Spring 6-2022

Signature Peptide Identification for Body Fluids in Sexual Assault Cases by Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS)

Kelci Somers
CUNY John Jay College, kelci.somers@jjay.cuny.edu

Signature Peptide Identification for Body Fluids in Sexual Assault Cases by Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS)

A Thesis Presented in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Forensic Science

John Jay College of Criminal Justice

The City University of New York

Kelci Somers

May 2022

Signature peptide identification for body fluids in sexual assault cases by liquid chromatography tandem mass spectrometry (LC-MS/MS)

Kelci Somers

This thesis has been presented to and accepted by the office of Graduate Studies, John Jay College of Criminal Justice in partial fulfillment of the requirements for the degree of Master of Science in Forensic Science.

Thesis Committee:

Thesis Advisor: Dr. Mechthild Prinz

Second Reader: Dr. Marta Concheiro-Guisan

Third Reader: Dr. Zhe Cheng

Table of Contents

	Pages
Acknowledgments	i
Abstract	ii
Introduction	1
Literature Review	2-15
Materials and Methods	16-22
Results	23-41
Discussion	42-47
Conclusions	47-48
Future Work	48-49
References	50-57

Acknowledgments

I would like to especially thank Dr. Prinz and Dr. Concheiro for working closely with me on this project and helping me every step of the way. I have learned a tremendous amount and appreciate that you both constantly encouraged me to push myself. I would also like to thank my third reader, Dr. Cheng, for taking the time to read my work. I would also like to thank my family and friends for bringing me joy, love, and endless support. Finally, I would like to thank my fiancé for being my support system throughout these two years. I have not only grown as a student but also as a leader and scientist.

Abstract

Body fluids contain proteins that perform functions specific to different types of body fluids. Therefore, the detection of signature peptides for these proteins can potentially identify a body fluid in a forensic investigation. This project aimed to validate a method to detect signature peptides in body fluids commonly found in sexual assault cases by LC-MS/MS. Signature peptides for semen and saliva fluids were combined with two signature peptides for vaginal fluids. Samples created using two donors each for saliva, semen, and vaginal fluids were extracted using a trypsin digest, with separation of the protein and DNA fractions. The LC-MS/MS was utilized in MRM mode to analyze the protein fractions and DNA fractions were quantified. In a sensitivity study, a minimum volume of 0.005 µL of semen and 0.01 µL vaginal fluid could still be detected. A study of stains on different substrates showed variability between protein and DNA recovery. Two and three-component mixtures were created to analyze the interferences between the body fluids. The semen peptides were detected in all but one mixture while saliva peptides were detected as only a major component in a mixture. The DNA fractions for semen stains or swabs contained sufficient male DNA for further analysis. One of the vaginal peptides was detected in all vaginal samples with good sensitivity, but the second vaginal peptide was not detected. This assay needs to be optimized further.

Introduction

Body fluids such as saliva, semen, and vaginal secretions are routinely found at crime scenes in connection with sexual assaults or other violent crimes. Determining what body fluid is present can be imperative to the case. The presence of semen, for example, can prove sexual activity and the presence of saliva could confirm a victim's description of an assault. The appearance of body fluid can be very deceiving (Orphanou, 2015). Human body fluid composition is very complex and forensic body fluid identification has many limitations. One of the main issues with identifying body fluids is deciding what body fluid is present and how to move forward.

Reliable body fluid identification helps with crime reconstruction and laboratories to target the most probative samples for DNA typing. Current body fluid identification assays can be tedious which has contributed to the fact that the backlog of sexual assault cases is increasing in crime laboratories (Quinlan, 2021). There is a need for fast, sensitive, and reliable methods to identify body fluids (McKiernan et al., 2021).

This research aimed to modify a method previously developed by Browne et al. (2020) for determining signature peptides in semen, and saliva using LC-MS/MS by adding peptide markers for vaginal secretions. This method incorporations a novel trypsin lysis procedure, which has the advantage that it provides a DNA fraction for STR analysis and a protein fraction to identify the body fluid(s). This method was tested for body fluid detection limits using a dilution series, and interference from other fluids in controlled mixtures containing varying ratios of semen, saliva, or vaginal secretions. Furthermore, different swabs and fabrics were studied to see if the method can be compatible with different substrates.

Literature Review

Body Fluids and their Compositions

Saliva, Semen, and vaginal secretions are body fluids of interest in forensic casework. Body fluids may be found at a crime scene or submitted as evidence after a hospital examination (e.g., sexual assault kit). Body fluids can be an important part of case reconstruction and can aid in verifying witness statements. Therefore, it is important to understand the composition of possible body fluids that can be encountered. Saliva has multiple functions including food digestion and acting as a protective mechanism for oral tissues. The proteins within saliva can be bacteria-killing and favor different environments (van Nieuw Amerongen et al., 2004). For example, histatins are more prevalent in an ionic environment. Furthermore, the three salivary glands (parotid, submandibular, and sublingual) introduce different proteins (van Nieuw Amerongen et al., 2004).

Another common body fluid found in forensic cases is semen and it is produced by the prostate gland and the seminal vesicles. The function of semen proteins is mainly for the formation of semen and reproduction (Uniprot, 2022). Semenogelin, a protein in semen, is a main component of semen coagulation and is formed in the seminal vesicle (de Lamirande, 2007). Another thing to consider is the protein composition after a vasectomy. For example, a vasectomy can alter the cysteine-rich secretory protein-1 proteins seen in semen (Légaré et al., 2013).

Finally, the vaginal canal comprises of mucosa, muscularis, and adventitia (The Human Protein Atlas, n.d.) and the pH is very acidic to protect the reproductive tract. The proteins in the vaginal canal that make up the function of the vaginal secretion carry semen up the vaginal tract and destroy bacteria or any foreign substance. Since body fluids have proteins that perform

specific functions, some of the proteins can be unique to that body fluid. The peptides that make up those unique proteins, are also specific and can be identified using mass spectrometry (Yang et al., 2013). Although, there are many other ways to analyze body fluids including presumptive testing and other confirmatory testing.

Presumptive Testing for Body Fluids

Presumptive testing is a screening tool to detect for the possible presence of the targeted substance. Presumptive tools can be used for body fluids like blood, semen, and saliva. A presumptive test can give an indication of the body fluid present, but a confirmatory test should confirm the presence using a more specific technique. Some presumptive tests for body fluids that will be mentioned below are color changing tests (amylase, acid phosphate), alternate light sources (Sperm Tracker), and antigen tests (prostate specific antigen).

Color Changing Tests

Acid phosphatase (AP) is a color test that turns violet in the presence of semen. AP is an enzyme present in many areas of the body but at much higher concentrations in the prostate gland. The color change occurs when α-Naphthyl phosphate is added to a seminal stain to produce α-Naphthol in the presence of the enzyme AP. Diazo blue B dye is then added and if α-Naphthol is produced, a violet color will appear (Laux, 2003). The color test that is most often used to determine the presence of saliva is based on the enzyme amylase. Amylase is present at high concentrations in saliva and the function of the enzyme is to digest starch (Hedman et al., 2011). To test for the presence of saliva, potassium iodine can be added to the sample to produce a purple/grey color. Or Benedict's solution can be used, which comprises of copper sulfate, and the blue solution will turn red in the presence of amylase (Cochran et al., 2008). An advantage to

a color test is it can give the analyst an idea of how to move forward with the sample. It also aids in case reconstruction. A disadvantage of color tests is that they aren't specific. For example, false positives and negatives can occur leading to possibly missing crucial evidence.

Alternate Light Sources

Alternate light sources can also be used to detect body fluid stains like semen. They provide a nondestructive and easy technique to detect body fluids for collection. Sperm Tracker is a spray that detects semen and acid phosphatase through a napththyl reaction. It provides a luminescent signal when semen is present and has shown to only dim with a high dilution and in the presence of feces (Borges et al., 2017). Compared to the mini-Crime Scope, the Sperm Tracker is superior because it is less easily inhibited. Advantages of this technique aid with crime reconstruction because the spray can provide an overall pattern. Also, it is very beneficial with showing body fluids or stains that weren't present to the human eye.

Antigen Tests

The final type of presumptive test that will be mentioned is an antigen test. The antigen test looks for the specific antigens like prostate-specific antigen (PSA) or semenogelin to presumptively identify the body fluid. PSA is an epithelial protein created in the prostate gland (Laffan et al., 2011). SERATEC PSA and RSID-Semen are tests that detect PSA and semenogelin. SERATEC PSA uses an immunoassay approach to detect PSA while RSID-Semen utilizes a screening strip to detect semenogelin. SERATEC PSA was found to be more sensitive than RSID-Semen by a factor of 10^{-1} and RSID-Semen produced a false positive result when in presence of Femodene, a female contraceptive (Laffan et al., 2011). Overall, presumptive tests can be very helpful in initially testing for a body fluid but can also be problematic if the sample

is degraded, or if in the presence of another body fluid that could possibly inhibit the result. Unlike a presumptive test, a confirmatory test is able to identify a body fluid confidently.

Confirmatory Testing for Body Fluids

DNA Methylation

Confirmatory testing is utilized to confidently identify body fluids. Some examples of confirmatory tests for body fluids are DNA methylation, RNA testing, and sperm search. DNA methylation detects the cytosine and guanine dinucleotides which are separated by a phosphate group (CpG) that are naturally methylated. Each type of tissue can be determined based on the level of methylation since it varies according to the body fluid (Frumkin et al., 2011). DNA methylation is capable of detecting a specific body fluid for aged samples and in mixtures. Furthermore, this process is more stable because DNA is more stable than RNA, which is helpful in forensic samples, since the sample may be subject to many environmental factors. (Choung et al., 2021).

RNA Methods

Although RNA is less stable than DNA, the confirmatory methods involving RNA are specific and reliable. RNA is an essential molecule with many varieties that have many functions within the body. Messenger RNA (mRNA), circular RNA (circRNA), and microRNA (miRNA) are just a few types that can be used to identify body fluids. There is also a future for next generation sequencing to detect single nucleotide polymorphisms that are present in mRNA and can potentially distinguish between different individuals in a mixed sample (Ingold et al., 2020).

mRNA is the molecule created from coding DNA regions that is responsible for protein production (Lynch & Fleming, 2021). Gene expression differs between each tissue in the body

which creates a unique composition of mRNA transcripts. Juusola & Ballantyne (2005) produced an octuplex system by identifying specific markers for blood, saliva, semen, and vaginal secretions and found that each body fluid could be identified in single or mixed stains. Albani & Fleming (2018) also identified four specific mRNAs for blood, semen, and menstrual fluid and all of the markers were found to be present for the specific body fluid.

CircRNA is produced through pre-mRNA back splicing. CircRNA markers have been added in addition to mRNA markers to increase the specificity and sensitivity of the experiment (Liu et al., 2020). How circRNA is expressed isn't completely understood so the analysis is dependent on the ratio between circRNA and linear RNA. CricRNA is not subject to degradation and is stable throughout the analysis due to its circular nature (Lynch & Fleming, 2021).

Finally, miRNA is a regulating, noncoding RNA sequence. miRNA is more stable than mRNA due to its short length, but it is not human specific. Since miRNA are connected to mRNA due to mRNA regulation, it can also be specific to certain body fluids. Although, miRNA is less explored since it is a recent interest for identifying body fluids (Lynch & Fleming, 2021).

Sperm Search

Sperm search is a method that can be used to specifically identify sperm in a sample by using a microscope. Typically, a Christmas Tree Stain with two dyes is used to dye the spermatozoa. The Nuclear Fast Red is a red dye that dyes the head of the spermatozoa and epithelial cell nuclei a red/pink color. The other dye is picroindigocarmine which dyes the epithelial cell plasma and spermatozoa tails green (Zeffer, 2014). This is completed in a two-step process with first adding the Nuclear Fast Red dye onto the microscope slide with the sample and waiting about 15 minutes. Then, distilled water is added to wash the sample before the picroindigocarmine dye is added. After a couple of seconds, the picroindigocarmine is washed

off with ethanol. Once the slide is dried, an analyst can view the slide and determine if spermatozoa are present (Zeffer, 2014). This technique is relatively specific expect the Nuclear Fast Red dye can also dye yeast cells a pink color.

Although all the above methods can work to identify body fluids, proteomics has shown to be a promising and specific way to identify body fluids. Proteins within the body can be specific to a body fluid or tissue due to their function. Browne et al. (2019); Danielson et al. (2018); Legg et al. (2014) (2017), and many others have shown success with identifying body fluids with proteins and their signature peptides.

Body Fluid Identification using Proteins

Proteomics is the study of the proteins within a tissue or body fluid and where the proteins are produced (Merkley et al., 2019). As stated previously, proteomics can be used to identify body fluids because proteins have specific functions. The function of a protein may be specific to that body fluid, or it may just be more abundant in that body fluid (Hu et al., 2011). Proteins are made up of peptides which are essentially the building blocks of proteins (Vitorino, 2018). Peptides are short amino acid sequences that are connected by peptide bonds to create a long amino acid sequence or a protein (Yang et al., 2013).

The use of proteomics in forensic science to identify body fluids can be highly advantageous to a case, especially when paired with mass spectrometry (Merkley et al., 2019). Also, evaluating multiple peptides (multiplex) in one confirmatory test creates a fast and efficient way to identify body fluids (Yang et al., 2013). But there can be issues with the interpretation that can impede the analysis. First, the analysis of peptides can be much more difficult on a mass spectrometer (MS) than for other substances. Also, the masses of peptides are very similar to their isomers

which can cause a misinterpretation (Merkley et al., 2019). It is recommended to use a more specialized software to analyze the data in conjunction with human interpretation. For example, Spectrum Mill MS Proteomics Workbench software suite can be used and set with modifications (Legg et al., 2014).

Several instruments have been used to identify and analyze proteins with and without a trypsin digest. Furthermore, peptides and proteins can be analyzed in different ways depending on what information needs to be collected. There is an untargeted study and a targeted study. An untargeted study is usually used when as many proteins or peptides as possible need be identified in one sample. An untargeted study by Legg et al. (2014), identified hundreds of proteins in the human body and over two hundred proteins were found in the body fluids studied by LC-MS/MS. A targeted analysis can be completed on specific peptides or proteins and is commonly performed with an instrument utilizing a triple quadrupole in multiple reaction monitoring (MRM) mode (Merkley et al., 2019). Some main instruments used to identify body fluids from proteins are high performance liquid chromatography (HPLC), LC-MS/MS, and the matrix assisted laser desorption ionization-time of flight-mass spectrometer (MALDI-TOF) (Legg et al., 2014; Yang et al., 2013).

Instrumentation for Protein-Based Body Fluid Identification

MALDI-TOF-MS

The matrix-assisted laser desorption ionization-time of flight-mass spectrometer (MALDI-TOF-MS) has three components: the ionization source, the analyzer, and the detector (Tuma, 2003). The sample is placed on a metal plate or a stainless-steel probe and then the laser is directed onto the sample. The laser creates an ion plume from the matrix and the analyte which

then goes into the mass analyzer, TOF (Skoog et al., 2007). The TOF separates the molecules depending on how long the molecules take to go through the tube when accelerated with the same energy. The molecules hit the detector and the time it took the molecule to pass will determine its mass-to-charge ratio (*m/z*) (Tuma, 2003).

The MALDI-TOF bottom-up approach can detect peptide markers from a trypsin digest (Kamanna et al., 2017). This type of instrumentation provides a high-resolution analysis with low detection limits. This instrumentation has, for example, been shown to have a high specificity for human blood (Kamanna et al., 2017). Another advantage to MALID-TOF-MS is the protein coverage it can provide with targeted and untargeted proteins (Yang et al., 2013). A disadvantage of using this instrument for detecting body fluid specific peptides/proteins is that the masses between the peptides could be similar to each other (Yang et al., 2013).

Raman Spectroscopy

Raman spectroscopy is the difference in scattering upon radiation of molecules depending on the chemical makeup. The sample can be gas, liquid, or solid and the laser source irradiates the sample creating the scattered radiation (Skoog et al., 2007). The amount of scattering can give information about the sample and its identification. Recently, there have been developments made to increase the sensitivity and selectivity specifically for protein studies (Kuhar et al., 2021). Raman spectroscopy can differentiate between non-human and human blood as well as discriminate between other body fluids. Furthermore, Raman spectroscopy is nondestructive and is even more robust in the presence of chemometric models that can interpret the data (Muro et al., 2016). Another advantage is that Raman can identify the proteins because the secondary structures are quite different from each other. A disadvantage to using Raman is

that the spectra are complex with only subtle differences so that body fluids can't be distinguished without sophisticated chemometrics software (Muro et al., 2016).

Attenuated Total Reflection-Fourier Transform-Infrared Spectroscopy (ATR-FT-IR)

ATR-FT-IR examines the reflection that occurs when the sample is hit with a beam of radiation (Skoog et al., 2007) A depth of penetration occurs when the sample is hit with the beam of radiation, and it penetrates the surface of the sample. This is also called an evanescent wave. The weakened energy then reaches the detector to create a spectrum (Hinton-Sheley, 2021).

ATR-FT-IR is nondestructive and requires minimal sample preparation, which can make it a better choice than FT-IR. ATR-FT-IR has been shown to qualitatively identify multiple body fluids: human blood, saliva, semen, and vaginal secretions. Each body fluid can be identified based on its characteristic peaks on the generate spectrum (Orphanou, 2015). ATR-FT-IR analysis, combined with partial least squares discriminate analysis (PLSDA), can determine a phenotype from blood (Mistek et al., 2019). This analysis is possible because of the difference in the lipids and proteins from person to person. The ATR-FT-IR showed a 92% accuracy in classifying for sex and race (Mistek et al., 2019). The advantages to using ATR-FT-IR is that there is not sample preparation, and it is nondestructive. A disadvantage to using ATR-FT-IR is that currently mixtures have not been studied and each body fluid can only be identified on its own (Orphanou, 2015). This may not be helpful in a forensic situation since most of the time, the analyst does not know what body fluid they are analyzing.

LC-MS/MS

LC-MS/MS is a powerful instrument that is known for its sensitivity, ease of use, and working well with non-volatile substances (Skoog et al., 2007). The first step in this instrument is the separation by the LC. Once the sample is injected, a high-pressure pump pushes the mobile

phase and the sample toward the guard column where any contaminants are trapped to protect the analytical column (Skoog et al., 2007). The sample is then introduced into the analytical column. Reversed phase is commonly used so that the nonpolar components in the samples will bind to the nonpolar stationary phase on the column. A gradient mobile phase is when the mobile phases composition changes throughout the run. The use of a gradient mobile phase will move the molecules through the column at different rates, depending on their polarity. The molecules will then move to the detector or the tandem mass spectrometer.

The sample is first evaporated and ionized to make the liquid sample suitable for the mass analyzer. The most common type of ionization source is electrospray ionization (ESI). The mass analyzer is comprised of a quadrupole. A quadrupole consists of four parallel rods that alternate between AC and DC voltages (Skoog et al., 2007) to separate the ions based on their *m/z*. The ions are introduced to the first mass analyzer, where the precursor ion is filtered. Then, the precursor ion is further broken down by application of collision energy and interaction with gas molecules in the collision cell (Skoog et al., 2007). Finally, specific product ions are filtered in the second quadrupole mass analyzer. Multiple reaction monitoring (MRM) will allow the user to select specific precursor and product ions whereas full scan mode will provide a full mass spectrum of the analyte. LC-MS/MS has been able to identify unique proteins in both MRM and full scan mode in conjunction with the Spectrum Mill MS Proteomics workbench software suite (Legg et al., 2014, 2017). MRM mode is used when specific peptides are chosen for analysis (target analysis). Browne et al. (2019) analyzed and chose specific peptides to analyze specific body fluids in MRM mode. They chose to analyze specifically for saliva with the peptides SMR3B-1&2 and semen with the peptides SEMG-1&2 in the sample. Full scan was utilized in Legg et al. (2014) when the goal was to detect multiple peptides in semen, saliva, blood, vagina

fluid, and menstrual blood (untargeted analysis). Both methods can be used to identify body fluids, but they are very different approaches in the manner the MS is acquired and processed.

All the instrumentation mentioned above is extremely helpful when identifying body fluids. Most of instruments can identify body fluids even in mixtures. Probative body fluid stains then are subject to DNA testing, here one approach is to identify male DNA using a Y-screen method. Some positive samples are processed with differential DNA extraction which will create a separate semen DNA fraction for easier analysis.

Direct to DNA or Y-Screen Assays

A Y-Screen can be used to detect male DNA present in a sample. The assay uses sodium hydroxide (NaOH) to lyse the epithelial and sperms cell to prepare for quantitative real time polymerase chain reaction (qPCR) quantification (Luyando, 2018). Typically, this assay is used after a body fluid stain was located and tested presumptively positive, omitting any confirmatory body fluid testing. If the sample contains enough Y-DNA, then the next step is differential analysis to separate the male fraction from the female fraction. Since it has been shown that fabric samples could interfere with this assay, the approach is better suited for sexual assault kits with vaginal or other orifice swabs (Thermo Fisher Scientific, 2015). The assay incorporates an internal PCR control (IPC) to indicate if inhibition is occurring.

A Y-screen kit used by the New York Office of Chief Medical Examiner, uses a lysis method without sodium dodecyl sulphate (SDS), mercaptoethanol, and dithiothreitol (DTT), all of which can inhibit qPCR. This assay was found to be sensitive enough to detect male DNA with a single sperm cell and worked successfully on soiled denim (Hickey & Mayall, 2019). Loop-mediated isothermal amplification (LAMP) is another technique used for male DNA

detection instead of PCR. Fluorescent primers are not needed which is more economical and allows for the signal to be detected by the naked eye. The colorimetric results from using LAMP show that this technique is comparable to the post-LAMP electrophoresis (Scott et al., 2019).

Utilizing the Y-screen can show if a mixture is present and if the Y-component can be recovered with STR analysis. If the mixture ratio between the small-autosomal component and the Y-component is less than 1:20, the male STR profile can be recovered. If the ratio is more than 1:20, then a differential extraction needs to be completed to try to separate the male fraction from the female fraction, an alternative is Y-chromosomal analysis (Prinz et al., 1997)**.**

Differential Extraction

One limitation of the Y-screen approach is that sexual assault kit samples routinely contain a combination of female and male cells. These mixed samples result in complex genotyping results, where the Short Tandem Repeat (STR) profile from the male perpetrator may be obscured by the victim's STR type. These results are more difficult to interpret and may results in fortuitous associations with a person of interest. Differential extraction is a procedure where an analyst first lyses the epithelial cells and pellets the sperm cells, then after several wash steps uses dithiothreitol (DTT) to lyse the sperm cells for DNA extraction (Butler, 2009) This separates the male fraction and female fraction in a sample. This extraction is time consuming since there are many wash cycles required (Clark et al., 2021). Also, some of the male fraction can be lost or stay mixed with the female fraction of DNA, which could be detrimental in a forensic case with a limited sample (Clark et al., 2021). Using the Y -screen approach, the analyst can determine if the male to female DNA ration and then complete a differential extraction if required.

Protein and DNA Co-Extraction

The Y-screen approach yields a DNA fraction but no body fluid information. To obtain a protein extract for body fluid testing and a DNA fraction could be tedious, time consuming, and consume a large amount of sample. Here simultaneous methods could reduce extraction time and to use less of the sample. Kranes et al. (2017) used a trypsin digestion extraction in conjunction with a microcon MW100 filter to separate the two fractions. The function of trypsin is to cleave proteins at Lys and Arg residues to separate the peptides with K or R at C-terminal (Burkhart et al., 2012). Kranes et al. (2017) compared trypsin and proteinase K as lysis enzymes and found trypsin to yield more DNA and proteins than proteinase K. A microcon MW100 filter doesn't allow product to go through the filter unless it is smaller than 100 kDa. Therefore, most peptides can flow through the filter, while the larger DNA molecules are retained. Once the peptide fractions are recovered, the filter is inverted to retrieve the DNA. Kranes et al. (2017) found that the peptides were particle free and ready for analysis on the mass spectrometry.

Peters et al. (2020) compared three simultaneous extraction methods including the one mentioned above. The Qiagen AllPrep DNA/RNA/Protein Mini Kit, the CFSRE method, and the Kranes et al. method were compared to see which method provided the best peptide peak area intensity by UPLC-MS/MS, DNA quantification values, overall quality of genetic profile, consistency among preparation replicates, and speed/cost and of workflow. Overall, the Qiagen AllPrep and CFSRE method were not as successful as the Kranes et al. (2017) method because the Qiagen AllPrep showed allele dropout and didn't meet the protein identification criteria. The CFSRE method was better than the Qiagen AllPrep but it still did not show interpretable results for some of the protein targets. Overall, Kranes et al., method was the most successful because it met all the criteria set (Peters et al., 2020).

Research Goals

The goals for this research are to validate the signature peptides, previously analyzed by Browne et al. (2019), SEMG-1 &2 for semen and SMR3B-1&2 for saliva by utilizing MRM mode on the LC-MS/MS. Also, to introduce signature peptides NGAL-1&2 to identify vaginal fluids. We hypothesize that SEMG-1&2 and SMR3B-1&2 will remain robust and reliable in identifying the semen and saliva, and that NGAL-1&2 will be able to identify vaginal fluids.

Materials and Methods

Sample Collection and Preparation

Three vaginal swabs, and two donors each for vaginal secretions, semen, and saliva were used for this research project. Vaginal swabs, liquid vaginal secretions, and semen samples were purchased from Lee BioSolutions (Maryland Heights, MO). To avoid freeze thaw cycles and preserve the integrity of the sample the two semen donor secretions of 1.0 mL and the two vaginal donor secretions of 100 µL were further aliquoted into four 1.5-mL Eppendorf tubes. Vaginal fluid 1 had to be diluted with nuclease-free water in order to pipette the vaginal secretion. Saliva was collected from two male volunteers, under a previously approved CUNY IRB (project #2017-0800). Each donor provided around 2 mL of oral fluid into a 50-mL Falcon tube. Oral fluid was collected from the male volunteers but will be referred to saliva for the rest of the study. The samples were aliquoted into four 1.5-mL Eppendorf tubes just as the semen and vaginal secretions had been.

For the substrate effect study, 2 μL of semen, saliva, and vaginal secretions were pipetted onto polyester swabs, cotton swabs, and a clean polyester T-shirt swatch. Each substrate was tested three times for each donor. For a sensitivity study, neat and serial dilutions of all six body fluid samples were added directly to an Eppendorf tube at the following volumes: 1.0 µL, 0.5 µL, 0.25 µL, 0.1 µL, 0.05 µL, and 0.025 µL. A further dilution series was completed to reach the detection limit for each peptide and donor at the following volumes: 0.01 µL, 0.005 µL, 0.0025 µL, 0.0001 µL, and 0.0005 µL. The saliva samples were spun down after the first dilution series. The body fluids were also mixed at various ratios in combinations of either two or three body fluids. Both vaginal secretions samples were used but for both semen and saliva, the respective donor whose samples consistently had higher concentrations in the previous studies was used.

Vaginal samples from Donors 1 and 2 were mixed at different ratios with saliva from Donor 2 and semen from Donor 2 on a polyester t-shirt cutting. The volume ratios of vaginal fluid to saliva or semen were 100:1, 50:1, 10:1, 5:1, 1:1, 1:5, 1:10, 1:50, 1:100. Then, all three of the body fluids were mixed on a polyester t-shirt cutting to determine whether each of the body fluids could still be detected in the presence of varying concentrations of the different body fluids. Table 1 shows the ratios created for the mixtures, using 4µL as the highest volume and ensuring all the body fluids were major components. This was prepared separately with vaginal fluid from Donors 1 and 2 with both semen Donor 2 and saliva from Donor 2.

Table 1. Mixture of three body fluids

Vaginal Samples	**Semen 2**	**Saliva 2**
4 µL	1 µL	4 µL
4 µL	4 µL	1 µL
4 µL	1 µL	1 µL
1 µL	1 µL	1 µL
1 µL	4 µL	4 µL
1 µL	1 µL	4 µL
1 µL	4 µL	1 µL

Standard Preparation

Vaginal signature peptides were added to the previously established assay (Browne, 2020) to create a more complex multiplex that would provide additional information in a sexual assault case. The vaginal signature peptides were chosen based on Legg et. al. (2018) and

Universal Protein Resource (UniProt, 2022). The protein chosen to detect vaginal secretion was neutrophil gelatinase-associated lipocalin (NGAL), and sequences for two signature peptides, NGAL-pep 1 and NGAL-pep 2, were provided using Genescript (Piscataway, NJ). Based on Browne's research, the semenogelin (SEMG) protein with SEMG1-pep1 and SEMG-1-pep2 and the saliva protein SMR3B and the peptides SMR3B-pep1 and SMR3B-pep2 (Browne, 2020) were kept as peptide targets. Table 2 lists the amino acid sequence for all the target peptides.

Table 2. Body Fluid Specific Proteins and Signature Peptides

Body Fluid	**Protein**	**Peptide Name**	**Sequence**	**Overall Charge**
Semen	Semenogelin-1	SEMG1-pep1	QGGSQSSYVLQTEELVANK	-1
		SEMG2-pep2	DIFSTQDELLVYNK	-2
Saliva	Submaxillary gland androgen regulated protein 3B	SMR3B-pep1	GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR	+1
		SMR3B-pep2	IPPPPPAPYGPGIFPPPPPQP	Neutral
Vaginal Secretion	neutrophil gelatinase-associated lipocalin	NGL-pep1	SYPGLTSYLVR	-1
		NGL-pep2	TFVPGCQPGEFTLGNIK	Neutral

The synthetic peptides NGAL-pep1 and NGAL-pep2 were dissolved in nuclease free water and optimized on the Shimadzu LC-MS/MS 8050 (Columbia, MD) to determine the quantifier and qualifier ions. A 1-µg/mL concentration was needed to optimize the peptide so a serial dilution starting at 1,000 µg/mL was completed. The other peptides for semen and saliva secretions were diluted as needed. All the synthetic peptides were mixed at a concentration of 50

nmol/mL in nuclease-free water. A dilution series was created from 50 nmol/mL down to 0.005 nmol/mL.

Trypsin Digestion and Microcon Filtration

The buffer ammonium bicarbonate was prepared by adding 10mL of nuclease-free water to 0.5g of pre-weighed ammonium bicarbonate. Each sample was placed in a 1.5 mL Eppendorf tube and 98 μL of the buffer, 1 μL of 1% Protease Max (Promega, Madison, WI), and 1 μL of 0.5M dithiothreitol (DTT) was added to the liquid or substrate sample. It was then incubated at 56° C while shaking at 1400 rpm for 20 minutes. Three microliters of indole acetic acid (IAA) were added to each sample to conserve the disulfide bonds from separation. This was then placed in the dark for 20 minutes at room temperature. To each sample, 1 μL of DTT was added to deactivate the IAA (Sigma Aldrich, St. Louis, MO) and then placed again in the dark for 20 minutes. One microliter of 0.1 μg/µL solution of trypsin was added to each sample and then the mixture was set to incubate for 3 hours at 37° C while shaking at 1400 rpm. The trypsin (Promega, Madison, WI) was then deactivated by placing the samples in a heat block at 99° C for 10 minutes. To extract the DNA and peptides from the substrate or swab, a dolphin tube with a spin basket was used. The sample was placed in a spin basket inside the dolphin tube (MIDScientific, Valley Park, MO) and the sample was spun down for 5 minutes at 1,500 rcf. Negative controls were included in each extraction completed to ensure the reagents were not contaminated.

A Microcon DNA Fast Flow (Millipore, Burlington, MA) filter unit was used for each sample to separate the protein fractions from the DNA. The sample from the dolphin tube was added to the Microcon filter unit and centrifuged for 20 minutes at 500 rcf. The flow through

contained the protein fractions and it was transferred into a protein low bind tube. The protein fractions were placed in the -80ºC freezer. The DNA fractions were recovered by inverting the Microcon filter and placing it into a new Microcon tube and centrifuging for 3 minutes at 1,000 rcf. The DNA fractions were then placed in a new Eppendorf tube and stored in the 20ºC freezer.

DNA Quantification

Quantification with the Quantifiler Trio DNA Quantification Kit from ThermoFisher was used to determine the human DNA concentration in the sample. Standards with predetermined concentrations were prepared to create a standard curve. The DNA standard in the kit was first diluted to a concentration of 50 ng/µL and a serial dilution was completed to 0.05 ng/µL following the manufacturer's instructions. The Quantifiler THP PCR Reaction Mix and the Quantifiler Trio Primer Mix were used as a master mix. Per each sample, 5 µL of Primer Mix and 4 µL of Reaction Mix were utilized. The reaction takes place in a 96-well plate, where 9 µL of master mix and 2 µL of each sample were added to each well based on the corresponding well on the plate map. The QuantStudio 5 (Thermo Fisher Scientific, Waltham, MA) was used for quantification and the software to collect and analyze the data was HID Real-Time PCR Software v1.3 (Thermo Fisher Scientific, Waltham, MA). Refer to Figure 1 below for the thermal cycle process. The small autosomal and the Y-chromosome targets were interpreted.

$$\frac{95^\circ\,\mathrm{C}}{2\ minutes} \rightarrow 40\ cycles\ of\ \left[\frac{95^\circ\,\mathrm{C}}{9\ seconds}, \frac{60^\circ\,\mathrm{C}}{30\ seconds}\right]$$

Figure 1. Thermal Cycling Conditions for Real-Time PCR System

Peptide Analysis

The Trypsin digested peptide fractions obtained after the Microcon filtration step are suitable for further analysis without additional treatment. The protein fractions were analyzed by a Liquid Chromatograph Tandem Mass Spectrometer (LC-MS/MS), Shimadzu 8050-LCMS triple quadrupole (Columbia, MD) with electrospray ionization in positive mode (ESI+) as the ion source. Multiple reaction monitoring (MRM) was used in order to pick specific ionizations and quantify the peptides. The injection volume was set to 20 μL. The samples were placed straight into LC-MS/MS vials and a calibration curve containing a mixture of all the synthetic peptides in the buffer of ammonium bicarbonate was also analyzed. The column employed for this experiment was the Agilent AdvanceBio Peptide mapping column (100 mm x 2.1 mm, 2.7 μm). The column conditions were as follows; the flow rate was 0.5 mL/min at 55° C for the column temperature; the MS was nebulizing gas was at 2 L/min; heat and drying gas at 10 L/min interface temperature 400° C; desolvation line (DL) temperature 250° C and a heat block temperature of 400° C. Mobile phase A contained water and 0.1% formic acid and mobile phase B contained acetonitrile and 0.1% formic acid. See Figure 2 below for the gradient scheme.

All of signature peptides were analyzed for each sample to ensure specificity between body fluids. The instrument analyzed each retention time and ion ratio for the peptides, but each sample was checked to ensure the correct parameters were set. The areas and concentrations were provided for each sample, but the concentration (nmol/mL) represents how much of that peptide is in the sample.

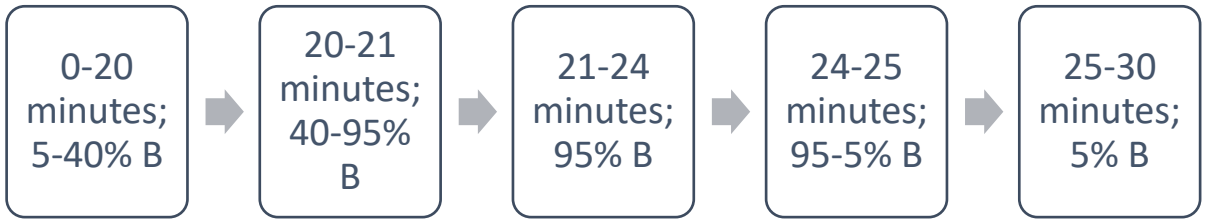

Figure 2. Steps for Gradient Scheme on the LC-MS/MS

Statistical Methods

The average, standard deviation, and relative standard deviation (RSD) were calculated from excel. The average was utilized to understand the triplicate replication on different substrates. The standard deviation was calculated to understand the differences in substrates and between donors. See Figure 3 for the equation used to calculate standard deviation. The RSD was used in order understand the percent variability of the donor for the substrates studied. See Figure 4 for the equation used to calculate RSD.

$$\sigma = \sqrt{\frac{\Sigma(x_i - \mu)^2}{N}}$$

Figure 3. Equation for Standard Deviation

$$RSD = \frac{\sigma}{\mu} x\ 100$$

Figure 4. Equation for Relative Standard Deviation

Results

Peptide Results

Optimization and Calibration for Multiplex for LC-MS/MS

In the previous study by Browne (2020) SEMG and SMR3B were optimized, and their precursor and product ions were determined. The precursor ion for SMR3B-1, a saliva signature peptide, is 1034.8 *m/z*. The quantifier product ion is 614.9 *m/z*, and the qualifier product ion is 517.8 *m/z*. The precursor ion for SMR3B-2 is 711.0 *m/z*, the quantifier product ion is 628.4 *m/z*, and the qualifier product ions are 614.4 *m/z* and 438.3 *m/z*. The precursor ion for SEMG 1, a semen signature peptide, is 680.0 *m/z*. The quantifier product ion is 136.2 *m/z*, and the qualifier product ion is 332.3 *m/z*. The precursor ion for SEMG-2 is 843.0 *m/z*, the quantifier product ion is 201.2 m/z, and the qualifier product is 424.3 *m/z*.

After NGAL-1 and NGAL-2, the vaginal signature peptides, were added to the multiplex a new optimization was required to determine the product and precursor ions for the additional peptides was required. To determine the transitions of NGAL-1 and NGAL-2, a serial dilution was performed from 1000 µg/mL to 1 µg/mL and run on the LC-MS/MS. NGAL-1 was found to have a precursor ion of 628.2 *m/z*, a quantifier ion of 503.0 *m/z*, and the qualifier ions are 908 *m/z* and 223 *m/z*. The precursor ion for NGAL-2 is 603.3 *m/z*. The quantifier ion is 221 *m/z*, and the qualifier ions are 229 *m/z* and 120 *m/z*. See Table 3 for the full report of the transitions for each peptide and their retention time.

To determine the cut off for each signature peptide, the following dilution were made for each peptide: 0.1 nmol/mL, 0.05 nmol/mL, 0.025 nmol/mL, 0.01 nmol/mL, 0.005 nmol/mL, and 0.0025 nmol/mL. See Table 4 below for the cut-offs for each peptide. Figure 5 represents a MRM chromatogram of the multiplex of the synthetic peptides.

Table 3. LC-MS/MS transitions for synthetic peptides

Peptide Name	Precursor *m/z*	Product *m/z*	Retention Time (min)
SMR3B-1	1034.8	614.9*	14.2
		517.8	
SMR3B-2	711.0	628.4*	13.0
		614.4	
		438.3	
SEMG-1	680.0	136.2*	10.3
		332.3	
SEMG-2	843.0	201.2*	12.2
		424.3	
NGAL-1	628.2	503*	10.9
		908	
		223	
NGAL-2	603.3	221*	12.3
		229	
		120	

**Quantifier Ions*

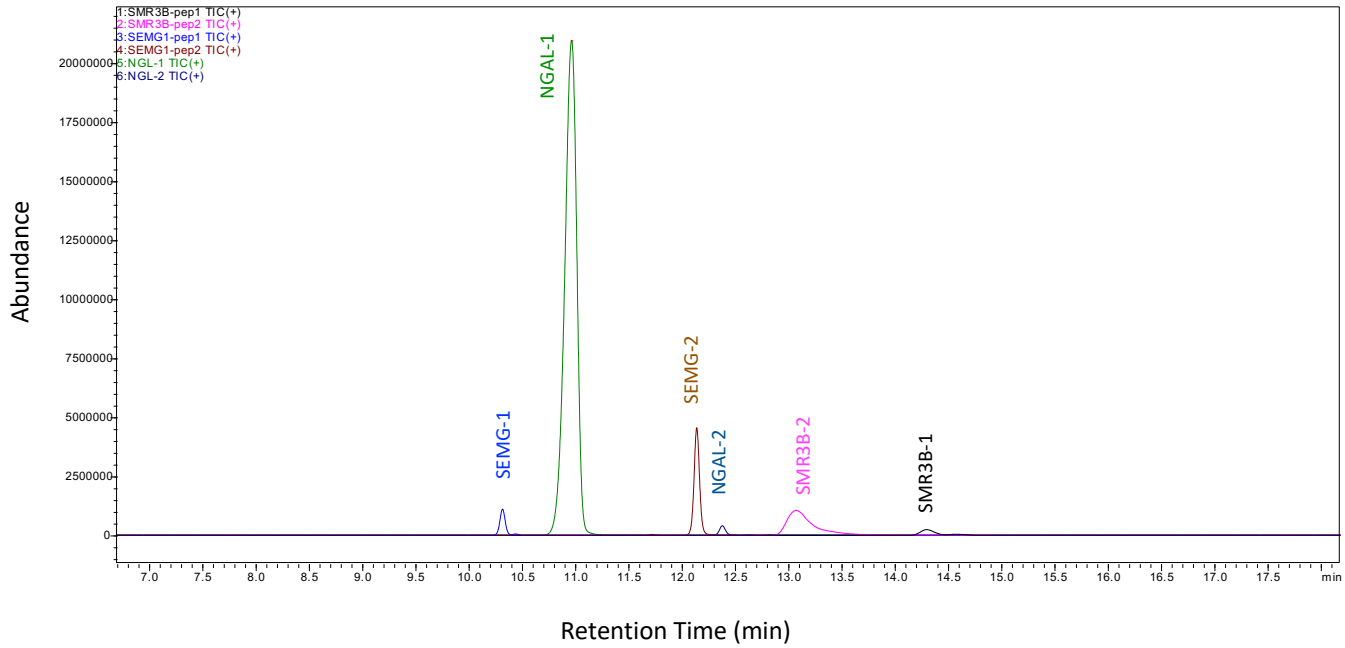

Figure 5. Chromatogram Showing the Multiplex with Synthetic Peptides

Table 4. Cut-Offs for each Synthetic Peptide in the Multiplex

Peptide	**Cut-Off (nmol/mL)**
NGAL-1	0.005
NGAL-2	0.05
SMR3B-1	0.05
SMR3B-2	0.005
SEMG-1	0.005
SEMG-2	0.0025

Substrate Study in Triplicate

All body fluids were tested in triplicate on cotton swabs, polyester swabs, and cuttings of a polyester t-shirt. All substrates were tested alone as a negative control, and no peptides or DNA were detected. Two microliters of each body fluid were pipetted onto each substrate. For the vaginal fluids, the substrate with the highest concentration for NGAL-1 was the cotton swab, then the polyester T-shirt cutting, and finally the polyester swab. Figure 6 shows a visual representation of the substrates with both vaginal fluids. Vaginal Fluid 1 & 2 yielded around the same concentration for each substrate. NGAL-2 was not detected for both donors on any of the substrates. This issue occurred throughout the entirety of the study.

Semen from both donors 1 & 2 was placed on each substrate. Figure 7 represents the concentration of SEMG-1 and SEMG-2 on each substrate. SEMG-1 had the highest concentration extracted from the polyester swab and the lowest concentration from the cotton swab. SEMG-2 showed this same pattern as SEMG-1. Donor 1 had an overall higher concentration of SEMG-1 and SEMG-2.

Finally, SMR3B-1 had a higher concentration with the cotton swab. The polyester swab then yielded a higher concentration than the polyester t-shirt. For SMR3B-2, the polyester swab showed the highest concentration, and then the cotton swab. Between the peptides, SMR3B-2 consistently had a higher concentration overall compared to SMR3B-1. Figure 8 demonstrates the difference between the two peptides.

The relative standard deviation was calculated to show the variability between the substrates. Figure 9 shows that the polyester t-shirt cuttings produce the highest variability between the peptides. SMR3B-1 and SMR3-2 have the lowest RSD values, under 40%. NGAL-1 consistently has the highest RSD, around 70%. The large RSD values could be due to only having two donors for each body fluid. A larger sample size would give a more accurate representation of variability between the substrates.

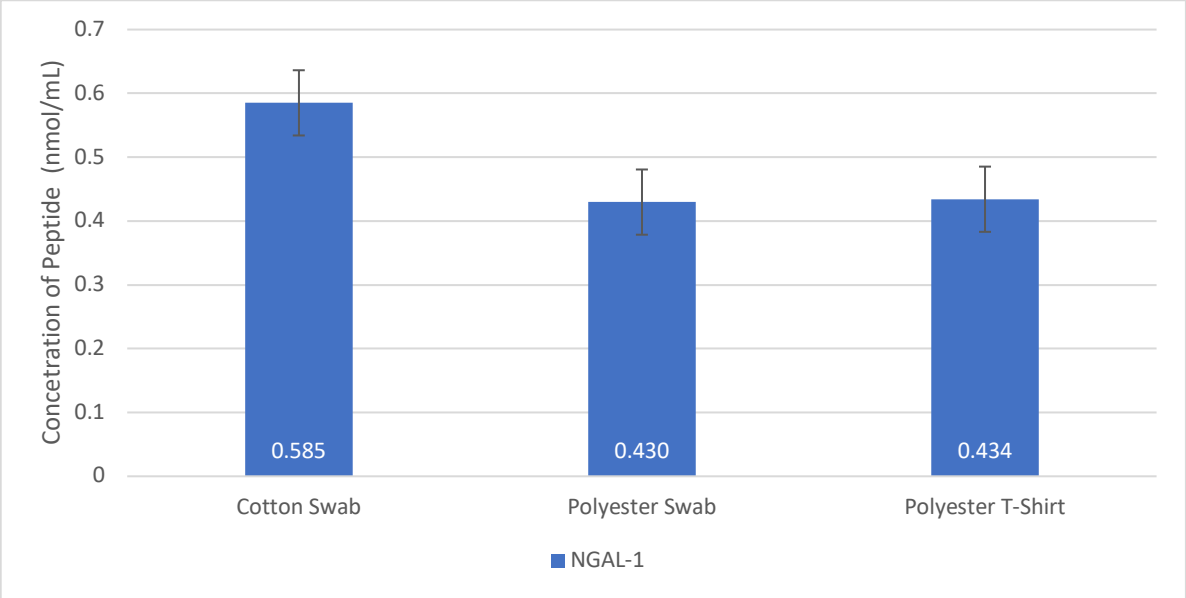

Figure 6. Vaginal fluid on all three substrates. Shown are the average vaginal peptide concentration for both donors and triplicate stains. N=6 for each substrate.

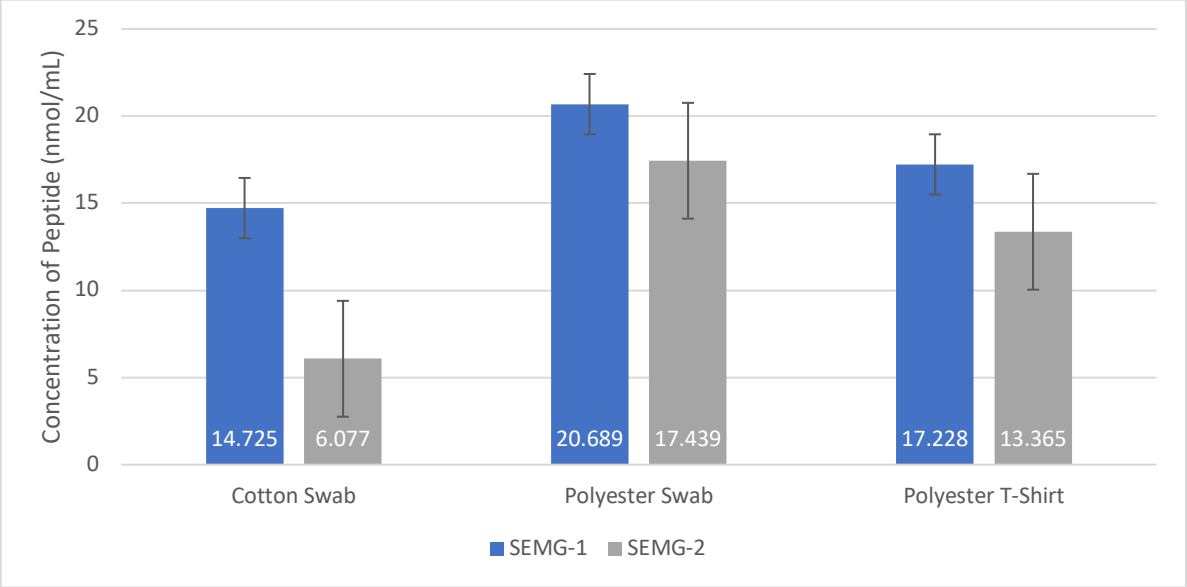

Figure 7. Semen on all three substrates. Shown are the average semen peptide concentrations for both donors and triplicate stains. N=6 for each substrate.

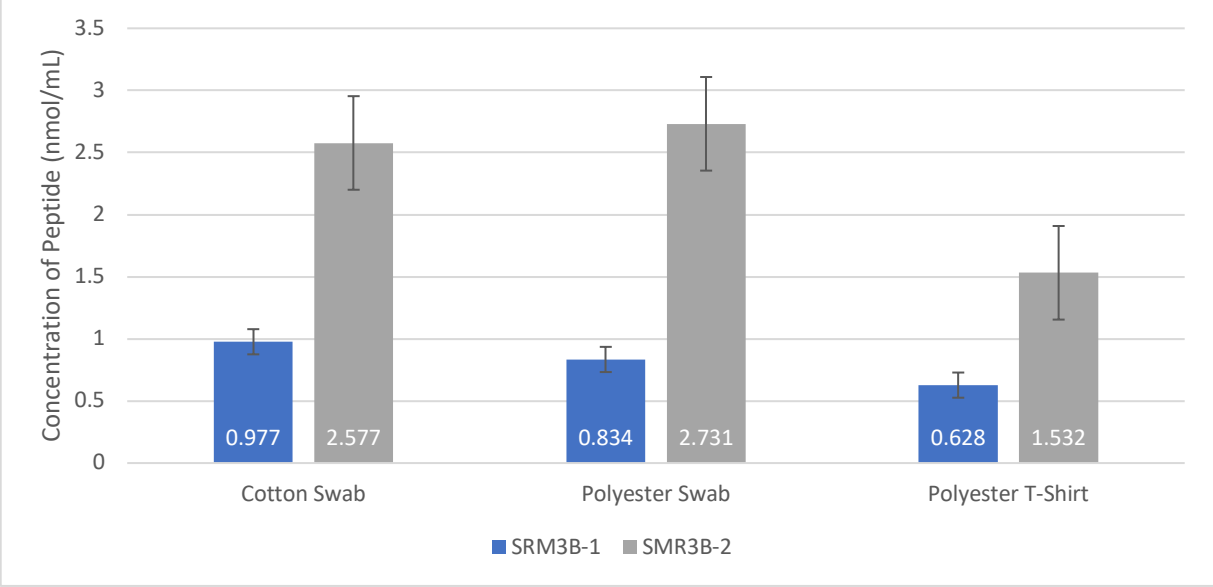

Figure 8. Saliva on all three substrates. Shown are the average saliva peptide concentrations for both donors and triplicate stains. N=6 for each substrate.

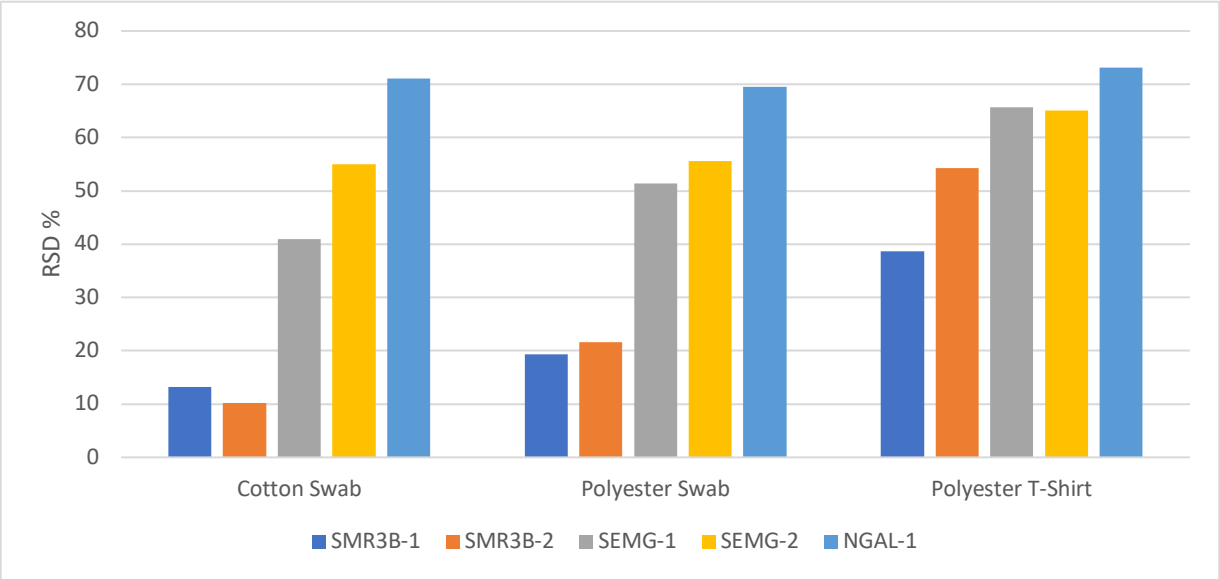

Figure 9. Relative Standard Deviation for each substrate. Shown is the percentage variability of the three substrates for each peptide across two donors and triplicate stains.

Table 5. Average peptide concentrations for both donors on different substrates

Peptide	Donor	Cotton Swab	Polyester Swab	Polyester T-Shirt
SMR3B-1	1	1.604	0.857	0.43
	2	0.889	0.811	0.825
SMR3B-2	1	2.524	2.677	0.891
	2	2.630	2.786	2.173
SEMG-1	1	9.253	11.31	7.170
	2	20.19	30.07	27.29
SEMG-2	1	3.178	9.155	5.712
	2	8.976	25.72	21.02
NGAL-1	1	0.951	0.701	0.715
	2	0.219	0.159	0.153

Table 5 shows averages over the triplicates for all donor substrate combinations. The body fluid specific peptides displayed different concentrations, with SEMG-1 and -2 having the strongest signal and NGAL-1 the lowest. These differences reflect the different concentrations of the target protein in the respective body fluid and were consistent throughout the rest of the study.

Sensitivity Study

To determine the minimum volume for each peptide to be detected, a sensitivity study was conducted. Each body fluid was first diluted to represent volumes from 1.0 µL to 0.025 µL. All the peptides besides SMR3B-1 reached the detection limit at 0.25 µL. Another dilution series was created to cover lower volumes from 0.01 µL to 0.0005 µL. It was determined that the cut-off for SMR3B-1 was 0.5 µL and SMR3B-2 was 0.025 µL. Both SEMG-1 and SEMG-2 had a detection limit of 0.005 µL. The cut-off for NGAL-1 was 0.01 µL.

SMR3B-2 was still detected at 0.025 µL but was not detected in the second dilution series. The saliva from both donors was centrifuged between these two studies because it was challenging to pipette. A considerable quantity of peptides could have been left in the mucous pellet created and discarded by centrifuging the saliva. Therefore, it is likely that the detection limit for SMR3B-2 is lower than 0.25 µL.

Figures 10 and 11 represents the dilution series for both semen peptides. Semen Donor 1 consistently had a higher concentration than Donor 2 for SEMG-1 and SEMG-2. Figure 12 represents NGAL, and it is evident that vaginal fluid 1 has a higher concentration than vaginal fluid 2, where no decrease was observed. The vaginal fluid from Donor 1 was diluted in a 1:1 ratio initially with nuclease-free water because it was a paste-like consistency. The dilution series for vaginal fluid 1 may have been more successful because it was easier to pipette. Finally, a

dramatic decline in concentration occurred for SMR3B-2 after 1 µL (see Figure 13) but the peptide was still detected at 0.025 µL.

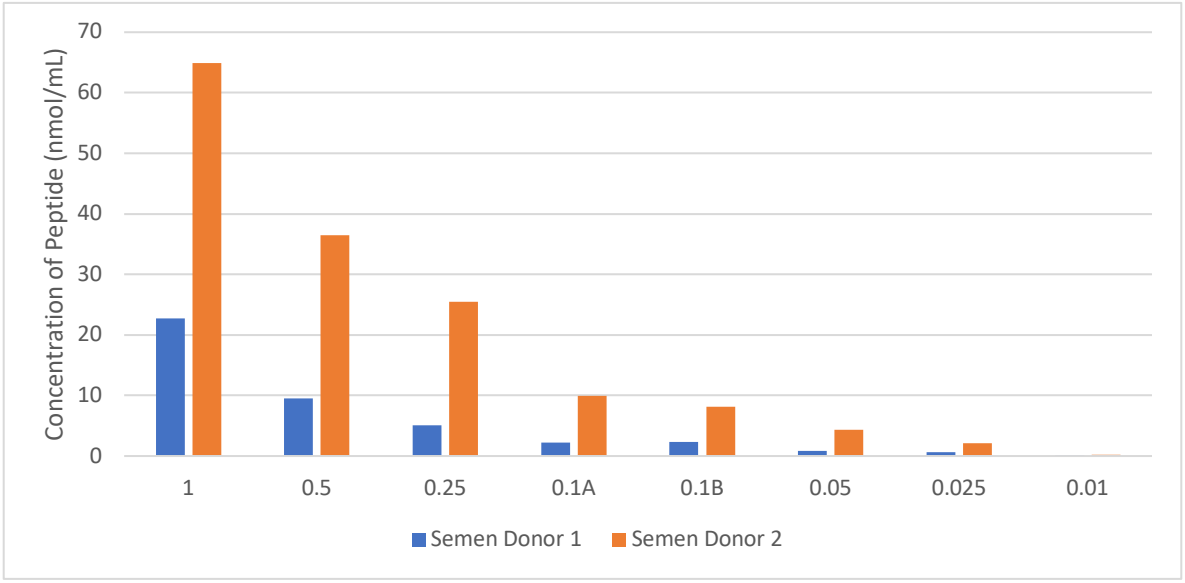

Figure 10. Semen dilution series. Shown are SEMG-1 concentrations for both semen samples.

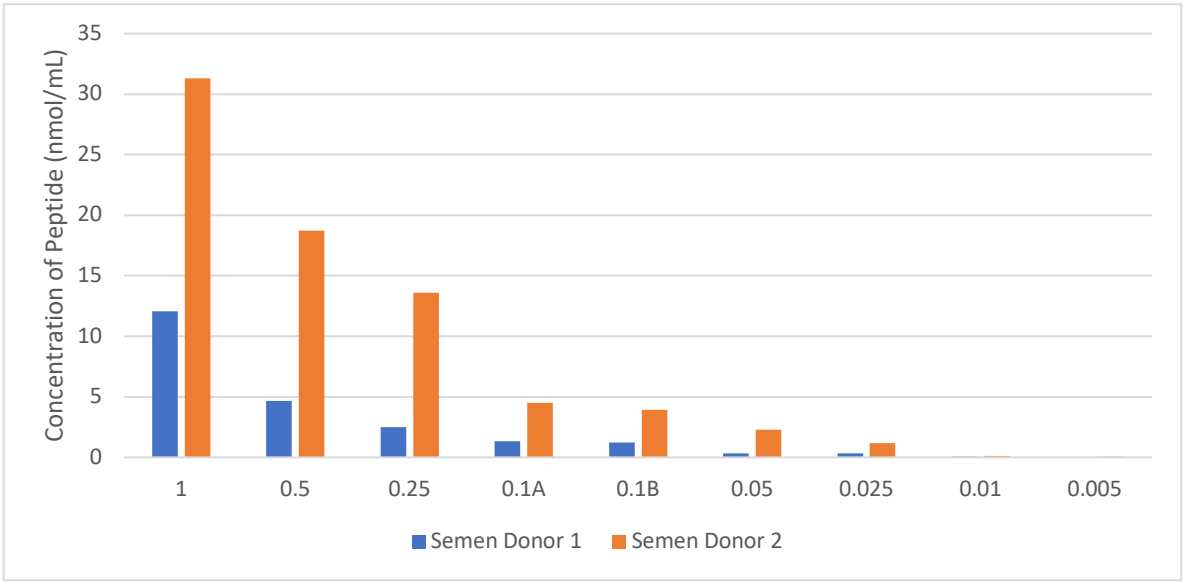

Figure 11. Semen dilution series. Shown are SEMG-2 concentrations for both semen samples.

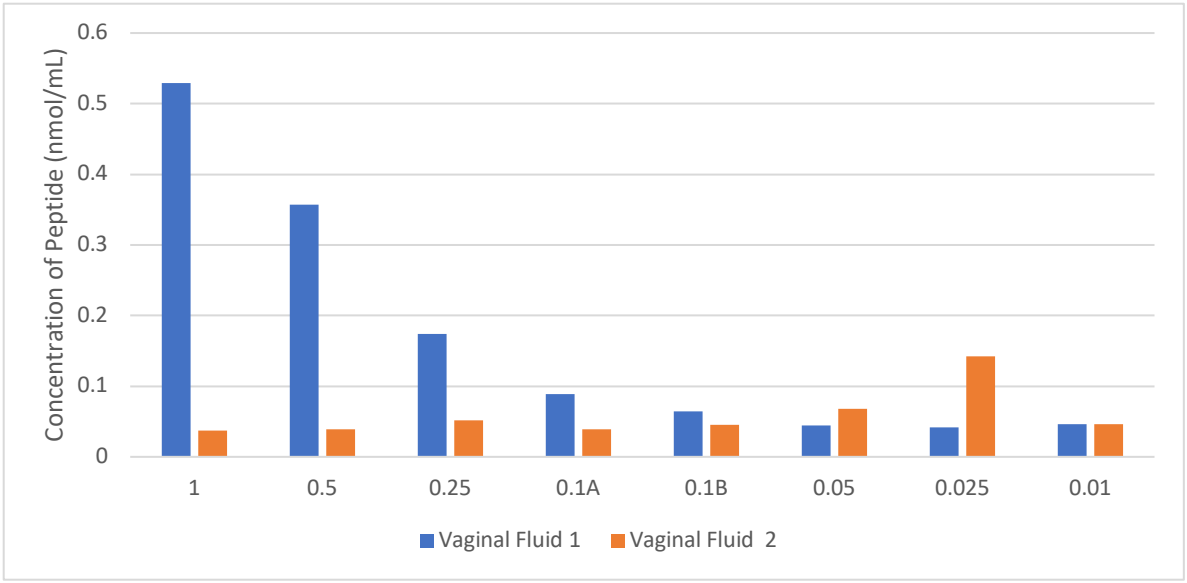

Figure 12. Vaginal fluid dilution series. Shown are NGAL-1 concentrations for both vaginal donors.

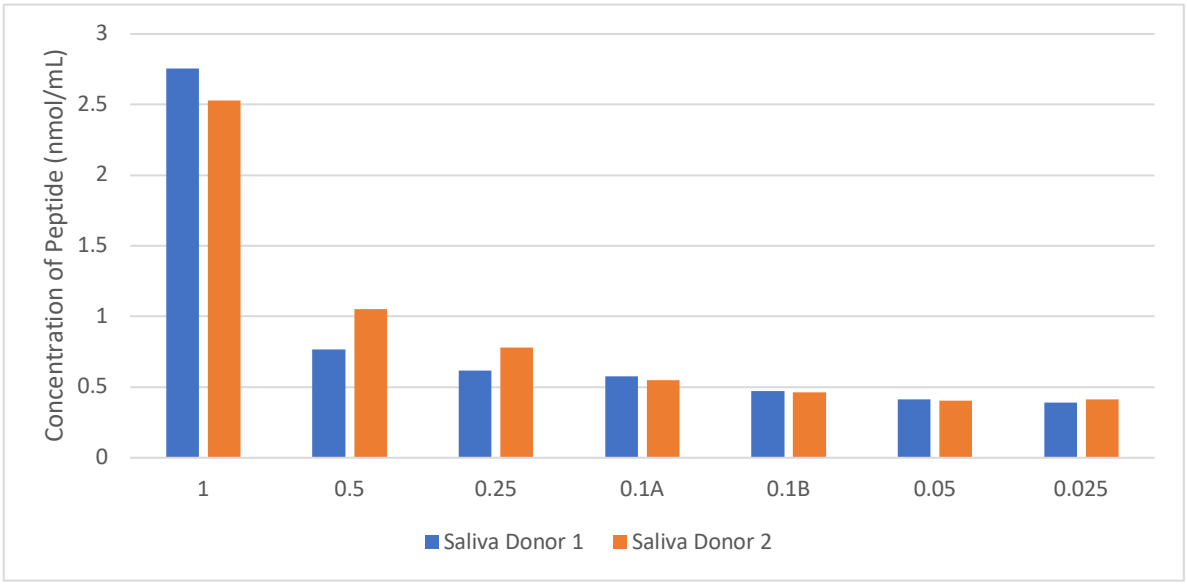

Figure 13. Saliva dilution series. Shown are SMR3B-2 concentrations for both saliva donors.

Mixtures

Body fluids were placed on polyester t-shirt at the following ratios of vaginal to semen and saliva fluids, 100:1, 50:1, 10:1, 5:1, 1:1, 1:5, 1:10, 1:50, and 1:100. The mixture series was

repeated for both vaginal fluid samples but using the same semen and saliva donor (donor 2 in each case). See Table 6 for the results of two-component mixtures. NGAL-1 is present in all ratios, and a consistent pattern is followed. SEMG-1 and SEMG-2 do not appear when present as a minor component in the 100:1 ratio. SMR3B-1 was only detected as a major component from 1:10 to 1:100. SMR3B-2 was not detected as a minor component. It was detected at 5:1 but not at the 1:1 ratio. Although, SMR3B-2 was detected as the major component from 1:5 to the 1:100 ratio.

Table 6. Signature peptide detection in two-component mixtures

Ratio (Vaginal fluid: Semen/Saliva)	**NGAL-1 (nmol/mL)**	**SEMG-1 (nmol/mL)**	**SEMG-2 (nmol/mL)**	**SMR3B-1 (nmol/mL)**	**SMR3B-2 (nmol/mL)**
100:1	0.5145				
50:1	0.473	0.17	0.067		
10:1	0.3295	0.777	0.489		
5:1	0.1935	0.978	0.615		0.441
1:1	0.0635	1.721	0.602		
1:5	0.075	6.881	2.268		0.52
1:10	0.0775	9.877	3.811	0.503	0.845
1:50	0.063	30.309	12.943	0.849	2.495
1:100	0.067	42.342	20.314	1.59	5.077

The values for NGAL-1 are averaged from all four samples mixed with semen and saliva

A three-component mixture series was included to test for any further interferences occurred with a more complicated mixture. See Table 11 in the Materials and Methods to see the

ratios created for this mixture series. Only 1:10 dilutions of each body fluid were used to create the ratios for the three-component mixture. Table 7 contains the concentrations in nmol/mL for each peptide in the mixture with vaginal fluid donor 1.

Table 7. Signature peptide detection in three-component mixtures

Vaginal: Semen: Saliva	**NGAL**	**SEMG-1**	**SEMG-2**	**SMR3B-1**	**SMR3B-2**
4:1:4 (C1)	0.12	1.171	0.901		0.304
4:4:1 (C2)	0.104	5.895	2.932		
4:1:1 (C3)	0.126	1.063	0.791		
1:1:1 (C4)	0.062	1.545	0.77		
1:4:4 (C5)	0.054	6.321	3.109		0.24
1:1:4 (C6)	0.066	1.3	0.749		0.258
1:4:1 (C7)	0.053	5.856	2.311		

For this mixture series, NGAL-1 was detected in all the ratios despite being a minor component. The values for NGAL-1 are consistent with the volume in the mixture. When NGAL-1 is the major component (0.4 µL), the average concentration is 0.117 nmol/mL. When NGAL-1 is the minor component (0.1 µL), the average concentration is 0.059 nmol/mL. SEMG-1 and SEMG-2 are both present in the mixture as major and minor components. The concentration for SEMG-1 is higher than SEMG-2 for all the mixtures. For SEMG-1, the average concentration as the major component is 6.024 nmol/mL. As the minor component, the average concentration of SEMG-1 is 1.270 nmol/mL. SEMG-2 has an average of 2.784 nmol/mL as the major component and 0.803 nmol/mL as the minor component. SMR3B-1 was not present in the

mixtures, and SMR3B-2 was only present as a major component. The average concentration for SMR3B-2 as the major component was 0.267 nmol/mL. Despite the differences in concentration between the body fluids, the ones detected were above the cut-off.

Mock Cases

Three 5-year-old vaginal swabs were collected from the -20° C freezer and tested in combination with semen Donor 1 and saliva Donor 1. Each swab was cut into three sections. One µL of fluid from Semen Donor 1 was added to 1/3 of each swab in an Eppendorf tube, and one µL of fluid from Saliva Donor 1 was added to the other 1/3 of each swab in an Eppendorf tube. The last third of the swab was used as a negative control to ensure no other body fluids were detected besides the vaginal fluid.

All peptides were detected besides NGAL-2. SMR3B-1 was detected at an average of 5.277 nmol/mL, and SMR3B-1 was detected at an average of 4.362 nmol/mL. SEMG-1 and SEMG-2 were detected at an average of 10.289 nmol/mL and 0.574 nmol/mL. NGAL-1 was detected in all swabs for an average of 0.382 nmol/mL. Figure 14 shows the multiplex of a mock case containing vaginal and saliva fluids and Figure 15 shows the multiplex with vaginal and semen fluids. Analyzing the mock cases was a successful experiment, and all the body fluids were detected.

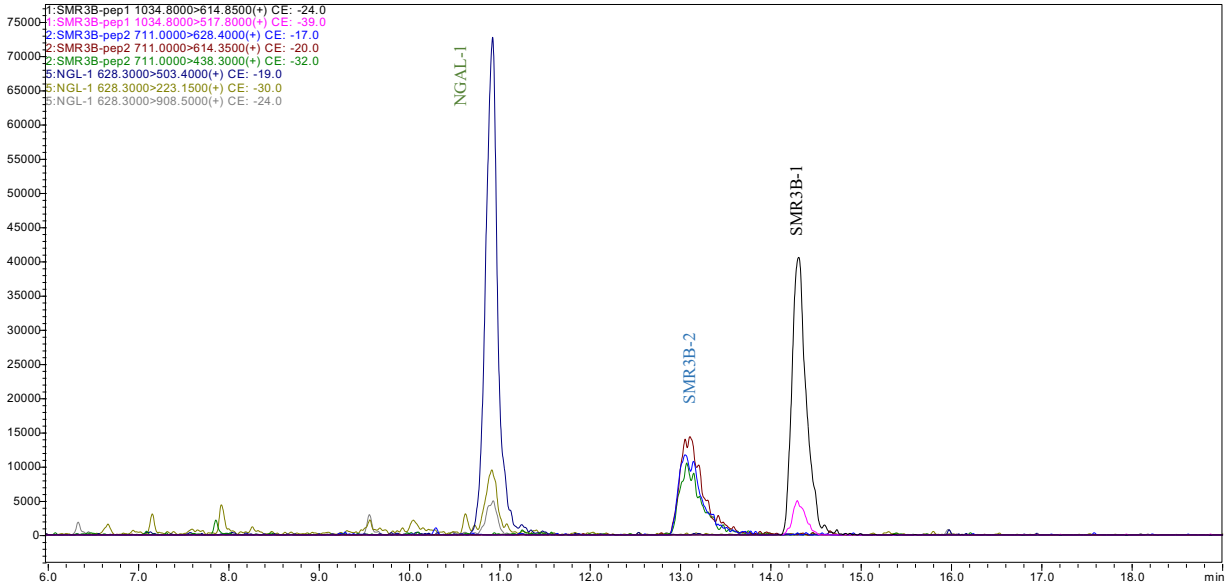

Figure 14. A chromatogram of the multiplex with a mock case of vaginal and saliva fluids

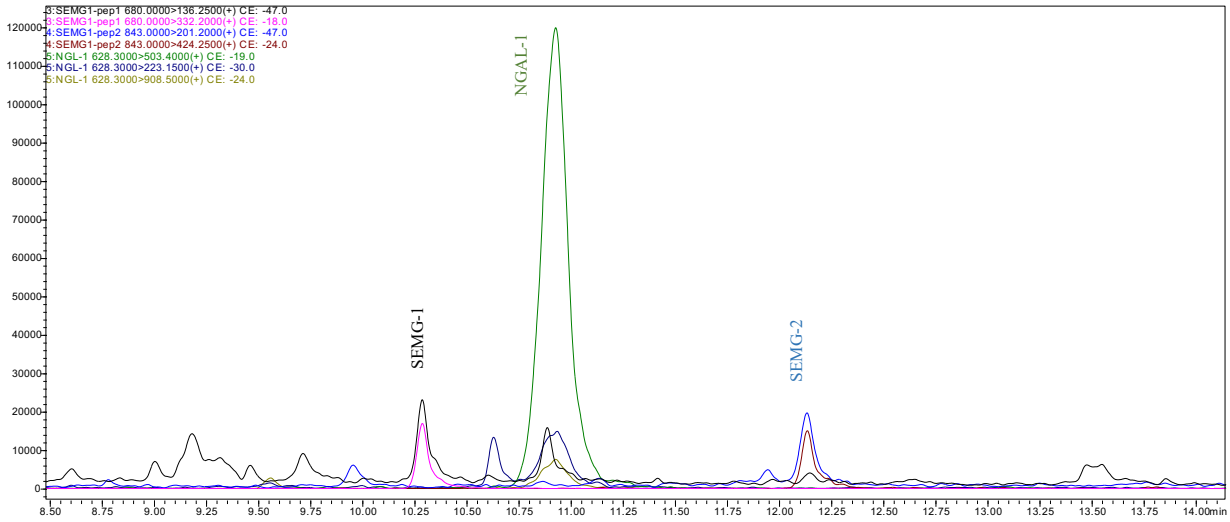

Figure 15. A chromatogram of the multiplex with a mock case of vaginal and semen fluids

DNA Results

The DNA fractions were extracted simultaneously with the peptide fractions. The DNA fractions were quantified, but no further analysis was completed due to time restraints. Although, completing the quantification step gave important insight into the possibility of downstream analysis.

Substrate Study

The substrate study again showed an effect of the substrate on DNA recovery. Table 8 shows that the two donors for each body fluid differ in their average DNA concentrations. This is caused by biological variation in cell counts in each body fluid. For all three body fluids, the polyester swab yielded the lowest DNA concentrations.

Table 8. Average DNA concentrations (ng/µL) for triplicate stains on different substrates

		Cotton Swab	**Polyester Swab**	**Polyester T-Shirt**
Saliva	1	0.30	0.22	0.40
	2	2.05	0.85	2.88
Semen	1	1.90	0.25	3.01
	2	7.05	2.50	5.96
Vaginal Fluid	1	4.86	0.51	4.98
	2	0.41	0.11	1.91

This substrate effect differed from the peptide results. Polyester swabs had the highest concentration for the semen peptides, SEMG-1 and 2. For saliva, the two peptides differed in

which substrate had the highest signal, but results were very similar for both swab types. Also, the vaginal fluids and NGAL-1 peptide concentration was similar between the polyester swab and the polyester t-shirt. See Table 9 to see the differences between the DNA fractions and the peptide fractions both vaginal fluids.

Table 9. Average peptide and DNA concentrations for both vaginal fluids

Vaginal Fluids	**DNA (ng/µL)**	**Peptide (nmol/mL)**
Polyester Swab	0.31	0.4296
Polyester T-Shirt	3.4	0.4341
Cotton Swab	2.61	0.5851

Sensitivity Study

The quantification of the sensitivity study showed a lower limit of DNA for vaginal fluids and semen fluids. DNA from vaginal Fluid 1 was still detected at 0.001 µL (See Figure 16), and the same was true for semen Donor 1, where this amount of semen showed a DNA concentration of 0.0002 ng/µL. Please note that these concentrations are not sufficient for a DNA profile. Based on, for example, the Globalfiler STR amplification kit validation, 60pg or 0.004 ng/µL of single source DNA are required for complete to almost complete genotypes (Ludeman et al., 2018). This concentration was obtained for 0.0025 µL of vaginal fluid and 0.01 µL of semen. For saliva the first set of dilution down to 0.025 µL yielded detectable DNA, but as with peptide, further dilutions were negative. These more diluted samples were created after the saliva was centrifuged. This may have pelleted buccal epithelial cells caused the loss of some of the DNA.

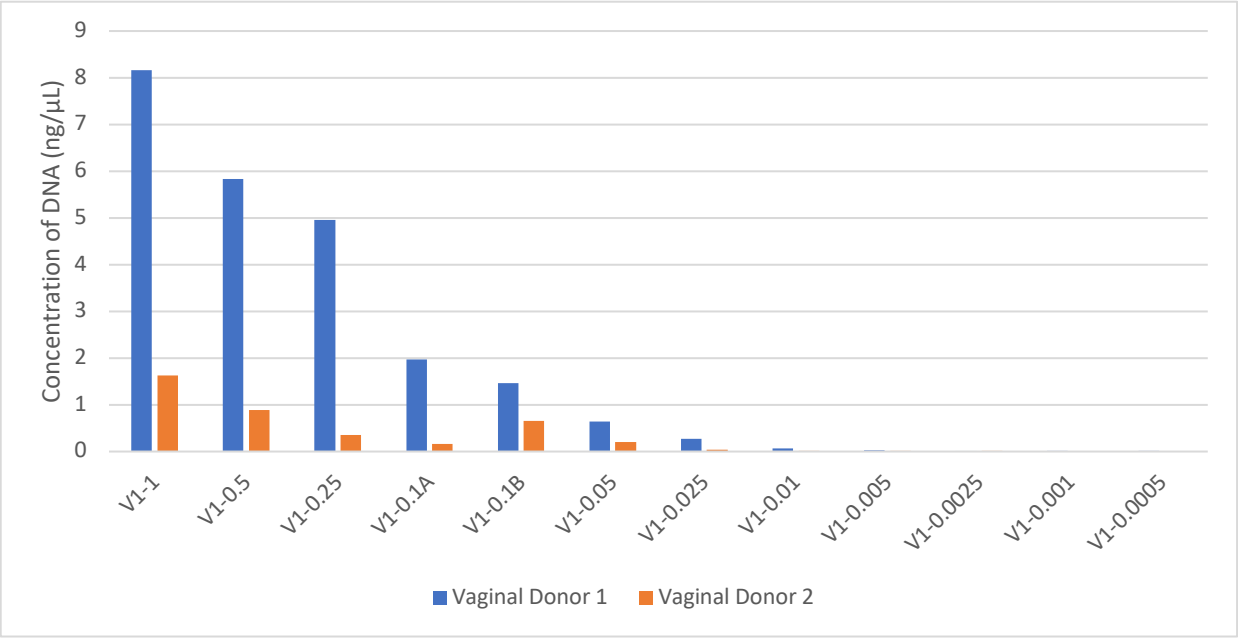

Figure 16. Vaginal dilution series. DNA concentrations of vaginal fluid samples 1 and 2.

Mixture Study

The vaginal fluid with saliva or semen mixtures were evaluated for the presence of male DNA to determine which samples would have yielded a male DNA profile. The male target in the current forensic human quantification kits is the key to the Y-screen casework concept. The male target will detect male DNA in a mixture and together with the autosomal targets provides information on how to proceed with a DNA sample. In the two-component mixture with the vaginal component and the semen or saliva, the Quantifier Trio still detected male DNA at the 100:1 vaginal: semen or saliva ratio. At the 100:1 for semen, the concentration for the Y-target was 0.01 ng/µL, which would be sufficient for a full Y-STR DNA profile. For saliva, the concentration was 0.0037 ng/µL, this amount may or may not yield a useable profile (Ludeman et al., 2018) See Figure 17 and Figure 18 for the concentration of semen and saliva at each mixture ratio.

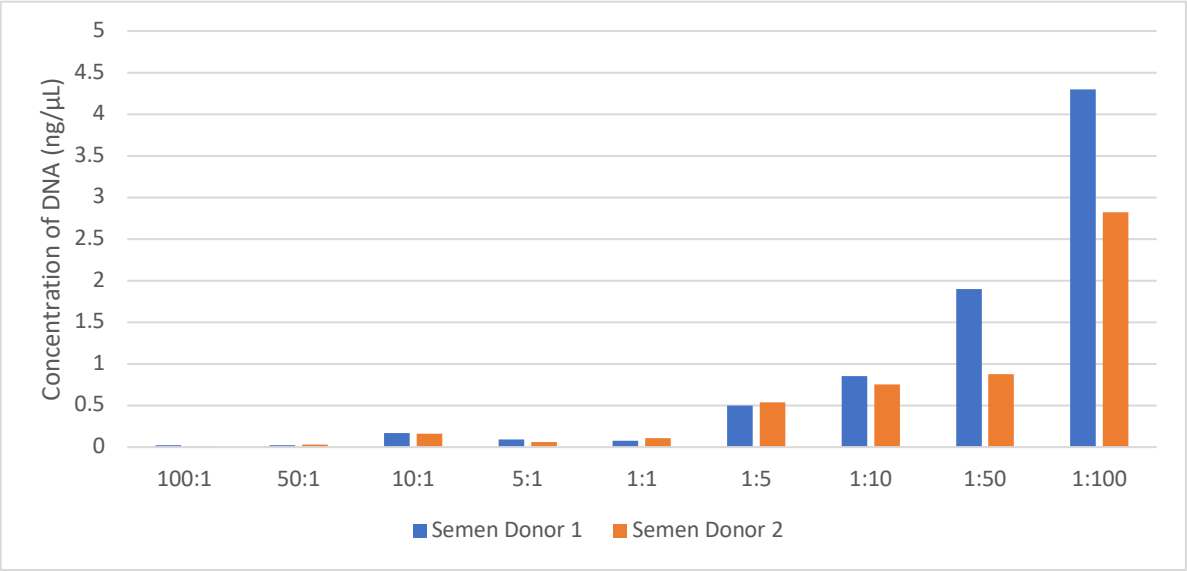

Figure 17. Two component mixture with semen. Shown is the male DNA of both semen samples at the different mixture ratios with vaginal fluid.

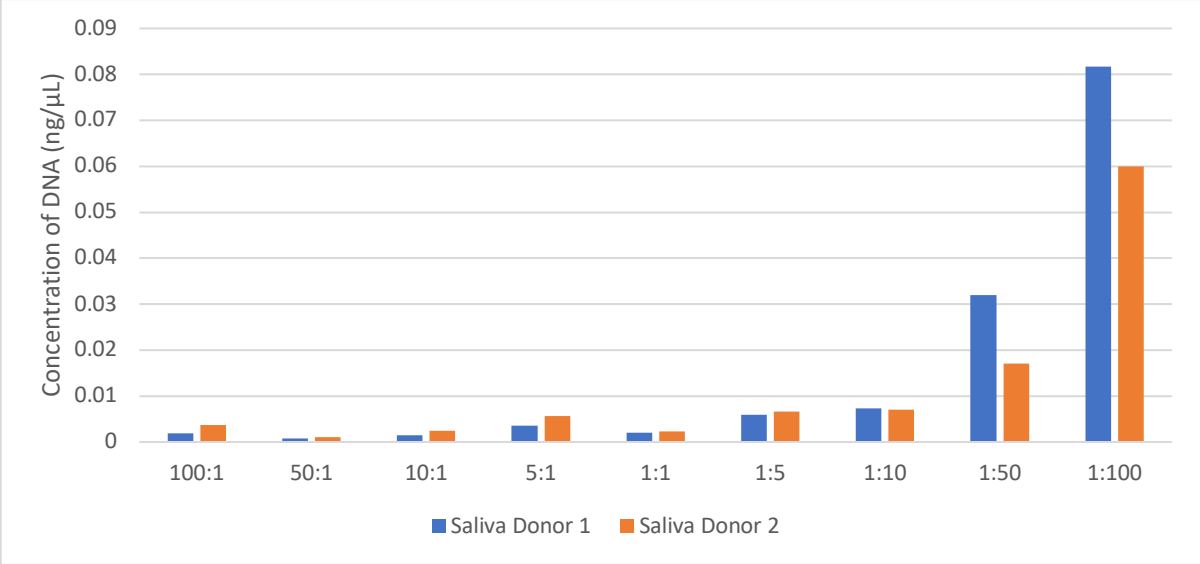

Figure 18. Two component mixture with saliva. Shown is the male DNA of the saliva samples at the different mixture ratios with vaginal fluid.

All the three-component mixtures yielded enough small-autosomal DNA for future DNA typing. The lowest amount of DNA detected was 0.08 ng/µL, and the highest amount was 2.46 ng/µL. More importantly, each mixture also yielded enough male DNA. When saliva is the

major component, the average Y-DNA concentration was 0.078 ng/µL. The average amount of Y-DNA detected when semen was the major component was 0.463 ng/µL. When both semen and saliva are the major components, the average Y DNA detected is 0.35 ng/µL. When saliva and/or semen were the minor components, an average of 0.123 ng/µL was detected as male DNA. In Figure 19, C2, C5, and C7 represent the highest concentration of Y-DNA. This is because semen is a major component in C2 and C7 and both semen and saliva are major components in C5.

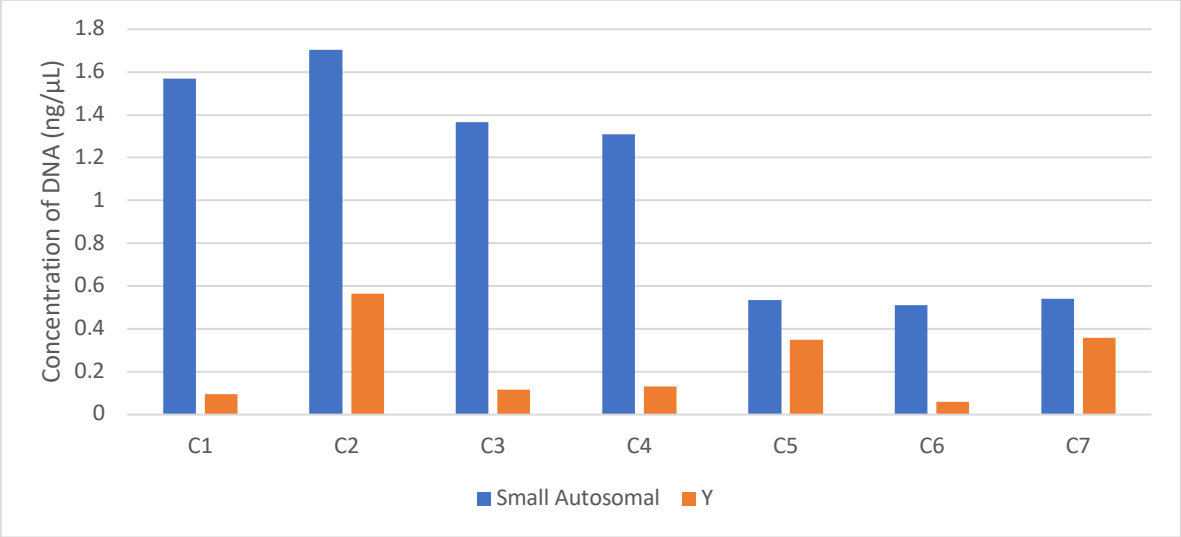

Figure 19. Three-component mixture of vaginal fluid, semen, and saliva. Shown is the comparison of the concentrations of small autosomal DNA to Y-DNA.

Mock Cases

The quantification on the mock cases determined that the amount of male DNA detected on all three swabs was sufficient for DNA typing, but the ratio between the total human DNA (small autosomal target) and the Y-chromosome target showed that only the semen samples had mixture ratios suitable for regular STR typing (see Table 10). The minor component is only detectable if the ratio is less than 1:20, for more extreme female to male DNA mixture only Y-

STR typing will reveal the male DNA genotype (Butler, 2009). Vaginal Swab 1 yielded the most DNA with both semen and saliva. On average, the vaginal swabs with the saliva contain 0.045 ng/µL of male DNA; the SA:Y ratios showed 600-1300x excess of autosomal DNA. The vaginal swabs with semen contained an average of 2.93 ng/µL of male DNA with a much smaller excess of autosomal DNA.

Table 10. DNA concentrations for mock casework samples

Vaginal Swab (VS) with Semen or Saliva	**Small auto**	**Y target**	**SA:Y ratio**
VS1 Saliva 1	104.36	0.08	1304.50
VS1 Semen 1	72.82	6.18	11.78
VS2 Saliva 1	2.84	0.0044	645.45
VS2 Semen 1	3.89	1.08	3.60
VS3 Saliva 1	64.73	0.05	1294.60
VS3 Semen 1	19.55	1.54	12.69

Discussion

This research aimed to further validate the semi-quantitative analysis of the semen (SEMG-1 and SEMG-2) and saliva (SMR3B-1 and SMR3B-2) signature peptides that were previously determined by Browne et al. (2019). The assay was modified by adding a vaginal marker, the protein NGAL, to identify vaginal fluids. NGAL-1 and NGAL-2 were selected to the assay due to the success in identifying vaginal fluids in other studies. Legg et al. (2017) verified NGAL as an appropriate vaginal marker for identifying solely vaginal fluids. NGAL was present in 100% of vaginal fluid samples and only in 6% of menstrual blood samples (Legg et al., 2017). Furthermore, NGAL was only present in human samples, and no non-human origins were found (Danielson et al., 2018). NGAL was also present in mixtures with saliva and semen, proving to be an abundant and reliable biomarker for vaginal fluid (Danielson et al., 2018).

NGAL-1 was successful in this study, but NGAL-2 was not detected in any of the conducted experiments. Danielson et al. (2018) and Legg et al. (2017) shared the same sequence for one of the NGAL peptides, SYPGLTSYLVR. For the other peptide, Legg et al. (2017) used TFVPGCQPGEFTLGNIK, and Danielson et al. (2018) used WYVVGLAGNAILR. Specific detection of TFVPGCQPGEFTLGNIK was never mentioned in the Legg et al. (2017) study. In this study, the peptide sequences used were SYPGLTSYLVR (NGAL-1) and TFVPGCQPGEFTLGNIK (NGAL-2). Throughout this study, SYPGLTSYLVR (NGAL-1) consistently was observed for each donor, and this specific peptide was reliable in detecting vaginal fluids. TFVPGCQPGEFTLGNIK (NGAL-2) was not detected in any of the experiments.

The signature peptides for semen fluid, SEMG-1, and SEMG-2, were found to be reliable, and abundant in all the studies. In this study, SEMG-1 and SEMG-2 had a minimum volume detection of 0.005 µL. SEMG-1 and SEMG-2 were found in all the mixtures but the

100:1 (vaginal: semen). Browne et al. (2019) found SEMG-1 and SEMG-2 to be detected at all volumes (0.1 µL minimum). Also, they found SEMG-1 and SEMG-2 to be present in all mixtures with saliva (Browne et al., 2019). Legg et al. (2017) detected SEMG-1 and SEMG-2 in all mixtures with different body fluids like menstrual blood, urine, vaginal fluid, and saliva. With the support from the previous studies and the findings in this study, one can state that SEMG-1 and SEMG-2 are abundant, robust, and can identify semen in a sample.

Looking at the saliva peptides, SMR3B overall was a reliable protein to use to identify saliva. SMR3B-2 was found to be more abundant overall and was detected at a lower volume than SMR3B-1. SMR3B-1 was detected until 0.5 µL while SMR3B-2 had a minimum detection volume of 0.025 µL. Also, SMR3B-2 was only detected in the three-component mixture and only when it was a major component while SMR3B-1 completely dropped out and no signal was detected. Furthermore, Legg et al. (2017) also determined SMR3B to be a reliable protein for the determination of saliva because it was found in 100% donors. SMR3B was detected in mixtures with saliva and vaginal fluids, making the peptide specific for casework. Danielson et al. (2018) also found that SMR3B-1 and SMR3B-2 can be detected with the required minimum transitions at a 1:32 dilution. Browne et al. (2019) determined SMR3B-1 and SMR3B-2 to be found in all donors and that SMR3B could be used to identify saliva. Although, it was not detected at 0.5 µL or 0.1 µL in the sensitivity study. Overall, SMR3B can identify saliva as a body fluid in most scenarios but only as a major component in a mixture.

The sensitivity studies determined the minimum volume at which each body fluid could be detected. Detection of SEMG-1, and SEMG-2 was successful and showed a continuous decline throughout each dilution series. Although, with semen Donor 1 and 2, the concentrations of DNA did not decrease. Also, the dilution series for saliva Donor 1 also did not show a

continuous decrease of DNA. The reason for this could be due to clumping or that the sample was not evenly distributed prior to the dilution step. This was also found regarding the peptide results in saliva. SMR3B-2 declined significantly after the first dilution. The dilutions were made with nuclease-free water, and it is possible that the saliva was not evenly distributed in the water due to the mucous and viscous nature of this body fluids. It was challenging to pipette from the first dilution to the next due to amount of mucous in the sample and the correct amount may have not been pipetted from the first to the second dilution. SMR3B-1 was not detected after 0.5 µL and Browne et al. (2019) had similar findings with SMR3B-1. They found that SMR3B-1 was not detected after 1µL.

After the first dilution series, NGAL-1, SEMG-1, SEMG-2, and SMR3B-2 were still detected at 0.025 µL and more dilutions had to be made to reach the detection limit for each peptide. Before this, both saliva samples were centrifuged to remove the mucous component, so it was easier to pipette. Saliva Donor 1 and Donor 2 had a concentration of 0.392 and 0.412 nmol/mL at 0.025 µL for SMR3B-2, so it was projected that this peptide would at least be detected at 0.01 µL. But no saliva peptides were detected in the second dilution series. The DNA results were also negative, and the Y-target did not detect any male DNA. This indicates that the centrifugations step removed all DNA containing buccal epithelial cells. The minimum volume determined for the saliva peptides needs to be reinvestigated, and new samples should be tested to determine if there is a lower limit of detection.

Different substrates were studied to ensure that this assay can work on similar substrates that may be involved in a sexual assault case. Cotton swabs were chosen because they are typically used for sample collection in sexual assault cases. Some sexual assault kits may also contain polyester swabs and finally, the polyester T-shirt fabric was chosen because it is similar

to the composition of some women's underwear. This study was successful in detecting both peptide and DNA fractions for all donors. Browne et al. (2019) also studied these substrates and both peptide and DNA fractions were successfully detected.

For this substrate study, the results varied between peptides and DNA. SEMG-1 and SEMG-2 showed the highest concentration with the polyester swab and NGAL-1, SMR3B-1 and SMR3B-2 had the highest concentration with cotton swabs. The DNA results for all three body fluids had the highest concentration on the polyester t-shirt cutting. The reason for this difference may be due to the structure of the substrate. Bruijns et al. (2018) studied the DNA recovery of cotton, polyester, rayon, foam, and nylon swabs. The authors mentioned that 20-76% of the DNA can be lost during the extraction step. The authors also found that the swabs that were more coiled, trapped the DNA containing cells in the swab and most of the DNA was unable to be extracted (Bruijns et al., 2018). In this study, it was found that DNA was better recovered from cotton swabs than the polyester swabs. The polyester swab may be more tightly coiled and therefore trapping the DNA in the swab. The thin polyester fabric easily released cellular DNA. The peptide fractions had more success with the swabs, and this makes sense because the target proteins are present in the body fluid and not in the cells and are smaller in size compared to DNA so they might not get as easily trapped.

The purpose of the repeatability study on different substrates was to see if the method gave reproducible results. With more donors added to the study, the variability of concentrations on each substrate may or may not follow more of a pattern since there is also variability between donors. It would be important to study this with a bigger sample size in order to determine which substrate would provide consistent results donor to donor. In this study, the polyester T-shirt showed the highest percentage of variability for all the peptides and the cotton swab showed the

smallest percentage of variability, especially with SMR3B-1 and SMR3B-2. It should also be noted that NGAL-1 had the highest variability throughout all three substrates. This is due to the large variation of concentrations between the two donors. Overall, the polyester t-shirt cutting showed the highest variability for all peptides.

The mixture study showed that all three body fluids can be detected despite the presence of one or two other body fluids. In this study, the two-component mixtures between vaginal secretions and semen showed almost no interference in this study. Only at the largest ratio (100:1) did SEMG-1 and SEMG-2 drop out. Also, SMR3B-1 and SMR3B-2 were only detected when saliva was a major component. Neither saliva peptides were detected at the 1:1 mixture and this could be due to the volume that was pipetted (1 μL). SMR3B-2 was also only detected as a major component and dropped out when only 1 μL of a 1:10 dilution of saliva was pipetted onto the polyester t-shirt cutting. Furthermore, this part of the study was completed after the saliva was centrifuged so this could have affected the results. Browne et al. (2019) studied body fluid mixtures on cotton swabs but also studied other substrates. It was found that the peptide signals were detected on all substrates. The polyester t-shirt was chosen as the substrate for the mixtures in this study since it was relatively successful with the peptides and very successful with DNA. Also, as stated previously, it mocks women's underwear that may be collected from a sexual assault. In Browne et al. (2019), it was found that semen and saliva could be both detected in a mixture except when saliva was the minor component.

Legg et al. (2017) found that relative abundance for the protein markers varies from person to person. In that study, they found that SEMG had the highest abundance, then SMR3B, and then NGAL (n=50). In this study, that same pattern was also found. This proves that the intensity or abundance of the biomarker will not cause interference. Only the volume of the body

fluid could cause interference. Furthermore, the level of abundance of the peptide in a specific body fluid speaks for the robustness of the peptide and this assay.

This assay was also capable of detecting body fluids in older samples. The mock case samples that were studied contained five-year-old vaginal swabs. Although not much degradation would be expected because the swabs were held in a -20°C freezer, it was still necessary to see if the assay could detect peptides in older samples. Danielson et al. (2018) successfully detected blood and semen protein biomarkers in aged samples. The assay in this study was successful in identifying vaginal fluid, semen, and saliva in all the vaginal swabs. Most importantly, NGAL-1 was abundant in all three swabs tested.

The total human and male DNA quantitation was an important comparison measure to ensure DNA fractions contained male DNA and to determine which samples would have yielded sufficient DNA for STR typing. Overall, the Y-DNA was detected for all the male samples. The Y-DNA had a higher concentration for samples with semen in it compared to samples with just saliva, which is consistent with semen containing a larger amount of DNA than saliva (Lee & Ladd, 2001). Although, it did depend on the donor because semen Donor 1 had a much higher concentration than semen Donor 2. In the mixtures, the Y-chromosome was detected even when the semen and saliva were minor components, but due to the differences in DNA content only the semen samples had enough DNA and a male to DNA ratio sufficient for STR typing.

Conclusions

This assay successfully identified vaginal fluids, saliva, and semen by using signature peptides. SEMG-1 and SEMG-2 are robust and abundant in samples with semen. SMR3B-1 & 2 are abundant but not always detected as the minor component in a mixture. Finally, NGAL-1 was

robust and abundant, but NGAL-2 was not detected in the vaginal samples. Furthermore, this assay provides a peptide and DNA fraction in unison, so less amount of the sample is used, which is practical for forensic cases. The DNA fractions were quantified, and it was determined that downstream analysis could be successfully completed due to sufficient DNA for more samples. Overall, this assay is robust and sensitive and can be successfully used to identify body fluids.

Future Work

Future work for this assay is required to make it more specific for vaginal fluids and to improve the sensitivity for the saliva peptides. An additional vaginal marker to replace NGAL-2 is needed to provide a confirmatory peptide to NGAL-1 to confidently identify a vaginal fluid. Another NGAL peptide sequence, like the one Danielson et al. (2018) used, is suggested before trying a different protein. This can be done by using the scan feature with the mass spectrometer to discover the peptides under these specific experiment conditions. A high-resolution MS experiment is needed in order to complete a full scan. Also, an internal standard could be added so that a more accurate quantification can be achieved. An internal standard is constant in each sample and would help the analyst resolve problems or strange results. Specifically for this study, the internal standard could have helped explain any unusual concentrations found with viscous body fluids. As suggested by Merkley et al. (2019), it would be beneficial to use a search database when searching the samples for the peptides. This would create less bias, and with a large multiplex, the search database could assist in identifications that would be hard for an analyst. Another suggestion for further validation would be to try degraded and older samples to see if the peptides can still identify body fluids under these circumstances. Finally, it would be

essential to generate the STR profiles for the samples to see if it is possible to interpret the alleles for the samples in the extreme mixtures.

References

Albani, P. P., & Fleming, R. (2018). Novel messenger RNAs for body fluid identification. *Science and Justice*, *58*(2), 145–152. https://doi.org/10.1016/j.scijus.2017.09.002

Borges, E., Degiuli, A., Desrentes, S., Popielarz, C., Blum, L. J., & Marquette, C. A. (2017). Evaluation of the SPERM TRACKER™ for Semen Stains Localization on Fabrics. *Journal of Forensic Research*, *8*(3), 1000380. https://doi.org/10.4172/2157-7145.1000380

Browne, T., Concheiro-Guisan, M., & Prinz, M. (2019). Semi quantitative detection of signature peptides in body fluids by liquid chromatography tandem mass spectrometry (LC–MS/MS). *Forensic Science International: Genetics Supplement Series*, *7*(1), 208-210. https://doi.org/10.1016/j.fsigss.2019.09.080

Browne, T. N. (2020). *Semi-quantitative Detection of Signature Peptides in Body Fluids by Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS)*. [MS Thesis]. John Jay College of Criminal Justice. https://academicworks.cuny.edu/jj_etds/148/

Bruijns, B. B., Tiggelaar, R. M., & Gardeniers, H. (2018). The Extraction and Recovery Efficiency of Pure DNA for Different Types of Swabs. *Journal of Forensic Sciences*, *63*(5), 1492–1499. https://doi.org/10.1111/1556-4029.13837

Burkhart, J. M., Schumbrutzki, C., Wortelkamp, S., Sickmann, A., & Zahedi, R. P. (2012). Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. *Journal of Proteomics*, *75*(4), 1454–1462. https://doi.org/10.1016/j.jprot.2011.11.016

Butler, J. (2009). *Fundamentals of Forensic DNA Typing* (1st ed.). Academic Press.

Choung, C. M., Lee, J. W., Park, J. H., Kim, C. H., Park, H. C., & Lim, S. K. (2021). A forensic case study for body fluid identification using DNA methylation analysis. *Legal Medicine*, *51*, 101872. https://doi.org/10.1016/j.legalmed.2021.101872

Clark, C., Turiello, R., Cotton, R., & Landers, J. P. (2021). Analytical approaches to differential extraction for sexual assault evidence. In *Analytica Chimica Acta, 1141*(2), 230–245. https://doi.org/10.1016/j.aca.2020.07.059

Cochran, B., Lunday, D., & Miskevich, F. (2008). Kinetic Analysis of Amylase Using Quantitative Benedict's and Iodine Starch Reagent. *Journal of Chemical Education, 85*(3), 401-403. http://www.jce.divched.org/Journal/Issues/2008/Mar/abs401.html

Danielson, P. B., Legg, K., & Mckiernan, H. F. S. (2018). *Developmental Validation of a High-Specificity Multiplex Assay for Human Body Fluid Identification*. (NCJ No. 251914). National Criminal Justice Reference Service. https://www.ojp.gov/library/publications/developmental-validation-high-specificity-multiplex-assay-human-body-fluid

de Lamirande, E. (2007). Semenogelin, the main protein of the human semen coagulum, regulates sperm function. *Seminars in Thrombosis and Hemostasis*, *33*(1), 60–68. https://doi.org/10.1055/s-2006-958463

Frumkin, D., Wasserstrom, A., Budowle, B., & Davidson, A. (2011). DNA methylation-based forensic tissue identification. *Forensic Science International: Genetics*, *5*(5), 517–524. https://doi.org/10.1016/j.fsigen.2010.12.001

Hedman, J., Dalin, E., Rasmusson, B., & Ansell, R. (2011). Evaluation of amylase testing as a tool for saliva screening of crime scene trace swabs. *Forensic Science International: Genetics*, *5*(3), 194–198. https://doi.org/10.1016/j.fsigen.2010.03.003

Hickey, J., & Mayall, K. (2019). *ForensicGEM Sperm Lysis on the ZyGEM PDQeX: validation of a rapid sperm lysis protocol for Y screening using real time PCR*. MicroGem Application Note. https://microgembio.com/wp-content/uploads/2019/03/forensicGEM_YScreening_AppNote_20191007.pdf

Hinton-Sheley, P. (2021). *ATR-FTIR-An-Overview.* AZoLifeSciences. https://www.azolifesciences.com/article/What-is-ATR-FTIR.aspx

Hu, L., Huang, T., Liu, X.-J., & Cai, Y.-D. (2011). Predicting Protein Phenotypes Based on Protein-Protein Interaction Network. *PLoS ONE*, *6*(3), e17668. https://doi.org/10.1371/journal.pone.0017668

Ingold, S., Dørum, G., Hanson, E., Ballantyne, J., & Haas, C. (2020). Assigning forensic body fluids to donors in mixed body fluids by targeted RNA/DNA deep sequencing of coding region SNPs. *International Journal of Legal Medicine*, *134*(2), 473–485. https://doi.org/10.1007/s00414-020-02252-w

Juusola, J., & Ballantyne, J. (2005). Multiplex mRNA profiling for the identification of body fluids. *Forensic Science International*, *152*(1), 1–12. https://doi.org/10.1016/j.forsciint.2005.02.020

Kamanna, S., Henry, J., Voelcker, N., Linacre, A., & Kirkbride, K. P. (2017). “Bottom-up” in situ proteomic differentiation of human and non-human haemoglobins for forensic purposes by matrix-assisted laser desorption/ionization time-of-flight tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, *31*(22), 1927–1937. https://doi.org/10.1002/rcm.7986

Kranes, S., Sterling, S. A., Mason, K., Anex, D., Hart, B., Parker, G., & Prinz, M. (2017). Simultaneous DNA and protein extraction using trypsin. *Forensic Science International: Genetics Supplement Series*, *6*, e203–e204. https://doi.org/10.1016/j.fsigss.2017.09.081

Kuhar, N., Sil, S., & Umapathy, S. (2021). Potential of Raman spectroscopic techniques to study proteins. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *258*(5), 19712. https://doi.org/10.1016/j.saa.2021.119712

Laffan, Á., Sawyer, I., Quinones, I., & Daniel, B. (2011). Evaluation of semen presumptive tests for use at crime scenes. *Medicine, Science and the Law*, *51*(1), 11–17. https://doi.org/10.1258/msl.2010.010040

Laux, D. L. (2003). *Forensic Detection of Semen I. The Acid Phosphatase Test*. Ohio Bureau of Criminal Identification. http://courseresources.mit.usf.edu/other/forensics/ang6515_dewitt/readings/m4/Laux.%20Forensic%20Detection%20of%20Semen%20I.%20The%20AP%20Test.pdf

Lee, H. C., & Ladd, C. (2001). Preservation and Collection of Biological Evidence. *Croatian Medical Journal*, *42*(3), 225–228. https://pubmed.ncbi.nlm.nih.gov/11387627/

Légaré, C., Cloutier, F., Makosso-Kallyth, S., Laflamme, N., Jarvi, K., Tremblay, R. R., & Sullivan, R. (2013). Cysteine-rich secretory protein 1 in seminal plasma: potential biomarker for the distinction between obstructive and nonobstructive azoospermia. *Fertility and Sterility*, *100*(5), 1253–1260. https://doi.org/10.1016/j.fertnstert.2013.07.1984

Legg, K. M., Powell, R., Reisdorph, N., Reisdorph, R., & Danielson, P. B. (2014). Discovery of highly specific protein markers for the identification of biological stains. *ELECTROPHORESIS*, *35*(21–22), 3069–3078. https://doi.org/10.1002/elps.201400125

Legg, K. M., Powell, R., Reisdorph, N., Reisdorph, R., & Danielson, P. B. (2017). Verification of protein biomarker specificity for the identification of biological stains by quadrupole time-of-flight mass spectrometry. *Electrophoresis*, *38*(6), 833–845. https://doi.org/10.1002/elps.201600352

Liu, B., Yang, Q., Meng, H., Shao, C., Jiang, J., Xu, H., Sun, K., Zhou, Y., Yao, Y., Zhou, Z., Li, H., Shen, Y., Zhao, Z., Tang, Q., & Xie, J. (2020). Development of a multiplex system for the identification of forensically relevant body fluids. *Forensic Science International: Genetics*, *47*, 102312. https://doi.org/10.1016/j.fsigen.2020.102312

Ludeman, M. J., Zhong, C., Mulero, J. J., Lagacé, R. E., Hennessy, L. K., Short, M. L., & Wang, D. Y. (2018). Developmental validation of GlobalFiler™ PCR amplification kit: a 6-dye multiplex assay designed for amplification of casework samples. *International Journal of Legal Medicine*, *132*(6), 1555–1573. https://doi.org/10.1007/s00414-018-1817-5

Luyando, K. (2018). *The Y-Screen Assay: Validation and Evaluation of Subsequent STR Success*. John Jay College of Criminal Justice [MS Thesis]. John Jay College of Criminal Justice. https://academicworks.cuny.edu/jj_etds/85/

Lynch, C., & Fleming, R. (2021). RNA -based approaches for body fluid identification in forensic science. *WIREs Forensic Science*, *3*(4), 1-23. https://doi.org/10.1002/wfs2.1407

McKiernan, H. E., Danielson, P. B., Brown, C. O., Signaevsky, M., Westring, C. G., & Legg, K. M. (2021). Developmental validation of a multiplex proteomic assay for the identification of forensically relevant biological fluids. *Forensic Science International*, *326*, 110908. https://doi.org/10.1016/j.forsciint.2021.110908

Merkley, E. D., Wunschel, D. S., Wahl, K. L., & Jarman, K. H. (2019). Applications and challenges of forensic proteomics. *Forensic Science International 297*, 350–363. https://doi.org/10.1016/j.forsciint.2019.01.022

Mistek, E., Halámková, L., & Lednev, I. K. (2019). Phenotype profiling for forensic purposes: Nondestructive potentially on scene attenuated total reflection Fourier transform-infrared (ATR FT-IR) spectroscopy of bloodstains. *Forensic Chemistry*, *16*, 1-11. https://doi.org/10.1016/j.forc.2019.100176

Muro, C. K., Doty, K. C., de Souza Fernandes, L., & Lednev, I. K. (2016). Forensic body fluid identification and differentiation by Raman spectroscopy. *Forensic Chemistry*, *1*, 31–38. https://doi.org/10.1016/j.forc.2016.06.003

Orphanou, C. M. (2015). The detection and discrimination of human body fluids using ATR FT-IR spectroscopy. *Forensic Science International*, *252*, e10–e16. https://doi.org/10.1016/j.forsciint.2015.04.020

Peters, M. S., Brown, C. O., Mckiernan, H. E., & de Luca, D. J. (2020). *Front-End Fractionation of DNA and Proteins for the Simultaneous Genetic and Serological Analysis of Sexual Assault Kit Samples*. [Poster]. Abstract B66, Proceedings of the American Academy of Forensic Sciences – 72nd Annual Scientific Meeting, San Diego, February 17-22, 2020. p. 291.

Prinz, M., Boll, K., Baum, H., & Shaler, B. (1997). Multiplexing of Y chromosome specific STRs and performance for mixed samples. *Forensic Science International*, *85*(3), 209–218. https://doi.org/10.1016/S0379-0738(96)02096-8

Quinlan, A. (2021). The Rape Kit's Promise: Techno-optimism in the Fight Against the Backlog. *Science as Culture*, *30*(3), 440–464. https://doi.org/10.1080/09505431.2020.1846696

Scott, A., Jackson, K., Carter, M., Comeau, R., Layne, T., & Landers, J. (2019). Rapid sperm lysis and novel screening approach for human male DNA via colorimetric loop-mediated isothermal amplification. *Forensic Science International: Genetics*, *43*. 102139 https://doi.org/10.1016/j.fsigen.2019.102139

Skoog, D., Holler, J., & Crouch, S. (2007). *Principles of Instrumental Analysis* (6th ed.). Thomson Brooks/Cole.

The Human Protein Atlas. (n.d.). *The vagina-specific proteome.* https://www.proteinatlas.org/humanproteome/tissue/vagina

Thermo Fisher Scientific. (2015). *Novel Y-screen assay for sexual assault samples using the Quantifiler Trio DNA Quantification Kit.* Appliedbiosystems. https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/Application-Note-Y-screen-assay-for-sexual-assault.pdf

Tuma, R. (2003). MALDI TOF Mass Spectrometry Getting a Feel for How It Works. *Oncology Times, 25*(19), 16. doi:10.1097/01.COT.0000290986.00178.61

UniProt. (n.d.). *UniProtKB-P80188 (NGAL_Human).* https://www.uniprot.org/uniprot/P80188

van Nieuw Amerongen, A., Bolscher, J. G. M., & Veerman, E. C. I. (2004). Salivary Proteins: Protective and Diagnostic Value in Cariology? *Caries Research*, *38*(3), 247–253. https://doi.org/10.1159/000077762

Vitorino, R. (2018). Digging Deep into Peptidomics Applied to Body Fluids. *PROTEOMICS*, *18*(2), 1700401. https://doi.org/10.1002/pmic.201700401

Yang, H., Zhou, B., Deng, H., Prinz, M., & Siegel, D. (2013). Body fluid identification by mass spectrometry. *International Journal of Legal Medicine*, *127*(6), 1065–1077. https://doi.org/10.1007/s00414-013-0848-1

Zeffer, J. (2014). FBS07- Microscopic Examination of Spermatozoa by Christmas Tree Stain. District of Columbia Department of Forensic Sciences. https://dfs.dc.gov/sites/default/files/dc/sites/dfs/page_content/attachments/FBS07%20Sprm%20Search.pdf