Spring 6-2020

Semi-quantitative Detection of Signature Peptides in Body Fluids by Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS)

Tebah N. Browne
CUNY John Jay College, tebah.browne@jjay.cuny.edu

Semi-quantitative Detection of Signature Peptides in Body Fluids by Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS)

Tebah N Browne

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.

Semi-quantitative detection of signature peptides in body fluids by liquid chromatography tandem mass spectrometry (LC-MS/MS)

Tebah N Browne

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. Grace Axler-DiPerte

Table of Contents **Pages**

Acknowledgments

I would like to thank everyone that has motivated and encouraged me along my journey. I would like to thank God for the strength, good health, and passion that kept me going. I would like to thank my family and friends for constant moral support and love. I would like to thank Dr. Mechthild Prinz and Dr. Marta Concheiro-Guisan for guiding and challenging me, throughout this journey. Because of them, I have grown into a mentally stronger, wiser, and better student. I would also like to thank Dr. Grace Axler-DiPerte for her time and guidance. I would like to extend gratitude to my lab mates in Dr. Prinz's lab. Thank you, guys, for the laughs and love. Lastly, I would like to thank all the volunteers that participated in this study.

Abstract

This study covers a modified semi-quantitative approach for liquid chromatography mass spectrometry (LC-MS/MS) signature peptide detection for body fluid identification. Peptide concentrations were measured based on synthetic peptide standards. Samples were processed with varying trypsin digestion and purification protocols, including a three-hour trypsin digestion and Microcon membrane filtration. The Microcon filtration method generates polymerase chain reaction (PCR) compatible DNA and peptide fractions that can be analyzed without any further purification. Preliminary validation tests covered stains on different substrates, semen/ saliva mixtures, minimum sample volume, and repeatability. All signature peptides in the multiplex were present at different concentrations and varied amongst donors. Saliva peptides were detected at lower concentrations and had a higher minimum sample volume. Semen peptides could be detected even as a minor component in a mixture. All semen and saliva peptides were detected on the various substrates. DNA fractions did not show signs of degradation or PCR inhibition.

Introduction

The identification of biological fluids at a crime scene can aide to determine the suspect or victim, exonerate an innocent person, provide clues for reconstruction by connecting the sample donor to the scene of the criminal act, corroborate witness testimony of the sequence of events, and assist further investigation (Yang, Zhou, Deng, Prinz & Siegel, 2013; An, Shin, Yang, & Lee, 2012; Virkler & Lednev, 2009). Furthermore, blood stains can be an indicator of a physical assault, struggle, or murder, while semen and/or vaginal fluid can provide leads to a sexual act or assault (Virkler & Lednev, 2009). Blood, semen, saliva, vaginal fluid, urine, and sweat are the body fluids that are common in forensic casework (Yang et al., 2013; An et al., 2012; Virkler & Lednev, 2009). Having identified a body fluid routinely leads to further laboratory testing, such as DNA analysis, which can be the source of highly probative information.

Body fluid identification can be a difficult task because some biological stains are similar in appearance to other stains or not visible to the naked eye; hence, an absolute and objective confirmation is mandated (Virkler & Lednev, 2009). This fact becomes critical when there is a possible occurrence of mixed stains (Virkler & Lednev, 2009). A stain can contain various biological fluids from multiple donors. In these cases, it is important to identify all possible body fluids and link all possible donors. The project presented here focused on an alternative method for detecting semen and saliva, two body fluids of interest in sexual assaults; background on both biological substances is discussed in detail below.

Literature Review

Saliva Physiology

Human saliva is a clear and slightly acidic substance that secretes from exocrine glands (Humphrey & Williamson, 2001). It is a complex mixture of fluids released from the major and minor salivary glands and the mucosal gingival crevicular fluids that contain oral bacteria and food debris (Humphrey & Williamson, 2001; Proctor, 2016). The major salivary glands include the parotid glands, the submandibular and sublingual glands (Humphrey & Williamson, 2001). Minor glands that secrete saliva are located in the lower lip, palate, tongue, cheeks, and pharynx (Humphrey & Williamson, 2001). Saliva is largely composed of amylase, lysozyme and mucin (Virkler & Lednev, 2008; Zapata, Fernández de La Ossa & García-Ruiz, 2015).

Semen Physiology

Human semen is a complex mixture of fluids secreted from several glands. These fluids are partially fused during ejaculation, creating an initially non-homogenous mixture (Owen & Katz, 2005). The first portion of the ejaculate is secreted from the bulbourethral and periurethral glands (Owen & Katz, 2005). The second portion is secreted from the prostate, and the majority of the ejaculate comes from seminal vesicles (Owen & Katz, 2005). The majority of seminal fluid is secreted from the seminal vesicles and this solution is composed of alkaline, fructose, citrate, and prostaglandins (Owen & Katz, 2005; Revenig, Leung, & Hsiao, 2014). More information on the composition of seminal fluid is given in Table 1.

Table 1: Seminal fluid composition adapted from Owen & Katz, 2005; Revenig et al., 2014.

Organ	**Contribution (%)**	**Composition**
Seminal vesicles	65-75	Alkaline, fructose, citrate, prostaglandins
Prostate	15-30	Acid phosphatase, citric acid, serine proteases, zinc
Vas deferens	5-10	Spermatozoa
Bulbourethral and periurethral glands	1-5	Pre-ejaculate, galactose, mucous

Traditional Body Fluid Identification Techniques and Limitations

Body fluid tests are classified into two groups: 1) presumptive and 2) confirmatory. Presumptive tests are utilized to provide an indication of the identity of the fluid. Presumptive testing may result in false positives and are not specific to a single body fluid; therefore, a confirmatory test must be conducted to validate the identity of the body fluid in the stain (Zapata et al., 2015; Yang et al., 2013; An et al., 2012; Vincini, 2010; Virkler & Lednev, 2009). Unlike presumptive tests, confirmatory tests are specific and a positive result for a particular body fluid indicates the presence of that fluid.

Methods employed for body fluid identification include spectroscopy, microscopy, chemical tests, immunological tests, and protein catalytic activity tests (Zapata et al., 2015). Some of these methods are presumptive, while others are confirmatory. Generally, chemical or enzymatic techniques are presumptive, while microscopic or immunological techniques are confirmatory (Harbison & Fleming, 2016). Catalytic, enzymatic, and immunological methods have low specificity, are destructive, not sensitive, and may interfere with other tests (Yang et al., 2013; An et al., 2012).

Spectroscopy

Alternate Light Source

The use of an alternate light source (ALS) is a straightforward, noninvasive, and nondestructive spectroscopic technique that's often used to analyze body fluids, fingerprints, and other trace evidence (Miranda, Prado, Delwing, & Daruge, 2014). It is routine to analyze an area using ALS before the application of any reagent (Miranda et al., 2014). ALS takes advantage of the light absorptive or photoluminescent properties of some substances when they are examined at a particular wavelength (Viner, Kagan, & Johnson, 2014). For semen, the excitation spectrum is broad ranging from 300 to 480 nm (Vandenberg & Van Oorschot, 2006). Nonetheless, semen was observed to undergo optimal visualization at an excitation wavelength of 455 nm, depending on the substrate (Sheppard, Cassella, Fieldhouse, & King, 2017). It was observed that saliva can be optimally visualized at 470 nm, on a cotton substrate (Tay et al., 2020). A summary of the light source devices that have been employed for semen and saliva identification can be found in Table 2. False positive results may occur with bacitracin zinc, barrier cream, hand cream, castile soap, Surgilube® lubricant, toothpaste and A&D ointment (Zapata et al., 2015).

Table 2: Various ALS techniques utilized to detect semen and saliva.

Light Source Device	Wavelength Range (nm)	Semen Detection	Saliva Detection	Source
Wood's Lamp	320-400	Yes	No	(Nelson & Santucci, 2002; An et al., 2015; Virkler & Lednev, 2009)
Blue- maxx™ BM500	390-500	Yes	No	(Nelson & Santucci, 2002; An et al., 2015; Virkler & Lednev, 2009)
Poliray	450	Yes	No	(Lincoln et al., 2006)
Polilight	310-650	Yes	Yes	(Tay et al., 2020; Vandenberg & Van Oorschot, 2006; An et al., 2015)
Lumatec® Superlight 400	320-700	Yes	Yes	(An et al., 2015; Virkler & Lednev, 2009; Zapata et al., 2015)

SEM-EDX

Electron microscopy is used to create high resolution imagery for biological and nonbiological substances. Scanning electron microscopy (SEM) is the most common type of electron microscopy. With SEM, the microscopic structure of an item is examined by scanning its surface using a high resolution and a large depth of field (Leng, 2013). SEM can be used to visualize spermatozoa (Nussdorfer, Cilenšek, Zorn, & Petrovič, 2018). When coupled with an energy dispersive X-ray (EDX) microanalysis detector, characteristic elemental components of a body fluid can be detected (Lászik et al., 1999). This method can be used to detect multiple body fluids, including semen and saliva (Quinton, 1978). Detection of zinc, sodium, phosphorus, sulfur, chlorine, potassium, calcium, etc., is used to indicate the presence of a semen or saliva and other body fluids (Lászik et al., 1999; Virkler & Lednev, 2009). The quantity of each element varies from body fluid to body fluid. The identification and distinction of an unknown stain will rely on the ratio of the elements (Virkler & Lednev, 2009). This method cannot be used for body fluid mixture analysis. This method requires a clean surface for optimal results, and it

experiences interference from the stain's substrate (Nussdorfer et al., 2018; Virkler & Lednev, 2009).

Choline Detection for Semen

Chemical and enzymatic tests used to determine the presence of semen and saliva are based on their composition (Harbison & Fleming, 2016). Choline is an essential nutrient that can be found in human semen at a concentration of 0.9 – 1.4 mg/ml (Takatori, Tomii & Tanaka, 1981). Choline is used as a biomarker for semen. The Florence test is used to indicate the presence of choline. The Florence test is a microchemical crystal test that utilizes a solution of iodine and potassium to form brown needle-like crystals in the presence of choline (Virkler & Lednev, 2009; Zapata et al., 2015). Due to the low sensitivity of this method, false negative results are common (Virkler & Lednev, 2009). Normal levels of choline in other body fluids, including vaginal fluid, and semen from other animals cannot be detected utilizing this method (Takatori et al., 1981; Virkler & Lednev, 2009).

A chemiluminescent test can also be used to detect choline (Virkler & Lednev, 2009; Zapata et al., 2015). This method involves the use of a choline oxidase and luminol solution. Luminol combined with choline oxidase becomes luminescent in the presence of choline. Isotachophoresis, a type of electrophoresis that implements separation based on ionic mobility, is also used to detect choline on a stain (Virkler & Lednev, 2009). This method produces no false positives, can be used for old stains, and can detect semen from vaginal swabs collected from deceased victims (Virkler & Lednev, 2009).

Acid phosphatase Detection for Semen

Acid phosphatase is an enzyme and major component of semen. It is used as a biomarker for semen. The traditional acid phosphate test involves a reaction between the semen stain, or an extract placed on a filter paper, a chemical substrate—monophenolic phosphoric acid or its derivatives—and a pH 5 acetate buffer (Raju & Iyengar, 1964). The acid phosphatase hydrolyzes the substrate, forming phenol and phosphate ions. The phenol formed reacts with a diazonium salt and produces a colored compound (Raju & Iyengar, 1964; Zapata et al., 2015).

Amylase Detection for Saliva

High levels of alpha-amylase 1 can be found in human saliva; hence, it is utilized for saliva detection. Amylase catalyzes the digestion of starch. Therefore, a starch-iodine test is used. Starch reacts with iodine to form a blue compound, but in the presence of salivary amylase, this color change does not occur because the amylase consumes the starch (Zapata et al., 2015; Hedman, Dalin, Rasmusson & Ansell, 2011; Virkler & Lednev, 2009). This test is not specific for saliva because it can provide positive results for blood and semen (Virkler & Lednev, 2009). Another amylase detection method involves the Phadebas® test reagent, which is largely composed of procion red amylopectin (Virkler & Lednev, 2009). Phadebas® takes advantage of amylase's starch digesting activity and includes a starch complex with an attached color molecule, which is released when the amylase breaks down the starch (Hedman, Gustavsson, & Ansell, 2008). This method can detect saliva that has been diluted 1:128; however, it is positive for urine, feces, face lotion, hand cream, and washing powders (Virkler & Lednev, 2009).

Prostate Specific Antigen Detection for Semen

Prostate specific antigen (PSA) or P30 is a glycoprotein secreted into the seminal plasma (Hochmeister et al., 1999). It is a biomarker for detecting human semen and is effective for semen deposited by individuals that are vasectomized or azoospermic (Hochmeister et al., 1999). Techniques used to detect PSA include rocket immunoelectrophoresis, enzyme-linked immunosorbent assay (ELISA), Ouchterlony double diffusion, radial immunodiffusion, crossover electrophoresis and immunochromatographic test strip (such as ABA card and Biosign® PSA) (An et al., 2015; Hochmeister et al., 1999; Zapata et al., 2015). All of these methods depend on the interaction of PSA and its antibody and they are used as confirmation (Virkler & Lednev, 2009; Zapata et al., 2015). Published works discuss the presence of PSA at low levels in other body fluids, such as urine and breast milk (An et al., 2015).

Microscopic Visualization of Spermatozoa

Microscopic visualization is used to detect spermatozoa by Christmas tree staining. Christmas tree staining involves the use of a nuclear fast red dye and picroindigocarmine (Bell, 2012). The nuclear fast red dye will turn the head red, while the picroindigocarmine will turn the tail green (Ispan, 2018; Virkler & Lednev, 2009; Zapata et al., 2015). Another method includes the addition of proteinase K, which denatures the epithelial cells and makes the head more vibrant (Virkler & Lednev, 2009). Christmas tree staining is limited in the fact that the semen donor must have sperm present; therefore, it does not work for donors that have been vasectomized or are azoospermic (Virkler & Lednev, 2009). Enhanced microscopy utilizes human spermatozoa specific antibodies coupled to a fluorescent dye (Miller et al., 2011). This very selective staining technique can be combined with laser microdissection for separating sperm cells from vaginal epithelial cells (Miller et al., 2011).

New Body Fluid Techniques

Raman

Raman spectroscopy is an analytical chemistry method that is rapidly gaining popularity in the forensic discipline (Schlagetter, Kammrath, & Glynn, 2017). Raman spectroscopy is a nondestructive method that is based on the inelastic scattering of a low-intensity laser light when it comes into contact with a sample (Harbison & Fleming, 2016; Virkler & Lednev, 2008). This technique requires little to no preparation and only picograms or femtoliters of the sample is needed (Virkler & Lednev, 2008). Raman spectroscopy shows little water inference, and therefore, can be used to analyze body fluids and their stains (Virkler & Lednev, 2008). The resultant Raman spectra are composed of multiple narrow bands that create a unique vibrational signature of the molecular structure of each body fluid (Harbison & Fleming, 2016; Virkler & Lednev, 2008). Body fluid analysis using Raman spectroscopy requires complex statistical analysis due to the heterogeneous nature of dry stains and variation among individuals (Harbison & Fleming, 2016). Unique vibrational Raman spectra have been determined for various body fluids, including semen and saliva (Sikirzhytski, Virkler, & Lednev, 2010).

X-ray fluorescence

X-ray fluorescence (XRF) is an imaging technique that is being implemented for forensic applications. XRF has been successfully utilized for the detection of semen (Zapata et al., 2015). XRF imaging involves the application of a micro-focused X-ray beam incident to different parts of the sample to develop an image (Pushie, Pickering, Korbas, Hackett, & George, 2014). The XRF is monitored by an energy dispersive detector (Pushie et al., 2014). National Aeronautics and Space Administration (NASA) technology is being used to create a portable XRF devise that can be used at the crime scene to detect abundant components of semen, as well as blood (Carr,

2009). This method is useful because it is nondestructive; however, it does not produce confirmatory results due to its lack of specificity (Carr, 2009).

Nuclear magnetic resonance

Nuclear magnetic resonance is an emerging analytical tool for identification and quantification of different metabolites, in body fluids and other complex mixtures, without the need for laborious sample preparation (Rai & Sinha, 2012). The metabolic composition of each body fluid yields a unique spectrum that can be used to distinguish one from the other (Harbison & Fleming, 2016). This method requires a series of complex statistical processes, considering donor to donor variability, in order to interpret the characteristic NMR profile produced for each body fluid (Harbison & Fleming, 2016). Nonetheless, NMR spectroscopy is advantageous because it doesn't focus on one chemical property of the body fluid but rather the metabolite profile as a whole, and therefore, gives rise to a 'metabolite fingerprint' (Scano et al., 2013). The unique spectra produced can be compared to spectra of samples from unknown origin for identification purposes and mixture analysis (Harbison & Fleming, 2016).

mRNA

Messenger ribonucleic acid (mRNA) profiling is a new method that has been investigated as an alternative to traditional body fluid identification assays (An et al., 2015; Harbison & Fleming, 2016; Ingold et al., 2018; Richard et al., 2012; Zapata et al., 2015). mRNA profiling depends on the individual expression of mRNAs that occurs due to the different mRNA sequences in different body fluid cell types (Ingold et al., 2018; Park et al., 2013; Richard et al., 2012; Sijen, 2015; Zapata et al., 2015). Targeting a mRNA transcript depends on the amount of transcript present, as well as the stability of each transcript in the body fluid cell; hence, different

mRNA markers have different sensitivities (Harbison & Fleming, 2016). mRNA profiling and traditional presumptive tests have similar limits of detection (Harbison & Fleming, 2016).

RNA is less stable, compared to DNA; however, studies have shown that mRNA is stable in body fluid stains found on various surfaces (An et al., 2015; Harbison & Fleming, 2016). A big advantage of mRNA profiling is that mRNA and DNA can be simultaneously extracted from the same stain (An et al., 2015; Harbison & Fleming, 2016; Lindenbergh, Maaskant, & Sijen, 2013; Richard et al., 2012). However, mRNA profiling is arduous and more expensive than other body fluid identification methods (Lindenbergh et al., 2013).

Reverse transcriptase polymerase chain reaction (RT-PCR) is a novel and sensitive method used to detect low-abundance mRNA isolated from biological stains (Harbison & Fleming, 2016). The method was first introduced in 2005 by Juusola and Ballantyne (2005; Harbison & Fleming, 2016). They proposed a RT-PCR based multiplex assay for the identification of body fluids common to forensic casework (An et al., 2015; Juusola & Ballantyne, 2005). The method exhibited sufficient sensitivity and detection was successful for 200 picograms to 12 nanograms of RNA (An et al., 2015; Juusola & Ballantyne, 2005). In 2007, Juusola and Ballantyne developed a multiplex utilizing a quantitative RT-PCR (qRT-PCR) assay (2007; An et al., 2015). The qRT-PCR assay calculates the difference between a reference housekeeping gene (ΔCT) and the desired RNA transcript (Harbison & Fleming, 2016; Juusola & Ballantyne, 2007) . This method uses numerical thresholds because it is highly sensitive (Harbison & Fleming, 2016; Juusola & Ballantyne, 2007). A problem that is commonly faced with this method is that there is a restriction on the dyes that can be used, and therefore, there is a limited number of markers that can be targeted in a single reaction (Harbison & Fleming, 2016).

High-resolution melting analysis has been proposed as a solution for the limitation (Harbison & Fleming, 2016).

In 2018, European Forensic Genetics (EUROFORGEN) and European DNA Profiling Group (EDNAP) worked on a collaborative method for mRNA-targeted body fluid identification utilizing massively parallel sequencing (MPS) (Ingold et al., 2018). With this sequencing method, a multiplex of markers can be detected in a single reaction, saving time and sample consumption. There is also potential for quantitative analysis due to a larger dynamic range, and the resolution of mixtures due to the ability to identify mRNA sequence variation (Ingold et al., 2018).

microRNA

mRNA markers can be used for successful and specific body fluid identification and have been proven to be stable for a long period of time. However, environmental conditions, such as humidity and hot climate, can influence mRNA stability and lead to degradation (An et al., 2015). Micro RNA (miRNA) markers have proposed as an alternative tool because they are smaller RNA molecules and less susceptible to degradation (An et al., 2015). miRNAs are non-coding molecules that are about 18-22 nucleotides long and regulate post-transcriptional gene expression (An et al., 2015). Body fluid identification utilizing mRNA relies on the fact that many miRNA molecules are tissue-specific; therefore, these molecules can be body fluid specific (An et al., 2015).

The first assay using miRNA as a biomarker was conducted by Hanson, Lubenow, & Ballantyne (2009; An et al., 2015; Van Steendam et al., 2013). In their study, they examined miRNA expression in body fluids common to forensic casework (Hanson et al., 2009). miRNA analyses are conducted using the same techniques as mRNA assays, i.e. RT-PCR (Sijen, 2015).

Creating a multiplex assay is very difficult because there are a small number of dyes available for miRNA-based RT-PCR assays (Van Steendam et al., 2013).

DNA methylation

DNA methylation is also being investigated for body fluid identification. DNA methylation is an epigenetic modification that consists of the addition of a methyl group to the 5' position of cytosine in CpG dinucleotide (Yuan, 2014; An et al., 2015; Harbison & Fleming, 2016). Whole-epigenome studies have shown that DNA undergoes tissue specific methylation patterns (An et al., 2015; Van Steendam et al., 2013). The analysis and detection of the tissue-specific methylation patterns can be applied to body fluid identification, in which body fluid specific cell methylation patterns are examined (An et al., 2015; Frumkin, Wasserman, Budowle, & Davidson, 2011; Van Steendam et al., 2013). The first reported method was developed by Frumkin et al. (An et al., 2015; Van Steendam et al., 2013). The detection of the methylation is conducted by using methylation-dependent restriction enzyme followed by PCR and/or bisulfite sequencing (Harbison & Fleming, 2016). This method can be combined with current STR typing assays to reduce sample consumption and time (Frumkin et al., 2011; Van Steendam et al., 2013). Nonetheless, in order to make DNA methylation applicable for casework, more markers need to be identified (An et al., 2015).

Proteomics

A proteomics-based method for body fluid identification has proven to be very promising because it is based on the detection of various high-specificity protein biomarkers (Legg et al., 2017). Unlike DNA, proteins can be found in abundance within the cell. They are also more stable than RNA. When creating a body fluid identification assay, the main goal would be to identify signature proteins, which are unique to the specific body fluid. Many studies have been

done to determine candidate biomarkers (Kamanna et al., 2016; Legg et al., 2014, 2017; Van Steendam et al., 2013; Yang et al., 2013). Table 3 presents a list of the most common protein biomarkers identified in published studies.

Identification of proteins is conducted using a separation method—high performance liquid chromatography—followed by an analytical method, such as mass spectrometry (Aebersold & Mann, 2003). There is a size limitation for mass spectrometry, so protein is usually digested to peptide fragments (Aebersold & Mann, 2003). Due to advances in mass spectrometry, proteins in complex samples can be identified, categorized, and quantified at high sensitivity (Domon & Aebersold, 2006). Additionally, mass spectrometry offers the advantage of multiplex analysis with high accuracy and specificity (Legg et al., 2017).

Table 3: Suggested protein biomarkers for semen and saliva identification*

Body Fluid	Expected Protein Biomarkers (Gene)	SwissProt/UniProt Number	IPI number
Semen	Semenogelin 1 (SEMG1)	P04279	IPI00023020
	Semenogelin 2 (SEMG2)	Q02383	IPI00025415
	Prostate-specific antigen (KLK3)	P07288	IPI00010858
	Prostatic acid phosphatase (ACPP)	P15309	IPI00396434
Saliva	Alpha-amylase 1 (AMY1A)	P04745	IPI00300786
	Histatin-1 (HTN1)	P15515	IPI00012024
	Cystatin_SA (CST2)	P09228	IPI00013382
	Cystatin_D (CST5)	P28325	IPI00002851
	Statherin	P02808	IPI00022990
	Submaxillary gland androgen regulated protein 3B (SMR3B_HUMAN)	P02814	IPI00023011

* information based on Kamanna et al., 2016; Legg et al., 2014, 2017; Van Steendam et al., 2013; Yang et al., 2013

Liquid Chromatography Mass Spectrometry

Reverse-Phase High Performance Liquid Chromatography

High-performance liquid chromatography (HPLC) is a separation method based on polarity (Aguilar, 2004). In reversed phase HPLC, the stationary phase is nonpolar, whereas the mobile phase is polar. Nonpolar molecules in the mobile phase will adsorb to the stationary phase, while polar molecules will have a stronger affinity for the mobile phase, and therefore, will pass through the column and elute first (Aguilar, 2004). Isocratic elution can occur, where the composition of the mobile phase is constant throughout the run. A gradient elution may also occur, in which the composition of the mobile phase varies increasing its elution strength

throughout the run. RP-HPLC is a very effective separation method for peptides and proteins and is commonly used in combination with mass spectrometry for body fluid identification (Domon & Aebersold, 2006; Legg et al., 2014, 2017; Van Steendam et al., 2013).

Mass Spectrometry

Mass spectrometry is a technique used to produce qualitative and quantitative data for an ionized analyte (Ho et al., 2015). The analyte is ionized by passing through an ionization source to acquire a negative or positive charge. The ions are then introduced to a mass analyzer and finally to the detector according to their mass-to-charge ratios (Ho et al., 2015).

Ionization

The development of ionization techniques such as electrospray and matrix assisted laser desorption ionization revolutionized mass spectrometry and made it applicable for the analysis of large biological molecules, such as proteins (Singhal, Kumar, Kanaujia, & Virdi, 2015). Both ionization techniques utilize soft ionization to convert peptides into ions by adding or removing one or more protons (Singhal et al., 2015). In soft ionization, ion formation does not result in significant fragmentation (Singhal et al., 2015).

Matrix assisted laser desorption/ionization (MALDI) is implemented for biological samples that cannot be analyzed using hard ionization methods. Initially, a solution of the sample and the matrix, an organic energy-absorbent compound, is created. The sample is ionized in the matrix with a laser beam. Desorption and ionization occur with the laser to produce protonated ions of the analytes (Singhal, Kumar, Kanaujia & Virdi 2015).

Electrospray ionization (ESI) is a soft ionization method used to produce gas phase ions without fragmentation (Banerjee & Mazumdar, 2012). In ESI, a continuous stream of the sample is passed through a capillary tube, which is maintained at a high potential (Ho et al., 2015). The

sample is nebulized using a flow of nitrogen gas. Ionization of the sample occurs in three steps: 1) dispersal of a mist composed of charged droplets, 2) evaporation of the solvent, and 3) ejection of the ion from the highly charged droplets (Ho et al., 2015). The charged droplets, produced at the end of the electrospray tip, travel along a pressure and potential gradient, towards the mass analyzer (Ho et al., 2015). With the assistance of a stream of nitrogen gas, the charged droplets undergo evaporation and decrease in size (Ho et al., 2015). This leads the droplets to increase in surface charge density and decrease in radius (Ho et al., 2015). When the electric field strength reaches a critical point, the ions at the surface of the droplets are ejected into the gaseous phase (Ho et al., 2015). The ions that are emitted are then accelerated into the mass analyzer (Ho et al., 2015).

Mass Analyzers

The mass analyzer is the component of a mass spectrometer that separates the ions based on mass-to-charge (m/z) ratios. There are two categories of mass analyzers: low and high resolution. Mass resolution describes the separation that occurs in a mass spectrum (Gross, 2011). Mass resolution is represented as the smallest difference in mass-to-charge that can be fully separated for a given signal (Gross, 2011). Higher resolution results in narrower and sharper peaks in the spectrum. Mass resolution is also a reflection of mass resolving power (Marshall & Hendrickson, 2008). Mass resolving power is expressed as a ratio between the mass (m) and the width of the peak at a height that is a certain fraction of the maximum peak height (dm), see Figure 1 (Macherone, 2013). Mass resolution and mass resolving power are directly proportional. Higher resolution instruments have higher resolving power.

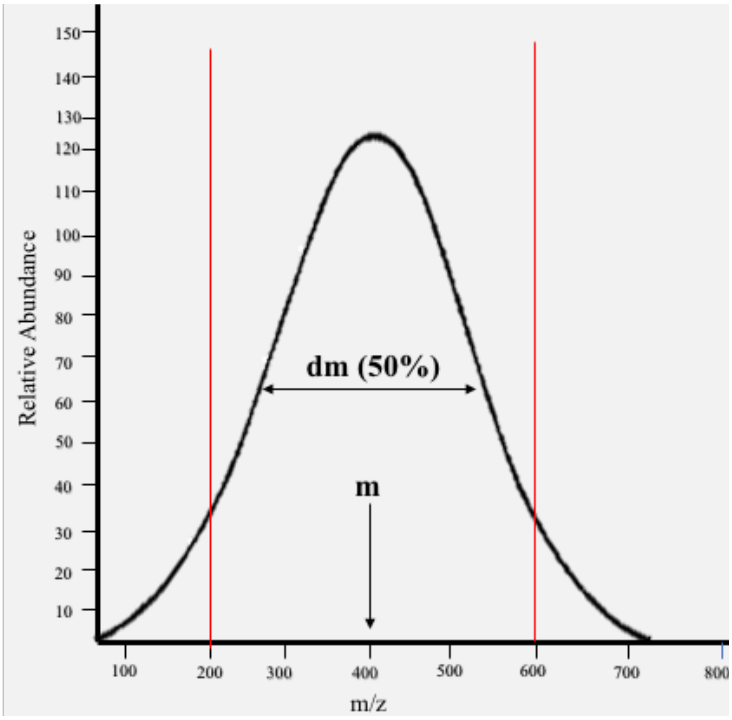

Figure 1: *A diagram displaying mass resolving power factors for determining resolution of a peak at a given ion, where m is the m/z 400 and dm is 50% of the maximum peak.*

Time-of-flight (ToF) is a high-resolution mass analyzer with a resolving power that usually exceeds 12,000 (Domon & Aebersold, 2006). ToF analyzers are usually used for proteomic, with MALDI as the ionization source. For ToF analyzers, ions are separated by the time they take to travel through a flight tube, which is under a vacuum with an applied fixed voltage (Domon & Aebersold, 2006). This technique operates in scan mode in which the full mass range is examined, and all mass-to-charge ratios are determined.

The quadrupole is a low-resolution mass analyzer that acts as a mass filter. The system is composed of four metal rods that are parallel and at equidistance (Ho et al., 2015). The rods have fixed direct current (DC) and alternating radiofrequency voltages applied to them (Ho et al., 2015). An electrical field is produced, and the ions are focused and passed through the middle of the quadrupole rod in an oscillating motion (Ho et al., 2015). The motion or oscillation of the ions will depend on the electric field. Only ions of a specified *m/z* will be in resonance with the

electrical field and permitted to pass through. This system can function in scan mode and selected-ion monitoring (SIM) mode. In SIM mode, only ions of a pre-determined *m*/*z* can pass through the quadrupole.

Tandem Mass Spectrometry

Tandem mass spectrometry (MS/MS) utilizes two mass analyzers. The most common system employs two quadrupoles with a collision cell separating both quadrupoles (Mittal, 2015). This technique is known as triple quadrupole (QqQ). The first quadrupole is used to determine or target the precursor ion. Next, the precursor ions migrate through the collision cell, where they are bombarded with an inert gas and undergo fragmentation to produce product ions. The product ions are then determined or targeted by the second quadrupole. This system can undergo a product scan when the first quadrupole is static, and the second quadrupole is in scanning mode (Ho et al., 2015). A precursor scan can be conducted when the first quadrupole is in scanning mode and the second quadrupole is static (i.e. in SIM mode) (Ho et al., 2015). The system can also undergo a neutral loss scan, in which both quadrupoles are in scanning mode. Multiple reaction monitoring (MRM) can be utilized when both quadrupoles are static (Ho et al., 2015).

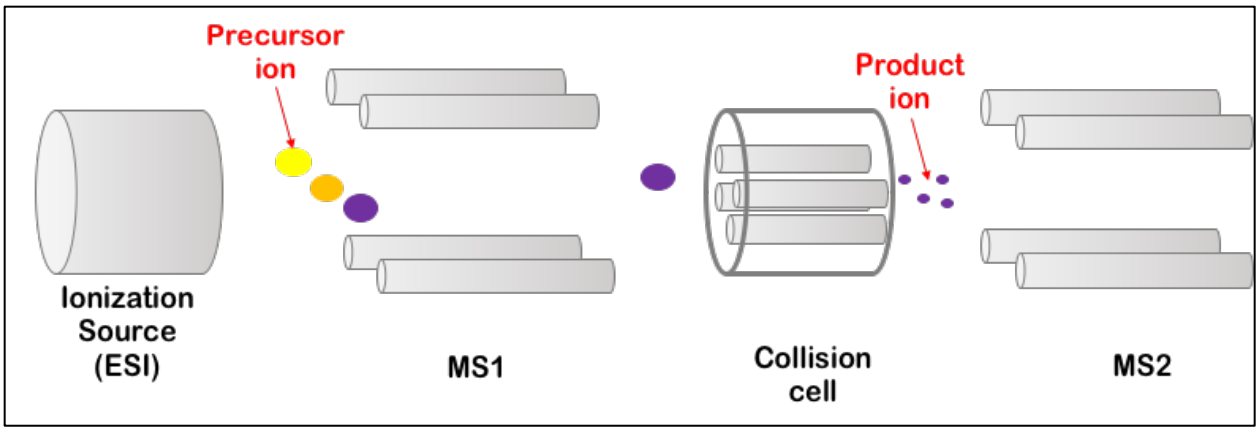

Figure 2: *Schematic diagram of a triple quadrupole system using electrospray ionization. This MS consists of two quadrupole mass analyzers, which act as mass filters and a radiofrequency only quadrupole, which causes fragmentation of the analyte through interaction with a collision gas.*

Aim of Research

The aim of this study is to develop a peptide-based method using LC-MS/MS for body fluid identification. We hypothesize that tandem low-resolution mass spectrometry can be implemented for reliable, semi-quantitative signature peptide detection. The specific goals are 1) to establish an assay using tandem low-resolution mass spectrometry in MRM mode, 2) to simultaneously analyze DNA, and 3) to validate the established assay. The MRM parameters for each peptide will be determined by analyzing synthetic peptides. Those parameters will then be applied to authentic samples. Method validation will include analyses to determine the peptides' minimum sample volume, repeatability of the peptide signal, assay compatibility with different substrates, and how the peptides can be detected in a mixture.

Methods and Materials

Sample Collection and Preparation

Semen was purchased from a commercial provider (Lee Biosolutions, Maryland Heights, MO). Saliva was collected from volunteers after CUNY IRB approval (#2017-0080). For the collection process, the volunteers were required to rinse their mouths only if they ate an hour prior to the collection process. Each volunteer was asked to provide at least 1 mL of saliva in 15 mL Eppendorf plastic centrifuge tubes (Hamburg, Germany).

Five µL of semen and saliva were spotted on swabs—cotton (Puritan), polyester (Fisherbrand, Thermo Fisher Scientific, Waltham, MA) and CEP (Fitzco, Spring Park, MN)—and white cotton t-shirt, black cotton t-shirt, blue denim, tissue paper, and polyester t-shirt. The tips of each swab were cut off. The other substrates were cut into circular pieces with diameters of approximately 0.5 cm. Cotton swabs and t-shirts stains containing 10 µL of semen, produced in 2014, were cut in half to estimate for 5 µL.

Liquid semen and saliva were used to create mixed stains on cotton swabs and t-shirts. To make mixed solutions, semen and saliva were added at the following ratios (v/v): 1 to 1, 1 to 2, 2 to 1, 1 to 5, and 5 to 1. Five µL of each mixture was placed on the appropriate substrate.

Standard Preparation

Saliva and semen signature peptides were chosen based on Legg et al. (2017), the Universal Protein Resource (UniProt), and PROWL Protein Info. Five peptides for saliva and semen were analyzed. For semen, semenogelin-1 was chosen as the biomarker and two signature peptides were targeted: SEMG1-pep1 and SEMG1-pep2. For saliva, submaxillary gland androgen regulated protein 3B (SMR3B) and histatin-1 were the biomarkers. One peptide was targeted for histatin-1, HIST 1, and two peptides were targeted for SMR3B, SMR3B-pep1 and

SMR3B-pep2. See Table 4 for the peptides' sequences and additional information. Synthetic peptides were obtained from Genscript (Piscataway, NJ). To prepare stock solutions of each peptide, 1 mg of the peptide was added to 1000 µL of the appropriate solvent. All of the synthetic peptides were dissolved in irradiated reversed osmosis water with the exception of SEMG1-pep2, which was dissolved in 0.1M ammonium bicarbonate.

Synthetic peptide standards were prepared by mixing all five peptides at a concentration of 50 nmol/mL and creating five 1:10 serial dilution steps down to 0.005 nmol/mL. A study to determine the cutoff for each synthetic peptide was conducted, in which mixtures containing all five peptides were prepared at the following concentrations: 0.001, 0.0025, 0.005, 0.01, 0.025, 0.05, and 0.1 nmol/mL.

Table 4: Semen and saliva signature peptides with corresponding sequences, charges, and stock solvents.

Tissue	Protein	Peptide Name	Sequence	Overall Charge	Solvent
Semen	Semenogelin-1	SEMG1-pep1	QGGSQSSYVLQTEELVANK	-1	Irradiated H_20
		SEMG1-pep2	DIFSTQDELLVYNK	-2	Irradiated H_20
Saliva	Histatin-1	HIST 1	EFPFYGDYGSNYLYDN	-3	Irradiated H_20
	submaxillary gland androgen regulated protein 3B	SMR3B-pep1	GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR	+1	Irradiated H_20
		SMR3B-pep2	IPPPPPAPYGPGIFPPPPPQP	Neutral	NH_4HCO_3

Trypsin Digestion

The incubation buffer consisted of 1% ProteaseMax™ surfactant (Promega, Madison, WI), 0.5 dithiothreitol (DTT) (Promega, Madison, WI), and freshly prepared 50 mM ammonium bicarbonate. ProteaseMax™ surfactant is a trypsin enhancer used to ensure maximum efficiency of the protease. DTT is used to disrupt the cell membrane and the protein disulfide bonds. The ammonium bicarbonate, which is commonly used to buffer trypsin digestions, must be fresh, (prepared within three days). The liquid (5 µL) or stain samples were placed in 100 µL of the buffer and incubated for 20 min at 56°C, while shaking at 1400 rpm. Three µL of iodoacetamide (IAA) (Sigma Aldrich, St. Louis, MO) was added to the solution. IAA is an alkylation base used to maintain the separation of the disulfide bonds. This mixture was then placed in the dark for 30 min at room temperature. One µL of DTT was placed into the solution to deactivate IAA and prevent overalkylation. One µL of a 0.1 µg/µL solution of trypsin (Promega, Madison, WI) was placed into the solution, which was incubated for 3 h at 37°C and 1400 rpm. Trypsin is an enzyme used to cleave other proteins after a lysine and arginine residue, if not followed by a proline residue. After the digestion, the trypsin was deactivated by incubating in a stationary heat block at 99°C for 10 min and cooling at 4°C for 10 min. See Table 5 for the main modifications that were tested.

To remove the substrates, each digest was decanted into a spin basket placed inside a dolphin tube (MIDScientific, Valley Park, MO) and spun down for 5 min at 1500 rcf. The samples were stored at -20°C.

Table 5: Trypsin digestion modifications that were tested and changed from Overnight with IAA protocol.

Removal of Alkylation Step	The addition of IAA and the extra DTT was omitted to shorten the digestion process.
Incubation Time Variation	The solutions were incubated in the trypsin digestion buffer overnight, approximately 16 h or for 3 h, at 37C and 1400 rpm.

HyperSep™ SpinTip Extraction and Concentration Method

The HyperSep™ C18 SpinTip (Thermo Fisher Scientific, Waltham, MA) was conditioned by dispersing the releasing (0.1% formic acid and 60% acetonitrile) and the binding (0.1% formic acid) solutions through the column using centrifugation at 1500 rcf. The substrate-free digests were passed through the columns using air pressure via a syringe. The columns with the immobilized peptides were washed using the binding solution. The peptides were released using the releasing solution. The final solution was evaporated using a TurboVap® LV Automated Evaporated System (Biotage, Charlotte, NC) and reconstituted in 50 L. To determine the best solvent for the instrumental analysis, C18 extractions were done utilizing water, 60%, 40%, 20% acetonitrile, and the digestion buffer solution as the reconstitution solvents.

Nanofiltration Method Purification Method

The digests were filtered with nano|Filter Vial PVDF 0.2µm (Thomson Instrument Company, Oceanside, CA). One hundred µL of the digest was placed into the vial. The screw-capped component of the nano|Filter was positioned into the vial and pressed down using hand pressure to slowly deposit the digest through the filter.

Peptide and DNA Co-Extraction Purification Method

The substrate-free digests were placed on Microcon® MW100 DNA Fast Flow tubes membrane units (Millipore, Burlington, MA) and centrifuged for 20 min at 500 rcf. The initial flow through (peptide fraction) was stored at -20ºC. The DNA fraction remains on top of the membrane. To recover the DNA fraction, 20 µL of sterile nuclease free water was placed into the inverted membrane unit, which was centrifuged for 3 min at 1000 rcf. The DNA fraction was immediately stored at 20ºC.

DNA Quantification

Extracted DNA was quantified using the Applied Biosystems® Quantifiler™ Trio DNA Quantification kit (Life Technologies, Thermo Fisher Scientific, Waltham, MA) on the QuantStudio™ 5 Real-Time PCR Systems (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA) utilizing a virtual curve. The virtual curve was prepared using the following standards and concentrations (Table 6).

Table 6: Quantifiler™ THP DNA quantification standard concentrations.

Standard	**Concentration (ng/µL)**
Standard 1	50
Standard 2	5
Standard 3	0.5
Standard 4	0.05
Standard 5	0.005

The standards were prepared through serial dilutions as instructed by the manufacturer. A master mix was prepared consisting of Quantifiler THP PCR Reaction Mix and Quantifiler Trio Primer Mix at a 5 to 4 ratio. Nine µL of the master mix was placed into the appropriate wells in a 96 well optical plate. Followingly, 2 µL of the negative controls and standards were added to the desired positions. The plate was sealed with an adhesive cover using a wedge tool and spun down before placing it in the QuantStudio and amplifying as shown in Table 7.

Table 7: Quantifiler Trio real time PCR cycling parameter for 30 cycles

Initial Incubation	**Denature**	**Anneal**	**Extend**	**Final Extension**	**Final Hold**
HOLD	CYCLE (30)			HOLD	HOLD
95 °C 11 min	94 °C 20 sec	59 °C 2 min	72 °C 1 min	60 °C 45 min	4 °C ∞

PCR Amplification

Polymerase chain reactions (PCR) were conducted utilizing GlobalFiler™ PCR Amplification Kit (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA). To conduct half volume reactions, at most 500 pg of the DNA extracts were utilized. The Master Mix and Primer Set were added at a 3.75 to 1.25 µL ratio. The positive control was diluted to 100 pg by adding 5 µL of the DNA control to 2.5 µL of 0.1X Tris-EDTA buffer. Seven and a half µL of the negative control, positive control, and DNA extracts were added to the Master Mix and Primer Set solution. A Veriti® 96-Well Thermal Cycler (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA) and conditions shown in Table 8 were used.

Table 8: Globalfiler Trio thermal cycler cycling parameter for 29 cycles.

Initial Incubation	**Denature**	**Anneal/Extend**	**Final Extension**	**Final Hold**
HOLD	CYCLE (29)		HOLD	HOLD
95 °C 1 min	94 °C 10 sec	59 °C 90 sec	60 °C 10 min	4 °C ∞

Capillary Electrophoresis and STR Analysis

An Applied Biosystems™ 3500 Genetic Analyzer was utilized for the capillary electrophoresis. A master mix consisting of 600 LIZ size standard v2.0 and Hi-Di™ formamide

(Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA) was made at a 0.36 to 10 µL ratio. Ten µL of the master mix was placed into pre-identified wells of an Applied Biosystems® MicroAmp® Optical 96-Well Reaction Plate (Thermo Fisher Scientific, Waltham, MA). Referring to a labelled plate map, 1.2 µL of PCR product, allelic ladder, as well as, positive and negative control were loaded into the designated wells. Empty wells within a column were filled with 10 µL of Hi-Di™ formamide. The reaction plate was sealed with a septum and briefly centrifuged to collect all of the content to the bottom. The plate was placed into the GeneAmp® 9700 PCR System (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA) and set to denature at 95^{o}C for 5 min and chill at 4^{o}C for 5 min. The reaction plate was centrifuged again and loaded on the 3500 Genetic Analyzer (Applied Biosystems®, Thermo Fisher Scientific, Waltham, MA). The length of the capillary was 36 cm using human identification (HID) and POP4 polymer for a 96 well reaction. The GF+Norm_POP4 assay was used. The raw sizing data were converted into allelic calls using GeneMarker® HID (SoftGenetics, State College, PA).

Peptide Analysis by LC-MS/MS

The purified or filtered peptide digests were analyzed using a Liquid Chromatograph Mass Spectrometer (LC-MS/MS), LCMS-8050 Triple Quadrupole, from Shimadzu (Columbia, MD). The ion source employed was electrospray ionization in positive mode (ESI+). Peptides' multiple reaction monitoring (MRM) optimization was performed by direct injection into the MS of 1-2 µL of each peptide individually at 5 pmol/ µL. The MRM transitions are shown in Table 14. The MS source conditions were as follows: nebulizing gas 2 L/min; heat and drying gas both at 10 L/min; interface temperature 400°C; DL temperature 250°C; heat block temperature 400°C. The chromatographic separation was performed onto an Agilent AdvanceBio Peptide mapping column (100 mm X 2.1 mm I.D., 2.1µM) protected with an Agilent guard column of similar

chemistry than the column (Santa Clara, CA) utilizing a flow rate of 0.5 mL/min at 55°C column temperature. A gradient consisting of water and 0.1% formic acid (mobile phase A) and acetonitrile and 0.1% formic acid (mobile phase B) was utilized. The gradient sequence was as follows: 0-20 min, 5-40% B; 20-21 mins, 40-95% B; 21-24 min, 95%B; 24-25 min, 95-5%B; 25-30 min, 5%B. The injection volume was 20 µL.

Statistical Methods

To determine the variation between the different digestion methods, variance values were obtained using the following equation:

$$\sigma^2 = \frac{\Sigma(x_i - \mu)^2}{n - 1}$$

where x equals the value, μ equals the mean, and n is the sample size. Variance can also be expressed as σ^2, where σ is the standard deviation. To test for repeatability in the peptide signals, relative standard deviation values were determined with a sample size of n=30. The following equation was used:

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

The programming language R was used to determine p values for the extraction modifications using t-distribution.

Results

Peptide Results

Trypsin Digestion Modifications

Different digest modifications were tested. The original procedure involved an overnight trypsin digestion based on a standard proteomics protocol provided by Dr. Zhe Cheng from the Weill Cornell Medicine Proteomics core facility in New York City (personal communication). That method was then modified by omitting the incubation with IAA, which alkylates free S-H groups and prevents reannealing of reduced disulfide bonds. For this modification we also added additional DTT to prevent overalkylation. In a third modification, the IAA step was left in place, but the duration of the trypsin digestion was reduced from overnight to 3-hour. Five donors were analyzed for saliva with each procedure. Table 9 displays a comparison of the three modifications for saliva.

The averages show that the removal of the alkylation step diminished the overall recovery of all three peptides. The values and standard deviations for each modification also reveal a high degree of variation from donor to donor. Donor 1 showed the highest values for most markers, especially for the 3-hour digestion methods, here the SMR3B values for both peptides were marked as outliers. Comparing the overnight and 3-hour digestion the HIST1 peptide only showed a small concentration increase. For both SMR3B values, the across donor average increase for the 3-hour digestion was above the upper standard deviation.

Table 9: Saliva peptide recovery for three trypsin digestion modifications*

SALIVA	DONORS	HIST1 (nmol/mL)	SMR3B-pep1 (nmol/mL)	SMR3B-pep2 (nmol/mL)
Overnight with IAA	DONOR 1	4.250	5.380	11.230
Overnight without IAA		0.911	2.390	4.610
3-hour without IAA		4.153	**77.365**	**34.851**
Overnight with IAA	DONOR 2	3.055	2.190	5.473
Overnight without IAA		0.480	0.850	2.090
3-hour without IAA		1.097	15.638	11.500
Overnight with IAA	DONOR 3	0.091	0.008	0.024
Overnight without IAA		0.094	0.022	0.030
3-hour without IAA		0.043	0.025	0.030
Overnight with IAA	DONOR 4	1.285	0.383	1.122
Overnight without IAA		0.438	0.815	0.501
3-hour without IAA		1.427	0.922	1.338
Overnight with IAA	DONOR 5	0.083	0.089	0.401
Overnight without IAA		N/A	N/A	N/A
3-hour without IAA		0.294	0.997	0.938
Overnight with IAA	Average	1.753+/-1.851	1.610+/-2.288	3.650+/-4.766
Overnight without IAA	Average	0.481+/-0.335	1.109+/-0.991	1.808+/-2.066
3-hour without IAA	Average	1.403+/-1.638	18.989+/-33.275	9.731+/-14.799

** outliers in bold*

Variance was calculated to determine how far apart the peptide recovery value was for each donor, when comparing the three methods. The variance for each donor can be found in Table 10. Donors 3, 4, and 5 have the lowest variance value, while Donor 1 has the highest value. SMR3B recovery for Donor 1 has the highest variance. As discussed previously, the 3-hour digestion resulted in a very high recovery for SMR3B peptides in all 5 donors. The recovery value for Donor 1 was significantly higher with the 3-hour digestion and were labelled as outliers. These outliers lead to a large variance for SMRB peptides in Donor 1.

Table 10: Variance between saliva peptide recovery for three trypsin digestion modifications

DONORS	HIST1	SMR3B-pep1	SMR3B-pep2
DONOR 1	3.611	1802.005	252.716
DONOR 2	1.808	66.888	22.720
DONOR 3	0.0008	0.00008	0.000009
DONOR 4	0.286	0.081	0.189
DONOR 5	0.022	0.413	0.144

The removal of the alkylation step also led to a decrease in semen peptide recovery. Donor 2 showed the highest recovery values for both markers. The values for SEMG1-pep1, for the digestion methods with IAA, were labelled as outliers. Additionally, for donor 4, the value for SEMG1-pep2 using the 3-hour digestion was also labeled as an outlier. Donor 5 showed the lowest recovery value, and no peptides were detected for one of the modifications, see Table 11. For SEMG1-pep2, the 3-hour digestion showed the highest peptide recovery for all donors.

The variance between the three methods was also calculated for the semen samples (see Table 12). For both peptides, donor 5 had the lowest variance, while Donor 2 had the highest variance value. The variance for Donor 2 is high because the recovery value for the digestion methods with alkylation were labelled as outliers and the digestion without alkylation resulted in a very low peptide signal.

Table 11: Semen peptide recovery for three trypsin digestion modifications*

SEMEN	DONORS	SEMG1-pep1 (nmol/mL)	SEMG1-pep2 (nmol/mL)
Overnight with IAA	DONOR 1	4.378	42.385
Overnight without IAA		0.578	1.998
3-hour without IAA		4.363	52.207
Overnight with IAA	DONOR 2	47.376	**245.601**
Overnight without IAA		4.642	3.279
3-hour without IAA		28.127	**324.064**
Overnight with IAA	DONOR 3	8.464	14.835
Overnight without IAA		1.392	0.525
3-hour without IAA		8.402	20.931
Overnight with IAA	DONOR 4	7.315	67.970
Overnight without IAA		0.954	2.976
3-hour without IAA		9.145	**255.866**
Overnight with IAA	DONOR 5	0.250	2.892
Overnight without IAA		ND^	ND^
3-hour without IAA		0.405	8.501
Overnight with IAA	Average	13.557+/-19.170	74.736+/98.789
Overnight without IAA	Average	1.891+/-1.863	2.194+/-1.240
3-hour without IAA	Average	10.088+/-10.673	132.314+/-146.787

** outliers in bold*
^ not detected

Table 12: Variance between semen peptide recovery for three trypsin digestion modifications

DONOR	SEMG1-pep1	SEMG1-pep2
DONOR 1	4.794	708.088
DONOR 2	458.055	27963.302
DONOR 3	16.526	109.722
DONOR 4	18.482	17247.133
DONOR 5	0.0417	18.682

Extraction/Purification Methods

Three different extraction/purification methods were tested on semen swab samples (Table 13). The first method utilized was HyperSep™ C18 SpinTip (C18) columns. The C18 method required multiple steps and was laborious and time-consuming; therefore, a more convenient method with a single pass through using nano|Filter vials was also tested. Lastly, in

order to conduct both protein and DNA analysis, a co-extraction method was tested using Microcon® MW100 membrane tubes. This membrane has a 100kD pore size, retaining the larger DNA molecules and any debris in the retentate above the membrane, while the smaller digestion products are in the flow-through.

Liquid and swab semen samples from two donors were used as a positive control to evaluate the procedures. A blank sample was set using a negative control that consisted of the buffer and extraction reagents. For both donors, the nano|Filter method resulted in a higher concentration of both peptides. However, for Donor A, the concentration of SEMG1-pep1 was slightly higher using the Microcon method. The nano|Filter and Microcon method produced very similar final concentrations of SEMG1-pep1 and SEMG1-pep2 for both donors. The p values for SEMG1-pep1 and SEMG1-pep2 are 0.00173 and 0.00058, respectively. Meaning, there was no significant difference between the two extraction methods. The C18 method produced peptide concentrations that were below the blank sample cut-off, classifying these samples as negative. The C18 extraction method lost a significant amount of peptide and led to false negative results.

Table 13: Nano|Filter , Microcon, and C18 extraction comparison*

Signature peptide	Donor	Liquid Semen NanoFilter (ng/mL)	Swab Nano|Filter (ng/mL)	Swab Microcon (ng/mL)	Swab C18 spin column (ng/mL)
SEMG1-pep1	Donor A	15.3	14.1	14.7	0*
	Donor B	64.8	59.1	54.0	0*
SEMG1-pep2	Donor A	286.9	274.2	237.5	0*
	Donor B	367.6	360.1	335.9	0*

** The C18 extraction method fell below the cutoff.*

Synthetic Peptide Calibration Curve and Multiplex Analysis by LC-MS/MS

Synthetic peptides and precursor and product scans were used to determine quantifier and qualifier transitions for each peptide (Table 14). For saliva, three peptides were analyzed: HIST1,

SMR3B-pep1 and SMR3B-pep2. Two peptides were analyzed for semen: SEMG1-pep1 and SEMG1-pep2. The quantifier and qualifier product ions for HIST1 had a *m/z* of 411 and 136, respectively. The *m/z* for the quantifier and qualifier for SMR3B-pep1 were 614 and 517, respectively. For SMR3B-pep2, one quantifier and two qualifiers were identified. The quantifier was *m/z* 628 and the qualifiers were *m/z* 624 and 438. Both semen peptides had one quantifier and one qualifier ions. For SEMG1-pep1, the quantifier was *m/z* 136 and the qualifier was *m/z* 332. Whereas, for SEMG1-pep2, the quantifier and qualifier were *m/z* 201 and *m/z* 424, respectively. Figure 3 displays the total ion chromatogram (TIC) of all 5 peptides.

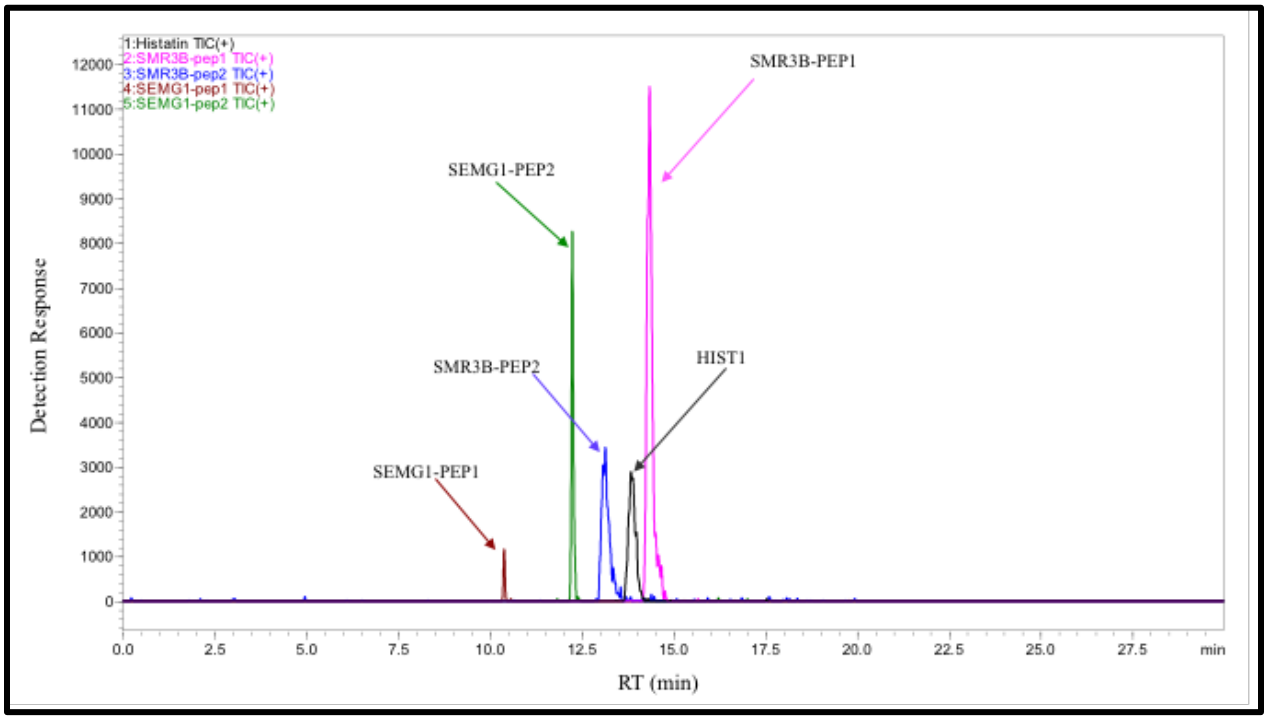

Figure 3*: An overlaid chromatogram of a mixed sample including all five peptides at 0.05 nmol/mL (in buffer).*

To determine the cutoff for the peptides, mixtures containing all five peptides were prepared at the following concentrations: 0.001, 0.0025, 0.005, 0.01, 0.025, 0.05, and 0.1 nmol/mL in the buffer. SMR3B-pep2 and SEMG1-pep2 had the lowest cutoff, at 0.0025 nmol/mL. The rest of the peptides showed a cutoff value at 0.005 nmol/mL (Table 15).

Table 14: LC-MS/MS transitions and retention times for synthetic peptides. Quantifier transitions are in bold.

Peptide name	Precursor m/z	Product m/z	Retention time (min)
Hist1	982.5	411.1	13.8
		136.2	13.8
SMR3B-pep1	1034.8	**614.9**	14.2
		517.8	14.2
SMR3B-pep2	711.0	**628.4**	13.0
		614.4	13.0
		438.3	13.0
SEMG1-pep1	680.0	136.3	10.3
		332.2	10.3
SEMG1-pep2	843.0	201.2	12.2
		424.3	12.2

Table 15: Cutoff values for each peptide.

Peptide	Cutoff (nmol/mL)
HIST1	0.0050
SMR3B-pep1	0.0050
SMR3B-pep2	0.0025
SEMG1-pep1	0.0050
SEMG1-pep2	0.0025

Determination of Minimum Saliva/Semen Volume

Liquid samples from the semen and saliva donors with the highest (Donor 1) and lowest (Donor 2) peptide concentrations as determined above (Tables 9 and 11) were used to test the

ability to detect target peptides in different volumes. This experiment establishes the minimum sample volume required to detect peptides in semen and saliva; results are shown in Table 16. The minimum sample volume for SMR3B-pep2 was not reached for Donor 1. In the case of Donor 2, the minimum sample volume was 0.5µL. For the donor with the highest peptide concentration the minimum sample volume for HIST1 and SMR3B-pep1 are 0.5µL and 1µL, respectively. For Donor 2, the minimum sample volume was 2µL.

HIST1 was problematic and provided unusual results for Donor 2. As the volume of the sample decreased, the concentration of the peptide decreased. This pattern was not observed for HIST1 in Donor 2. HIST1 was not detected in the sample volume of 1µL but was detected in the sample volume of 0.5µL. Additionally, the concentration of HIST1 was greater in the 0.5µL volume than in the 2µL volume.

For semen, the sample with the highest (Donor 1) and lowest (Donor 2) concentration of peptides was detected for all tested volumes. SEMG1-pep1 was detected in all the volumes, for both donors. The minimum sample volume for SEMG1-pep2 was not reached for Donor 1. Whereas, for Donor 2, the minimum sample volume was 0.5µL.

Table 16: Concentration (nmol/mL) of product ions for signature peptides in extracts from 2, 1, 0.5, and 0.1 µL of saliva or semen.

Sample Description	HIST1	SMR3B-pep1	SMR3B-pep2	SEMG1-pep1	SEMG1-pep2
2 µL of saliva Donor 1	0.267	0.847	4.728	ND	ND
1 µL of saliva Donor 1	0.176	0.285	3.154	ND	ND
0.5 µL of saliva Donor 1	0.050	ND	0.321	ND	ND
0.1 µL of saliva Donor 1	ND	ND	0.079	ND	ND
2 µL of saliva Donor 2	0.050	0.0425	0.438	ND	ND
1 µL of saliva Donor 2	ND	ND	0.134	ND	ND
0.5 µL of saliva Donor 2	0.122	ND	0.079	ND	ND
0.1 µL of saliva Donor 2	ND	ND	ND	ND	ND
2 µL of semen Donor 1	ND	ND	ND	9.618	112.67
1 µL of semen Donor 1	ND	ND	ND	8.335	87.403
0.5 µL of semen Donor 1	ND	ND	ND	1.965	18.739
0.1 µL of semen Donor 1	ND	ND	ND	0.639	5.058
2 µL of semen Donor 2	ND	ND	ND	3.779	2.552
1 µL of semen Donor 2	ND	ND	ND	2.360	1.445
0.5 µL of semen Donor 2	ND	ND	ND	0.811	0.608
0.1 µL of semen Donor 2	ND	ND	ND	0.247	ND

ND: not detected.

Repeatability of Peptide Signal in Saliva and Semen

Five microliter aliquots of liquid semen and saliva were extracted and analyzed in triplicate for a total of 30 samples. The relative standard deviation (RSD) was calculated over all concentrations. The results are shown in Figure 4. Values over all three extractions were closer for semen than saliva, with all RSD values lower than 20% for the triplicates. HIST1 had the highest RSD. An RSD above 60% was observed for HIST1 concentrations detected in donors B and C. High RSD values close the 40% were also observed for the SMR3B-pep2 concentrations in donors C and D. Overall the detection of semen peptides was more consistent within triplicates per donor than the detection of saliva peptides.

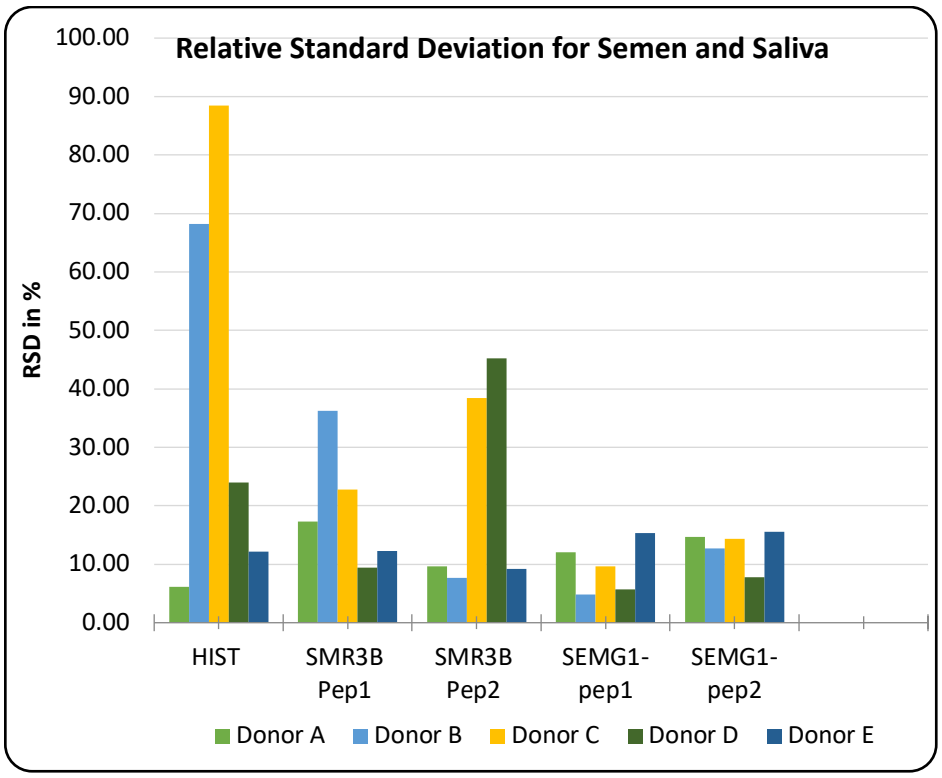

Figure 4: *A bar graph displaying the relative standard deviation in percent (%RSD) for 5 semen and 5 saliva donors (peptide fraction). Please note that Donor A, B, C, D, and E refers to different individuals for each sample type.*

Saliva and Semen Peptide Detection on Stains on Different Substrates

The current assay was tested on various substrates in triplicates (n=48) for both semen and saliva peptide detection. Aged semen stains (n=4) were also analyzed to determine if the desired peptides would be detected. Five fabric substrates and three swabs were analyzed after spotting 5 µL of saliva or semen from different donors. Results are shown in Table 17. This multiplex contains three saliva and two semen signature peptides. Two semen and three saliva peptides were detected on all fabric and swab substrates, with the exception of blue denim. For the first trial, only 2 out of the 3 saliva peptides were detected on blue denim. Both semen peptides were detected in the aged stains on cotton t-shirt and cotton swabs.

Table 17: Semen and saliva peptide detection on eight different stain substrates.

Substrate	Average number of detected Saliva peptides	Average number of detected Semen peptides
Blue Denim	2.7	2
Black Cotton	3	2
Tissue Paper	3	2
White Cotton	3	2
Polyester T-Shirt	3	2
Polyester Swab	3	2
Cotton Swab	3	2
CEP Swab	3	2
2014 Cotton T-Shirt	n/a	2
2014 Cotton Swab	n/a	2

Saliva and Semen Peptide Detection in Body Fluid Mixtures

Stains were made on cotton t-shirts and swabs with 5 L of semen and saliva mixtures in the following ratios: 1:5, 1:2, 1:1, 2:1, and 5:1. The mixture study results on white cotton were as expected based on the different signature peptides concentrations in saliva and semen (Table 18). Higher peptide signals were observed for both body fluids when they were the major component of the mixture, and a decrease in signal was observed when they were the minor component. The two semen markers were detected for all samples even as a minor component, while HIST and SMR3B-pep1 dropped out for the 2:1 (HIST1) and 5:1 (both) mixtures. The same mixture series spotted on cotton swabs had different results. No HIST1 target ions were detected for the 1:1 mixture, all other samples had all expected signals present. The reason for this is unclear. HIST1 was previously labelled as a problematic marker due to observed inconsistency in its signal. HIST1 expression varies in individuals; however, the same saliva donor was used to make the mixtures for cotton t-shirt and swab. Therefore, the HIST1 concentration should have remained the same for both sample types.

Table 18: Concentration (nmol/mL) of signature peptides in saliva and semen mixture at different ratios on cotton swabs and cotton t-shirt.

COTTON T-SHIRT					
Semen: Saliva	**HIST1**	**SMR3B-pep1**	**SMR3B-pep2**	**SEMG1-pep1**	**SEMG1-pep2**
1:5	0.047	0.079	0.446	0.63	6.265
1:2	0.071	0.172	1.21	1.938	27.228
1:1	0.064	0.091	0.651	2.176	29.016
2:1	ND	0.057	0.291	1.994	28.099
5:1	ND	ND	0.18	2.313	26.114
COTTON SWABS					
Semen: Saliva	**HIST1**	**SMR3B-pep1**	**SMR3B-pep2**	**SEMG1-pep1**	**SEMG1-pep2**
1:5	0.308	0.611	6.000	4.36	45.857
1:2	0.202	0.362	4.571	5.245	65.07
1:1	ND	0.151	1.876	7.669	95.154
2:1	0.092	0.096	1.648	8.225	113.58
5:1	0.058	0.048	0.741	8.427	129.813

ND: not detected.

DNA Results

For samples processed with the protein/DNA co-extraction method DNA analyses were conducted in parallel to the proteomic analyses. The protein fractions of the extracts were analyzed on the LC-MS/MS, while some DNA fractions were analyzed by quantification, amplification, and capillary electrophoresis. Not all of the DNA fractions were analyzed by PCR-STR genotyping. Having established that the resulting DNA extracts were of sufficient quality and compatible with PCR, the remaining samples were evaluated based on the quantitation results.

The DNA fractions from the 3-hour with IAA extraction were quantified, amplified, and underwent capillary electrophoresis to produce STR profiles (see Figure 5). The initial concentration of the samples can be found in Table 18. All samples had DNA concentrations sufficient for amplifying the recommended target amount of 500 pg of DNA.

Table 18: Human DNA concentrations for semen and saliva liquid samples using 3-hour with IAA co-extraction method.

Sample Type	Donor	DNA in ng/uL
Semen	A	27.162
	B	3.705
	C	175.022
	D	18.564
	E	19.810
Saliva	A	0.065
	B	0.460
	C	0.300
	D	3.882
	E	1.287

Full profiles were observed for all saliva and semen samples, with the exception of semen Donor C, which had an allelic dropout at TPOX (Figure 5). There were no full locus dropouts. Yindel and DYS391 were part of the STR marker panel but they were excluded from Figure 5 because they are not applicable to females. Yindel and DYS391 were present in all male donors. The STR profiles of all the samples show signs of degradation, with a ski slope shape.

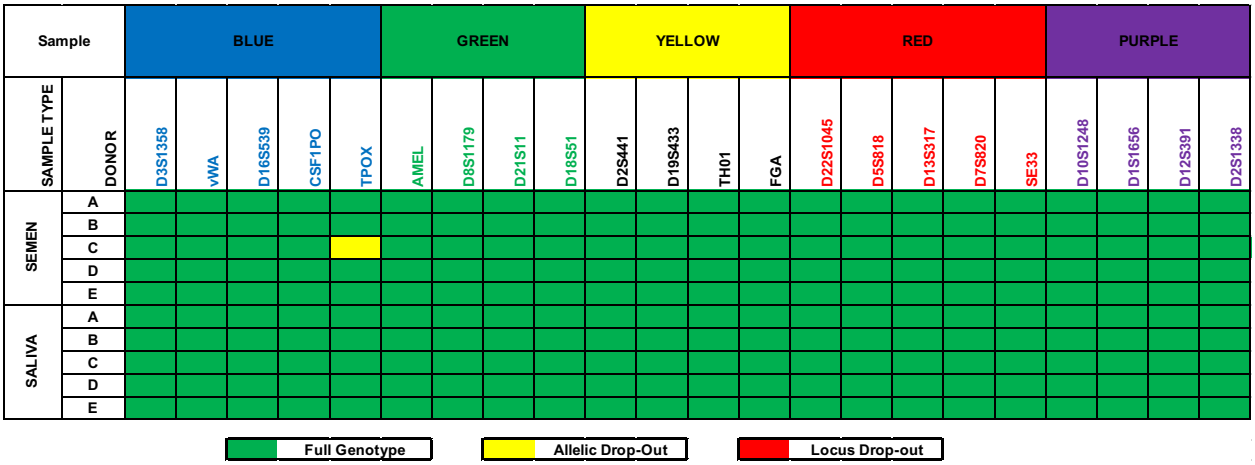

Figure 5: *Heat map from DNA co-extracted from five liquid semen and five liquid saliva samples after a 3-hour trypsin digestion with IAA and additional DTT.*

Human DNA concentrations were established for the DNA fractions from the minimum sample volume analysis. DNA was isolated from different volumes of body fluids: 2, 1, 0.5 and 0.1 µL. The samples were extracted once but quantified twice. The semen samples had higher DNA concentrations than the saliva samples. The initial volume of the body fluid and the concentration of the extracted DNA should be directly proportional. For example, the DNA concentration of a sample that has an initial volume of 2 µL should be twice that of the DNA concentration of a sample that has an initial volume of 1 µL. Figure 6 displays the relationship between the initial volume of liquid saliva and DNA concentration. For saliva, there was a reliable correlation between the two variables (R^2= .9538), see Figure 6. This correlation was not as reliable for the semen sample (R^2=0.6948), see Figure 7. For semen, the DNA extracted from 2 µL of semen was 6.5X more concentrated than the DNA extracted from 1 µL. Moreover, 0.5 µL of semen resulted in a slightly higher concentration of DNA than 1 µL.

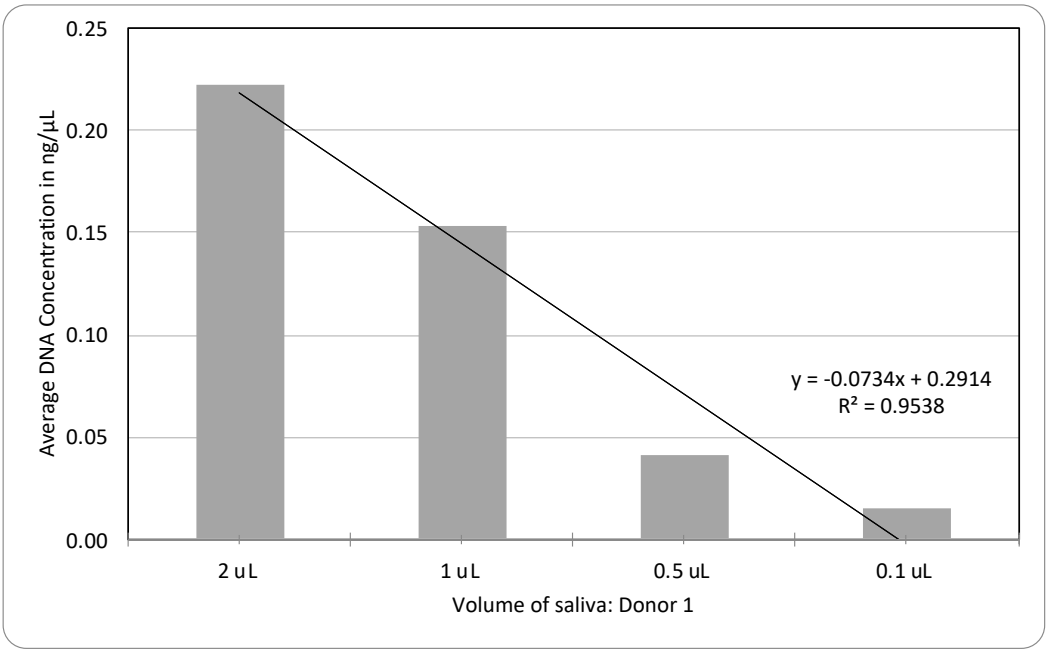

Figure 6: *Average DNA concentrations for saliva dilution series.*

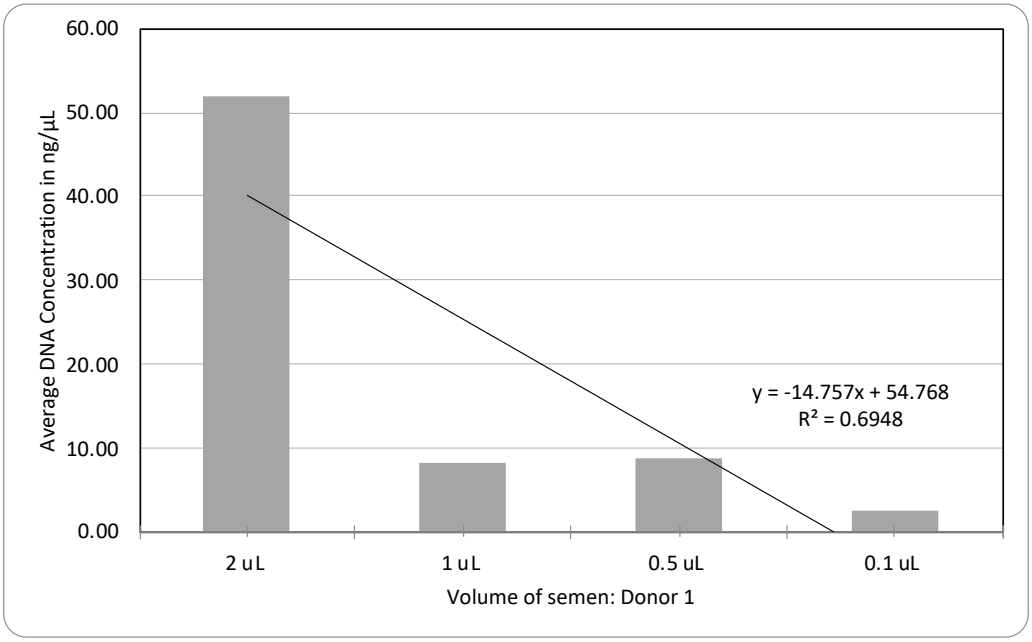

Figure 7: *Average DNA concentrations for semen dilution series.*

The DNA fractions of the repeatability analysis were also examined. Liquid samples were analyzed in triplicates for a total of n=30 samples. The relative standard deviation was calculated over all DNA concentrations, see Figure 8. Like with the protein fractions, the values over all three extractions were closer for semen than saliva. For saliva, Donor E had the lowest RSD, below 20%, while Donor C had the highest RSD, above 90%. For semen, Donor C had an RSD below 10%, whereas Donor A had an RSD above 80%. Semen donors C, D, and E as well as saliva Donor E had RSD values below 20%. Whereas semen donors A and B and saliva donors A, B, C, and D had RSD values above 60%. The DNA results for both semen and saliva lack consistency within the triplicates per donor.

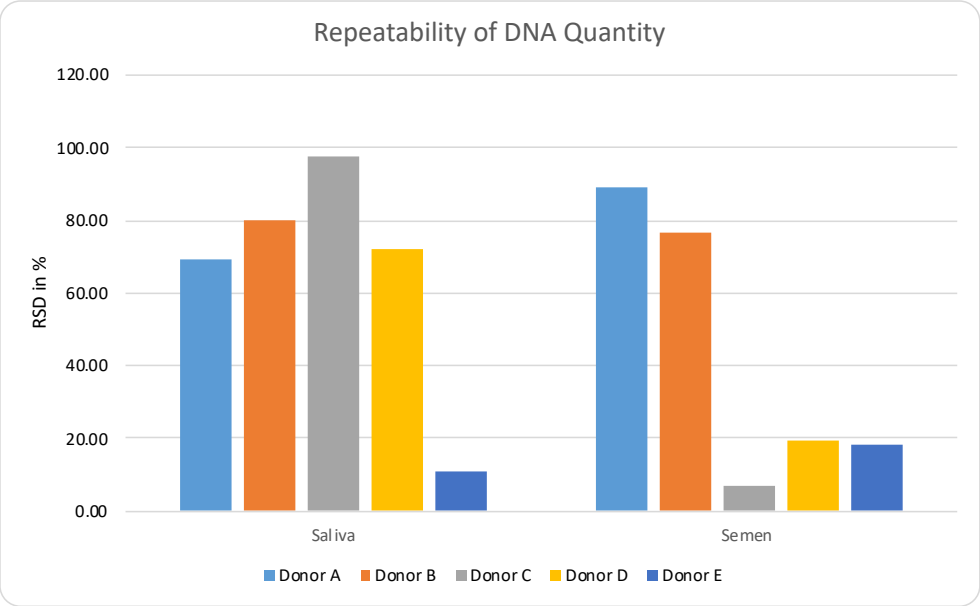

Figure 8: A bar graph displaying the relative standard deviation in percent (%RSD) for 5 semen and 5 saliva donors (DNA fraction).

The last set of DNA samples analyzed were extracted from the substrate and aged stain analysis, in which 5 µL of saliva or semen were spotted on various substrates. Table 19 provides a summary of the results. Again, saliva samples have lower DNA concentrations than semen samples, reflecting the lower amount of cellular material present (Butler, 2011). For the saliva samples, the stain on the polyester t-shirt produced the highest DNA concentration, followed by blue denim and white cotton. The semen stain on black cotton had the highest DNA concentration for semen, followed by polyester t-shirt and cotton swab. Polyester swab had low DNA concentrations for both saliva and semen. However, the CEP swab had even lower DNA concentrations, for saliva. Two different donors had been the source for the 2014 cotton t-shirt and swab stains. DNA yields for both donors were higher than for the contemporary samples, a finding most likely based on biological variation of DNA carrying sperm cells counts amongst males. There was no systematic difference between the two donors. For the cotton t-shirt, Donor 2 had the highest DNA concentration; whereas, for cotton swab, Donor 1 had the highest concentration.

Table 19: Human DNA concentrations for semen and saliva stains on different stain substrates.

Substrate		Saliva DNA in ng/uL	Semen DNA in ng/uL
Blue Denim *		2.599	9.098
Black Cotton*		1.553	26.402
Tissue Paper*		1.442	12.679
White Cotton*		2.420	11.459
Polyester T-Shirt*		4.173	18.610
Polyester Swab#		0.027	1.172
Cotton Swab#		0.109	13.346
CEP Swab#		0.011	5.016
2014 Cotton T-Shirt^	Donor 1		74.61
	Donor 2		96.05
2014 Cotton Swab^	Donor 1		88.31
	Donor 2		85.97

* two donors, one for semen, one for saliva
# two different donors, one for semen, one for saliva
^ third and fourth semen donor

Discussion

The goal of this project was to create a semi-quantitative assay for body fluid identification, using liquid chromatography tandem mass spectrometry. The LCMS triple quadrupole mass spectrometer was successfully adapted for proteomic analysis of body fluids using five signature peptides—three for saliva and two for semen. The saliva specific HIST1 protein was selected for preliminary testing. However, a published study examined histatin-1 as a saliva biomarker and observed that it was only found in 30% of donors (Legg et al., 2017). In that study, five additional saliva markers were analyzed (cystatin SA, cystatin D, submaxillary gland androgen-regulated protein, statherin and mucin 5B) and the sample size for saliva donors was n=50. The submaxillary gland androgen-regulated protein (SMR3B) marker was the only one that was successfully detected in 100% of the saliva samples tested by Legg et al. (2017). The study concluded that the histatin-1 marker was not robust in terms of its detection; therefore,

cannot serve as a reliable saliva marker alone (Legg et al., 2017). In our study, we used the same marker for histatin-1 as Legg et al. (2017) and also faced difficulties due to the markers lack of robustness. HIST1 was proven to be problematic with a low detection sensitivity. For submaxillary gland androgen-regulated protein 3B, we used the markers (SMR3B-pep1 and SMR3B-pep2) that were utilized in the Legg et al. (2017) saliva panel, and identified as confirmatory biomarkers (Legg et al., 2017). For our study, the SMR3B markers were detected even in low saliva volumes, 1µL for SMR3B-pep1 and <0.1µL for SMR3B-pep2. However, the sensitivity of the markers varies from donor to donor due to the different peptide concentrations. No inconsistencies were observed with both markers, e.g. as the volume of saliva decreased, the concentration of the peptide decreased. Additionally, SMR3B-pep1 and SMR3B-pep2 were successfully detected on all substrates analyzed.

The semen markers, SEMG1-pep1 and SEMG1-pep2 both originate from the semenogelin-1 protein. The markers utilized in this study was part of the Legg et al. semen panel (2017). In that study, the SEMG1 markers were detected in 100% of the donors (n=50) (Legg et al., 2017). The SEMG1 markers were identified as confirmatory biomarkers for semen, along with semenogelin-2 markers (Legg et al., 2017). The SEMG1 peptides displayed high sensitivity in our method. Each peptide was detected in 0.1 µL samples, although the sensitivity was donor-based. They were also detected on all substrates and showed low signal variability.

Authentic samples were used to determine the minimum sample volume for the peptides. The minimum sample volume of the authentic peptides should be similar to the cutoffs observed for the synthetic peptides. This was the case for SMR3B-pep2, which was the most sensitive peptide for the saliva samples. For the authentic semen sample, SEMG1-pep1 and SEMG1-pep2

showed similar sensitivities. However, higher concentrations of SEMG1-pep2 were found in most donors.

This assay can be modified by the addition of more biomarkers. The LCMS that was utilized for this project can target up to fifty *m/z*. More research can be done on alternative biomarkers. Additional saliva biomarkers can be analyzed such as cystatin SA and statherin, which were found in 94% and 90%, respectively, of the samples tested according to the Legg et al. study (2017). For semen, semenogelin-2 markers can also be examined.

Prior to the instrumental analysis of these signature peptides, samples needed to be processed to generate defined size peptides in a particle free solution compatible with LC-MS/MS separation and ionization. The required trypsin digestion is commonly allowed to incubate overnight (Legg et al., 2017; Van Steendam et al., 2013; Yang et al., 2013). Therefore, this method comparison included overnight trypsin digestion as one of the tested procedures. However, overnight procedures are time-consuming and cause problematic delays in forensic casework. For a faster turn-around, the digestion was reduced to 3-hour. There was little to no difference in the peptide recovery for most donors, when comparing the overnight and 3-hour durations. For some donors, the 3-hour digestion actually improved peptide recovery. Hence, three-hour digestions were utilized for proceeding analyses.

To reduce the overall time of the analysis, the alkylation step was removed, but this led to a loss in peptide signal and concentration. The alkylation step is very important to the recovery of the peptides, especially when using a highly sensitive method such as LCMS (Boja & Fales, 2001). For thorough peptide recovery, disulfide bonds must undergo reduction and sulfhydryl groups should undergo alkylation (Suttapitugsakul, Xiao, Smeekens, & Wu, 2017). Without these steps, there will be difficulties detecting peptides that are involved in the disulfide bonds

(Suttapitugsakul et al., 2017). Additionally, it has been suggested that treating samples with IAA may help proteins overcome enzymatic degradation (Boja & Fales, 2001). However, overalkylation can occur at a pH near 7 (Boja & Fales, 2001).

It is very important to purify the final peptide digests. Sample debris, such as dye or fiber, can hinder the peptide signal and clog the LCMS. The nano|filter and Microcon methods proved to be the most successful (Table 11). The nano|filters make the purification process very simple for liquid sample analysis. No *a priori* knowledge is necessary to use the vials, and there is little to no sample loss. However, when working with substrate samples, a press device may be needed to completely filter the digests and the filters are easily clogged, which makes substrate analysis tedious. Kranes et al. combined trypsin digestion and Microcon filtration to simultaneously extract DNA and peptides (2017). The Microcon method is more time-consuming than the nano|filter method, but the Microcon filters do not clog as easily as the nano|filters. Using the Microcon method, DNA can also be extracted, and further DNA analysis can be done. Hence, the Microcon method was used for further studies. In the original paper, the trypsin digestion and Microcon co-extraction method had been utilized to isolate protein and DNA components from fingerprints. They found that the DNA yields were similar to a proteinase K extraction (Kranes et al., 2017).

The establishment of peptide separation and detection modes on an instrument previously used for toxicology required the purchase of a peptide separation column like the Agilent AdvanceBio Peptide mapping column and a peptide calibration standard like the Pierce™ Retention Time Calibration Mixture (Thermo Scientific™, Waltham, MA) (results not shown). The next step was to determine the precursor and product ions to target for MRM mode specific to the selected body fluid markers. This was achieved using synthetic peptides for each target.

The retention times of the peptides ranged from approximately 10 to 14 minutes. Although HIST1 and SMR3B-pep2 showed very similar retention times, they have different transitions; therefore, they could be correctly identified and quantified.

The repeatability study showed that the quantification of the semen peptides is repeatable, but this was not the case for saliva peptides. There was high intra-variability with the saliva peptide signal even though these tests were performed using the same donor. Since multiple aliquots were digested, a possible cause could be uneven distribution of the target proteins in the sample. A reproducibility study still needs to be conducted to determine the inter-variability of each signal. The addition of a labeled peptide as internal standard may improve these results.

Using the described methodology, saliva and semen peptides were detected on multiple substrates. The substrates examined are commonly used material and fabrics as well as swabs used for evidence collection. The peptides signal was not hindered, and peptides were detected on all substrates. The most challenging substrate was denim because it was highly absorptive, soaking up all the reagents. However, peptides were detected for all denim stains, with only a single signal dropping out. Moreover, the mixture analysis shows that the presence of one peptide does not hinder the signal of another. Semen and saliva peptides can be simultaneously detected in a mixture and a relative quantification can be made. With the saliva proteins, especially histatin 1 being present at lower levels and resulting in less intense signals, as expected saliva detection was partial for the saliva as the minor component in the 5:1 mixture. In order to determine if the presence of semen affects the signal of the saliva peptides and vice versa, an analysis needs to be conducted that uses ratios in which the volume of one of the body

fluids remains constant as the volume of the other changes. Then, it can be determined if the presence and quantity of one body fluid depresses the signal of the peptides in the other.

Lastly, resulting DNA fractions were compatible with qPCR and STR typing with no indication of inhibition. The STR electropherograms had ski slope shapes, which is an indicator of degradation. The reason for the degradation is unknown and needs to be explored. As expected, the semen DNA fractions had an overall greater concentration of DNA than the saliva DNA fraction. However, the correlation between the initial sample volume and the DNA concentration is stronger for saliva than semen. As the volume of saliva decreases, the concentration of the DNA decreases as well. This may be related to uneven distribution of cellular material with the liquid samples, but in terms of repeatability, more semen samples were consistent within the replicates. Moreover, DNA was sufficiently extracted from various substrates. DNA extracted from polyester t-shirt, blue denim, and white cotton had the highest DNA concentration for saliva. DNA extracted from the black cotton, polyester t-shirt, and cotton swabs gave the highest DNA concentration for semen. A high quantity of DNA was extracted from the aged semen stains. For future analysis, STR profiles should be produced for the substrate samples to determine if there are signs of inhibition and increased degradation.

Overall, this test design was successful. There were no nonspecific signals for these two fluids. The semen peptides were only detected in semen samples and the saliva peptides were only detected in saliva samples. The final sample preparation method is less time consuming than previously published methods (Legg et al., 2017; Van Steendam et al., 2013; Yang et al., 2013) due to the reduced digestion times. and has the advantage of also generating a DNA fraction, thus reducing evidence consumption. Future work will focus on adding additional semen and saliva markers, as well as increasing sensitivity. With the use of a co-extraction

method, body fluid and DNA testing can be combined and if necessary conducted alongside genetically variant peptide analysis, which can produce valuable information on the identity of the donor (Kranes et al., 2017; Sterling et al., 2019). The method also shows promise for toxicology applications where signature peptide quantitation could serve as an internal marker for biological matrix variation and proteomic analysis can be used to identify and examine drug resistance-associated proteins (Cruz et al., 2017). The developed method for body fluid detection can be used to provide specific and sensitive results and isn't time-consuming, unlike enzymatic, antibody-based and chemical techniques.

References

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics: Abstract: Nature. *Nature*, *422*(6928), 198–207. https://doi.org/10.1038/nature01511

Aguilar, M.-I. (2004). *Reversed-Phase High-Performance Liquid Chromatography BT - HPLC of Peptides and Proteins: Methods and Protocols* (M.-I. Aguilar (ed.); pp. 9–22). Springer New York. https://doi.org/10.1385/1-59259-742-4:9

An, J. H., Shin, K.-J., Yang, W. I., & Lee, H. Y. (2015). Body fluid identification in forensics. *The Korean Society for Biochemisty and Molecular Biology*, *11*(Complete), 188–201. http://creativecommons.org/li- censes/by-nc/3.0

Banerjee, S., & Mazumdar, S. (2012). Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry*, *2012*, 1–40. https://doi.org/10.1155/2012/282574

Bell, S. (2012). Christmas tree stain. In *A Dictionary of Forensic Science* (1st ed., p. 45). Oxford University Press.

Boja, E. S., & Fales, H. M. (2001). Overalkylation of a protein digest with iodoacetamide. *Analytical Chemistry*, *73*(15), 3576–3582. https://doi.org/10.1021/ac0103423

Butler, J. (2011). *Advanced Topics in Forensic DNA Typing: Methodology.* (3rd ed.).

Carr, S. (2009). Forensic science: The future of body fluid identification. *Measurement and Control*, *42*(10), 310–313. https://doi.org/10.1177/002029400904201004

Cruz, I. N., Coley, H. M., Kramer, H. B., Madhuri, T. K., Safuwan, N. A. M., Angelino, A. R., & Yang, M. (2017). Proteomics analysis of ovarian cancer cell lines and tissues reveals drug resistance-associated proteins. *Cancer Genomics and Proteomics*, *14*(1), 35–52. https://doi.org/10.21873/cgp.20017

Domon, B., & Aebersold, R. (2006). Mass spectrometry and protein analysis. *Science*, *312*(5771), 212–217. https://doi.org/10.1126/science.1124619

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

Gross, J. H. (2011). Isotopic Composition and Accurate Mass. In *Mass Spectrometry* (2nd ed., pp. 88–91). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-10711-5

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

Harbison, S., & Fleming, R. (2016). Forensic body fluid identification: state of the art. *Research and Reports in Forensic Medical Science*, 11–23. https://doi.org/10.2147/RRFMS.S57994

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

Hedman, J., Gustavsson, K., & Ansell, R. (2008). Using the new Phadebas® Forensic Press test to find crime scene saliva stains suitable for DNA analysis. *Forensic Science International: Genetics Supplement Series*, *1*(1), 430–432. https://doi.org/10.1016/j.fsigss.2007.10.205

Ho, C. S., Lam, C. W. K., Chan, M. H. M., Cheung, R. C. K., Law, L. K., Lit, L. C. W., Ng, K. F., Suen, M. W. M., & Tai, H. L. (2015). Electrospray ionisation mass spectrometry: principles and clinical applications. *The Clinical Biochemist. Reviews*, *24*(February), 3–12.

Hochmeister, M. N., Budowle, B., Rudin, O., Gehrig, C., Borer, U., Thali, M., & Dirnhofer, R. (1999). Evaluation of Prostate-Specific Antigen (PSA) Membrane Test Assays for the

Forensic Identification of Seminal Fluid. *Journal of Forensic Sciences*, *44*(5), 12042J. https://doi.org/10.1520/jfs12042j

Humphrey, S. P., & Williamson, R. T. (2001). A review of saliva Normal composition, flow, and function. Humphrey, Williamson. 2001. Journal of Prosthetic Dentistry.pdf. *The Journal of Prosthetic Dentistry*, *85*(2), 162–169. https://doi.org/10.1067/mpr.2001.113778

Ingold, S., Dørum, G., Hanson, E., Berti, A., Branicki, W., Brito, P., Elsmore, P., Gettings, K. B., Giangasparo, F., Gross, T. E., Hansen, S., Hanssen, E. N., Kampmann, M. L., Kayser, M., Laurent, F. X., Morling, N., Mosquera-Miguel, A., Parson, W., Phillips, C., … Haas, C. (2018). Body fluid identification using a targeted mRNA massively parallel sequencing approach – results of a EUROFORGEN/EDNAP collaborative exercise. In *Forensic Science International: Genetics* 34, 105–115. https://doi.org/10.1016/j.fsigen.2018.01.002

Ispan, D. A. (2018). *Validation of "Christmas tree " staining method for microscopic observation of spermatozoa*. Universidade de Santiago de Compostela.

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

Juusola, J., & Ballantyne, J. (2007). mRNA Profiling for Body Fluid Identification by Multiplex Quantitative RT-PCR. *Journal of Forensic Sciences*, *52*(6), 1252–1262. https://doi.org/10.1111/j.1556-4029.2007.00550.x

Kamanna, S., Henry, J., Voelcker, N. H., Linacre, A., & Kirkbride, K. P. (2016). Direct identification of forensic body fluids using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *International Journal of Mass Spectrometry*, *397–398*, 18–26. https://doi.org/10.1016/j.ijms.2016.01.002

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

Lászik, A., Weisser, H. J., Keresztury, L., Pollak, S., Papp, G., & Pozsgai, I. (1999). PCR typing of human semen stains after SEM-EDX examination. *International Journal of Legal Medicine*, *112*(6), 376–379. https://doi.org/10.1007/s004140050017

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

Leng, Y. (2013). Scanning Electron Microscopy. In *Materials Characterization* (pp. 127–161). Wiley, New York.

Lincoln, C. A., McBride, P. M., Turbett, G. R., Garbin, C. D., & MacDonald, E. J. (2006). The use of an alternative light source to detect semen in clinical forensic medical practice. *Journal of Clinical Forensic Medicine*, *13*(4), 215–218. https://doi.org/10.1016/j.jcfm.2006.02.016

Lindenbergh, A., Maaskant, P., & Sijen, T. (2013). Implementation of RNA profiling in forensic casework. *Forensic Science International: Genetics*, *7*(1), 159–166. https://doi.org/10.1016/j.fsigen.2012.09.003

Macherone, A. (2013). The Future of GC/Q-TOF in Environmental Analysis. *Comprehensive*

Analytical Chemistry, *61*, 471–490. https://doi.org/10.1016/B978-0-444-62623-3.00020-4

Marshall, A. G., & Hendrickson, C. L. (2008). High-Resolution Mass Spectrometers. *Annual Review of Analytical Chemistry*, *1*(1), 579–599. https://doi.org/10.1146/annurev.anchem.1.031207.112945

Miller, K. W. P., Old, J., Fischer, B. R., Schweers, B., Stipinaite, S., & Reich, K. (2011). Developmental validation of the SPERM HY-LITER TM Kit for the identification of human spermatozoa in forensic samples. *Journal of Forensic Sciences*, *56*(4), 853–865. https://doi.org/10.1111/j.1556-4029.2011.01796.x

Miranda, G. E., Prado, F. B., Delwing, F., & Daruge, E. (2014). Analysis of the fluorescence of body fluids on different surfaces and times. *Science and Justice*, *54*(6), 427–431. https://doi.org/10.1016/j.scijus.2014.10.002

Mittal, R. D. (2015). Tandem Mass Spectroscopy in Diagnosis and Clinical Research. *Indian Journal of Clinical Biochemistry*, *30*(2), 121–123. https://doi.org/10.1007/s12291-015-0498-9

Nelson, D. G., & Santucci, K. A. (2002). An alternate light source to detect semen. *Academic Emergency Medicine*, *9*(10), 1045–1048. https://doi.org/10.1197/aemj.9.10.1045

Nussdorfer, P., Cilenšek, I., Zorn, B., & Petrovič, D. (2018). Adapted methods for scanning electron microscopy (SEM) in assessment of human sperm morphology. *Bosnian Journal of Basic Medical Sciences*, *18*(1), 43–48. https://doi.org/10.17305/bjbms.2017.2173

Owen, D. H., & Katz, D. F. (2005). A review of the physical and chemical properties of human semen and the formulation of a semen simulant. *Journal of Andrology*, *26*(4), 459–469. https://doi.org/10.2164/jandrol.04104

Park, S. M., Park, S. Y., Kim, J. H., Kang, T. W., Park, J. L., Woo, K. M., Kim, J. S., Lee, H. C., Kim, S. Y., & Lee, S. H. (2013). Genome-wide mRNA profiling and multiplex quantitative RT-PCR for forensic body fluid identification. *Forensic Science International: Genetics*, *7*(1), 143–150. https://doi.org/10.1016/j.fsigen.2012.09.001

Proctor, G. B. (2016). The physiology of salivary secretion. *Periodontology 2000*, *70*(1), 11–25. https://doi.org/10.1111/prd.12116

Pushie, M. J., Pickering, I. J., Korbas, M., Hackett, M. J., & George, G. N. (2014). Elemental and chemically specific x-ray fluorescence imaging of biological systems. *Chemical Reviews*, *114*(17), 8499–8541. https://doi.org/10.1021/cr4007297

Quinton, P. M. (1978). Techniques for microdrop analysis of fluids (sweat, saliva, urine) with an energy-dispersive X-ray spectrometer on a scanning electron microscope. *American Journal of Physiology-Renal Physiology*, *234*(3), F255–F259. https://doi.org/10.1152/ajprenal.1978.234.3.F255

Rai, R. K., & Sinha, N. (2012). Fast and accurate quantitative metabolic profiling of body fluids by nonlinear sampling of 1H-13C Two-dimensional nuclear magnetic resonance spectroscopy. *Analytical Chemistry*, *84*(22), 10005–10011. https://doi.org/10.1021/ac302457s

Raju, P. S., & Iyengar, N. K. (1964). Acid Phosphatase Reaction as a Specific Test for the Identification of Seminal Stains. *The Journal of Criminal Law, Criminology, and Police Science*, *55*(4), 522. https://doi.org/10.2307/1140908

Revenig, L., Leung, A., & Hsiao, W. (2014). Ejaculatory physiology and pathophysiology: Assessment and treatment in male infertility. *Translational Andrology and Urology*, *3*(1), 41–49. https://doi.org/10.3978/j.issn.2223-4683.2014.02.02

Richard, M. L. L., Harper, K. A., Craig, R. L., Onorato, A. J., Robertson, J. M., & Donfack, J. (2012). Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis. *Forensic Science International: Genetics*, *6*(4), 452–460. https://doi.org/10.1016/j.fsigen.2011.09.007

Scano, P., Locci, E., Noto, A., Navarra, G., Murgia, F., Lussu, M., Barberini, L., Atzori, L., De Giorgio, F., Rosa, M. F., & D'Aloja, E. (2013). 1H NMR metabolite fingerprinting as a new tool for body fluid identification in forensic science. *Magnetic Resonance in Chemistry*, *51*(8), 454–462. https://doi.org/10.1002/mrc.3966

Schlagetter, T., Kammrath, B. W., & Glynn, C. L. (2017). The use of raman spectroscopy for the identification of forensically relevant body fluid stains. *Spectroscopy (Santa Monica)*, *32*(12), 19–24.

Sheppard, K., Cassella, J. P., Fieldhouse, S., & King, R. (2017). The adaptation of a 360° camera utilising an alternate light source (ALS) for the detection of biological fluids at crime scenes. *Science and Justice*, *57*(4), 239–249. https://doi.org/10.1016/j.scijus.2017.04.004

Sijen, T. (2015). Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers. *Forensic Science International: Genetics*, *18*, 21–32. https://doi.org/10.1016/j.fsigen.2014.11.015

Sikirzhytski, V., Virkler, K., & Lednev, I. K. (2010). Discriminant Analysis of Raman Spectra for Body Fluid Identification for Forensic Purposes. *Sensors*, *10*(4), 2869–2884. https://doi.org/10.3390/s100402869

Singhal, N., Kumar, M., Kanaujia, P. K., & Virdi, J. S. (2015). MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology*, *6*(AUG), 1–16. https://doi.org/10.3389/fmicb.2015.00791

Sterling, S. A., Mason, K. E., Anex, D. S., Parker, G. J., Hart, B., & Prinz, M. (2019). Combined DNA Typing and Protein Identification from Unfired Brass Cartridges. *Journal of Forensic Sciences*, *64*(5), 1475–1481. https://doi.org/10.1111/1556-4029.14042

Suttapitugsakul, S., Xiao, H., Smeekens, J., & Wu, R. (2017). Evaluation and optimization of reduction and alkylation methods to maximize peptide identification with MS-based proteomics. *Molecular BioSystems*, *13*(12), 2574–2582. https://doi.org/10.1039/C7MB00393E

Takatori, T., Tomii, S., & Tanaka, T. (1981). The determination of choline in human semen by the enzymic method. *Forensic Science International*, *17*(1), 79–84. https://doi.org/0379073881901924

Tay, J. W., Joudo, J., Tran, T., Ta, H., Botting, J. L., Liew, Y. C., Cooper, P. L., & Rye, M. S. (2020). Comparison of Crime-lite® 82S, Polilight® PL400 and Polilight® PL500 for the detection of semen and saliva stains. *Australian Journal of Forensic Sciences*, 1–11. https://doi.org/10.1080/00450618.2020.1718208

Van Steendam, K., De Ceuleneer, M., Dhaenens, M., Van Hoofstat, D., & Deforce, D. (2013). Mass spectrometry-based proteomics as a tool to identify biological matrices in forensic science. *International Journal of Legal Medicine*, *127*(2), 287–298. https://doi.org/10.1007/s00414-012-0747-x

Vandenberg, N., & Van Oorschot, R. A. H. (2006). The use of Polilights® in the detection of seminal fluid, saliva, and bloodstains and comparison with conventional chemical-based screening tests. *Journal of Forensic Sciences*, *51*(2), 361–370. https://doi.org/10.1111/j.1556-4029.2006.00065.x

Vincini, L. (2010). *The Characterisation and Identification of Body Fluid Proteins for Forensic*

Purposes [Robert Gordon University]. http://openair.rgu.ac.uk

Viner, T. C., Kagan, R. A., & Johnson, J. L. (2014). Using an alternate light source to detect electrically singed feathers and hair in a forensic setting. *Forensic Science International*, *234*, e25–e29. https://doi.org/10.1016/j.forsciint.2013.10.033

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

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

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

Yuan, B.-F. (2014). 5-Methylcytosine and Its Derivatives. *Advances in Clinical Chemistry*, *67*(1), 151–187.

Zapata, F., Fernández de la Ossa, M. Á., & García-Ruiz, C. (2015). Emerging spectrometric techniques for the forensic analysis of body fluids. *TrAC - Trends in Analytical Chemistry*, *64*, 53–63. https://doi.org/10.1016/j.trac.2014.08.011