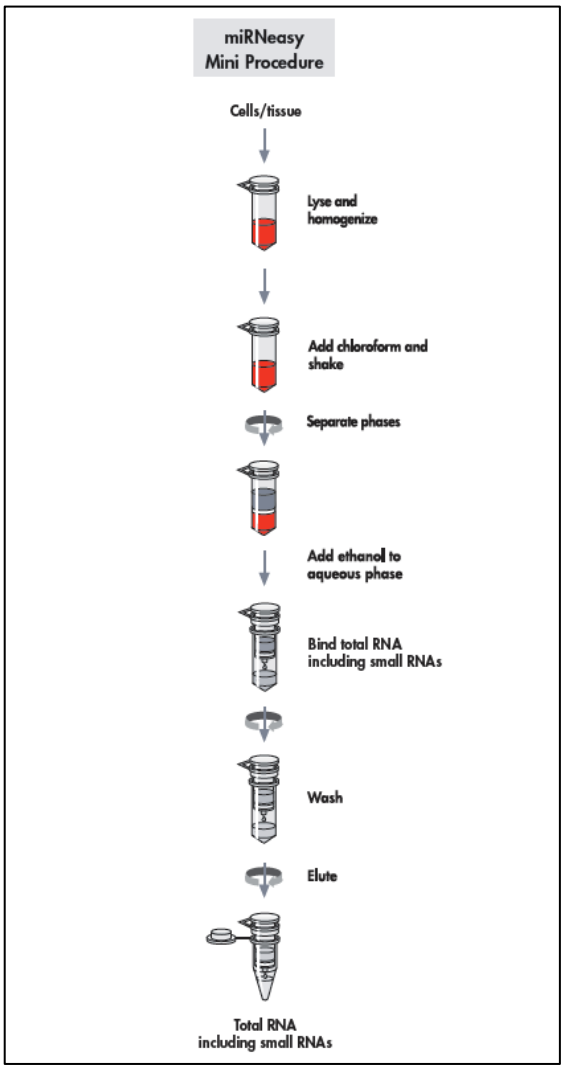

Figure 4: miRNeasy Mini Procedure

Adapted from Qiagen (2012). RNeasy® Mini Handbook. *Qiagen.* Hilden, Germany

depicts the general procedure for RNA isolation using the miRNeasy Mini Kit (Qiagen®).

Using this kit, RNA isolation was performed via a combination of a guanidine-isothiocyanate lysis step and a silica-membrane purification step. Following cell lysis and disruption, ethanol was added to the samples to enhance binding interactions between the RNA and silica

membrane. This ethanol/lysate mixture was then loaded into the silica membrane and purified using a series of wash buffers, including a purification step using DNase. This step was used to eradicate any DNA from the samples. The concentrated RNA was then eluted into 40 µL of RNase free water. Following miRNA isolation, the extracts were stored at -20˚C until further use. Refer to Appendix A for more detailed instructions on the proper protocol.

3.3.2 RNA Quantification and Quality Assessment

3.3.2.1 NanoDrop One^{C}

The NanoDrop One^{C}, which is an ultraviolet-visual (UV-Vis) spectrophotometer, was the first of four instruments utilized. Samples were placed on a collection pedestal (1-2 µL each) where UV-Vis light was passed through each sample. The wavelength at which this light was absorbed was utilized by the instrument to formulate both concentration and purity measurements. Beer's Law (as depicted in Equation 1) was used to find the actual concentration.

$$A = \varepsilon b C$$

$\mathrm{A} = \text{absorbance}$
$\mathrm{\varepsilon} = \text{molar absorptivity}$
$b = \text{path length}$
$C = \text{concentration}$

Equation 1: Beer's Law

This instrument can detect concentrations of RNA from 1.2 ng/µL to 22,000 ng/µL. It is important to note, however, that this instrument is not human specific.

3.3.2.2 Quibit 3 Fluorometer

Secondly, the Qubit 3 Fluorometer using the RNA HS Assay kit was utilized as another method for the quantification of the RNA extracts. This method utilizes highly specific

quantification assays using fluorescent probes. It is considered to be more comprehensive and reliable than UV-Vis methods. Additionally, it is considered more human specific. Refer to Appendix B for more detailed instructions on the proper protocol.

3.3.2.3 Agilent Bioanalyzer

Lastly, the Agilent 2100 Bioanalyzer was utilized using both the Small RNA kit and the RNA 6000 Nano kit. The chips utilized within these kits essentially perform small-scale capillary electrophoresis. The Small RNA kit was utilized to determine the percentage of miRNA within each sample, compared to the other small RNA content. The Nano 6000 kit served two purposes: the first purpose was to determine the human specific concentration of RNA within each extract and the second purpose was to provide a quality assessment of the of the sample in the form of an RNA integrity number (RIN). Refer to Appendix C and D for more detailed instructions on the proper protocol for each kit.

3.3.3 Next Generation Sequencing via MiSeq FGx Platform

3.3.3.1 Sample Preparation Using TruSeq Small RNA Protocol

After quantification of the samples was complete, several rounds of sequencing were performed using venous blood, semen, vaginal fluid, menstrual blood, and saliva.

The TruSeq Small Library Prep Reference Guide was followed for sample preparation. On Day 1 of sample preparation, 3’ and 5’ ends of the miRNA strands within the sample were ligated with RNA 3’ Adapter and RNA 5’ Adapter, respectively. This was done by pipetting and mixing specific volumes of the samples and Illumina® provided consumables together and subjecting the samples to specific temperatures at pre-determined periods of time. Then, the

samples were reverse transcribed into complementary DNA (cDNA) using SuperScript II Reverse Transcriptase provided by Thermo Fisher Scientific®. A thermal cycler was used for incubating the samples at 50˚C for one hour. After reverse transcription was completed, the individual samples were indexed with a specific RNA PCR Primer Index. Each sample received a different index, which allowed the samples to be easily identified. Amplification was then performed on a thermal cycler for 15 cycles, which was the maximum number of cycles recommended (see Appendix E for full details). At this point the samples were stored at -20˚C until further used.

On Day 2 of the sample preparation, the cDNA samples were pooled into one microcentrifuge tube, mixed together and purified via gel electrophoresis. A 1X Novex TBE Buffer, the XCell SureLock System (as depicted in Figure 5), Novex TBE gels (6%, 10 wells,

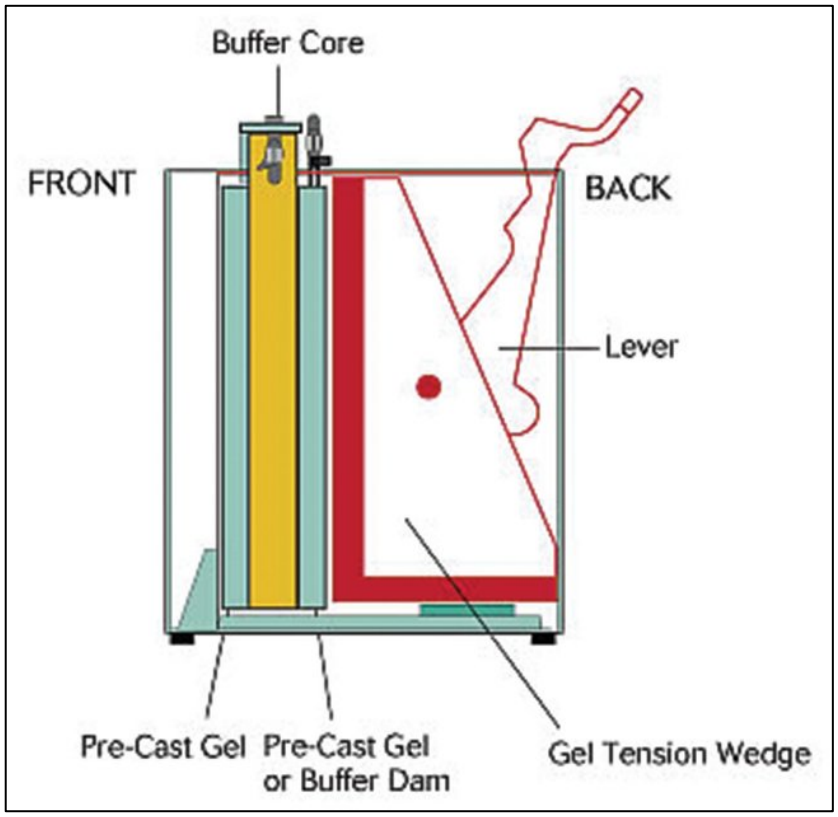

Figure 5: XCell SureLock System Set Up
As adapted from fishersci.com

pre-set) were used for this method. Each was provided by all provided by Thermo Fisher Scientific®

The pooled samples were stained with a loading dye and then loaded into two wells (25 µL each). On either side of the samples, a Custom RNA Ladder and High Resolution Ladder (both stained with the same loading dye) were loaded into the wells. The gel was subjected to 145 V of electricity for approximately 60 minutes. After the gel completed running, the gel was stained with ethidium bromide (0.5 µL/mL in water) for several minutes. The gel was then visualized with a UV-transilluminator. The bands located at the 147 nt area were the mature miRNA and were cut out for further DNA purification. An example of what the purified cDNA content after gel electrophoresis purification is depicted in Figure 6.

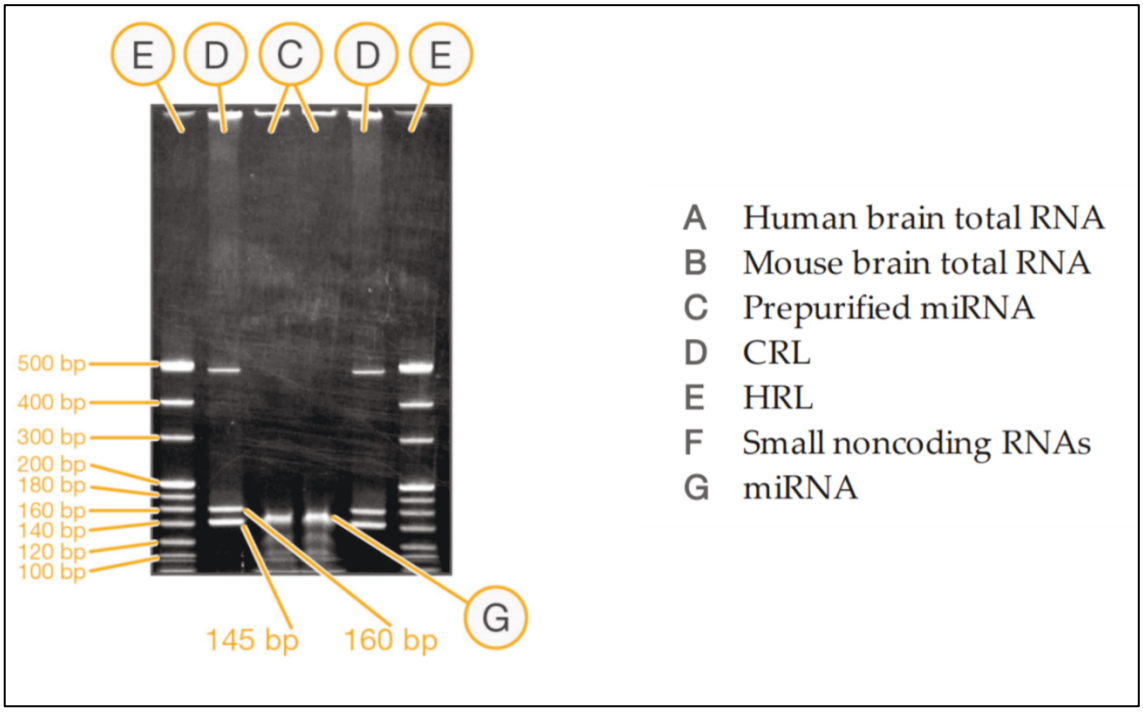

Figure 6: Example of Purified cDNA after Electrophoresis

As adapted from Illumina. (2015). TruSeq® Nano DNA Library Prep Reference Guide. *Illumina*, 1–38. https://doi.org/FC-121-9006DOC

The bands were placed in a 0.5 mL gel breaker tube nested in a 2.0 mL microcentrifuge tube and spun down. Ultrapure water (300 µL) was added to the gel debris within the 2.0 mL microcentrifuge tube and shaken at room temperature for four hours. The final sample was concentrated using a series of washes with glycogen, 3M NaOAc, and ethanol (for a full protocol, please see Appendix). The resulting pellet was resuspended in 10 uL of 10 mM Tris-HCl, pH 8.5. The quality and quantity of the purified DNA was analyzed using the DNA 1000 kit in conjunction with the 2100 Bioanalyzer provided by Agilent Technologies®. The sample was then stored at -20˚C until further use. Please refer to Appendices E-G for full protocols.

3.3.3.2 Sequencing of Prepared Samples with MiSeqFGx

Before sample preparation started, the MiSeqFGx was prepared for sequencing. A sample sheet was created detailing the necessary information for each sample including what index paired with each sample.

The final sample, or Library, was diluted to a 4 nM concentration based on the concentration obtained from the 2100 Bioanalyzer. Then, the 4 nM library was denatured and brought to a concentration of 12 pM.

The PhiX control was then diluted down to a concentration of 4 nM. The diluted PhiX control was then denatured and diluted down to a concentration of 12 pM. The 12 pM library was then spiked with a 25% volume of the 12 pM PhiX control.

The KV 3 reagent cartridge was thawed and then prepared for sequencing. Then, 600 µL of the final library (spiked with 25% 12 pM PhiX) was loaded into the cartridge. Then the cartridge was inserted into the MiSeqFGx and a sequencing run was started. After the run

completed, a post-run wash was performed. Refer to Appendixes E-G for more detailed instructions on the proper protocols for Next Generation Sequencing sample preparation.

3.3.3.3 Data Analysis

The data files from the MiSeq FGx were taken from the instrument and uploaded to an Illumina sponsored website called BaseSpace, which analyzed the raw sequencing data. This is also the website where statistical analysis could be done on said data.

3.3.4 Literature Review for Unique miRNAs

A literature search was performed to determine which miRNAs have been shown to be specific to the five body fluids being studied (menstrual blood, semen, vaginal material, menstrual blood, and saliva). A panel of miRNAs was chosen based on the literature review for further validation.

3.3.5 RT-qPCR Validation of miRNAs

3.3.5.1 Reverse Transcription

Reverse transcription (RT) was performed to transform the extracted miRNA into its complementary DNA (cDNA). For each sample, the reaction was performed using the following amounts of reagents represented in Table 1.

Table 1: Reagent Volumes for Reverse Transcription

Reagent	Volume Added to Master Mix per sample
Nuclease Free water	4.57 µL
dNTP mix (100nM)	0.17 µL
10x RT Buffer	1.65 µL
miRNA Specific Stem Loop Primer (50nM)	3.1 µL
Multiscribe (50 U/µL)	1.1 µL
RNase Inhibitor (20 U/µL)	0.21 µL
Premix Volume	**10.0 µL**
Total/miRNA (1-10 ng)	5.0 µL
Total Reaction Volume	**15.0 µL**

Reverse transcription was performed in a hood with functioning laminar flow. The hood and items within the hood were cleaned with 70% Industrial Reagent Alcohol and underwent fifteen minutes of UV exposure before its use. This was done to maintain a sterile environment. The components of the Master Mix were allowed to thaw on ice before use. The Master Mix was prepared in brown 1.5 mL microcentrifuge tubes to protect the miRNA-specific Stem Loop Primers from the light. A Master Mix was prepared for each individual miRNA for the samples chosen for analysis.

miR-16 was chosen as an endogenous control for all five body fluids and, therefore, was reverse transcribed for every single sample chosen for analysis.

Each prepared sample was centrifuged prior to being loaded onto the thermal cycler. The reverse transcription reaction was performed using a GeneAMP® PCR System 9700 Thermal

Cycler (Applied Biosystems). For each miRNA target group, a negative RT blank control was prepared. These blanks consisted of the RT master mix and biology grade water. This was included to ensure no contamination occurred during the reverse transcription step. Table 2 depicts the parameters that the Thermal Cycler was run at.

Table 2: Thermal Cycler Parameters for RT Reaction

	Time	**Temperature**
Hold	30 Minutes	16˚C
Hold	30 Minutes	42˚C
Hold	5 Minutes	85˚C
Hold	∞	4˚C

After the RT reactions were completed, the cDNA samples were centrifuged and the placed into storage at -20˚C until further use.

3.3.5.2 Relative Quantification – Polymerase Chain Reaction (qPCR)

After proper thawing on ice, the cDNA samples were amplified using the ABI 7500 Real Time PCR System (Applied Biosystems). Relative quantification PCR involves the real time monitoring of the progress of the amplification. The quantities of the target molecules are measured after each cycle completes. Each cycle consists of a denaturing step, a primer annealing step, and a template extension step. The process as a whole consists of three separate phases. The first phase is the exponential phase, where the amount of product is doubled during each cycle. The next two phases, the linear phase and plateau phase, have a reduced detection limit and, therefore, are less important for the calculation of relative quantification. In the current

procedure, the PCR reaction ran for 40 cycles. If the target miRNA was in a particular sample, it would be expected that the amplified products within the sample would reach the Cycle Threshold (C_t) by the end of the 40 cycles. The Cycle Threshold is the point at which the amplified product produces a detectable fluorescent signal. For example, a lower amount of starting material will take longer to accumulate to a detectable level and will, therefore, result in a higher C_t-value. The $C_t$value can then be used to make conclusions on the relative quantities of the target within each sample.

All reactions were prepared in a vented hood that was cleaned with 70% Industrial Reagent Alcohol and treated with UV light for 15 minutes before use. Throughout the preparation process, a clean and sterile environment was maintained. The cDNA samples that had been previously prepared were allowed to thaw on ice. A master mix for each miRNA target were prepared in 1.5 mL brown centrifuge tubes to protect the light sensitive miRNA target reagents. Each sample was prepared in triplicate using a master mix containing the reagents in Table 3.

Table 3: Master Mix for Real Time PCR

Reagents	**Master Mix Volume (per one sample)**
Taqman® Universal PCR Master Mix	5.0 µL
Nuclease Free Water	3.8 µL
miRNA PDAR	0.5 µL
Total Master Mix Volume (Premix)	9.3 µL
cDNA	0.7 µL
Total Volume of Reaction	10.0 µL

The master mix and samples were dispensed in a 96-well plate. Along with the samples, a no target control (NTC) was included within the plate. The NTC consisted of the PCR Master mix and nuclease free water, with no cDNA present. This was included to determine if any contamination occurred during this step. The Reverse Transcription blank (RT) for each miRNA target was also included with each of the reactions. In each of these instances, contamination would be seen as amplified cDNA product. Additionally, each plate included an inter assay control (IAC) to ensure that each separate plate was performing in a similar manner. The IAC was determined to be the sample with the highest original RNA concentration and was reverse transcribed using the miRNA-16 target. Figure 7 depicts an example of a potential set-up of a

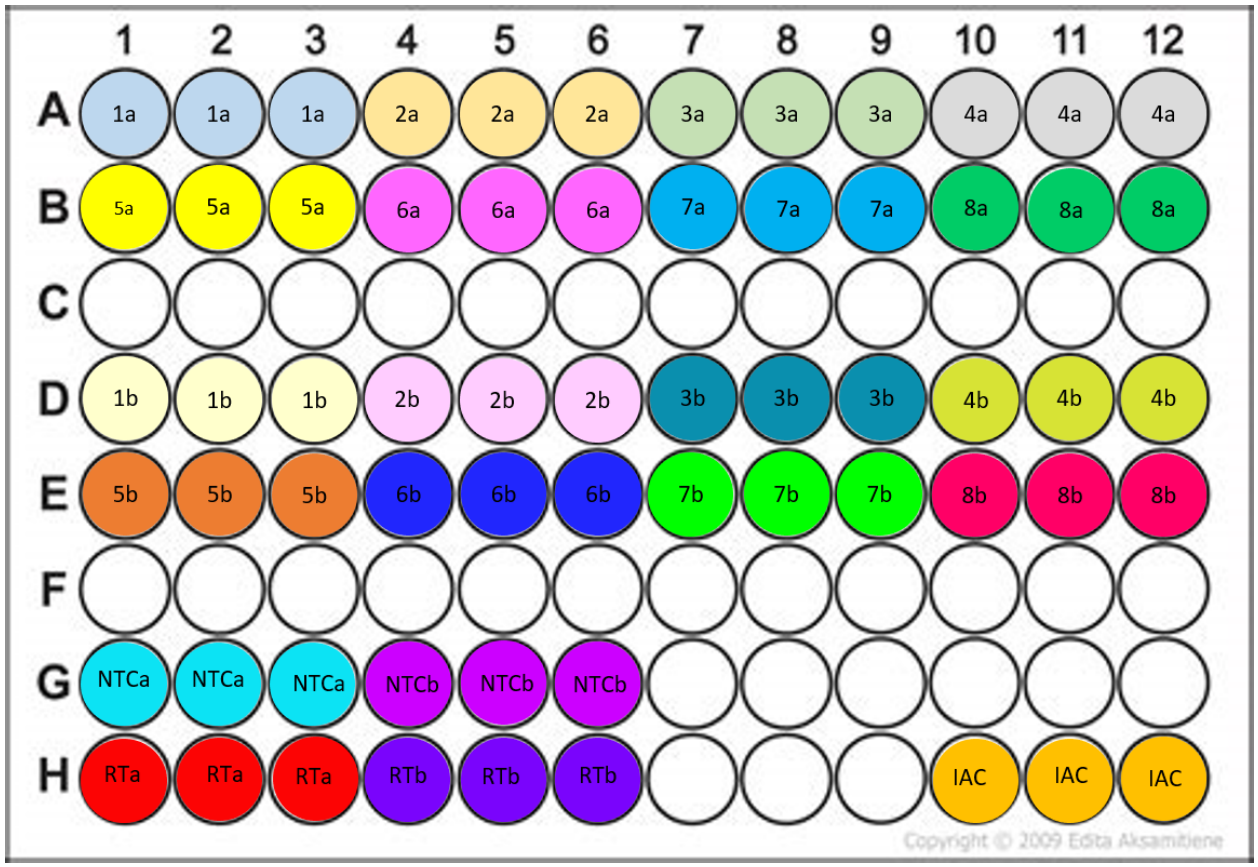

Figure 7: Example of a 96-well Plate Set Up

96-well plate that holds eight samples of two different miRNA targets (a and b), the appropriate NTC, the appropriate, RT blanks, and the IAC.

After all of the components of a plate reaction were prepared, the plate was centrifuged for approximately 30-60 seconds. It was then loaded into the ABI 7500 Real-Time PCR system and run at the following setting with a final reaction volume of 10 µL (Table 4).

Table 4: The Running Parameters of the ABI 7500 Real-Time PCR System

	Time	**Temperature**
Hold	2 Minutes	50˚C
Hold	10 Minutes	95˚C
40 Cycles	15 Seconds	95˚C
	60 Seconds	60˚C

After the run was complete, the plate and its components were disposed of in biohazardous waste.

3.3.6 RT-qPCR Data Statistical Analysis

miRNA expression analysis requires the use of an endogenous control. In this thesis, an endogenous control is a miRNA that has expression levels that are known to be relatively stable throughout the body fluids being tested. Based on a literature review, miR-16 was chosen as the endogenous control. Therefore, each miRNA target in each sample being studied was expressed relative to the miR-16 expression in order to normalize the results. An average C_t value was calculated based on the triplicate of each sample. The average C_t value for the endogenous

control for that sample was subtracted from the target miRNA of interest. This value was considered the ΔC_t value, as seen in Equation 2.

$$\Delta Ct = C_{T(target\ miRNA)} - C_{T(endogenous\ control)}$$

Equation 2: ΔC_t Value

Statistical analysis was carried out using the ΔC_t-values of each sample using the statistical software package, Minitab 18.0 (Mintab Ltd). Two sample T-tests and one-way ANOVA was used to determine association and comparison between different target miRNA groups and determine the possible statistical significance. Results with a p-value less than 0.05 were considered statistically significant.

CHAPTER 4: RESULTS

4 CHAPTER 4: RESULTS

4.1 RNA Quantification Results

4.1.1 Venous Blood

All venous blood samples had measurable RNA content with each method used. The average concentrations for the total RNA content using NanoDrop One^{C} method, Qubit method, and the RNA 6000 Nano kit/2100 Bioanalyzer method were 20.02 ng/µL, 27.934 ng/µL, and 9.67 ng/µL respectfully. The average RNA Integrity number (RIN) was 4.7. It is important to note that with both the Qubit method and RNA 6000 Nano kit/2100 Bioanalyzer, several samples were too low in concentration for the instrument to determine the concentration or the RIN. According to the Small RNA/2100 Bioanalyzer method, the average concentration for small RNA within the sample was 17.78 ng/µL and the average miRNA concentration was 4.76 ng/µL. A full representation of the data from each of these methods can be found in the Appendix.

4.1.2 Menstrual Blood

All menstrual blood samples had measurable RNA content with each method used. The average concentrations for the total RNA content using NanoDrop One^{C} method, Qubit method, and the RNA 6000 Nano kit/2100 Bioanalyzer method were 120.20 ng/µL, 150.06 ng/µL, and 99.23 ng/µL respectfully. The average RNA Integrity number (RIN) was 4.4. It is important to note that with the Qubit, several samples were too low in concentration for the instrument to determine the concentration or the RIN. According to the Small RNA/2100 Bioanalyzer method, the average concentration for small RNA within the sample was 362.23 ng/µL and the average miRNA concentration was 98.44 ng/µL. A full representation of the data from each of these methods can be found in the Appendix.

4.1.3 Semen

All semen samples had measurable RNA content with each method used. The average concentrations for the total RNA content using NanoDrop One^{C} method, Qubit method, and the RNA 6000 Nano kit/2100 Bioanalyzer method were 54.16 ng/µL, 47.91 ng/µL, and 27.38 ng/µL respectfully. The average RNA Integrity number (RIN) was 2.93. It is important to note that with both the Qubit method and RNA 6000 Nano kit/2100 Bioanalyzer, several samples were too low in concentration for the instrument to determine the concentration or the RIN. According to the Small RNA/2100 Bioanalyzer method, the average concentration for small RNA within the sample was 189.51 ng/µL and the average miRNA concentration was 90.023 ng/µL. A full representation of the data from each of these methods can be found in the Appendix.

4.1.4 Vaginal Material

All vaginal materials samples had measurable RNA content with each method used. The average concentrations for the total RNA content using NanoDrop One^{C} method, Qubit method, and the RNA 6000 Nano kit/2100 Bioanalyzer method were 372.26 ng/µL, 413.84 ng/µL, and 314.77 ng/µL respectfully. The average RNA Integrity number (RIN) was 3.88. According to the Small RNA/2100 Bioanalyzer method, the average concentration for small RNA within the sample was 1241.65 ng/µL and the average miRNA concentration was 162.77 ng/µL. A full representation of the data from each of these methods can be found in the Appendix.

4.1.5 Saliva

All saliva samples had measurable RNA content with each method used. The average concentrations for the total RNA content using NanoDrop One^{C} method, Qubit method, and the RNA 6000 Nano kit/2100 Bioanalyzer method were 76.92 ng/µL, 112.87 ng/µL, and 86.82 ng/µL respectfully. The average RNA Integrity number (RIN) was 2.64. It is important to note that with both the Qubit method and RNA 6000 Nano kit/2100 Bioanalyzer, several samples were too low in concentration for the instrument to determine the concentration or the RIN. According to the Small RNA/2100 Bioanalyzer method, the average concentration for small RNA within the sample was 23.22 ng/µL and the average miRNA concentration was 11.23 ng/µL. A full representation of the data from each of these methods can be found in the Appendix.

4.2 Sequencing Results

The first several attempts at sequencing various samples, the sequencing run failed without collecting any data. After consulting with Illumina representatives and an on-site visit from an Illumina engineer, it was determined that the MiSeq FGx sequencing platform had been improperly aligned and would not have produced any successful results.

After the instrument was properly calibrated, another sample run was attempted. While the run completed without errors, no miRNAs were identified as being sequenced. Other small RNAs, however, were successfully sequenced. Through consultations with Illumina representatives, one last sequencing attempt was performed with two body fluid samples that had the highest RIN to determine if sequencing would be successful in this case. The reasoning for this was that if the sequencing was not successful with only two sample of high quality, then

there was little chance that it would be successful with more samples of lower quality. The sequencing run completed without errors, but no miRNAs were successfully sequenced. Like previous runs, other small RNAs were sequenced. Figure 8 shows a representative pie chart

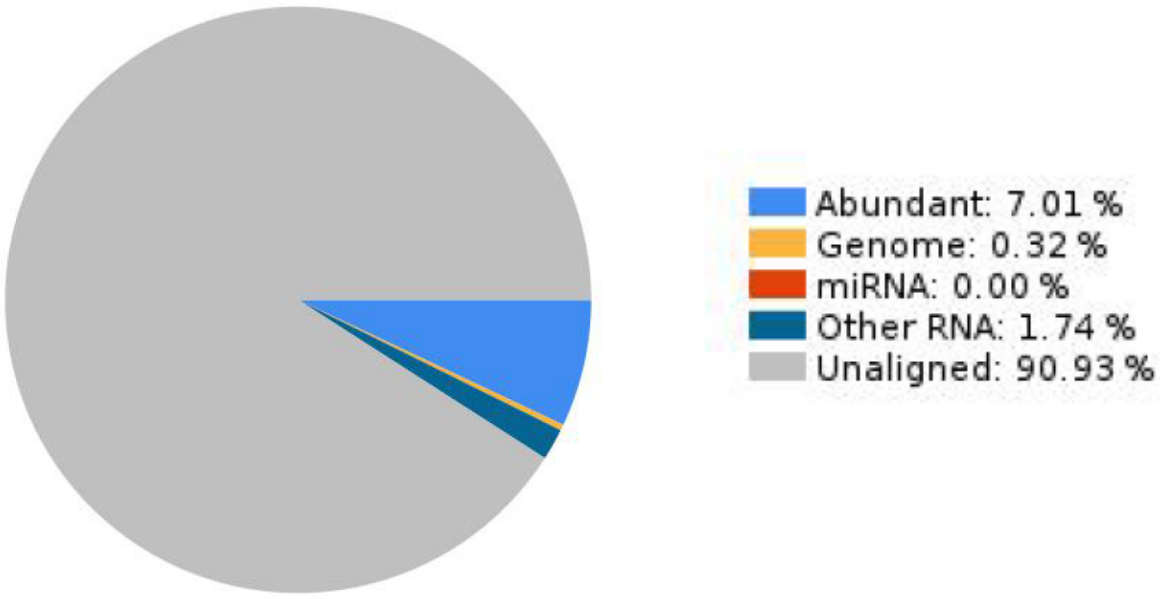

Figure 8: Sequencing Results

showing the final sequencing results from these runs. After the last, unsuccessful sequencing results, miRNAs were chosen for further validation based on an extensive literature review.

4.3 miRNAs Chosen from Extensive Literature Review

After an extensive literature review, several miRNAs were chosen for further validation based on their previously reported specificity for specific body fluids. Additionally, a miRNA was chosen for normalization purposes. Table 5 shows the miRNAs chosen for validation and for which body fluid they were reported to be specific for.

Table 5: miRNAs Chosen for Further Validation

miRNA Chosen for Validation	Body Fluid Biomarker
miR-451	Venous Blood
miR-412	Menstrual Blood
miR-891a	Semen
miR-10b	Semen
miR-205	Saliva
miR-124a	Vaginal Fluid
miR-16	Normalizer

The miRNAs were chosen for further validation because they were the most commonly associated miRNAs to the particular body fluid they were chosen for. While they were not the only miRNAs found in the literature search, they appeared the most and seemed to be the most specific to the chosen body fluids.

4.4 miRNA Expression via RT-qPCR Results

4.4.1 Statistical Parameters

For each miRNA, the difference in the expression levels across all five body fluids was assessed with a one-way ANOVA test. The p-value required for a statistically significant difference in expression was less than 0.05. For each miRNA, the expression level of the target body fluid was compared to the other four body fluids to determine if the difference was

statistically significant. The p-value required for a statistically significant difference in expression was less than 0.05.

4.4.2 miR-451 Expression Results

Figure 9 shows the results of the statistical analysis of the expression of miR-451 represented in the form of boxplots.

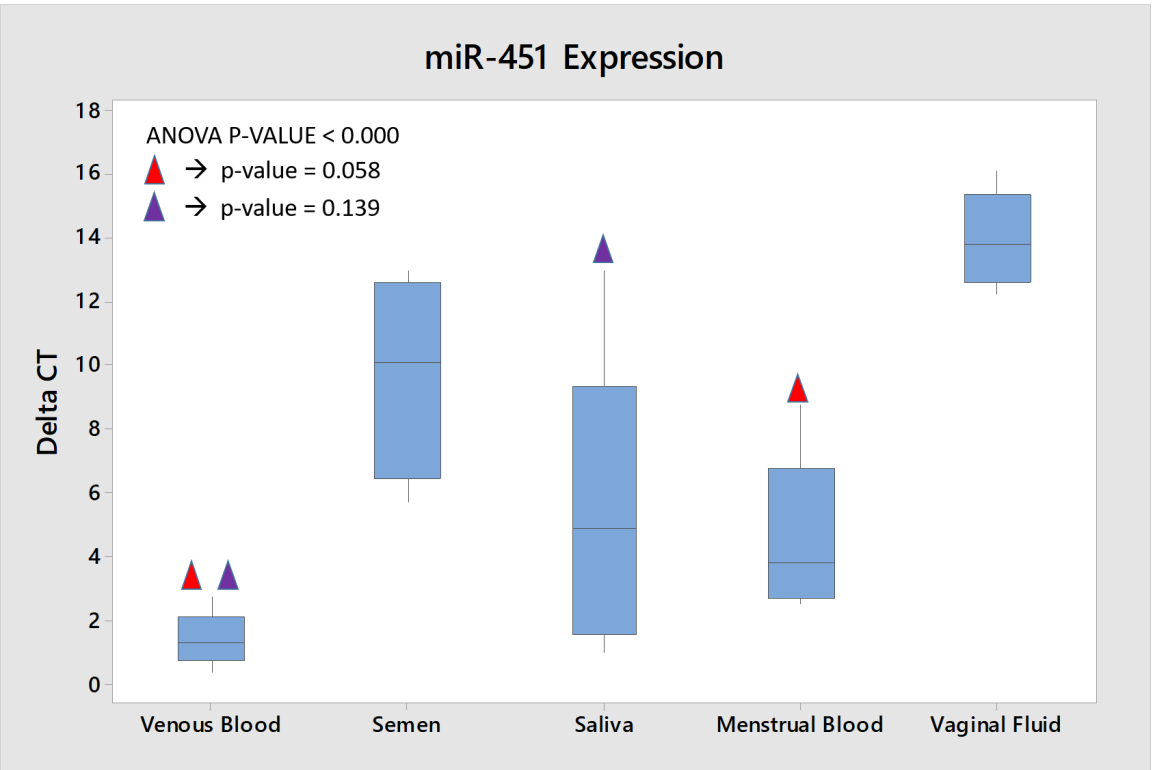

Figure 9: miR-451 Expression Boxplots

An ANOVA test performed in Minitab 18 showed that there was a statistically significant difference in the expression of miR-451 across all five body fluids after normalization. As the statistical results for miR-451 expression shows, when comparing venous blood to saliva, the p-value was 0.139, which was above the p-value required for a statistically significant difference. When comparing venous blood to menstrual blood, the p-value was 0.058, which was also above

the required p-value for a statistically significant difference. When comparing venous blood to semen and vaginal fluid, however, the p-values < 0.05, indicating that these differences in expression levels were statistically significant. It is important to note that the range of expression levels for both saliva and semen does appear be large.

4.4.3 miR-412 Expression Results

Figure 10 shows the results of the statistical analysis of the expression of miR-412 represented in the form of boxplots.

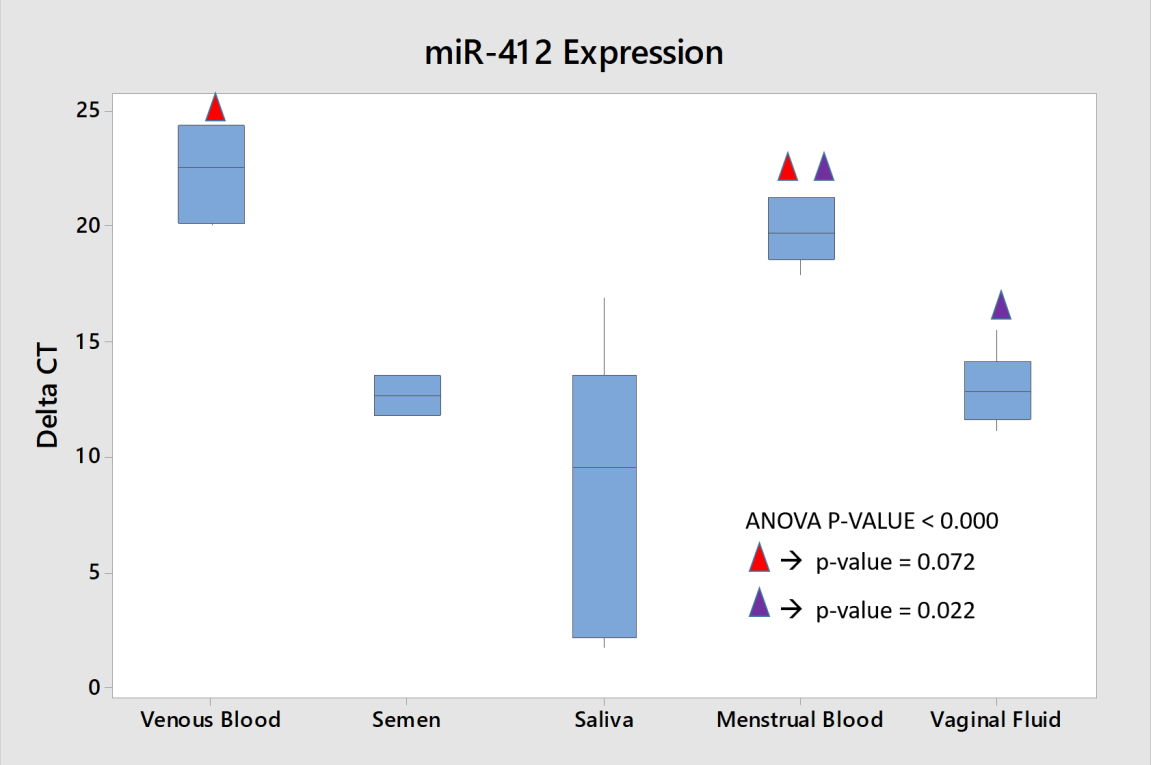

Figure 10: miR- 412 Expression Boxplots

An ANOVA test performed in Minitab 18 showed that there was a statistically significant difference in the expression of miR-412 across all five body fluids after normalization. As the results show, the p-value was 0.072 was when comparing the expression levels of miR-412 between menstrual blood and venous blood. Because this p-value was above the p-value required

for a statistically significant difference in expression levels, this difference cannot be considered significant. When comparing menstrual blood to vaginal fluid, the expression levels appeared to overlap. When statistical testing was performed, however, the p-value was 0.022, which was below the required p-value for a statistically significant difference in expression levels. This indicated that the difference in expression levels between these two body fluids was statistically significant. When comparing menstrual blood to saliva and semen, the p-values < 0.05, indicating that these differences in expression levels were statistically significant. It is important to note that the range of expression levels for saliva does appear to be large.

4.4.4 miR-891a Expression Results

Figure 11 shows the results of the statistical analysis of the expression of miR-891a represented in the form of boxplots.

Figure 11: miR-891a Expression Boxplots

An ANOVA test performed in Minitab 18 showed that there was a statistically significant difference in the expression of miR-412 across all five body fluids after normalization. As the results for miR-891a expression shows, the expression levels for semen and saliva appear to overlap. The statistical testing, however, showed that the p-value was 0.026, which was lower than the p-value required for a statistically significant difference between the expression levels. Therefore, semen and saliva showed a statistically significance difference in their expression. When comparing semen to venous blood, menstrual blood, and vaginal fluid, the p-values were also less than the p-value required for statistically significant difference in expression levels. This indicated that the difference between these body fluids was statistically significant.

4.4.5 miR-10b Expression Results

Figure 12 shows the results of the statistical analysis of the expression of miR-891a represented in the form of boxplots.

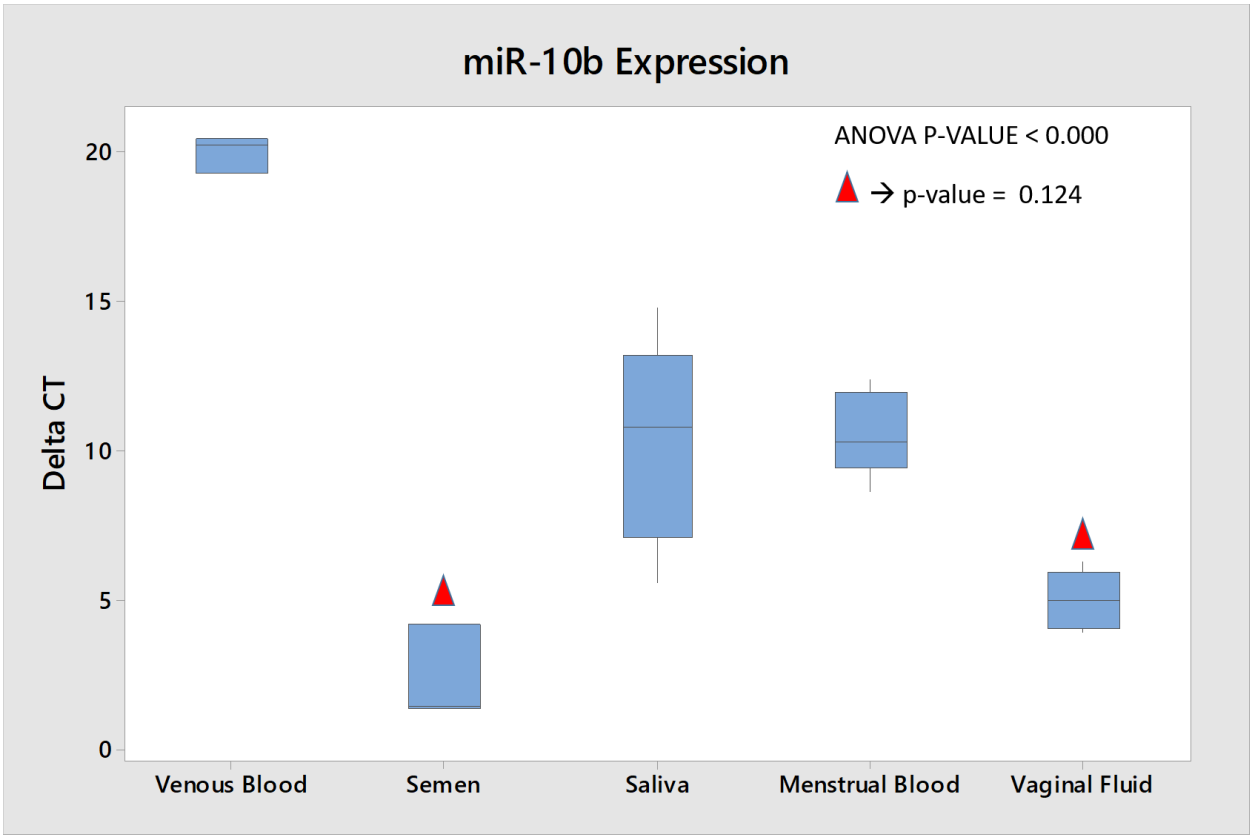

Figure 12: miR-10b Expression Boxplots

An ANOVA test performed in Minitab 18 showed that there was a statistically significant difference in the expression of miR-10b across all five body fluids after normalization. As the results for miR10b expression shows, the p-value attributed to the difference between semen and vaginal fluid was 0.124. Because this p-value exceeds 0.05, this difference was not statistically significant. The results also show that the differences in expression levels between semen and venous blood, saliva, and menstrual blood have p-values less than 0.05. This indicates that these expression differences are statistically significant. The results also show that the expression differences between venous blood and all other body fluids is also statistically significant.

4.4.6 miR-205 Expression Results

Figure 12 shows the results of the statistical analysis of the expression of miR-891a represented in the form of boxplots.

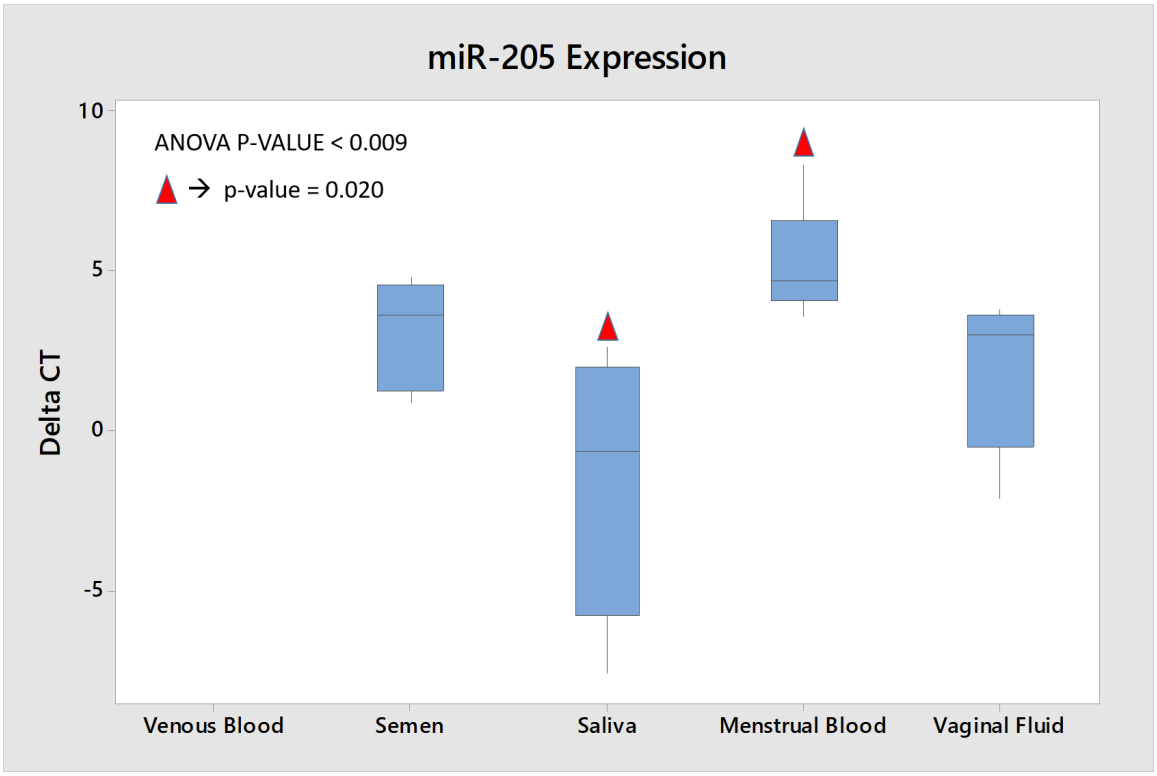

Figure 13: miR-205 Expression Boxplots

An ANOVA test performed in Minitab 18 showed that there was a statistically significant difference in the expression of miR-10b across all five body fluids after normalization. When comparing the expression levels of saliva to semen and vaginal fluid, the p-values were both greater than 0.05 (0.066 and 0.156, respectively). This indicated that the difference between the expression levels was not statistically significant. The expression results also shows that when comparing saliva expression levels to the expression levels of menstrual blood, the p-value was 0.020, which indicated that the difference between the expression levels was statistically

significant. Additionally, venous blood was shown to have no miR-205 expression. It is important to note that the range of expression levels for saliva was extensive.

CHAPTER 5: CONCLUSIONS

5 CHAPTER 5: CONCLUSIONS

5.1 Sequencing via MiSeq FGx Platform

After several attempts of sequencing the miRNA content within the body fluid samples using the MiSeq FGx Platform in conjunction with the TruSeq Small RNA Kit, there was no usable miRNA data. Because other types of small RNA were successfully sequenced within the samples, the sequencer is capable of sequencing RNA content. Based on these results, it appears that utilizing the MiSeq FGx platform with the TruSeq Small RNA kit may not be the optimal technique for sequencing the miRNA content in body fluid samples.

5.2 miRNA Expression via RT-qPCR

5.2.1 miR-451 Expression Results

According to studies performed by Hansen *et al*, Courts *et al,* Sirker *et al*, and Mayes *et al* mir-451 has showed greater expression levels in venous blood as compared to other forensically relevant body fluids. It was the goal of this study to determine if miR-451 could be used to not only differentiate between blood and other body fluids, but also differentiate between venous blood and menstrual blood. Based on the statistical analysis of miR-451 the current study, showed that miR-451 couldn't be used to differentiate venous blood from menstrual blood and saliva because the differences between the expression levels was not statistically significant. Venous blood, however, could be differentiated from both semen and vaginal fluid using miR-451 because the differences in the expression levels was statistically significant. More specifically, venous blood was more highly expressed compared to that of semen and vaginal fluid.

These results are inconsistent with the studies found in the literature. The majority of the studies found that miR-451 could, at the very least, be used to differentiate between blood and all other body fluids. The fact that in this study venous blood could not be used to differentiate between venous blood and saliva contradicts these findings.

The range of expression levels of miR-451 within semen and saliva is quite large. This can be problematic because it shows great variability of expression within one body fluid. This may make it more difficult to make body fluid identification using miR-451. Ideally, the expression levels of a miRNA should be small so the results can be trusted more. A larger sample size may minimize this effect.

5.2.2 miR-412 Expression Results

It was previously reported by Hansen *et al* and Mayes *et al,* that miR-412 showed increased expression in menstrual blood compared to other forensically relevant body fluids, including menstrual blood. This differential expression allowed researchers to identify menstrual blood compared to other body fluids. It was the goal of this study to determine if miR-412 could be used to differentiate between menstrual blood and other body fluids, especially venous blood. Based on the statistical analysis of the expression levels of miR-412 within this study, it cannot be used to differentiate between menstrual blood and venous blood because the difference between the expression levels was not statistically significant. These results are inconsistent with that found within the published literature, which indicated miR-412 expression should significantly different between menstrual and venous blood.

The results from this study do, however, show that miR-412 can be used to differentiate menstrual blood from semen, saliva, and vaginal fluid because the difference between the expression levels was statistically significant. More specifically, the expression levels of miR-

412 are more highly expressed in menstrual blood compared to semen, saliva, and vaginal fluid. This increased expression is compatible with what is seen in the published literature. Overall, based on the results of this study, miR-412 can be used to differentiate between blood (venous and menstrual) and other forensically relevant body fluids.

It is important to note that the range of expression levels of miR-412 in saliva was large. As stated before, this can be problematic because this shows great variability of expression of miR-412 within saliva, which could make it more difficult to successfully differentiate saliva from menstrual blood using this miRNA.

5.2.3 miR-891a Expression Results

It was previously reported by Sauer *et al,* Mayes *et al,* Seashols-Williams *et al,* and Tian *et al,* that miR-891a was more abundantly expressed within semen compared to other body fluids. This differential expression lead to the ability of researchers to identify semen using miR-891a. It was the goal of this study to validate the identification capabilities of miR-891a based on differential expression. Based on the statistical analysis of the expression of miR-891a across the body fluids, miR-891a can be used to differentiate semen from all other body fluids because the difference between the expression levels was statistically significant. More specifically, semen was more highly expressed compared to all other body fluids. These results are consistent with that found in the literature, with the exception of Saur *et al* which showed that miR-891a was only expressed in semen (Saur *et al,* 2016). This would indicate that miR-891a is not truly specific to semen as indicated in Saur *et al*.

The range of expression levels of miR-891a in saliva, however, is large. Although difference the expression levels between semen and saliva is statistically significant, the high variation of miR-891a expression in saliva could potentially encroach into the expression levels

of semen. This could make a body fluid identification slightly more difficult. However, more experiments should be done to determine if this could truly become an issue.

5.2.4 miR-10b Expression Results

According to previous studies performed by Hanson *et al*, Sauer *et al*, Sirker *et al*, Mayes *et al*, and Tian *et al*, the expression of miR-10b was shown to be more abundant in semen compared to other forensically relevant body fluids. The differentiation and identification capabilities of miR-10b was the basis of why it was chosen for further validation in this study. It was the goal of this study to determine if miR-10b could be used to differentiate between semen and other body fluids. Based on the statistical analysis of miR-10b, the p-value attributed to the difference in expression between semen and vaginal fluid was 0.124. Because this value was greater than 0.05, this difference was not statistically significant. Therefore, miR-10b could not be used to differentiate between semen and vaginal fluid, which contradicts previously reported findings.

The results also show that miR-10b can be used to differentiate semen from venous blood, saliva, and menstrual blood because the expression level differences were statistically significant. These findings do comply with previously reported findings.

5.2.5 miR-205 Expression Results

miR-2015 was chosen for validation because a previous studies performed by Hanson *et al,* Wang *et al,* Mayes *et al,* and Van der Meer *et al* showed that it appeared to be differentially expressed in saliva compared to other forensically relevant body fluids. It was the goal of this study to determine if miR-205 could be used to differentiate between saliva and other body

fluids. Based on the statistical analysis of the expression levels of miR-205 within this study, the identification capabilities of miR-205 were not particularly strong. The expression levels show that miR-205 could not be used to differentiate saliva from semen or vaginal fluid because the differences in the expression levels were not statistically significant. This lack of differentiation abilities of miR-205 directly contradicts previously published results.

The expression levels of miR-205, however, did show that it could be used to differentiate between saliva and menstrual blood because the difference between the two expression levels was statistically significant. This does agree with the previously published results mentioned above. The expression levels of miR-205 also shows that miR-205 was not expressed in venous blood. This lack of expression in venous blood was also seen in Wang *et al* and Mayes *et al*). Technically, these results do indicate that miR-205 can be used to differentiate between venous blood and saliva based on the lack of expression in venous blood. Overall, the ability to utilize miR-205 to differentiate between saliva and menstrual/venous blood is consistent with the literature search.

Like with the other miRNAs, the range of expression levels of miR-205 in saliva was large. In this case, this causes an issue being able to differentiate between saliva and the other body fluids because the large expression levels of saliva overlap that of semen and vaginal fluid. Therefore, the identification of salvia becomes more difficult because of this large expression range. The large expression range may be due the small sample size and increasing the sample size may correct this issue slightly.

CHAPTER 6: DISCUSSION

6 CHAPTER 6: DISCUSION

6.1 Next Generation Sequencing

There may be several reasons for this sequencing failure. An initial concern for the failure of the sequencing was errors during sample preparation. However, because Illumina representatives were present several times during sample preparation and did not point out any mistakes, this seems like an unlikely option. Another potential point of error was that the quality of the RNA samples being sequenced was not optimal for sequencing. Because the last sequencing run consisted of the two samples of the best quality and the sequencing still failed, this also seems to be unlikely. A more likely reason for the sequencing failure is that the MiSeq FGx may not have extensive enough sequencing capabilities to sequence the miRNAs within the samples. In other words, the MiSeq FGx may not produce enough sequencing reads per sample for the analytical software within the instrument to sequence the miRNA content in the sample. This seems like the most likely scenario as no body fluid sample reached the required amount of reads for successful sequencing, which was 1-2 million reads. Without more experimentation, however, there is no way to determine the true reason or combination of reasons why the miRNA content in the samples were not able to be sequenced.

6.2 miRNA Expression Analysis

The results from this study show both agreement and disagreement with previously published studies evaluating miRNAs for body fluid identification. This illustrates the need for more extensive research to determine an appropriate panel of miRNAs for body fluid identification. Ideally, the expression levels of the miRNAs within the panel ultimately chosen for body fluid identification needs to be reliably reproduced by several different people working

independently of one another. Additionally, there should be little to no overlap in expression of each miRNA in the different body fluids. This ensures that proper identification can be performed for each body fluid. Because this was not achieved within in this research study or any other ones, there needs to be more research done.

6.3 Future Works

The range of expression of saliva within each miRNA was quite extensive. This made it more difficult for saliva to be distinguished from other body fluids. The reason for this large expression range for saliva is unknown. It is possible that this could be mediated by increasing the number of saliva samples used within RT-qPCR. A larger sample size may be enough to make the expression levels more conducive for identification. Another possible reason this could have occurred was due to degradation of some of the saliva samples, which could have resulted in the large expression range. More research, however, needs to be done to determine if these are viable options.

Additionally, body fluids can be subjected to different experimental conditions to determine if the miRNA expression changes. For example, a time study could be performed to see if miRNA expression is truly stable over time. Additionally, body fluids could be subjected to different environmental conditions like sunlight exposure or extreme cold to determine if miRNA expression remains stable throughout these changes. Throughout each of these experiments, more samples should be used and the samples should be from a diverse set of people. This would allow for the results to be trusted when being applied to a larger population and many different demographic groups.

Overall, more research needs to be done to assess the expression levels of miRNAs in each body fluids over a wide range of environmental conditions, demographics, and time periods to ensure that they can be reliably used in many different situations. Ultimately, any panel of miRNAs that are chosen for body fluid identification will be used within forensic investigations. Because the nature for forensic investigations is unpredictable, the expression levels of the miRNAs need to be studied extensively so they can be reliably used for forensic purposes.

CHAPTER 7: Appendix

7 Chapter 7: Appendices

7.1 **Appendix A – miRNeasy Extraction Protocol**

This protocol was used for all samples. As stated before, liquid samples were extracted using 500 uL. For samples collected via sterile swabs, the swab head was removed and extracted was performed on the swab head.

Place sample in 700 uL QIAzol Lysis Reagent. Vortex or pipet to mix. Incubate sample at room temp for five minutes. Add 140 uL chloroform to the tube containing sample and shake vigorously for fifteen seconds. Incubate sample at room temperature for 2-3 minutes. Centrifuge sample for fifteen minutes (12,000 x g at 4˚C). Transfer the upper aqueous layer to a new collection tube. Add 1.5 volumes of 100% ethanol and mix thoroughly by pipetting up and down several times. Do not centrifuge the sample at this time. Pipet up to 700 uL of the sample, including any precipitate that formed into spin column in a 2 mL collection tube. Centrifuge for 15 seconds (≥ 8000 x g at room temperature). Discard flow through. Repeat this step using the remainder of the sample. Pipet 350 uL of Buffer RWI into the spin column and centrifuge for 15 seconds at ≥ 8000 x g. Discard the flow through. At this time, prepare the DNase solution by pipetting 10 uL DNase stock solution to 70 uL of Buffer RDD. Mix gently by inverting tube and centrifuge briefly to collect residual liquids from side of tube. Add the 80 uL of prepared DNase solution to the spin column and incubate at room temperature for fifteen minutes. Add 350 uL of RWI to the spin column. Centrifuge for 15 seconds at ≥ 8000 x g. Discard the flow through. Pipet 500 uL of RPE into the spin column. Centrifuge for 15 seconds at ≥ 8000 x g to wash column. Discard the flow through. Add another 500 uL of RPE into column and centrifuge for two minutes at ≥ 8000 x g to dry the column. Discard the flow through. Place the spin column into new 2 mL collection tube and discard old collection tube with flow through. Centrifuge in a

microcentrifuge at full speed for one minute. Transfer RNeasy spin column to a new 1.5 mL collection tube. Pipet 30-50 uL RNase-free water directly onto RNeasy spin column membrane. Centrifuge for one minute at ≥ 8000 x g to elute the RNA.

The samples were then stored at -20˚C in an area that was isolated from the collected samples until further use.

7.2 **Appendix B – Qubit 3 Fluorometer**

For optimal results, all reagents must be stored at room temperature. Additionally, the sample preparation and testing should be done at room temperature as well. After the samples are prepared, they are only stable for three hours and must be tested within that time frame.

First, one must prepare the samples for testing. Set up the required number of 0.5 mL tubes for the number of samples and standards being used, keeping in mind that the Qubit RNA HS Assay has 2 standards. It's important to note that only thin-walled, clear tubes can be used. Prepare the working solution by diluting the HS Reagent 1:200 in the HS Buffer. Add 190 µL of the working solution into each of the tubes used for standards. Add 10 µL of each standard into the appropriate tube and vortex for 2-3 seconds. Add 199 µL of the working solution to the tubes for the samples. Add 1 µL of the samples to the appropriate tubes and vortex for 2-3 seconds. Let the mixtures within the tubes at room temperatures for 2 minutes.

After the incubation period has ended, it is necessary to set the Qubit 3.0 Fluorometer for reading standards and samples. On the home screen, choose "RNA" and then select the assay type by choosing "RNA: High Sensitivity". Choose "Read Standard" to continue on to processing the standards. Insert the tube containing Standard 1 into the sample chamber, close the lid, and then choose "Read Standard". The reading will be complete in approximately three

seconds. Repeat these steps with the tube containing Standard 2. After the second standard is read, the instrument should show calibration results. If the calibration results are within the guidelines, the samples can be processed.

When the samples are ready to be read, choose “Run Samples” and select the sample volume added to the working solution and the unit of measurement preferred. Insert a sample tube into the sample chamber, close the lid, and then choose “Read tube”. The concentration of the sample will be displayed on the screen. Repeat these steps until all of the samples have been read.

7.3 **Appendix C – 2100 Bioanalyzer Small RNA Kit**

Before preparing samples for the 2100 Bioanalyzer and the Small RNA Kit, the chip priming station and reagents should be prepared. Regarding the chip priming station, the base plate should be placed into position C. Regarding the syringe clip, the lever must be placed in the lowest possible position. The kit reagents must be kept at 4˚C, apart from the Small RNA Ladder which should be kept at -70˚C. It is important to use a new syringe and new electrode cleaners with every new kit being used. Regular maintenance should also be performed on the 2100 Bioanalyzer, including cleaning the electrodes before and after sample runs.

When the samples are ready to be prepared for running, take them out and allow them to thaw on ice. The Small RNA Ladder should also be taken out of -70˚C and thawed on ice. The kit reagents should be taken out of the 4˚C and allow to equilibrate at room temperature for at least thirty minutes. The complete volume of the Small RNA gel matrix should be transferred to the top of the spin filter. The spin filter should be centrifuged for fifteen minutes at room temperature at 10,000 g ±. After centrifugation, the filter can be removed and the gel is ready for

sample preparation. It will be used for the preparation of the gel-dye mix. Vortex the Small RNA dye concentrate for approximately ten seconds. Pipette 2 µL of the prepared die into a 0.5 mL microcentrifuge tube that is RNase free. Add 40 µL of the previously prepared Small RNA gel matrix and mix the solution bey pipetting. The mixture should be spun in a microcentrifuge at 13,000 g for ten minutes at room temperature.

After the gel-dye mix has been centrifuged, it needs to equilibrate to room temperature for thirty minutes before being loaded into the Small RNA chip. Place a Small RNA chip on the chip priming station and slowly pipette 9.0 µL of the gel-dye mix into the well that is marked with a bold "G". Lock the latch of the chip priming station and press the plunger of the syringe all the way down. Allow the plunger to stay in this position for 60 seconds and then release the plunger. Pipette 9.0 µL of the gel-dye mix into the two well marked with a non-bold "G". Pipette 9 µL of the Small RNA conditioning solution into the well that is marked with "CS". Pipette 5 µL of the Small RNA marker into the well that is marked with a ladder symbol and each of the eleven sample wells. Pipette 1 µL of the Small RNA Ladder into the well that is marked with the ladder symbol. Pipette 1 µL of each sample into each of the sample wells. If well is not being used for a sample, 1 µL of the Small RNA maker or RNase free water can be utilized. Vortex the chip for 60 seconds at 2400 rpm in an IKA vortex mixer. After the chip has been vortexed, it must be placed into the 2100 Bioanalyzer and processing must commence within five minutes.

Open the 2100 Bioanalyzer software and place the chip within the electrode slot within the 2100 Bioanalyzer. Choose the "Instrument" tab and choose the Small RNA assay in the "Assays" folder. Fill in any sample information in the table that is provided. When ready, choose "Start" to start the run. Allow the assay to run and the data can be obtained in the "Data" tab.

7.4 Appendix D – 2100 Bioanalyzer RNA 6000 Nano Kit

Like with the Small RNA Kit, the chip priming station and reagents for the RNA 6000 Nano Kit must be prepared before samples can be prepared. Regarding the base plate of the chip priming station, the base should be placed in position C. Regarding the syringe clip of the chip priming station, the lever of the clip should be in the top most position. The kit reagents must be kept at 4˚C, apart from the RNA Ladder which should be kept at -70˚C. It is important to use a new syringe and new electrode cleaners with every new kit being used. Regular maintenance should also be performed on the 2100 Bioanalyzer, including cleaning the electrodes before and after sample runs.

Before samples can be prepared, the gel-dye mix must be made. All reagents must be taken out of storage and allowed to equilibrate to room temperature for at least half an hour. The RNA Ladder, however, needs to be thawed on ice. Place 550 µL of the RNA 6000 Nano gel matrix into the top of the provided spin filter and centrifuge the spin filter for ten minutes at 1,500 g ± 20%. Aliquot 65 µL of the filtered gel into a 05 mL microcentrifuge tube that is RNase-free. Vortex the RNA 6000 Nano dye concentrate for approximately ten seconds. At 1 µL of RNA 6000 Nano dye concentrate to the 65 µL aliquot of the filtered gel. Vortex the mixture and then centrifuge the mixture for ten minutes at room temperature at 13,000 g.

After the gel-dye mixture is done being centrifuged, it must be allowed to equilibrate to room temperature for approximately thirty minutes. The gel-dye mixture must be protected from light during this time as well. Place a new RNA Nano chip on the chip priming station and pipette 9.0 µL of the gel-dye mixture in the well that is marked the a bold “G”. Close the chip priming station and push the plunger all the way down. Allow the plunger to stay in that position for thirty seconds and then release. Pipette 9.0 µL of the gel-dye mixture into the two wells that

are marked with the non-bolded “G”. Pipette 5 µL of the RNA 6000 Nano marker into the well that is marked with the ladder symbol and in each of the 12 sample wells. Pipette 1 µL of the RNA ladder into the well that is marked with the ladder symbol. Pipette 1 µL of each of the samples into each of the 12 sample wells. Place the prepared chip in the IKA vortex mixer the vortex the chip for sixty seconds at 2400 rpm. Once this is complete, the processing of the chip must start within five minutes.

Open the 2100 Bioanalyzer software and place the chip within the electrode slot within the 2100 Bioanalyzer. Choose the “Instrument” tab and choose the RNA Nano 6000 assay in the “Assays” folder. Fill in any sample information in the table that is provided. When ready, choose “Start” to start the run. Allow the assay to run and the data can be obtained in the “Data” tab.

7.5 Appendix E – Day 1 of TruSeq Small RNA Protocol

7.5.1 Ligate Adapters

Thaw all consumables (10 mM ATP, HML, RA3, RA5, RNase inhibitor, STP, T4 RNA Ligase, Ultrapure water, T4 RNA Ligase 2/Deletion Mutant) on ice. Centrifuge each consumable at 600 x g for five seconds and set aside on ice.

Take out samples being sequenced and thaw on ice. Once thawed, prepare the samples appropriately in order to get 1 µg of total RNA in 5 µL of Nuclease Free water. Combine the 5 µL of sample with 1 µL of RA3. Pipette to mix and then centrifuge briefly. Place prepared samples on a thermal cycler preheated to 70˚C. Incubate for 2 minutes. Remove the samples from the thermo cycler and place on ice. Preheat the thermo cycler to 28˚C. In a new 200 µL PCR tube, combine the following volumes on ice (multiply each volume by the number of samples being prepared):

- HML (2 µL)
- RNase Inhibitor (2 µL)
- T4 RNA Ligase 2, Deletion Mutant (1 µL)

Pipette to mix and centrifuge briefly. Add 4 µL of the mixture to the tubes of RA3/total RNA mixture and pipetted to mix. Place these samples on the thermocycler preheated to 28˚C and incubate for 1 hour. After 1 hour, add 1 µL of STP and pipette to mix. Return samples to thermo cycler and continue incubating at 28˚C for 15 minutes. At the end of 15 minutes, remove the samples from the thermal cycler and place on ice. At this time, the 3' Adapters have been ligated.

Next, the 5' Adapters need to be ligated. Preheat the thermo cycler to 70˚C. Add 1.1 x N µL (where N = number of samples) RA% to a new 200 µL PCR tube. Place this tube on the preheated thermo cycler and incubate for 2 minutes. Remove this tube from the thermocycler and place on ice. Preheat the thermo cycler to 28˚C. Add 1.1 x N µL 10 mM ATP to the tube of RA5 and pipette to mix. Add 1.1 x N µL T4 RNA Ligase to the RA5/ATP mixture and pipette to mix. Add 3 µL of this mixture to the tube containing the RA3 mixture and pipette to mix. Place the prepared samples on the preheated thermo cycler and incubate at 28˚C for one hour. After an hour passes, remove the samples from the thermal cycler and place on ice.

7.5.2 Reverse Transcribe and Amplify Libraries

7.5.2.1 Preparation

Thaw all consumables (25 mM dNTP mix), PML, RPI, RNA PCR Primer Indices, RTP, RNase Inhibitor, ultrapure water, 5X First Strand Buffer, 100 mM DTT, High Sensitivity DNA

chip, SuperScript II Reverse Transcriptase) on ice. Centrifuge all consumables at 600 x g for five seconds and set aside on ice. Preheat the thermocycler to 70˚C.

7.5.2.2 Dilute 25 mM dNTP Mix

Combine the following volumes in a new 200 µL PCR tube labeled 12.5 mM dNTP Mix. Make sure to multiply each volume by the number of samples being prepared and prepare 10% extra.

- 25 mM dNTP Mix (0.5 µL)
- Ultrapure Water (0.5 µL)

Pipette to mix and set aside on ice.

7.5.2.3 Reverse Transcription

Add 6 µL of each adapter-ligated RNA library to new 200 µL PCR tubes. Add 1 µL RNA RT Primer to the tube of adapter-ligated RNA at pipette to mix. Centrifuge the mixture briefly. Place the tubes on the preheated thermo cycler and incubate at 70˚C for 2 minutes. After 2 minutes, remove from the thermo cycler and place on ice. Preheat the thermo cycler to 50˚C. Combine the following volumes of reagents in a new 200 µL PCR tube on ice, making sure to multiply each volume by the number of libraries (samples) being prepared. Make 10% extra reagent.

- 5X First Strand Buffer (2 µL)
- 12.5 mM dNTP Mix (0.5 µL)
- 100 mM DTT (1µL)
- RNase Inhibitor (1 µL)
- SuperScript II Reverse Transcriptase (1 µL)

Pipette to mix and then centrifuge briefly. Add 5.5 µL of this mixture to the tube of the adapter ligated RNA/primer mix and pipette briefly. Centrifuge this mixture briefly. Place the samples on the preheated thermocycler and incubate at 50˚C for 1 hour. After 1 hour, remove the samples from the thermal cycler and place on ice.

7.5.3 Amplify Libraries

Combine the following reagents in a new 200 µL PCR tube on ice, making sure to multiply each volume by the number of libraries (samples) being prepared. Make 10% extra reagent. This is the PCR master mix.

- Ultrapure water (8.5 µL)
- PML (25 µL)
- RP1 (2 µL)

Pipette to mix and centrifuge briefly. Add 35.7 µL of the Master Mix to the adapter-ligated RNA mixture tubes. Pipette to mix. Add 2 µL of a pre-determined RPIX to each mixture. Each different tube/sample should get a different RPIX. Pipette to mix and then centrifuge briefly. Incubate the samples in a thermocycler using the following program.

- Choose the preheat lid option and set to 100˚C
- 98˚C for 30 seconds
- 11-15 cycles of:
    - 98˚C for 10 seconds
    - 60˚C for 15 seconds
    - 72˚C for 10 minutes

- 4˚C hold

Run each library (sample) on a High Sensitivity DNA chip (as seen in Appendices F).

This is a safe stopping point. The tubes can be capped and stored at -25˚C to -15˚C for up to 7 days.

7.6 Appendix F – 2100 Bioanalyzer DNA 1000 Chip Kit

Like with the Small RNA Kit and RNA 6000 Nano Kit, the chip priming station and reagents for the DNA 1000 Chip Kit must be prepared before samples can be prepared. Regarding the base plate of the chip priming station, the base should be placed in position C. Regarding the syringe clip of the chip priming station, the lever of the clip should be in the top most position. syringe clip of the chip priming station, the lever of the clip should be in the top most position. The kit reagents must be kept at 4˚C. It is important to use a new syringe and new electrode cleaners with every new kit being used. Regular maintenance should also be performed on the 2100 Bioanalyzer, including cleaning the electrodes before and after sample runs.

First, all of the kit reagents must be taken out of 4˚C storage and allowed to equilibrate to room temperature for at least 30 minutes. Vortex the blue-caped DNA dye concentrate for 10 seconds and centrifuge briefly. Pipette 15 µL of the dye concentrate into the red-capped DNA gel matrix vial. Vortex for 10 seconds to properly mix the two components. Transfer the gel-dye mix to the top receptacle of a spin filer and place the spin filter into a microcentrifuge and centrifuge for `5 minutes at room temperature at 1400 g ± 20%. After centrifugation, discard the filter according to good laboratory practices and label the gel-dye mixture with the date of preparation. The prepared gel-dye mix must be used within 4 weeks of preparation.

7.7 Appendix G – Day 2 of TruSeq Small RNA Protocol

7.7.1 Preparing Consumables

Thaw all consumables on ice and centrifuge at 600 x g for five seconds and then set the consumables aside on ice. The consumables include the Custom RNA Ladder (CRL), High Resolution Ladder (HRL), 5 µm filter tubes, amplified cDNA construct, DNA loading dye, 3 M NaOAc (pH 5.2), 10 mM Tris-HCL (pH 8.5), and ethanol. Additionally, set aside the following consumables: gel breaker tubes, High Sensitivity DNA Kit, Novex TBE gels (6%, 10 wells), Novex TBE running buffer (5x), nuclease-free 200 µL PCR tubes, nuclease-free 1.5 mL microcentrifuge tubes, a razor blade, ultrapure ethidium bromide (0.5 µL/mL in water), and 100% ethanol.

7.7.2 Preparing Running Buffer and Electrophoresis Chamber

Dilute the 5x Novex TBE Running Buffer to a concentration of 1x. Approximately 800-1000 µL of the 1x running buffer will be needed for the running of the gel. Place the 6% Novex TBE gel into the gel electrophoresis chamber (Novex SureLock System) per manufacturer instructions. Fill the running chamber with the running buffer until the 6% Novex TBE gel is fully covered.

7.7.3 Run Gel Electrophoresis

Combine 2 µL of the CRL with 2 µL DNA loading dye in a new 200 µL PCR tube and pipette to mix. Combine 2 µL of the HRL and 2 µL DNA loading dye in a new 200 µL PCR and pipette to mix.

Combine all amplified cDNA content (48-50 µL per sample) and 10 µL of DNA loading dye in a 1.5 microcentrifuge tube. Pipette to mix thoroughly. Load 2 gel lanes with the 2 µL HRL mixture in the two most outer lanes. Load 2 gel lanes with 2 µL CRL mixture into the lanes next to the HRL (moving toward the middle of the gel). Load 2 gel lanes with 25 µL of the amplified cDNA content mixture into the middle lanes of the gel. Run the gel for 60 minutes at 145 V or until the dye front leaves the gel. Once the gel is finished running, remove the gel from the unit.

7.7.4 Recover the Purified cDNA Construct from the Gel

Open the gel cassette carefully, making sure not to break the gel. Stain the gel with the ethidium bromide in a clean container for two to three minutes. Properly dispose of the ethidium bromide once the staining has completed.

Place a gel breaker tube into a 2.0 mL microcentrifuge tube. View the gel on a UV transilluminator. Cut out the miRNA bands in the sample lanes using a razor blade. The bands should be approximately at the 147 nt band region. Use the HRL and CRL bands to determine the sizing of the sample bands to ensure that the correct miRNA bands are cut out. Place the cut-out bands into the 0.5 mL gel breaker tube that is nested in the 2.0 mL microcentrifuge tube. Centrifuge the nested tubes at 20,000 x g for 2 minutes. This allows the gel through the holes into the 2 mL microcentrifuge tube. If one so chooses, the final library can be concentrated. If this optional procedure is not done, add 200 µL of ultrapure water to the gel debris in the 2.0 mL tube. Rotate or shake the tube for at least 2 hours to elute the DNA. Transfer the eluate and gel debris to the top of the 5 µm filter and centrifuge at 10 seconds at 600 x g.

7.7.5 Concentrate Final Library (Optional)

Add 300 µL ultrapure water to the gel debris in the 2.0 mL microcentrifuge tube. Rotate or shake the microcentrifuge tube for at least 2 hours to elute the DNA. Transfer the eluate and gel debris to the top of a 5 µm filter and centrifuge at 600 x g for 10 seconds. Discard the filter. Add the following volumes to the eluate:

- 2 µL glycogen
- 30 µL 3M NaOAc
- 975 µL 100% ethanol
- 2 µL 0.1 X Pellet Paint (optional)

Pipette to mix thoroughly. Centrifuge at 20,000 x g on a benchtop microcentrifuge at 20 minutes at 4˚C. Remove and discard the supernatant and leave the pellet intact. If the pellet becomes loose, centrifuge at 20,000 x g for 2 minutes to re-create the pellet. Wash the pellet with 500 µL of 70% ethanol. Centrifuge at 20,000 x g for 2 minutes. Remove and discard the supernatant, leaving the pellet intact. Place the microcentrifuge tube with the lid open in a 37˚C heat block until the pellet is dry (approximately 7 minutes). Resuspend the pellet in 10 µL of 10 mM Tris-HCl, pH 8.5.

7.7.6 Check Libraries

Load 1 µL of the purified library/sample onto an Agilent Technologies 2100 Bioanalyzer using a DNA-specific chip, such as the DNA 1000 or High Sensitivity DNA. The current study used a DNA 1000 chip/kit and instructions can be found in Appendices F.

Check the size, purity, and concentration of the library. For clustering, use the total of all molarities from the results. For example, if there are three peaks, add the molarity of each of the peaks to determine the overall molarity.

7.7.7 Normalize Libraries

Normalize the library concentration to 2 nM using Tris-HCl 10mM, pH 8.5. This is a safe stopping point. Cap the tube and store at -25˚C to -15˚C for up to 7 days.

7.8 Appendix H – Day 3 of TruSeq Small RNA Protocol

7.8.1 Prepare the Reagent Cartridge

Remove the reagent cartridge from -25˚C to -15˚C storage and place the reagent cartridge in a water bath of room temperature deionized water. Ensure that the water does not go above the marked line on the reagent cartridge. Allow the reagent cartridge to thaw in the room temperature water for approximately 60-90 minutes or until the liquid within the cartridge is thawed completely. Remove the cartridge from the water bath and tap it on a bench. This dislodges water from the base of the cartridge. The cartridge base should be dried and no water should be left on the top of the reagent cartridge. Invert the reagent cartridge ten times to mix the reagents and then visually ensure that all of the reagents are thawed completely. Visually ensure that the reagents in positions 1, 2, and 4 are fully mixed and contain no precipitates. Place the reagent cartridge on ice or store it at 2˚C to 8˚C until ready to set up a run.

7.8.2 Load Sample Reagents into Reagent Cartridge

Use a low-lint lab tissue to clean the foil seal that is covering the “Load Samples” reservoir. Use a clean 1 mL pipette tip to pierce the foil seal that is covering the “Load Samples” reservoir. Pipette 600 µL of the prepared sample libraries into the “Load Samples” reservoir, making sure not to touch the foil seal as the sample is dispensed. If air bubbles are present, gently tap the cartridge on the bench to release the bubbles.

7.8.3 Run Samples on MiSeq FGx

Log into the MiSeq FGx with the proper credentials. From the welcome screen, select “Sequence” and then select “Research Use Only Run” when asked to select run type. Follow the prompts to load the flow cell and the reagents in order to set up the run.

The flow cell is immersed in a storage buffer in a flow cell container and must be removed and cleaned. Using gloves, grip the base of the plastic cartridge with plastic forceps and remove it from the flow cell container. Rinse the flow cell with nuclease-free water. Both the glass and plastic cartridge are thoroughly rinsed. Dry the flow cell and cartridge using a lint-free lens cleaning tissue, making sure to pat dry the area of the gasket and the adjacent glass. Clean the flow cell glass using an alcohol wipe or ethanol. Visually inspect the glass to make sure it is free of streaks, fingerprints, and lint. Dry any excess alcohol with a lint-free lens cleaning tissue. To load the flow cell into the instrument, raise the flow cell compartment door and press the release button to the right of the flow cell latch to open it. Ensure that the flow cell stage is free of lint or other debris. Place the flow cell on the flow cell stage and gently press down on the flow cell latch to close it over the flow cell. Close the flow cell compartment door and then click Next on the instrument’s screen.

Remove the bottle of SBS solution (PR2) from 2˚C to 8˚C storage and gently invert the bottle to mix the solution. Remove the lid from the bottle. Open the reagent compartment door and raise the sipper handle until it locks in place. Place the SBS solution bottle in the indentation to the right of the reagent chiller. Ensure that the waste bottle is empty. Lower the sipper handle slowly into the SBS solution. Select Next on the instruments' screen.

Open the reagent chiller door. Hold the reagent cartridge on the end with the Illumina label (there is a handle on that side) and slide the reagent cartridge into the reagent chiller until it is fitted into the instrument. Close the reagent chiller door and then close the compartment door. Select Next on the instrument's screen.

Load the appropriate sample sheet and ensure that the worklist name, analysis workflow, and read length are correct. A pre-run check will be performed. Wait until this is complete and then select "Start Run". While the instrument is running, the progress can be monitored.

7.9 Appendix I – Quantification Raw Data

Table 6: Quantification Raw Data

Sample	Nanodrop Conc. (ng/uL)	Quibit 3 Conc. (ng/uL)	2100 Bioanalyzer RNA 6000 Nano Kit		2100 Bioanalyzer Small RNA Kit		
			RNA Conc. (ng/uL)	RIN	Small RNA Conc. (ng/uL)	miRNA Conc. (ng/uL)	% miRNA (%)
1	6.8	out of range (too low)	2	NA	6.4873	2.3645	36
2	4.8	out of range (too low)	3	NA	4.8766	1.5806	32
3	6.4	out of range (too low)	1	NA	8.1917	3.3584	41
4	5.7	out of range (too low)	4	NA	4.1155	0.9282	23
5	23.7	18.8	13	4.9	19.3239	6.2536	32

6	69	71.8	34	3.4	70.1771	20.6854	29
7	33.2	28.4	15	5.8	19.7044	2.8303	14
9	18	13.4	9	NA	14.8675	2.2461	15
10	12.6	7.27	6	NA	12.2604	2.5645	21
11	161.9	180	87	2.5	170.2607	64.7692	38
12	38.2	2.85	NA	NA	NA	NA	NA
13	54	55.4	30	2.1	434.0601	195.4688	45
14	7.4	6.58	6	NA	4.9274	3.5204	71
15	124.4	80	NA	NA	NA	NA	NA
16	60.9	64	46	2.4	681.7783	355.316	52
18	8	7.2	7	NA	5.4764	3.764	68
21	37.6	43.4	26	4.9	22.349	10.3997	47
22	19.5	25.2	15	1.2	13.1301	7.0153	53
23	307.9	318	403	2.5	94.9288	44.0502	46
24	5.2	out of range (too low)	3	NA	2.5854	1.3397	52
25	317.1	360	392	2.6	40.4944	19.7728	49
26	7.2	5.58	3	NA	4.4362	1.8467	42
28	2.7	out of range (too low)	2	NA	1.2057	0.4256	35
29	23.9	25.6	21	1.5	31.8625	15.5918	49
30	1.8	out of range (too low)	3	NA	0.6092	0.285	47
32	81.6	76.2	65	3.5	13.4182	5.9051	44
66	41.6	49	22	2.3	30.3993	16.8577	55
046A	48.7	47.4	28	1.9	16.439	9.0877	55
046B	162.7	160	123	2.9	12.7216	7.5229	59
046C	25.7	31.4	22	3	730.3137	36.8845	5
047A	12.4	10.1	6	NA	8.9582	4.0692	45
047B	61.4	46.6	29	3.9	10.8808	6.9785	64
047C	120.4	91.6	92	5.5	NA	NA	NA
048A	215	216	248	4.1	2294.012	615.2069	27
048B	456	300	549	4.1	33.0115	15.0418	46
048C	409.8	385	319	3.8	2060.285	546.8156	27
051A	9.3	5.32	3	NA	3.2389	1.3601	42
051B	21.8	14.8	8	NA	4.2941	2.3676	55
051C	6.4	64.2	3	NA	2.5545	1.2816	50
052A	105	112	94	2.8	444.9739	189.8946	43

052B	90.8	92.6	44	2.1	20.1176	8.7834	44
052C	59.3	63.6	32	1.5	21.7038	10.689	49
053A	78.7	80.4	60	6.4	21.9673	10.7679	49
053B	29.2	27.4	16	7.1	12.434	6.3014	51
053C	54.8	57.6	36	6.8	16.9728	8.6363	51
054A	135.4	128	104	5.4	15.1064	7.8462	52
054B	228.2	157.2	155	3.1	97.2388	1.1383	1
054C	124.2	116	75	6.1	1569.583	515.1483	33
055A	142.2	134	200	6.3	27.5086	11.5487	42
055B		104	78	4.8	1164.185	320.1063	27
055C	22.3	18.7	9	NA	13.4845	5.3348	40
068A	237.2	148	132	3.5	796.6704	206.2282	26
068B	166.8	134	135	6.5	10.7397	5.9621	56
068C	101.6	90.6	80	5.6	8.5031	4.3395	51
056A	154	136.4	96	3.3	13.2023	6.6022	50
056B	75.6	64.8	43	2.7	12.6614	5.5767	44
056C	149.5	158	74	2.8	15.0189	6.9614	46
057A	190	188	208	3.9	17.8175	10.0118	56
057B	34.7	10.5	6	NA	6.5182	3.4726	53
057C	28.3	28	417	2.8	999.5266	273.6595	27
058A	239.7	200	222	2.6	20.4912	9.2986	45
058B	354.6	356	179	2.8	4660.195	1227.423	26
058C	84.5	152	71	2.8	1061.699	160.0607	15
060A	49	55.4	34	8.1	11.008	6.861	62
060B	9.4	82.8	6	NA	6.3077	3.3912	54
060C	47.6	56	30	7.8	9.9987	5.9162	59
061A	236.7	291	225	4.4	1526.636	304.748	20
061B	232.6	200	148	5.5	19.1369	7.2987	38
061C	310	350.4	423	5.7	1217.632	201.8665	17
062A	416.3	399	58	4.4	254.6584	109.4042	43
062B	320.1	375.2	386	5.6	3511.156	525.811	15
062C	628.8	574.4	397	5.9	2420.294	40.0086	2
063A	271.3	out of range (too low)	310	2.8	2358.337	150.5322	6
063B	285.2	221.4	367	4.2	2036.866	90.3311	4
063C	223.3	207	133	2.9	2018.678	225.259	11

064A	166.8	160	88	2.6	20.9915	13.5142	64
064B	279.8	174	116	2.9	1505.714	492.1097	33
064C	181.9	166	113	4	11.0752	6.7715	6
065A	841.9	728	609	3.1	1623.173	238.2669	15
065B	1870.5	1140	1704	2.7	5136.211	11.7639	0
065C	1218	2955	754	2.7	29.4481	10.557	36
069A	863.2	818	688	3.4	813.8969	95.664	12
069B	787.7	1010	876	2.8	3519.86	373.816	1
069C	616.9	744	662	3.4	2391.162	266.1963	11

7.10 Appendix J – Raw Data from RT-qPCR

7.10.1 C_t Values and ΔC_t Values for miR-451

Table 7: miR-451 C_t Value Data

Sample	miR-16	miR-451	Δ Ct
1	14.12	15.468	1.348
4	15.717	16.912	1.195
6	14.023	16.778	2.755
9	14.196	15.692	1.496
10	16.091	16.445	0.354
12	31.8	38.054	6.254
13	20.62	32.004	11.384
15	20.146	29.02	8.874
16	24.305	37.316	13.011
18	24.486	30.196	5.71
22	26.323	32.039	5.716
23	34.267	36.462	2.195
29	23.08	36.101	13.021
32	31.388	32.376	0.988
66	27.252	32.155	4.903
47C	15.28	18.204	2.924
53A	17.343	21.199	3.856
54C	17.358	26.178	8.82
55A	14.529	19.288	4.759
68B	16.022	18.547	2.525
57A	24.344	36.607	12.263
61C	24.412	37.466	13.054
62C	23.78	38.481	14.701

63B	22.424	38.553	16.129
69C	23.232	37.047	13.815

7.10.2 C_t Values and ΔC_t Values for miR-412

Table 8: miR-412 C_t Value Data

Sample	miR-16	miR-412	Δ Ct
1	14.12	38.552	24.432
2	14.57	37.115	22.545
4	15.717	40	24.283
6	14.023	34.274	20.251
10	16.091	36.109	20.018
12	31.8	33.089	1.289
15	20.146	33.695	13.549
16	24.305	36.138	11.833
22	26.323	36.536	10.213
23	34.267	36.905	2.638
29	23.08	40	16.92
32	31.388	33.1	1.712
66	27.252	36.764	9.512
47C	15.28	36.472	21.192
53A	17.343	35.184	17.841
54C	17.358	36.684	19.326
55A	14.529	35.881	21.352
68B	16.022	35.702	19.68
57A	24.344	37.207	12.863
61C	24.412	35.532	11.12
62C	23.78	36.583	12.803
63B	22.424	34.614	12.19
69C	23.232	38.721	15.489

7.10.3 C_t Values and ΔC_t Values for miR-891a

Table 9: miR-891a C_t Value Data

Sample	miR-16	miR-891a	Δ Ct
1	14.12	37.81	23.69
4	15.717	37.101	21.384
6	14.023	34.878	20.855
9	14.196	40	25.804

10	16.091	37.526	21.435
12	31.8	26.549	-5.251
13	20.62	27.489	6.869
14	22.183	25.318	3.135
15	20.146	23.755	3.609
16	24.305	22.887	-1.418
18	24.486	24.317	-0.169
22	26.323	38.866	12.543
23	34.267	38.311	4.044
29	23.08	38.872	15.792
32	31.388	37.465	6.077
66	27.252	40	12.748
47C	15.28	39.518	24.238
53A	17.343	40	22.657
54C	17.358	40	22.642
55A	14.529	40	25.471
68B	16.022	37.057	21.035
57A	24.344	38.607	14.263
61C	24.412	40	15.588
62C	23.78	40	16.22
69C	23.232	37.864	14.632

7.10.4 C_t Values and ΔC_t Values for miR-205

Table 10: miR-205 C_t Value Data

Sample	miR-16	miR-205	Δ Ct
4	15.717	40	24.283
5	16.02	40	23.98
6	14.023	40	25.977
9	14.196	40	25.804
10	16.091	40	23.909
12	31.8	29.06	-2.74
13	20.62	25.385	4.765
14	22.183	25.759	3.576
15	20.146	24.515	4.369
16	24.305	25.157	0.852
18	24.486	26.09	1.604

22	26.323	27.625	1.302
23	34.267	30.32	-3.947
29	23.08	25.69	2.61
32	31.388	23.789	-7.599
66	27.252	26.59	-0.662
47C	15.28	23.598	8.318
53A	17.343	20.866	3.523
54C	17.358	22.074	4.716
55A	14.529	19.176	4.647
68B	16.022	20.539	4.517
57A	24.344	22.195	-2.149
61C	24.412	27.376	2.964
62C	23.78	**27.169**	3.389
63B	22.424	26.217	3.793
69C	23.232	24.348	1.116

8 References

Agilent Technologies, Inc. (2013). Agilent Small RNA. Agilent Technologies, Inc. Waldbronn, Germany

Bartel, D. P. (2004, January 23). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell. https://doi.org/10.1016/S0092-8674(04)00045-5

Caratti, S., Turrina, S., Ferrian, M., Cosentino, E., & De Leo, D. (2015). MiSeq FGx sequencing system: A new platform for forensic genetics. Forensic Science International: Genetics Supplement Series, 5, e98–e100. https://doi.org/10.1016/j.fsigss.2015.09.040

Courts, C., & Madea, B. (2011). Specific micro-RNA signatures for the detection of saliva and blood in forensic body-fluid identification. Journal of Forensic Sciences, 56(6), 1464–1470. https://doi.org/10.1111/j.1556-4029.2011.01894.x

Giampaoli, S., DeVittori, E., Valeriani, F., Berti, A., & Romano Spica, V. (2017). Informativeness of NGS Analysis for Vaginal Fluid Identification. Journal of Forensic Sciences, 62(1), 192–196. https://doi.org/10.1111/1556-4029.13222

Hanson, E., Lubenow, H., & Ballantyne, J. (2009). Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. Forensic Science International: Genetics Supplement Series, 2(1), 503–504. https://doi.org/10.1016/j.fsigss.2009.08.184

Illumina. (2015). An introduction to next generation sequencing technology. San Diego, CA: Illumina

Illumina. (2015). Targeted Next-Generation Sequencing for Forensic Genomics Table of Contents, 12.

Illumina. (2015). TruSeq® Nano DNA Library Prep Reference Guide. Illumina, 1–38. https://doi.org/FC-121-9006DOC

Jäger, A. C., Alvarez, M. L., Davis, C. P., Guzmán, E., Han, Y., Way, L., … Holt, C. L. (2017). Developmental validation of the MiSeq FGx Forensic Genomics System for Targeted Next Generation Sequencing in Forensic DNA Casework and Database Laboratories. Forensic Science International: Genetics, 28, 52–70. https://doi.org/10.1016/j.fsigen.2017.01.011

James, S., Nordby, J., & Bell, S. (2014). No Title. In Forensic Science: An Introduction to Scientific and Investigative Techniques (pp. 205–228).

Lech, K., Ackermann, K., Wollstein, A., Revell, V. L., Skene, D. J., & Kayser, M. (2014). Assessing the suitability of miRNA-142-5p and miRNA-541 for bloodstain deposition timing. Forensic Science International: Genetics, 12, 181–184. https://doi.org/10.1016/j.fsigen.2014.06.008

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). the C. elegans\rheterochronic gene lin-4 encodes small RNAs with antisense\rcomplementarity to lin-14. Cell , 75: 843–85, 843–854. https://doi.org/10.1016/0092-8674(93)90529-Y

Li, Z., Bai, P., Peng, D., Wang, H., Guo, Y., Jiang, Y., … Zhang, L. (2017). Forensic Science International : Genetics Screening and confirmation of microRNA markers for distinguishing between menstrual and peripheral blood, 30, 24–33.

Mclaughlin, G., & Lednev, I. K. (2015). In situ identification of semen stains on common substrates via Raman spectroscopy. Journal of Forensic Sciences, 60(3), 595–604. https://doi.org/10.1111/1556-4029.12708

Murakami, Y., Tanahashi, T., Okada, R., Toyoda, H., Kumada, T., Enomoto, M., … Azuma, T. (2014). Comparison of hepatocellular carcinoma miRNA expression profiling as evaluated by next generation sequencing and microarray. PLoS ONE, 9(9), 1–9. https://doi.org/10.1371/journal.pone.0106314

Muzzey, D., Evans, E. A., & Lieber, C. (2015). Understanding the Basics of NGS: From Mechanism to Variant Calling. Current Genetic Medicine Reports, 3(4), 158–165. http://doi.org/10.1007/s40142-015-0076-8

Nussbaumer, C., Gharehbaghi-Schnell, E., & Korschineck, I. (2006). Messenger RNA profiling: A novel method for body fluid identification by Real-Time PCR. Forensic Science International, 157(2–3), 181–186. https://doi.org/10.1016/j.forsciint.2005.10.009

Old, J. B., Schweers, B. A., Boonlayangoor, P. W., & Reich, K. A. (2009). Developmental Validation of RSID™-Saliva: A Lateral Flow Immunochromatographic Strip Test for the Forensic Detection of Saliva. Journal of Forensic Sciences, 54(4), 866-873. doi:10.1111/j.1556-4029.2009.01055.x

Park, J., Kwon, O., Kim, J., Yoo, H., Lee, H., Woo, K., . . . Kim, Y. (2014). Identification of body fluid-specific DNA methylation markers for use in forensic science. Forensic Science International: Genetics, 13, 147-153.

Qiagen (2012). RNeasy® Mini Handbook. Qiagen. Hilden, Germany

Quail, M., Smith, M. E., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., … Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. BMC Genomics, 13(1), 341. https://doi.org/10.1186/1471-2164-13-341

Sauer, E., Madea, B., & Courts, C. (2014). An evidence based strategy for normalization of quantitative PCR data from miRNA expression analysis in forensically relevant body fluids. Forensic Science International: Genetics, 11(1), 174–181. https://doi.org/10.1016/j.fsigen.2014.03.011

Sauer, E., Reinke, A. K., & Courts, C. (2016). Differentiation of five body fluids from forensic samples by expression analysis of four microRNAs using quantitative PCR. Forensic Science International: Genetics, 22, 89–99. https://doi.org/10.1016/j.fsigen.2016.01.018

Schweighardt, A. J., Tate, C. M., Scott, K. A., Harper, K. A., & Robertson, J. M. (2015). Evaluation of Commercial Kits for Dual Extraction of DNA and RNA from Human Body Fluids,,. Journal of Forensic Sciences, 60(1), 157–165. https://doi.org/10.1111/1556-4029.12586

Seashols-Williams, S., Lewis, C., Calloway, C., Peace, N., Harrison, A., Hayes-Nash, C., … Zehner, Z. E. (2016). High-throughput miRNA sequencing and identification of biomarkers for forensically relevant biological fluids. Electrophoresis, 37(21), 2780–2788. https://doi.org/10.1002/elps.201600258

Shi, S. Q., Ke, J. J., Wu, W. Q., & Xu, Q. S. (2017). Serum miRNA-203 expression is associated with chemo-response to standard FOLFOX treatment of patients with colorectal cancer. International Journal of Clinical and Experimental Pathology, 10(1), 105–116.

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

Thermo Fischer Scientific (2016). Basic Principles of RT-qPCR. Retrieved from https://www.thermofisher.com/us/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/basic-principles-rt-qpcr.html

Thermo Fischer Scientific (2017). NanoDrop One User Guide. Thermo Fischer Scientific. Wilmingtion, DE

van der Meer, D., Uchimoto, M. L., & Williams, G. (2013). Simultaneous analysis of micro-RNA and DNA for determining the body fluid origin of DNA profiles. Journal of Forensic Sciences, 58(4), 967–971. https://doi.org/10.1111/1556-4029.12160

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

Virkler, K., & Lednev, I. K. (2008). Raman spectroscopy offers great potential for the nondestructive confirmatory identification of body fluids. Forensic Science International, 181(1–3), 1–5. https://doi.org/10.1016/j.forsciint.2008.08.004

Visone, R., & Croce, C. M. (2009). MiRNAs and Cancer. The American Journal of Pathology, 174(4), 1131–1138. https://doi.org/10.2353/ajpath.2009.080794

Wang, Z., Zhang, J., Wei, W., Zhou, D., Luo, H., Chen, X., & Hou, Y. (2015). Identification of Saliva Using MicroRNA Biomarkers for Forensic Purpose. Journal of Forensic Sciences, 60(3), 702–706. https://doi.org/10.1111/1556-4029.12730

Wehling, M. (2015). Principles of translational science in medicine: From bench to bedside (2nd ed.). Amsterdam: Elsevier/AP, Academic Press is an imprint of Elsevier. doi:https://doi.org/10.1016/B978-0-12-800687-0.00003-7

Winter, J., Jung, S., Keller, S., Gregory, R. I., & Diederichs, S. (2009). Many roads to maturity: MicroRNA biogenesis pathways and their regulation. Nature Cell Biology, 11(3), 228-234. doi:10.1038/ncb0309-228

Yagi, Y., Ohkubo, T., Kawaji, H., Machida, A., Miyata, H., Goda, S., … Yokota, T. (2017). Next-generation sequencing-based small RNA profiling of cerebrospinal fluid exosomes. Neuroscience Letters, 636, 48–57. https://doi.org/10.1016/j.neulet.2016.10.042

Yang, Y., Xie, B., & Yan, J. (2014). Application of Next-generation Sequencing Technology in Forensic Science. Genomics, Proteomics & Bioinformatics, 12(5), 190–197. https://doi.org/10.1016/j.gpb.2014.09.001

Zheng, G., Du, L., Yang, X., Zhang, X., Wang, L., Yang, Y., … Wang, C. (2014). Serum microRNA panel as biomarkers for early diagnosis of colorectal adenocarcinoma. British Journal of Cancer, 111(10), 1985–1992. https://doi.org/10.1038/bjc.2014.489

Zubakov, D., Boersma, A. W. M., Choi, Y., Van Kuijk, P. F., Wiemer, E. A. C., & Kayser, M. (2010). MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation. International Journal of Legal Medicine, 124(3), 217–226. https://doi.org/10.1007/s00414-009-0402-3