Final Technical Report

REPORT TITLE: **Developmental Validation of a High-Specificity Multiplex Assay for Human Body Fluid Identification**

AWARD NUMBER: **2012-DN-BX-K035**

AUTHORS: **Phillip B. Danielson, PhD; Kevin Legg, PhD and Heather McKiernan, MSFS**

ABSTRACT

Overview – The advent of DNA profiling has transformed the field of forensic serology by making it possible to individualize biological stains. While DNA analysis of an evidentiary swab may reveal the presence of a DNA profile consistent with an alleged victim, the DNA profile cannot indicate whether the DNA came from saliva, vaginal fluid, urine or a host of other sources. The ability to confidently associate a DNA extract with a specific tissue source or to accurately characterize mixed stains, however, can provide criminal investigators with critical information.

The use of protein biomarkers for confirmatory body fluid identification has attracted significant interest due advances in mass spectrometry. Versatility, sensitivity and ease of use make triple quadrupole mass spectrometers in multiple reaction monitoring mode (QQQ-MRM) the "gold standard" for the analysis of complex samples. Building on the results of previous NIJ supported biomarker research, a targeted ion QQQ-MRM assay for the confirmatory identification of six human body fluids of forensic utility (*i.e.*, saliva, semen, peripheral blood, menstrual fluid, vaginal fluid, and urine) was developed and optimized. The applicability of mass-spectrometry based body fluid identification to samples encountered in a forensic context was rigorously validated in accordance with 2012 SWGDAM Validation Guidelines.

Project Objectives **-** The following four core research objectives were central to successfully achieving this goal.

1) **Select diagnostic target ions for existing biomarkers** to complete a six-stain multiplex QQQ-MRM assay
2) **Optimize the performance of the six-stain multiplex QQQ assay** using synthetic standards, as well as single- and mixed-source samples.
3) **Conduct a rigorous developmental validation** of the multiplex assay.
4) **Develop appropriate Standard Operating Procedures and Interpretation Guidelines**.

Results and Conclusions **-** All core objectives were successfully completed. QQQ-MRM assay development required rigorous characterization of individual precursor-product ion pairs to identify

those best suited for use in a multiplex assay. Biomarker proteins, peptides and transitions were evaluated to identify those that had a unique fragmentation pattern, were abundant, efficiently ionized and had a mass to charge ratio greater than that of the peptide. The targeted-ion inclusion lists compiled for each individual fluid were then used to build a dynamic, retention time restricted, QQQ-MRM assay for each body fluid. To achieve the shortest run time without sacrificing assay sensitivity or reliability, chromatographic separation (HPLC chip and gradient profile) and collision energy were rigorously optimized. To circumvent the potential impact of retention time variability on assay interpretation, internal reference standards were developed to normalize for any retention time shift. Similarly, an Internal Protein Control (IPC) was developed to detect possible inhibition and/or matrix-associated ion suppression. Based on the results of empirical testing, the optimized settings and controls resulted in the unambiguous detection of all targeted protein biomarkers.

Developmental validation studies were carried out to demonstrate the quality and robustness of the multiplex assay. These studies, encompassed assessments of assay sensitivity, repeatability, reproducibility, species specificity, mixtures and a range of casework type samples and were designed to meet Standard 8.2 of the FBI's "Quality Assurance Standards for Forensic DNA Testing Laboratories". Based on the results of the developmental validation studies, a mass spectrometry-based workflow offers significant advantages compared to existing serological methods. QQQ-MRM assays are not easily compromised by interfering chemicals; dependence on a single protein target; potential antibody cross reactivity, false negative results from degraded samples or undetected hook effects – all of which are intrinsic limitations of the immunochromatographic methods currently in common use.

Standard Operating Procedures (SOPs) have been drafted consistent with ISO17025 and ASCLD/LAB-International standards. These SOPs include instructions on sample preparation and instrumental method parameters along with data interpretation guidelines.

In toto, assays based on the use of mass spectrometry offer an approach to the identification of those human body fluids that are most frequently encountered by serologists in forensic caseworking laboratories that is confirmatory and thus more reliable than existing presumptive assays.

Disclaimer: This project is supported by Award No. 2012-DN-BX-K035, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in connection with this project are those of the researchers and do not necessarily reflect those of the Department of Justice.

Table of Contents

EXECUTIVE SUMMARY

Introduction and Statement of Problem

Blood and semen factors that once held promise as discriminatory instruments for individualizing biological stains have been supplanted by DNA markers, which can be amplified from tiny amounts of biological material. While DNA analysis of an evidentiary swab may reveal the presence of a DNA profile consistent with an alleged victim, the DNA profile cannot indicate whether the DNA came from saliva, vaginal fluid, urine or a host of other sources. The ability to confidently associate a DNA extract with a specific tissue source or to accurately characterize mixed stains, however, can provide criminal investigators with critical information.

Consider the case of an alleged sexual assault where a DNA profile consistent with the victim is found on the mouth of a bottle in the suspect's possession. The victim states that the suspect used the bottle as a foreign object to penetrate her vaginally. The suspect counters that the alleged victim had drunk from the bottle and that no sexual contact occurred. Both stories may explain the presence of the victim's DNA on the bottle. The ability to reliably detect traces of vaginal fluid or potentially a mixture of both vaginal fluid and saliva in this case could help to either confirm or refute these opposing claims.

Review of Relevant Literature

While tests for the presence of blood, semen, saliva and urine exist[1-3], some are laborious (*e.g.*, creatinine test for urine). Others require that serologists be proficient at a variety of methodologies, some of which employ reagents that pose health and safety risks. For example, the chemical instability of picric acid (used by some labs to test for urine) presents an explosion hazard and is toxic to liver and kidney tissue. Other serological tests consume significant amounts of a valuable sample while yielding only presumptive results.

Tests for evidence of vaginal contact have proven extremely challenging. The iodine-based Lugol's test which detects glycogenated cells held promise for identifying vaginal cells[4, 5]. More rigorous studies, however, revealed that Lugol's positive cells were also present in the male urethra[6], male urine deposits[7] and on >50% of penile swabs from males who had abstained from sex for several days[7]. Modifications to improve the reliability of Lugol's test have been suggested[8] but are not conducive for use with casework.

The routine testing for blood and seminal fluid by forensic laboratories has been greatly facilitated by the development of rapid immmunochromatographic assays[3, 9, 10]. As with any antibody-based assay, however, results are "presumptive by definition" because the potential for antibody cross-reactivity with non-target molecules can never be eliminated[11]. Moreover, casework-type samples may include environmental contaminants that can interfere with antibody binding, thereby reducing assay sensitivity[12]. For a range of other body fluids, forensically-validated commercial kits based on body fluid specific antigens are lacking entirely and this often leaves the forensic analyst without the ability to make a substantive statement about the potential tissue source of a DNA profile.

Due in part to the limitations associated with existing methods of stain identification, several novel approaches to biological stain identification are now being explored. These research efforts have as their goal the development of a more sensitive and uniform strategy for analyzing body fluids capable of providing analysts with confirmatory results. Emerging approaches include biological stain identification based on messenger- and micro-RNA expression profiles[13-18], epigenetic[19-22] modifications, Raman spectroscopy[23-25] and protein-biomarker detection by mass-spectrometry[26, 27]. Each of these proposed methods have their own strengths and weaknesses. Accordingly, they are not so much competing or mutually exclusive technologies, rather, they are potentially complementary technologies that will make it possible for analysts to obtain useful information from a much larger range of casework samples. These emerging strategies also offer an opportunity for greater standardization and automation of biological stain analysis as well as the incorporation of additional tests for body fluids which are not covered by existing methods. The potential to bring greater uniformity, standardization and thus automation to forensic serological testing would be akin to the type of progress that has been achieved over the past couple of decades in DNA profiling.

Protein biomarkers have attracted significant interest in recent years due in large part due to the strides that have been made in the tools to identify and characterize them. It is now possible to rigorously map entire proteomes with high reproducibility using automated 2-dimensional HPLC systems or MudPIT (multidimensional protein identification technology) to identify potentially useful biomarkers. Once identified, mass-spectrometry-based targeted-ion assays can facilitate the unambiguous detection and quantitation of even low abundance proteins, against a background of other non-target molecules. This has resulted in a wealth of new opportunities to develop protein-based assays for medical and forensic applications such as body fluid identification.

One of the significant advantages of a protein biomarker approach is the tremendous diversity of potential targets that are made possible due to post-translational modification in different tissues. Another key advantage is the stability of many proteins under conditions that lead to degradation of other molecules. Proteins are among the most long-lasting of all biological molecules having been routinely isolated from even ancient biological material[28, 29] and post-mortem tissue[30].

Core Research Objectives

The central goal of the current research project was to developmentally validate the performance parameters of a targeted ion mass-spectrometry assay (triple quadrupole or QQQ) for the confirmatory identification of six human body fluids of forensic utility (*i.e.*, saliva, semen, peripheral blood, menstrual fluid, vaginal fluid, and urine). The following four core research objectives were central to successfully achieving this goal.

1) **Select diagnostic target ions for existing biomarkers** to complete a six-stain multiplex QQQ assay and demonstrate their accurate detection using single source reference samples.
2) **Optimize the performance of the six-stain multiplex QQQ assay** using synthetic standards as well as single-source and mixed-source reference samples.

3) **Conduct a rigorous developmental validation** of the multiplex assay that meets Standard 8.2 of the FBI's "Quality Assurance Standards for Forensic DNA Testing Laboratories".

4) **Develop appropriate Standard Operating Procedures and Interpretation Guidelines** for use of the multiplex QQQ assay for casework samples.

The successful completion of these objectives makes it possible to more accurately and confidently associate a DNA sample with a specific type of biological stain. This can complement the use of DNA profiling and help to pave the way for interlaboratory evaluation and adoption by practitioners.

Methods

Human Subjects

All research was IRB reviewed, approved and conducted in full compliance with U.S. Federal Policy for the Protection of Human Subjects (Basic DHHS Policy for Protection of Human Research Subjects; 56 FR 28003). Body fluid samples were collected from a total of 43 adult (>18 y.o.) human volunteers (20 males; 23 females) recruited from among the Arcadia University student population and employees of NMS Labs. All participants signed a statement of informed consent to participate in the research.

Body Fluid Collection and Protein Extraction

A total of 20-40 samples of each of six forensically-relevant body fluids (*i.e.*, peripheral and menstrual fluid, vaginal fluid semen, saliva and urine) were collected for protein analysis. The specific choice of the bodily fluids to be analyzed and the size of the study population reflected discussions with forensic practitioners at caseworking laboratories.

It was determined that a sample sizes of 10 body fluid samples per experiment would make it make it possible to reliably detect significant variability in study results. This is consistent with the recommendations of the European Network of Forensic Science Institutes (ENFSI) DNA Working Group on the minimum number of samples for forensic validation studies[31]. The statistical basis for this is that based on a student's t-distribution, ±2.576 standard deviations around the mean from a sample size of 10 captures 96.72% of the variability of the population. By comparison, an infinite sample size would capture 99% of the variability at ±2.576 standard deviations.

It should be emphasized the variability that is being assessed is that associated with the performance of the mass spectrometry assay and not the interindividual variability associated with biomarker expression in human populations. Interindividual variability was previously assessed under NIJ award 2009-DN-BX-K165.

Protein Quantization

Protein quantification is valuable to ensure that optimal quantities of protein are available for downstream analyses by mass spectrometry. The Thermo Scientific Pierce Micro BCA Protein Assay (Thermo Fisher Scientific, Rockford, IL) was used to determine final protein concentration (in duplicate) of solubilized samples of whole body fluids. Manufacturer recommended protocols were

followed for sample set up and analysis on a BioTek Instruments 96-well microplate reader set at 562 nm.

Mass Spectrometry

Dried stains were resolubilized in diH_2O. Following protein quantification, up to 20 µg total protein was lyophilized and then reconstituted in a solution of ammonium bicarbonate (ABC). The proteins were then alkylated with iodoacetamide and digested overnight with trypsin. Digested samples were lyophilized and resuspended in acetonitrile and formic acid.

A liquid chromatography (LC) mass spectrometer assay was developed to simultaneously scan for the presence of six different body fluids. Mass spectrometry was performed on an Agilent Technologies HPLC-chip/MS system coupled to an Agilent 6430 Quadrupole Mass Spectrometer. The HPLC chip column used was a 150mm 300 A C18 Analytical with a 160 nl enrichment column.

Data Analyses

Data analysis was performed using Skyline software suite (University of Washington). To enhance the specificity of the method, two MRM transitions for each targeted peptide were employed. Called peptide peaks are those that met the following criteria: (1) a signal-to-noise ratio > 3; (2) the same retention time ± 1% as the heavy labeled reference peptide. (3) the same ion response ratio ± 20% as the heavy labeled reference peptide.

Casework Type Samples

The applicability of a mass-spectrometry-based assay to samples in a forensic context was assessed using a series of casework type samples (Appendix C). This included samples deposited on a variety of substrates (*e.g.*, cotton, denim, leather, metal, glass, plastic, sanitary napkins and polystyrene) or exposure to environmental contaminants (*e.g.*, bleach, soil, detergent, chewing tobacco). Swabs to simulate sexual assault type evidence and a series of aged body fluids were also assessed.

Results and Discussion

Selection of Optimal Diagnostic Target Ions for Existing Biomarkers

Although QQQ-MRM is the gold standard of modern quantitative analyses[32], development requires rigorous characterization of each precursor-product ion pair to identify those best suited for use in a multiplex assay. Optimal transitions had a unique fragmentation pattern; were abundant; efficiently ionized and had a mass to charge ratio greater than that of the peptide. Table 1 presents a summary of the target biomarker and peptides for each of six human body fluids (*i.e.*, urine, semen, saliva, vaginal fluid, menstrual fluid and peripheral blood). These target biomarkers were incorporated into a targeted-ion inclusion list to build a dynamic, retention time restricted, QQQ-MRM assay for each body fluid. Note that while menstrual and vaginal fluids are distinct body fluids in theory, menstrual fluid is are always mixed with vaginal fluid in practice. Thus, a single vaginal/menstrual fluid panel was used.

Executive Summary Table 1

Body Fluid, Target Protein Biomarker, Precursor Peptide Sequence

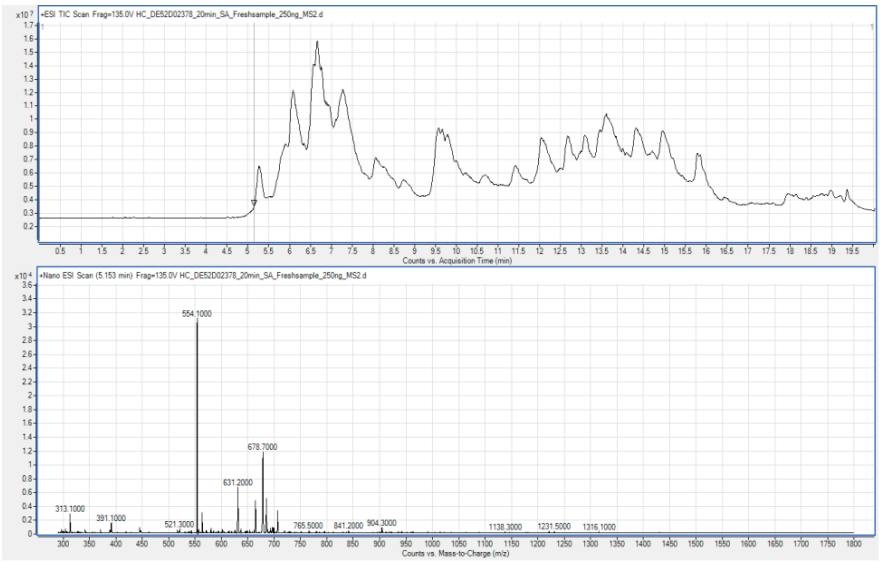
Fluid	Target BioMarker	Target Peptide Sequence
Urine	Uromodulin	TLDEYWR
		STEYGEGYA[Cys(CAM)]DTDLR
Seminal Fluid	Prostatic Acid Phosphatase	ELSELSLLSLYGIHK
		FQELESETLK
	Prostate Specific Antigen (PSA)	LSEPAELTDAVK
		IVGGWE[Cys(CAM)]EK
	Semenogelin-2	DIFTTQDELLVYNK
		DVSQSSISFQIEK
Saliva	Statherin	FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF
	Submaxillary gland androgen regulated protein	IPPPPPAPYGPGIFPPPPPQP
		GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR
	Amylase	LSGLLDLALGK
		IAEYMNHLIDIGVAGFR
Vaginal/ Menstrual Fluid*	Cornulin	ISPQIQLSGQTEQTQK
		GQNRPGVQTQGQATGSAWVSSYDR
	Matrigel-induced Gene C4 protein	DGVTGPGFTLSGSC[Cys(CAM)]C[Cys(CAM)]QGSR
		GC[Cys(CAM)]VQDEFC[Cys(CAM)]TR
	Suprabasin	ALDGINSGITHAGR
	Neutrophil gelatinase-associated lipocalin	SYPGLTSYLVR
		WYVVGLAGNAILR
Peripheral Blood	Alpha-1 Antitrypsin	LSITGTYDLK
		SVLGQLGITK
	Hemopexin	NFPSPVDAAFR
		GGYTLVSGYPK
	Hemoglobin subunit beta	GTFATLSELH[Cys(CAM)]DK
		SAVTALWGK

* Although menstrual and vaginal fluids are distinct body fluids in theory, menstrual fluids markers are always mixed with vaginal fluid markers in practice. Thus the protein biomarkers for these two fluids are combined into a single vaginal/menstrual fluid panel.

Performance Optimization of the Multiplex QQQ Assay

Sample turnaround time has a profound effect on the ability of forensic testing labs to take advantage of the analytical strengths of this assay. Therefore, a front-end optimization sought to achieve the shortest possible run time without sacrificing sensitivity or reliability.

Chromatographic Separation: Several HPLC Chip options was conducted to identify the best chip in terms of reproducibility, separation efficiency and detection sensitivity. Compared to both the ProtID-Chip-150 and the Polaris-HR-Chip, the Large Capacity Chip was identified as having the best overall performance for development of the multiplex body fluid assay in terms of sample partitioning and usable chromatographic data (Figure 1).

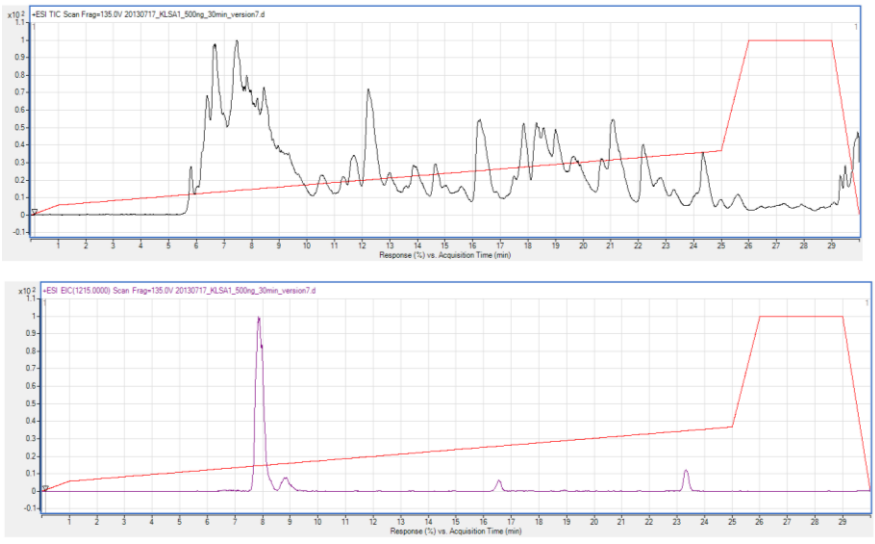


Executive Summary Figure 1

(Top) MS2 scan of salivary proteins on the Large Capacity Chip displaying an absence of overloaded peaks.

(Bottom) The higher quality chromatography results in a more readily resolvable mass spectral data**.**

HPLC Gradient: Using extracts of salivary proteins as a model, an optimal mobile phase gradient was identified (3%-8% organic mobile phase in 1 min; 8%-35% in 24 min; 90% flush). This gradient yielded excellent separation across the run while minimizing assay time (Figure 2). Nanoflow LC runs using microfluidic chip technology generally exceed one hour which makes the current method of only 24 minutes quite fast by comparison.



Executive Summary Figure 2

(Top) MS scan of a reference saliva sample analyzed using the optimized gradient with the % organic phase overlaid in red.

(Bottom) Extracted ion chromatogram showing the elution of the saliva biomarker statherin prior to the organic phase flush in red.

Collision Energy: Instrument settings for collision energy and fragmenter voltage were optimized *in silico* and then evaluated empirically to achieve performance gains by maximizing ion detection. Synthetic peptides were used to avoid possible matrix effects from the biological fluids of interest.

Confirmation of Optimized Assay Parameters: In combination, the assays for the six biomarker master mixes identified a total of sixteen body fluid-specific proteins based on the detection of 26 peptides and 88 transitions. To test the specificity of the optimized assay parameters, synthetic peptides for all body fluid

biomarkers were combined into a target biomarker master mix which was then analyzed using the optimized instrument/assay parameters.

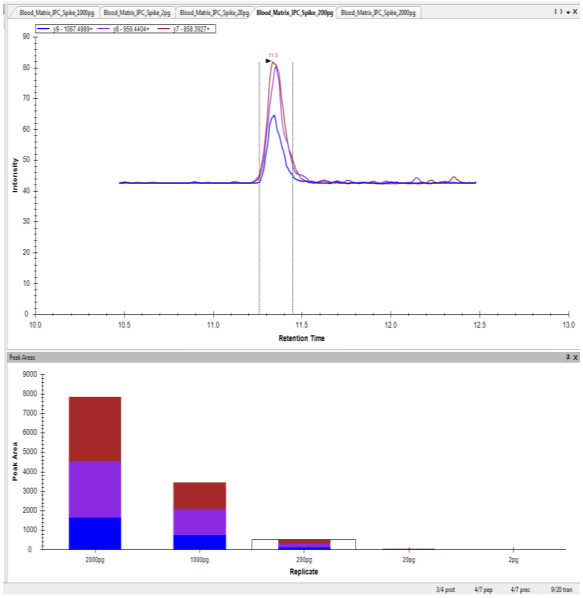
Internal Reference Standards: To circumvent the potential impact of retention time variability on assay interpretation, internal reference standards were designed and purchased from New England Peptide. These reference standards were identical to the peptides present in a biological fluid of interest except that stable "heavy isotope" labels are used to produce a shift in mass.

Internal Positive Control: An Internal Protein Control (IPC) was developed to detect inhibition and/or matrix-associated ion suppression. The IPC consists of two trypsin-cleavable proteins added to each sample and cleaved during sample preparation. A second set of peptides having the same amino acid sequence as the cleaved precursor peptides but labeled with a non-radioactive "heavy" carbon isotope allows the signal intensity ratio between the normal and heavy peptides to be monitored during sample injections (Figure 3).

Executive Summary Figure 3

(Top) Representative MRM chromatogram showing reliable peak shape for a 200 pg input of the IPC in the presence of 10 µg of peripheral blood.

(Bottom) and consistent ion ratios across three serial dilutions of the IPC (2000 pg, 1000 pg and 200 pg).



Developmental Validation Studies

A series of developmental validation studies were conducted based on the 2012 Scientific Working Group on DNA Analysis Methods (SWGDAM) Validation Guidelines for DNA Analysis Methods[33]. These were designed to meet Standard 8.2 of the FBI's "Quality Assurance Standards for Forensic DNA Testing Laboratories".

Sensitivity: The sensitivity limit for each of the six biological fluids was determined based on serial dilutions of body fluid samples pooled from ten individuals so as to obtain an averaged indication of assay sensitivity. Results from these studies are presented in Table 2 where protein biomarkers are organized by target body fluid panel; dark green cells indicate detection of all target transitions and light green cells indicate the detection of a minimum of one transition for the target peptide. Dilutions are expressed in terms of microliters of body fluid. For example, all transitions for uromodulin (the urine biomarker) can be detected in as little as 1.5625 µl of urine and at least one transition can be detected in 390.6 nl of urine.

Executive Summary Table 2

Sensitivity Limits for Each of the Six Biological Matrices

µl of Target Body Fluid Detected at Each Dilution Level

		Dilution Series																		
		1:1	1:2	1:4	1:8	1:16	1:32	1:64	1:128	1:256	1:512	1:1,024	1:2,048	1:4,096	1:8,192	1:16,384	1:32,768	1:65,536	1:131,072	1:262,144
Menstrual Fluid	Alpha 1 Antitrypsin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031					
	Hemopexin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031					
	Hemoglobin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031	0.0015	0.0008	0.0004		
	Cornulin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906												
	LY6	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953											
	NGAL	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953											
	Suprabasin	25.0	12.5	6.25	3.125	1.5625	0.7813													
Peripheral Blood	Alpha 1 Antitrypsin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031	0.0015	0.0008			
	Hemopexin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031	0.0015				
	Hemoglobin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031	0.0015	0.0008	0.0004	0.0002	0.0001
Vaginal Fluid	Cornulin	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488								
	LY6	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906											
	NGAL	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977									
	Suprabasin	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906											
Saliva	Statherin	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813												
	SubMax	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488								
	Amylase	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244							
Semen	Acid Phosphatase	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061					
	Prostate Specific Antigen	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061					
	Semenogelin 2	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031				
Urine	Uromodulin	50	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906											

Repeatability and Reproducibility: Reproducibility as indicated by the percent coefficient of variation (%CV) values was determined for sample and technical replicates. The maximum allowable coefficient of variation percentage was 25% based on common industry practice.

Protein extraction reproducibility was assessed on the basis of variation in average BCA protein quantitation values as a function of two or more people performing the same extraction. The majority of assay results were characterized by excellent reproducibility. The only exception of note was the lower reproducibility associated with urine extraction due to the presence of urea which reacts with the BCA agent resulting in a artificial elevation in the apparent protein content of a sample. The use of inflated protein content values increases the probability of false negative results. The removal of urea is achieved with an 80% acetone precipitation step. The %CV values for each fluid for the overall analytical method in terms of the peak area ratios as well as the peak retention times and the ion response ratios (for native and AQUA peptides) are indicated in Table 3. Aside from urine samples, reduced reproducibility was associated with the suboptimal peak morphologies of extremely high abundance targets (*i.e.*, some hemoglobin and semenogelin peptides) or extremely low abundance targets (*i.e.*, suprabasin).

Executive Summary Table 3

Repeatability (Analyst 1 and Analyst 2) and Reproducibility (Overall) of the Analytical Method

			Area Ratio Natural v AQUA			Retention Time Natural v AQUA			Ion Response Ratio Natural			Ion Response Ratio AQUA		
			Analyst 1 %CV	Analyst 2 %CV	Overall %CV	Analyst 1 %CV	Analyst 2 %CV	Overall %CV	Analyst 1 %CV	Analyst 2 %CV	Overall %CV	Analyst 1 %CV	Analyst 2 %CV	Overall %CV
Menstrual Blood	Alpha 1 Antitrypsin	LSITGTYDLK	8.90	7.70	8.30	0.09	0.08	0.09	3.10	4.00	3.80	29.50	19.10	26.90
		SVLGQLGITK	10.10	5.70	8.30	0.06	0.08	0.07	1.60	1.80	1.80	2.80	3.20	3.00
	Hemopexin	NFPSPVDAAFR	9.00	7.90	8.40	0.12	0.14	0.13	13.00	9.10	11.10	3.20	3.50	3.40
		GGYTLVSGYPK	23.50	20.50	21.80	0.56	0.44	0.51	NR	NR	NR	6.50	5.90	6.40
	Hemoglobin	SAVTALWGK	9.30	4.20	7.20	0.08	0.05	0.07	2.60	2.70	2.70	5.10	6.10	5.50
	Cornulin	ISPQIQLSGQTEQTQK	21.90	20.80	21.30	0.35	0.32	0.33	NR	NR	NR	5.50	6.50	6.00
	LY6	GCVQDEFCTR	13.50	12.50	13.00	0.20	0.24	0.22	15.00	16.90	15.80	3.50	3.30	3.40
	NGAL	WYVVGLAGNAILR	26.50	24.40	26.70	0.18	0.21	0.21	NR	NR	NR	7.10	7.30	7.10
		SYPGLTSYLVR	7.00	6.80	7.10	0.14	0.13	0.13	15.10	16.90	16.00	2.80	2.90	2.90
	Suprabasin	ALDGINSGITHAGR	28.60	21.00	25.70	0.24	0.23	0.24	NR	NR	NR	4.70	3.40	4.10
Peripheral Blood	Alpha 1 Antitrypsin	LSITGTYDLK	16.40	17.30	16.70	0.23	0.17	0.20	3.90	3.80	3.80	21.60	14.60	18.20
		SVLGQLGITK	15.30	17.20	16.10	0.08	0.08	0.08	1.40	3.20	2.40	3.80	4.00	3.90
	Hemopexin	NFPSPVDAAFR	13.10	10.70	11.80	0.10	0.09	0.10	3.50	3.70	3.70	2.70	3.70	3.20
		GGYTLVSGYPK	10.50	8.80	9.70	0.24	0.29	0.26	9.00	12.15	10.60	5.50	6.80	6.10
	Hemoglobin	SAVTALWGK	46.20	27.10	39.90	0.09	0.09	0.09	3.50	3.30	3.30	6.60	4.50	5.60
		GTFATLSELHCDK	44.20	29.20	29.20	0.82	0.17	0.17	12.30	11.30	11.30	11.00	20.70	16.40
Vaginal Fluid	Cornulin	GQNRPGVQTQGQATGSAWVSSYDR	7.80	9.20	8.60	0.19	0.22	0.20	17.10	19.40	18.10	6.40	8.20	7.30
		ISPQIQLSGQTEQTQK	7.40	7.50	7.40	0.16	0.18	0.17	13.50	8.40	7.60	17.30	5.20	5.60
	LY6	GCVQDEFCTR	7.50	9.50	8.50	0.14	0.22	0.18	6.10	7.30	6.70	2.90	2.80	2.90
	NGAL	WYVVGLAGNAILR	6.20	4.30	5.30	0.11	0.12	0.11	7.20	4.00	5.80	3.80	3.80	3.90
		SYPGLTSYLVR	4.30	3.00	3.70	0.06	0.08	0.07	5.50	5.30	5.50	3.60	3.00	3.30
	Suprabasin	ALDGINSGITHAGR	12.20	12.10	12.20	0.16	0.20	0.18	26.60	27.10	26.70	3.40	4.00	3.70
Saliva	Statherine	FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF	6.10	5.00	5.60	0.09	0.10	0.10	8.10	7.10	8.00	15.00	12.00	14.10
	SubMax	GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR	3.70	2.60	3.40	0.13	0.14	0.14	1.90	1.50	1.70	9.40	7.40	8.80
		IPPPPPAPYGPGIFPPPPPQP	6.70	7.50	7.10	0.25	0.33	0.29	2.60	2.50	2.50	3.60	7.60	6.00
	Amylase	LSGLLDLALGK	6.80	7.80	7.20	0.11	0.05	0.08	1.80	1.80	1.80	27.00	21.40	24.60
		IAEYMNHLIDIGVAGFR	9.10	6.70	8.00	0.03	0.04	0.03	1.80	1.60	1.70	5.10	4.70	4.80
Semen	Prostatic AP	ELSELSLLSLYGIHK	25.40	20.20	22.70	0.03	0.03	0.03	2.80	1.70	2.40	4.60	2.20	3.60
	Prostate Specific Antigen	IVGGWECEK	24.10	17.30	20.70	0.09	0.00	0.06	1.90	1.90	2.00	4.40	3.80	4.20
		LSEPAELTDAVK	21.10	16.30	18.60	0.06	0.09	0.08	3.90	2.70	3.30	3.90	4.10	4.10
	Semenogelin 2	DIFTTQDELLVYNK	34.50	31.30	32.70	0.06	0.04	0.05	1.30	1.20	1.30	7.50	5.60	6.50
		DVSQSSISFQIEK	24.30	18.70	21.50	0.16	0.16	0.15	2.50	1.60	2.10	5.30	4.90	5.10
Urine	Uromodulin	TLDEYWR	46.90	45.30	46.50	0.14	0.14	0.14	14.90	21.10	18.10	49.90	48.70	48.90
		STEYGEGYACDTDLR	69.80	81.80	79.10	0.23	0.19	0.22	6.70	7.50	7.20	5.60	4.70	4.70

* Measured as a coefficient of variation for area ratios of the natural peptide as compared to the Aqua peptide, retention time for the natural peptide as compared to the Aqua peptide, ion response ratios for the natural peptide and ion response ratios for the Aqua peptide.

Species Specificity: The human-specificity of the assay biomarkers was demonstrated using both *in silico* and empirical methods. As biomarker identification is based on the amino acid sequence of the target peptides, non-human peptides would have to possess the exact same amino acid primary sequence as a target peptide to be mistaken for a human target. The amino acid sequences of the target peptides were first screened against the SWISS-PROT databases of 550,116 distinct proteins from 13,257 species. In addition, the target peptide sequences were also screened against the conceptual amino acid translations of major genomic databases (*i.e.* GenBank and NCBI RefSeq). These searches demonstrated the human specificity for the target biomarker peptides. This finding was confirmed empirically using a series of non-human blood samples. As anticipated, those non-human peptides that were found to be identical to target human biomarker peptides were associated with higher order primates. This fact has been taken into consideration in the interpretation guidelines for the assay.

Mixtures: Competition for peptide detection between different contributors is not a concern with mass spectrometry based body fluid identification. Rather, mixture studies are conducted to assess potential matrix effects. All possible pair-wise combinations of the six body fluids targeted by the QQQ-MRM multiplex assay were evaluated using equal volumes of corresponding body fluids. The results were compared to previously described sensitivity studies to detect potential matrix effects.

In total, 45 mixed samples were tested. In all, 42 of the 45 samples that were tested, accurately characterized each fluid in the mixture. Urine was the only fluid which yielded suboptimal results. Uromodulin was successfully identified across all replicates when mixed with saliva or vaginal fluid but not detected when mixed with semen, menstrual or peripheral blood. This most likely reflects the very large difference in protein content associated with equal volumes of urine versus semen, menstrual or peripheral blood

Casework-Type Samples: Among the most important of the developmental validation studies, are those that replicate the types of challenging samples encountered by forensic practitioners. By simulating the characteristics of authentic forensic samples, it is often possible to identify areas for improvement in the assay and the Standard Operating Procedures (SOPs) and interpretation guidelines.

Over 100 casework-type samples were tested to assess the potential impact of a broad range of sample parameters. Casework-type samples (Appendix C) included single-source and sexual assault type stains recovered from a variety of substrates (*e.g.*, cotton, leather, skin, latex, styrofoam, denim, sanitary pad). The potential impact of environmental contaminants (*e.g.*, spermicides, personal lubricants, detergent, fecal matter) and including several of those that have been previously identified as having inhibitory impacts on DNA profiling technologies (*e.g.*, soil, leather, indigo dyes in denim, bleach and tobacco juice). The results of these tests are presented in Table 4 where dark green cells indicate detection of all target peptides; light green cells indicate the detection of a minimum of one target peptide and red cells indicate a failure to detect any of the expected target peptides. All casework-type type samples prepared with urine, vaginal fluid, semen or saliva were unambiguously and confidently identified. From the menstrual fluid samples, the blood component was always successfully identified. The identification of the non-blood components of menstrual fluid necessary to link the sample to the female reproductive system, however, proved to be more variable (*i.e.*, they were not reproducibly detected in all samples).

Among samples containing potential inhibitors, blood biomarkers in peripheral blood and menstrual fluids were successfully identified when mixed with soil, bleach or when extracted from leather or denim. However, mixing blood with household laundry detergent or 10% bleach resulted in a failure to detect any blood-specific proteins. As is the case for DNA analysis of samples containing *Taq* polymerase inhibitors, the development of additional front-end sample preparation protocols may enable the successful processing of these samples.

Executive Summary Table 4

QQQ-MRM Detection of Body Fluid Biomarkers in Forensic Casework-Type Samples

		Fluid Confirmation					
		Menstrual Blood	Peripheral Blood	Vaginal Fluid	Saliva	Semen	Urine
Menstrual Blood	Cotton		ND	ND	ND	ND	ND
	Denim		ND	ND	ND	ND	ND
	Pad		ND	ND	ND	ND	ND
	Rectal Swab		ND	ND	ND	ND	ND
	Spermicide		ND	ND	ND	ND	ND
	Lubricant		ND	ND	ND	ND	ND
	Vaginal Swab+Lube		ND	ND	ND	ND	ND
Peripheral Blood	Soil	ND		ND	ND	ND	ND
	10% Bleach	ND		ND	ND	ND	ND
	Leather	ND		ND	ND	ND	ND
	Detergent	ND		ND	ND	ND	ND
	Cotton	ND		ND	ND	ND	ND
Vaginal Fluid Case Samples	Cotton	ND	ND		ND	ND	ND
	Finger Swab	ND	ND		ND	ND	ND
	Penile Swab	ND	ND		ND	ND	ND
Saliva Case Samples	Styrofoam	ND	ND	ND		ND	ND
	Gum	ND	ND	ND		ND	ND
	Cup	ND	ND	ND		ND	ND
	Cotton	ND	ND	ND		ND	ND
	Condom	ND	ND	ND		ND	ND
	Tobacco	ND	ND	ND		ND	ND
Semen Case Samples	Condom	ND	ND	ND	ND		ND
	Cotton	ND	ND	ND	ND		ND
	Denim	ND	ND	ND	ND		ND
	Oral Swab	ND	ND	ND	ND		ND
	Rectal Swab	ND	ND	ND	ND		ND
	Spermicide	ND	ND	ND	ND		ND
	Lubricant	ND	ND	ND	ND		ND
Urine Case Samples	Cotton	ND	ND	ND	ND	ND	
	Soda	ND	ND	ND	ND	ND	
	Ceramic Cup	ND	ND	ND	ND	ND	
	Styrofoam	ND	ND	ND	ND	ND	

*Dark green cells indicate detection of all target peptides; light green cells indicate the detection of a minimum of one target peptide and red cells indicate a failure to detect any of the target peptides.

Given the frequency with which degraded samples are encountered by forensic practitioners, protein degradation may also be a concern. While there is abundant evidence based work in the fields of ancient DNA[34] and ancient proteomics[35] that proteins are generally more stable over time than nucleic acids, extensive degradation may still adversely impact the ability to obtain interpretable data from aged or weathered materials. Thus, a series of saliva, peripheral blood, semen, and urine samples which had been aged at room temperature from 2 to 7 years were analyzed.

As exemplified by the results obtained with aged blood and semen samples, the QQQ-MRM multiplex assay is able to obtain confident body fluid identification based on at least one, and often multiple, target protein biomarkers in all aged samples tested. This is shown in Table 5 where dark green cells indicate detection of all target transitions and light green cells indicate the detection of a minimum of one transition for the target peptide. Thus, extensive protein degradation can take place before trypsin cleavage recognition sequences or the peptides themselves are lost.

Executive Summary Table 5

QQQ-MRM Detection of Peripheral Blood Biomarkers in Aged Blood and Semen Stains

			Blood				
			7 years	5 years	4 years	2 years	2 years
Peripheral Blood	Biomarkers	Alpha 1 Antitrypsin					
		Hemopexin					
		Hemoglobin					

			Semen						
			5 years	5 years	4 years	4 years	2 years	2 years	2 years
Semen	Biomarkers	Acid Phosphatase		ND	ND		ND	ND	ND
		Prostate Specific Antigen		ND	ND		ND		ND
		Semenogelin 2							

*dark green cells indicate detection of all target transitions; light green cells indicate the detection of a minimum of one transition

Standard Operating Procedures and Interpretation Guidelines

Based on the results of the method development and optimization work and on the results from the developmental validation studies, Standard Operating Procedures (SOPs) have been drafted using document templates from an ISO17025 and ASCLD/LAB-International accredited caseworking laboratory (see Appendices A and B). These SOPs include instructions and information on sample extraction/preparation, instrumental method parameters and data interpretation guidelines. While these documents will provide valuable guidance to analysts in an operational caseworking environment, it is recognized that these are "living documents" which may and should evolve based on the experiential input of practitioners.

Conclusions and Implications for Policy and Practice

Targeted-ion mass spectrometry-based body fluid identification offers significant advantages compared to the existing serological methods. For example, assay results are not easily compromised by interfering chemicals; dependence on a single protein target; potential antibody cross reactivity, weak test lines in low abundance samples or false negative results from degraded samples or undetected hook effects that are intrinsic to widely used immunochromatographic methods.

Because of the superior sensitivity of the QQQ-MRM assay, however, it would not be unreasonable to anticipate that some high-specificity protein biomarkers may be expressed at low but potentially detectible levels in non target tissues. This has necessitated that the interpretation guidelines take into consideration the relative expression patterns of biomarkers across tissues and the overall profile of detected protein biomarkers in a questioned sample. As a result, highly degraded or trace samples where only a subset of targeted biomarkers are detected may (in some cases) still limit analysts to reporting only a presumptive result. For the majority of samples, however, the sensitivity and accuracy of the QQQ-MRM assay will provide analysts with confirmatory identifications of questioned stains.

In toto, assays based on the use of mass spectrometry offer a more reliable approach to the confirmatory identification of those human body fluids that are most frequently encountered by serologists in forensic caseworking laboratories.

Implications for Policy and Practice

The current research applies cutting edge proteomic technologies (triple quadrupole multiple reaction monitoring or QQQ-MRM) to the development of a sensitive, reliable, human-specific assay for the confirmatory identification of questioned biological stains. This will directly address the inherent limitations of the approaches currently employed in case-working laboratories. Nonetheless, some casework samples will still present challenges that complicate interpretations. The release of small quantities of blood into the oral cavity as a result of using dental floss or a minor injury to the inside of the mouth may be detected as a mixed stain – which it is. In such cases, the experienced judgment of a serologist will be needed to assess the potential significance of the mixture.

Implementing this approach does involve a significant up-front investment. However, the cost-to-benefit ratio of the instrument is substantial. QQQ-MRM assays are unequaled for detecting trace quantities of target compounds and by multiplexing the assays, the cost/test will drop below that of many existing immunoassays. On the legal side, the study results reported here coupled with publication in peer-reviewed journals, interlaboratory validation studies and adoption by serologists (even as an investigational tool), will help to place the findings of this research on sound legal footing.

Future Research

As with any new analytical technique, additional improvements in cost, sensitivity, and throughput should be explored. For example, implementation of multi-capillary genetic analyzers for coupled with sample automation systems have led to tremendous improvements in casework throughput over the years. Likewise, a promising direction for future proteomics based serology research would be to modify the current mass spectrometry-based workflow for high throughput analysis. To do this, manual sample preparation could be automated and the frontend nanoflow liquid chromatography system could be replaced with a high flow ultra-performance liquid chromatography interface.

Greater use of automation solutions for this application is another promising direction for future research. For example, the Beckman Biomek® NXP Laboratory Automation Workstation, the Agilent Technologies Bravo platform, or Tecan Freedom EVO® all offer "off the shelf" proteomics automation solutions. For the improved liquid chromatography system, replacing the 1200 series nanoflow chip cube employed during this project with a high flow 1290 series UHPLC, run times of 30 minutes/sample could be reduced to 3-10 minutes/sample – ideal for the high-throughput demands of forensic laboratories.

Executive Summary References Cited

1. *Biology Methods Manual*. 1978: Metropolitan Police Forensic Science Laboratory.
2. *Protocol Manual*. 1989: FBI Laboratory Serology Unit.
3. Hochmeister, M.N., et al., *Evaluation of prostate-specific antigen (PSA) membrane test assays for the forensic identification of seminal fluid.* J Forensic Sci, 1999. **44**(5): p. 1057-60.
4. Rees, B. and T.J. Rothwell, *The identification of phosphoglucomutase isoenzymes in semen stains and its use in forensic casework investigation.* Med Sci Law, 1975. **15**(4): p. 284-93.
5. Thomas, F.a.v.H., W., *The demonstration of recent sexual intercourse in the male by the Lugol method.* . Medicine, Science and the Law, 1963. **3**: p. 169-171.
6. Rothwell, T.J. and K.J. Harvey, *The limitations of the Lugol's iodine staining technique for the identification of vaginal epithelial cells.* J Forensic Sci Soc, 1978. **18**(3-4): p. 181-4.
7. Hausmann, R., C. Pregler, and B. Schellmann, *The value of the Lugol's iodine staining technique for the identification of vaginal epithelial cells.* Int J Legal Med, 1994. **106**(6): p. 298-301.
8. Jones, E.L., Jr. and J.A. Leon, *Lugol's test reexamined again: buccal cells.* J Forensic Sci, 2004. **49**(1): p. 64-7.
9. Laux, D.L., A.J. Tambasco, and E.A. Benzinger. *Forensic Detection of Semen II. Comparison of the Abacus Diagnostics OneStep ABAcard p30 Test and the Seratec PSA Semiquant Kit for the Determination of the Presence of Semen in Forensic Cases*. [cited; Available from: http://mafs.net/pdf/laux2.pdf.
10. Schweers, B.A., et al., *Developmental validation of a novel lateral flow strip test for rapid identification of human blood (Rapid Stain Identification--Blood).* Forensic Sci Int Genet, 2008. **2**(3): p. 243-7.
11. Laffan, A., et al., *Evaluation of semen presumptive tests for use at crime scenes.* Med Sci Law, 2011. **51**(1): p. 11-7.
12. Hobbs, M.M., et al., *Vaginal swab specimen processing methods influence performance of rapid semen detection tests: a cautionary tale.* Contraception, 2010. **82**(3): p. 291-5.
13. Juusola, J. and J. Ballantyne, *Multiplex mRNA profiling for the identification of body fluids.* Forensic Sci Int, 2005. **152**(1): p. 1-12.
14. Hanson, E.K., H. Lubenow, and J. Ballantyne, *Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs.* Anal Biochem, 2009. **387**(2): p. 303-14.
15. Danaher, P., et al., *Facile semi-automated forensic body fluid identification by multiplex solution hybridization of NanoString(R) barcode probes to specific mRNA targets.* Forensic Sci Int Genet, 2015. **14**: p. 18-30.
16. Erin Hanson, S.I., Cordula Haas, Jack Ballantyne, *Targeted multiplexed next generation RNA sequencing assay for tissue source determination of forensic samples.* Forensic Science International: Genetics Supplement Series, 2015. **5**: p. 441-443.
17. Luo, X.Y., et al., *MicroRNA markers for forensic body fluid identification obtained from miRCURY™ LNA array.* Forensic Science International: Genetics Supplement Series, 2015. **5**: p. e630-e632.
18. Silva, S.S., et al., *Forensic miRNA: potential biomarker for body fluids?* Forensic Sci Int Genet, 2015. **14**: p. 1-10.

19. Frumkin, D., et al., *DNA methylation-based forensic tissue identification.* Forensic science international Genetics, 2011. **5**(5): p. 517-24.

20. Lee, H.Y., et al., *Potential forensic application of DNA methylation profiling to body fluid identification.* International journal of legal medicine, 2012. **126**(1): p. 55-62.

21. Antunes, J., et al., *Tissue-Specific DNA Methylation Patterns in Forensic Samples Detected by Pyrosequencing(R).* Methods Mol Biol, 2015. **1315**: p. 397-409.

22. Antunes, J., et al., *High-resolution melt analysis of DNA methylation to discriminate semen in biological stains.* Anal Biochem, 2015. **494**: p. 40-45.

23. Sikirzhytski, V., A. Sikirzhytskaya, and I.K. Lednev, *Multidimensional Raman spectroscopic signature of sweat and its potential application to forensic body fluid identification.* Analytica chimica acta, 2012. **718**: p. 78-83.

24. McLaughlin, G. and I.K. Lednev, *In Situ Identification of Semen Stains on Common Substrates via Raman Spectroscopy.* J Forensic Sci, 2015. **60**(3): p. 595-604.

25. Zapata, F.l., M.Ã.n. FernÃ¡ndez de la Ossa, and C. GarcÃa-Ruiz, *Emerging spectrometric techniques for the forensic analysis of body fluids.* TrAC Trends in Analytical Chemistry, 2015. **64**: p. 53-63.

26. Mechthild Prinz., Y.T., Donald Siegel, Heyi Yang, Bo Zhou, Haiteng Deng. *Establishment of a Fast and Accurate Proteomic Method for Body Fluid/Cell Type Identification*. 2011 [cited 2012; Available from: https://www.ncjrs.gov/pdffiles1/nij/grants/236538.pdf.

27. Legg, K.M., et al., *Discovery of highly specific protein markers for the identification of biological stains.* Electrophoresis, 2014. **35**(21-22): p. 3069-78.

28. Cappellini, E., et al., *Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins.* Journal of proteome research, 2012. **11**(2): p. 917-26.

29. Lindgren, J., et al., *Microspectroscopic evidence of cretaceous bone proteins.* PLoS One, 2011. **6**(4): p. e19445.

30. Johnston, N.L., et al., *Multivariate analysis of RNA levels from postmortem human brains as measured by three different methods of RT-PCR.* Journal of Neuroscience Methods, 1997. **77**(1): p. 83-92.

31. *Recommended minimum criteria for the validation of various aspects of the DNA profiling process.* 2010, ENFSI DNA Working Group

32. Boyd, R., Basic, C, & Bethem, R, *Trace quantitative analysis by mass spectrometry*. 20008, West Sussex, England: Wiley.

33. *SWGDAM Validation Guidelines for DNA Analysis Methods*. 2012, Scientific Working Group on DNA Analysis Methods.

FINAL TECHNICAL REPORT (MAIN BODY)

I. Introduction

Statement of the Problem

Blood and semen factors that once held promise as discriminatory instruments for individualizing biological stains have been supplanted by DNA markers, which can be amplified from tiny amounts of biological material. While DNA analysis of an evidentiary swab may reveal the presence of a DNA profile consistent with an alleged victim, the DNA profile cannot indicate whether the DNA came from saliva, vaginal fluid, urine or a host of other sources. The ability to confidently associate a DNA extract with a specific tissue source or to accurately characterize mixed stains, however, can provide criminal investigators with critical information.

Consider the case of an alleged sexual assault where a DNA profile consistent with the victim is found on the mouth of a bottle in the suspect's possession. The victim states that the suspect used the bottle as a foreign object to penetrate her vaginally. The suspect counters that the alleged victim had drunk from the bottle and that no sexual contact occurred. Both stories may explain the presence of the victim's DNA on the bottle. The ability to reliably detect traces of vaginal fluid or potentially a mixture of both vaginal fluid and saliva in this case could help to either confirm or refute these opposing claims. A broad variety of other scenarios can easily be imagined where the ability to differentiate between menstrual and peripheral blood, or urine and saliva would have equally important probative value.

In short, a sensitive, reliable and human-specific approach for the confirmatory identification of both single-source and mixed-source biological stains could overcome the limitations of existing serological methods. This can be achieved by the development and validation of a multiplex assay for human body fluids based on a comprehensive panel of high-specificity protein biomarkers. That the assay employs a single, well-established analytical technology – targeted-ion mass spectrometry which is already employed in many forensic toxicology labs - will help to facilitate its adoption by caseworking laboratories.

Review of the Relevant Literature: Current Approaches to Stain Identification

While tests for the presence of blood, semen, saliva and urine have long existed[1-3], issues surrounding specificity and sensitivity have been long-standing concerns among forensic serologists. Some contemporary serological tests consume significant amounts of a valuable sample while failing to provide adequate sensitivity or specificity. For example, some tests for saliva may consume half of an evidentiary swab. Moreover, the detection of saliva is generally based on assays for the presence of the enzyme α-amylase (*i.e.*, salivary amylase)[36] activity. This requires the preservation of enzyme function – a factor that makes it difficult, and often impossible, to test aged and weathered material or items contaminated with substances that inhibit enzyme activity. Additionally, α-amylase activity is also present in a variety of non-salivary body fluids including human blood serum, urine and cervical mucus[37-39], albeit at much lower levels than in saliva. Being aware of the presumptive nature of this

test, forensic analysts are cautious and typically limit their interpretation to stating that "a positive amylase result is consistent with saliva".

Tests for evidence of vaginal contact have proven even more challenging. Over the years, this has involved attempts to identify vaginal epithelial cells in evidentiary samples. While the use of histochemical stains to detect sperm cells[40] is routine, staining to differentiate epithelial cells types (*e.g.*, skin, buccal and vaginal cells) has not been as successful. In the 1960s, the iodine-based Lugol's test held promise for identifying vaginal cells[4, 5]. This was based on studies suggesting that vaginal cells contained more glycogen than other epithelial cells. Rigorous studies, however, revealed that Lugol's positive cells were also present in the male urethra[6], male urine deposits[7] on >50% of penile swabs from males who had abstained from sex for several days and the oral mucosa [7]. Modifications to improve the reliability of Lugol's test have been suggested[8] but are not conducive for use with casework. Similarly, a modified Dane's staining technique is able to differentiate pure samples of vaginal, buccal and skin cells[41] but was unable to distinguish between a pure buccal cell sample and a mixed preparation of vaginal and skin cells. Given that forensic samples often contain cell mixtures, this limits the forensic utility of this approach. We are not aware of any public forensic laboratory that currently employs histological staining to reliably identify vaginal epithelial cells.

In contrast, the routine testing for blood and seminal fluid by forensic laboratories has been greatly facilitated by the development of commercial one-step immunoassay tests with good specificity and sensitivity. For example, the ABAcard® (Abacus Diagnostics) and HemeDirect (Seratec®) kits use the protein hemoglobin while the RSID™-Blood (Independent Forensics) uses the protein glycophorin A as markers for the presence of blood[3, 9, 10]. Similarly, the p30 protein serves as a marker for the presence of seminal fluid[3, 9, 42]. However, p30 can also be found in female ejaculate[43], breast milk[44], urine[45] and other non-target fluids (albeit at lower concentrations). Semenogelin is employed as a high-specificity marker by the RSID™-Semen kit[46]. As with any immunoassay, however, results are "presumptive by definition" because the potential for antibody cross-reactivity with non-target molecules (although remote) can never be eliminated[11]. Moreover, casework-type samples may include environmental contaminants or other factors related to sample processing that can either inhibit antibody binding, thereby reducing assay sensitivity[12] or promote non-specific antibody binding resulting in a false positive result[47]. Only the direct visual identification of sperm cells by microscopy enables an analyst to report a confirmatory result[48], It is often difficult and laborious, however, to locate sperm cells in close association with epithelial cells or non-cellular debris. Fluorescence microscopy can facilitate sperm identification [48] but microscopy in general is useless for analyzing samples from vasectomized or otherwise aspermatic males or with degraded material lacking detectible sperm heads.

For a range of other body fluids, forensically-validated commercial kits based on body fluid specific antigens are lacking entirely. This often leaves the forensic analyst without the ability to make a substantive statement about the potential tissue source of a DNA profile. Part of the reason for this is that much less has historically been known about proteins that might have potential utility as markers for other forensically-relevant body fluids but which are not abundant. Since traditional protein detection

methods often lacked the sensitivity required to detect low-abundance biomarkers in casework samples, the forensic serology community has not been able to take advantage of these potential biomarkers.

Review of the Relevant Literature: Emerging Approaches to Stain Identification

Several novel approaches to biological stain identification are have been explored over the past several years. These research efforts have as their goal the development of a more sensitive and uniform strategy for analyzing body fluids capable of providing analysts with confirmatory results. Emerging approaches include biological stain identification based on messenger- and micro-RNA expression profiles, epigenomic modifications, Raman spectroscopy and protein-biomarker detection by mass-spectrometry. Each of these proposed methods have their own strengths and weaknesses. Accordingly, some of these are potentially complementary technologies that will make it possible for analysts to obtain useful information from a much larger range of casework samples[49]. These emerging strategies also offer an opportunity for greater standardization and automation to forensic serological testing – an advance that is akin to the progress achieved over the past couple of decades in DNA profiling.

mRNA Markers amplified by reverse transcription PCR and detected by capillary electrophoresis are being studied as a means of identifying body fluids on the basis of differential expression profiles[13]. For example, matrix metalloproteinase mRNA transcripts from the endometrium have been investigated as a marker for menstrual blood[50]. In 2007, a multiplex assay for identifying blood, saliva, semen, and menstrual blood was developed based on mRNA markers[51]. Because of its compatibility with existing DNA amplification technology, mRNA profiling as a means of identifying body fluids has attracted significant research interest in recent years[15, 16, 52]. The presumed sensitivity of mRNA to degradation has often been raised as potential concern with this approach. An in-depth study of RNA recovery under a variety of conditions, however, found that RNA remained stable in samples that had been kept dry and could be recovered after 180 days of storage[53] while samples exposed to rain were unrecoverable after one to seven days. Thus, the utility of RNA-based markers in an operational environment is largely a function of the original integrity and storage conditions of the evidentiary material.

microRNA Markers are non-coding molecules involved in post-transcriptional regulation of gene expression. Because of their short lengths (generally <25nt) and evidence of tissue-specific expression patterns[54], they have been explored as promising markers for the characterization of more highly degraded samples where longer mRNA targets might be difficult to amplify. Subsequent studies of candidate miRNA markers have often revealed low-level expression in non-target tissues or lack of tissue-specific reproducibility between studies[14, 54], researchers have shifted their attention to the use of quantitative PCR combined with mathematical approaches that may allow a target stain to be identified on the basis of a broader miRNA expression profile rather than on the absolute presence or absence of a given marker[18, 55]. While continuing to hold promise, it has also been pointed out that the use of miRNAs for multiplex biological stain assays may be technically difficult due to limitations on the number of fluorescent tags currently available for quantitative PCR assays.

Epigenetic Markers rely on tissue-associated differences in DNA methylation patterns[21, 22, 56, 57] as a means of identifying different biological stains. As with RNA markers this approach employs pattern

analysis but has the advantage of making it possible to directly "query" the DNA in a sample to determine the tissue from which it originated. Initial studies of epigenetic markers have demonstrated the potential utility of the approach using semen, saliva and skin tissue[19, 20]. Potentially complicating the use of epigenetic assays, though, is the observation that while tissue-specific methylation differences can be identified within an individual, significant inter-individual epigenomic variation in these tissue-specific patterns also exists[58]. In addition, global changes in DNA methylation are associated with cancers[59] and other diseases[60] and an emerging body of evidence points to environmental factors that may impact DNA methylation patterns[61]. Finally, similarities in developmentally related tissues (*e.g.*, the male prostate and female periurethral glands[43]) may complicate the interpretation of some assay results.

Raman Spectroscopy is an approach to body fluid identification based on the inelastic scattering of laser light as it interacts with proteins and other molecules present in a sample[24, 25]. In an effort to accommodate sample heterogeneity, a multidimensional "spectroscopic signature" is created and advanced statistical analysis is used to search for the best match between an expected "spectroscopic signature" and that of a questioned sample. Promising results with single source stains have been reported using this approach[23] which has the advantage of being non-destructive and rapid. Because of its reliance on statistical pattern fitting, however, it is unclear to what extent this strategy can accommodate more challenging mixed stains such as those containing contaminants or that are degraded, *i.e.*, any forensic type sample that deviates substantially from the reference "spectroscopic signature".

Protein Biomarkers have attracted significant interest in recent years due in large part to the strides that have been made in the tools to identify and characterize them. It is now possible to rigorously map entire proteomes with high reproducibility using automated 2-dimensional HPLC systems or MudPIT (multidimensional protein identification technology) to identify potentially useful biomarkers. Once identified, mass-spectrometry-based targeted-ion assays facilitate the unambiguous detection and quantitation of even low abundance proteins, against a background of other non-target molecules. This has resulted in a wealth of new opportunities to develop protein-based assays for medical and forensic applications such as body fluid identification.

In addition to the protein biomarker-based multiplex assay described in this application, other researchers have also reported success in this area. Using a panel of biomarkers identified through a literature search and empirical studies, a multiplex assay for blood, saliva and semen was developed on a MALDI-TOF mass spectrometer. The assay characterized single and mixed biological stains in the nanoliter range and worked well with forensic type samples aged up to 20 months[26]. Unfortunately, the candidate biomarkers for detection of menstrual blood and vaginal fluid were not detected by MALDI-TOF. Ongoing efforts are directed at identifying new markers for these fluids. As with other types of markers, it is also recognized that protein profiles may be altered by biological perturbations due to disease and interindividual variability.

One of the significant advantages of a protein biomarker approach is the tremendous diversity of potential targets that are made possible due to post-translational modification in different tissues. As a

result, a single protein may be differentially modified by one's metabolism in two different body fluids, making it possible to develop highly specific assays in cases where epigenomic patterns or mRNA expression profiles might not differ. Another key advantage is the stability of many proteins under conditions that lead to degradation of other molecules. Proteins are among the most long-lasting of all biological molecules having been routinely isolated from ancient biological material[28, 29]. The oldest recognized protein sequence is that of a collagen protein isolated from a 3.4 million year old high arctic camel[35] while the oldest reported DNA sequence is from a 700,000-year-old horse[34]. In a more forensically applicable study, a 99.5% decrease in mRNA levels was observed in post-mortem brain tissue while protein levels remained relatively constant[30]. Still, as is the case with all biological molecules, proteins do fragment and degrade over time. The use of protein biomarkers, however, can be readily adapted to detect protein fragments. Thus even partially degraded target biomarkers may be detected[62].

Statement of Fundamental Hypotheses and Core Research Objectives

Fundamental Hypotheses: Major advances in versatility, detection sensitivity and ease of use make triple quadrupole (QQQ) mass spectrometers the instruments of choice for clinical and analytical testing laboratories that require high-throughput analysis of complex or environmentally "dirty" samples. The current research was initiated to determine the extent to which this technology might prove equally useful to forensic case-working laboratories. The central goal of the current research project was to developmentally validate the performance parameters of a targeted ion mass-spectrometry assay for the confirmatory identification of six human body fluids of forensic utility (*i.e.*, saliva, semen, peripheral blood, menstrual/vaginal fluids, and urine). This can complement the use of DNA profiling by making it possible to more accurately and confidently associate a DNA sample with a specific type of biological stain. The lack of this type of confirmatory assay has long presented forensic serologists with significant challenges in their effort to bring greater clarity and objective confidence to the evidence that they are asked to analyze in many criminal cases.

It was specifically hypothesized, therefore, that:

- A prototype multiplex QQQ mass spectrometry-based assay, which has already been developed, can be fully optimized to accurately detect the presence of up to six human body fluids (saliva, semen, peripheral blood, menstrual blood, urine and vaginal fluid).
- Through a rigorous series of developmental validation studies, the accuracy, reliability, repeatability, human specificity and sensitivity limits of this multiplex assay will exceed the performance of biological stain characterization assays currently employed in forensic laboratories.

Core Research Objectives: The current research program has built on the successful completion of previous NIJ funded projects that have charted a course from basic research to practical application. This work began with the rigorous comparative proteomic mapping of thousands of proteins from six forensically relevant human body fluids which made it possible to identify a panel of candidate high-specificity protein biomarkers for each stain. Following biomarker discovery the specificity of each

candidate protein, the consistency with which it can be detected and the degree of interindividual variability in its expression was evaluated across a larger population of human subjects. For these early analytical studies, a quadrupole time of flight (Q-TOF) mass spectrometer was used. This platform, however, would be unacceptably slow for most practical applications in a case-working environment. At the inception of this project, it was thought that shifting to a higher-sensitivity QQQ platform, however, would resulted in both higher-quality results and faster assay times. This was demonstrated in pilot studies where a three-stain (*i.e.*, saliva, semen, and vaginal fluid) multiplex assay was developed. The central goal of the current research project was to fully develop a QQQ multiplex by incorporating the biomarkers for all six body fluid into a single unified assay; to thoroughly assess its performance limits and thus its potential applicability to casework. The following four core research objectives were central to successfully achieving this goal.

5) **Select diagnostic target ions for existing biomarkers** to complete a six-stain multiplex QQQ assay and demonstrate their accurate detection using single source reference samples.
6) **Optimize the performance of the six-stain multiplex QQQ assay** using synthetic standards as well as single-source and mixed-source reference samples.
7) **Conduct a rigorous developmental validation** of the multiplex assay that meets Standard 8.2 of the FBI's "Quality Assurance Standards for Forensic DNA Testing Laboratories".
8) **Develop appropriate Standard Operating Procedures and Interpretation Guidelines** for use of the multiplex QQQ assay for casework samples.

The successful completion of these objectives not only represents the culmination of work completed under previous NIJ Research and Development awards but it has now helped to pave the way for interlaboratory evaluation and adoption by forensic practitioners.

II. Methods

Human Subjects

The University of Denver Institution Review Board for Research Involving Human Subjects (IRB) reviews all research involving human subjects, regardless of funding source, to ascertain that the rights and welfare of subjects are being protected. The IRB is responsible for assuring that recruitment advertising is not misleading or coercive to the research subject. All projects using human subjects are reviewed no less than annually.

All research conducted under DNA Research and Development Award 2012-DN-BX-K035 was IRB reviewed, approved and conducted in full compliance with U.S. Federal Policy for the Protection of Human Subjects (Basic DHHS Policy for Protection of Human Research Subjects; 56 FR 28003). This applies to all research and data analysis activities conducted either at the University of Denver or at the laboratory at the Center for Forensic Science Research and Education in Willow Grove, Pennsylvania.

The purpose and significance of the research and the methods that would be used to collect body fluid samples was thoroughly explained to each volunteer. All participants then signed a statement of informed consent to participate in the research. Recruitment notices were posted at the Center for Forensic Science Research & Education where Arcadia University graduate students in forensic science and NMS Labs employees share a common cafeteria space. The student traffic in this area consists primarily of forensic science-oriented graduate and undergraduate students. As no health care associated information was collected, HIPAA authorization was not required.

Body fluid samples were collected from a total of 43 adult (>18 y.o.) human volunteers (20 males; 23 females). Study participants were recruited from among the Arcadia University student population and NMS Labs employees. While the study participants reflected the ethnic and age diversity of the Arcadia University student population and NMS Labs employees, there were an insufficient number of study participants to enable a statistically substantive partitioning of the sample on the basis of biogeographic origin or broad age cohorts. It should be emphasized, however, that while such detail was beyond the scope of the current project, there may be value in addressing this in the future.

An important consideration in determining an appropriate sample size for individual experiments was the impact of the number of body fluid samples analyzed and the ability to reliably capture any significant variability in assay performance. For an infinite population under an assumption of a standard normal distribution, the 95% confidence interval is 1.96 standard deviations. By comparison, the 95% confidence interval for a sample size of 5, 10, and 50 would be 2.78, 2.26 and 2.01 respectively. Balancing the importance of capturing statistical variability with the time and financial limitations of the project, therefore, it was determined that a sample sizes of 10 body fluid samples per experiments would make it make it possible to reliably detect significant variability in study results. This is also consistent with the recommendations of the European Network of Forensic Science Institutes (ENFSI) DNA Working Group on the recommended minimum number of samples for forensic validation studies[31].

Body Fluid Collection and Protein Extraction

A total of 20-40 samples of each of six forensically-relevant body fluids (*i.e.*, peripheral and vaginal/menstrual fluids, semen, saliva and urine) were collected for protein analysis. It was determined that a sample sizes of 10 body fluid samples per experiment would make it make it possible to reliably detect significant variability in study results. This is consistent with the recommendations of the European Network of Forensic Science Institutes (ENFSI) DNA Working Group on the minimum number of samples for forensic validation studies[31].) The statistical basis for this is that based on a student's t-distribution, ±2.576 standard deviations around the mean from a sample size of 10 captures 96.72% of the variability of the population. By comparison, an infinite sample size would capture 99% of the variability at ±2.576 standard deviations.

It should be emphasized the variability that is being assessed is that associated with the performance of the mass spectrometry assay and not the interindividual variability associated with biomarker expression in human populations. Prior NIJ award 2009-DN-BX-K165 documented and evaluated interindividual variability in protein biomarker detection and expression levels using 50 human subjects

for each of six body fluids. In that study, a sample size of 50 was considered statistically sufficient because: 1) all biomarkers for each body fluid were detected in all 50 subjects and 2) based on a student's t-distribution, ±2.576 standard deviations around the mean from a sample size of 50 captures 98.69% of the variability of the population. An infinite sample size by comparison would capture 99% of the variability at ±2.576 standard deviations. By contrast, many studies of DNA markers are performed to determine the frequency of allelic markers in the general population. DNA population studies, therefore, typically employ larger samples sizes because not all alleles are expressed in all individuals.

The choice of the bodily fluids to be analyzed and the size of the study population reflected discussions with forensic practitioners at caseworking laboratories including forensic serologists from NMS Labs which is an ASCLD\LAB-*International* accredited facility. In addition, independent experts in sexual assault examination have helped to guide this research to best meet the needs of the practitioner community. The procedures employed for sample collection were in accordance with the NIH guidelines.

Saliva: Donors were directed to gently brush their teeth and thoroughly rinse their mouth with sterile water to remove residual food particles. After 5 minutes to allow secretion of saliva, the donor was instructed to place a Sarstedt Salivette™ saliva collection sponge into their mouth and to gently chew and roll the sponge around in their mouth for 3-4 minutes. The sponge was then placed into a sterile plastic conical tube. This allowed for the collection of large quantities of relatively pure saliva while reducing protein contamination from food items. Salivette™ sponges were centrifuged for 2 min at 1500 x *g* at 4°C to recover saliva which was transferred to 15 ml conical vials.

Seminal Fluid: Donors were directed to refrain from sexual activity for a minimum of 24 hours and then to obtain a 3-6ml sample of seminal fluid by masturbation in the privacy of their home. The subject was requested to directly deposit the fluid into a sterile plastic collection cup provided by the laboratory and then to refrigerate the sample until it could be transported to the lab at the donor's earliest convenience (within 1 hour). Semen was then incubated at room temperature for at least 30 minutes to allow it to liquefy and aliquoted to a 15 ml conical vial.

Peripheral Blood: Donors were escorted to the Student Health Center where a 10-15ml sample of whole blood was obtained by a certified nurse using venipuncture. The blood was drawn into a sterile vacuum tube containing an anticoagulant.

Urine: Donors were directed to deposit a midstream urine sample (>50ml) into a sterile collection cup provided by the laboratory.

Vaginal Fluid: Following clinically accepted procedures, vaginal secretions were self-collected by study participants. The collection protocol employed an FDA-approved over-the-counter, hypoallergenic cup (SoftCup™). The device is similar to the hypoallergenic menstrual cup which is used as a tampon replacement during menses. For the collection of vaginal secretions, donors were instructed to insert the Softcup™ for periods of up to 12 hours and then transfer the secretions into a 50 ml sterile container.

Donors were directed to refrigerate the sample until it could be transported to the lab at their earliest convenience (typically within 1 hour). Upon receipt, the liquid was transferred to 50 ml conical vials.

Menstrual Fluid: Following clinically accepted procedures, menstrual blood was self-collected by study participants in the privacy of their home. The collection protocol employed an FDA-approved over-the-counter latex-free, hypoallergenic cup (DivaCup™) for the collection of menstrual flow. The donor was directed to insert the cup into the vagina for the first evening of menses in addition to the day before and the day after. After menses had started, the cup remained in place for up to one hour at a time. The cup was then gently removed; the contents were poured into a sterile 50ml conical tube and refrigerated until delivered to lab (within 1 hour). Blood serum was removed to a 15 ml conical vial. It should be noted that while some menstrual blood samples were still liquid and relatively uncoagulated upon their delivery to the lab, other samples contained significant clots. These samples were homogenized to break up the clots so as to facilitate sample processing. The disaggregation of the clotted material did not have an effect on the outcome of the assay.

Protein Quantization

Protein quantization is valuable to ensure that optimal quantities of protein are available for downstream analyses by mass spectrometry. Samples with too much protein or too little protein will compromise the quality and reliability of the assay results. The Thermo Scientific Pierce Micro BCA Protein Assay (Thermo Fisher Scientific, Rockford, IL) was used to determine protein concentration of each solubilized sample of whole body fluids. Manufacturer recommended protocols were followed for sample set up and analysis. Seven standards (1,500, 1,000, 750, 500, 250, 125, and 25 µg/ml) were analyzed in duplicate and averaged to create a standard quadratic curve. The absorbance of standards and research samples was measured on a BioTek Instruments 96 well microplate reader set to analyze at a single wavelength of 562 nm.

Mass Spectrometry

All biological samples were deposited directly into 1.5ml microcentrifuge tubes or onto sterile fiber tipped swabs or other substrates for analysis. Up to 50 µl of blood and up to 125 µl of remaining biological fluids were used per sample. Dried stains were resolubilized by soaking in 400 µl of diH_2O for 30 minutes. Samples were vortexed frequently to facilitate extraction of biological material from the substrate. Sample substrates were then transferred into clean spin baskets and centrifuged at 16,800 x *g* for 10 minutes. For samples containing excessive quantities of blood, 400 µl of HemogloBind was added. Samples were vortexed for 30 seconds and mixed via inversion for 15 minutes prior to two centrifugation steps at 4200 x *g* for 2 minutes each. For samples containing suspected denaturants as indicated by a failed IPC (i.*e*. urea, detergents, etc.), 1.2 ml of acetone were added. Samples were vortexed, stored at -20 °C for 30 minutes then centrifuged in a refrigerated microcentrifuge at 4 °C at 12,400 x *g* for 10 minutes. An additional 600 µl of acetone was added to pelleted material and samples were stored at -20 °C for 15 minutes prior to centrifugation in a refrigerated microcentrifuge at 4 °C at 12,400 x *g* for 10 minutes. To resolubilized pelleted protein, 150 µl of 50 mM ABC was added and samples were placed in a thermomixer set at 30 °C and 850 RPM for 15 minutes. Samples underwent a

final centrifugation step in a refrigerated microcentrifuge at 4 °C at 12,400 x *g* for 10 minutes and resulting supernatant was transferred to a clean 1.5 ml microcentrifuge tube for analysis.

Based on protein quantization results, up to 20 µg total protein was transferred to a 1.5 ml low retention microcentrifuge tube and lyophilized in a vacuum evaporator to dryness. Dried protein samples were reconstituted in 15 µl of 50 mM ammonium bicarbonate (ABC), 15 µl neat 2,2,2-triflouroethanol (TFE) and 1 µl 200 mM dithiothreitol (DTT) reducing agent and then shaken in a thermomixer set at 60 °C and 850 RPM for 30 minutes. The proteins were then alkylated by the addition of 1.5 µl of 200 mM IAA (Iodoacetamide) and the sample was shaken in the dark for 30 minutes at room temperature. The proteins were brought to volume by the addition of 250 µl 50 mM ABC and digested overnight with trypsin at 37°C. Digested samples were then lyophilized in a vacuum evaporator and resuspended in 3% acetonitrile and 0.1% formic acid.

A liquid chromatography (LC) mass spectrometer assay was developed to simultaneously scan for the presence of six different body fluids in 30-minutes. This assay targets a total of 26 individual precursor ions consisting of 6 peripheral blood peptides, 5 saliva peptides, 6 seminal fluid peptides, 2 urine peptides, and 7 vaginal/menstrual blood peptides.

Mass spectrometry was performed on an Agilent Technologies HPLC-chip/MS system (http://www.agilent.com/en-us/video/hplc-chip-technology) coupled to an Agilent 6430 Quadrupole Mass Spectrometer. The HPLC chip column used was a 150mm 300 A C18 Analytical with a 160 nl enrichment column. Samples run on the chip system were first injected onto the trap/enrichment column where they were washed and concentrated before they were back flushed onto the analytical/running column for analysis. Columns were equilibrated in 0.1% Formic acid in water. Run conditions employed buffer A (0.1% formic acid in water) and B (90% Acetonitile, 10% water, 0.1% formic acid). An initial 30 minute run employed a gradient of 3% B to 35% B over 24 minutes. This was followed by 90% B to flush the column and then reequilibration at 3% A. Following sample acquisition, the column was equilibrated for 6 minutes to initial conditions. Nanoflow LC runs using microfluidic chip technology generally exceed one hour which makes the current method of only 24 minutes quite fast by comparison.

Data Analyses

Data analysis was performed using Skyline software suite distributed freely by the MacCoss Lab at the University of Washington. Proteins are composed of a sequence of amino acids arranged in a linear order. This allows for the prediction, to a certain degree of confidence, the fragmentation pattern and MS/MS spectra that will be produced. To enhance the specificity of the method, two MRM transitions for each peptide are employed. Detectable peptide peaks are those that meet the following criteria: (1) The peak has a signal to noise ratio greater than 3. (2) The peak for the natural peptide should have the same retention time as the corresponding heavy labeled peptide ± 1%. (3) The ion response ratio for the natural peptide should be equal to that of the ion response ratio of the corresponding heavy labeled internal standard peptide ± 20%.

Casework Type Samples

The applicability of a mass-spectrometry based body fluid assay to samples encountered in a forensic context was assessed using a series of casework type samples (Appendix C). Specifically, the ability of the biomarkers to be detected in body fluid samples recovered from a variety of substrates including cotton, denim, leather, metal, glass, plastic, sanitary napkins and styrofoam was tested. Similarly the impact of exposure to environmental contaminants/insults was assessed. For these assays, 5-10µl aliquots of bodily fluids applied to sterile cotton tipped applicators that had previously been dipped in such agents as 10% bleach, neat bleach, soil, detergent, spermicidal lubricants, chewing tobacco, soda and lotion were used. Swabs to simulate sexual assault type evidence were assessed including oral swabs, rectal swabs, vaginal swabs, penile swabs and finger swabs. Finally, a series of aged body fluids were analyzed ranging from 2 to 7 years.

III. Results and Discussion

Selection of Optimal Diagnostic Target Ions for Existing Biomarkers

QQQ-MRM is considered to be the gold standard for modern quantitative analyses[32] particularly in the context of high-throughput protein analysis. The use of the QQQ-MRM platform (Figure 1) provides analysts with high confidence in the accuracy of the results obtained for a given stain. This is because each individual body fluid body is identified based on the presence of multiple biomarker proteins (*e.g.*, Statherin and Submaxillary Gland Androgen Regulated Protein for saliva). The presence of each biomarker protein, in turn, is based on the isolation of multiple peptide cleavage products (*i.e.*, precursor ions). The presence/identity of each precursor ion, in turn, is confirmed by detection of its fragmentation products (*i.e.*, product ions). This internal confirmation and reconfirmation stands in contrast to existing forensic assays where identification is typically based on a single binding event between an antibody and its presumed target protein.

QQQ-MRM, assay development requires the rigorous characterization of each precursor-product ion pair to identify those best suited for use in a multiplex assay. This process employed a database of preexisting targeted-ion Q-TOF data collected under award 2009-DN-BX-K165. Using data generated from that study, optimal biomarker proteins, peptides and transitions were evaluated to identify those that has a unique fragmentation pattern, were abundant, efficiently ionized and had a mass to charge ratio greater than that of the peptide. Tables 1-5 present a comprehensive list of target biomarker peptides and transitions for each of six human body fluids (*i.e.*, urine, semen, saliva, vaginal fluid, menstrual fluid and peripheral blood). It should be noted that while menstrual and vaginal fluids are distinct body fluids in theory, menstrual fluid is are always mixed with vaginal fluid in practice. Thus, a single vaginal/menstrual fluid panel was used.

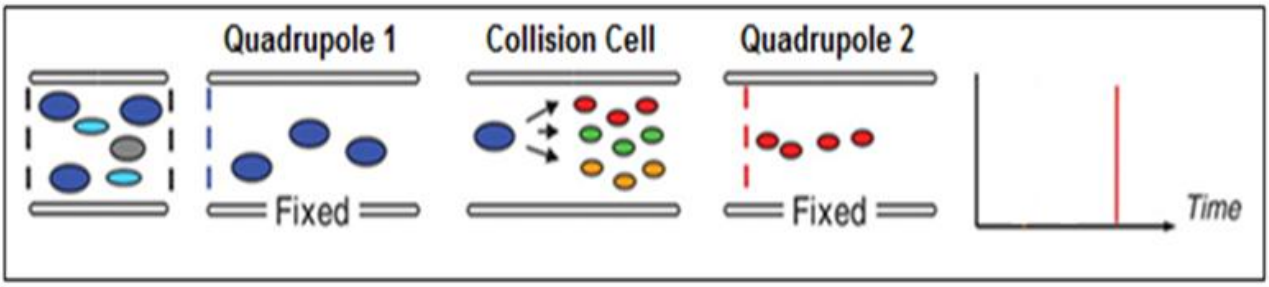


Figure 1 Multiple Reaction Monitoring (MRM) technique on a triple quadrupole mass spectrometer (QQQ). Precursor ions are targeted in the first quadrupole mass filter (Q1), fragmented in the collision cell, and then individual product ions are isolated in the second quadrupole mass filter (Q2). (Image modified from Domon, B. & Aebersold, R. (2006) Mass spectrometry and protein analysis. Science, 312, 212-217).

The targeted-ion inclusion lists compiled for each individual fluid were then used to build a dynamic, retention time restricted, QQQ-MRM assay for each body fluid. Five single-source references samples for each target body fluid were then prepared and analyzed using the cognate targeted-ion inclusion lists. Peak shape, abundance and retention time were monitored over in order to confirm the unambiguous detection of the each precursor-product ion pair. This information was used to evaluate the reliability with which transitions were detected. Peptides with interfering signals or those with low response were eliminated. In some cases, this necessitated the selection of additional protein markers that were more reliable.

Table 1 **Urine Biomarker, Peptide and Transition List**

Fluid	BioMarker	Peptide Sequence	Charge State	m/z	Targeted Ions
Urine	Uromodulin	TLDEYWR	2	491.7	
					[y5] - 768.3311
					[y4] - 653.3042
		STEYGEGYA[Cys(CAM)]DTDLR	2	868.9	
					[y11] - 1256.5212
					[y10] - 1199.4997
					[y9] - 1070.4571

Table 2 Seminal Fluid Biomarker, Peptide and Transition List

Fluid	BioMarker	Peptide Sequence	Charge State	m/z	Targeted Ions
Seminal Fluid	Prostatic Acid Phosphatase	ELSELSLLSLYGIHK	3	568	
					[y8] - 930.5407
					[y7] - 817.4567
					[y6] - 730.4246
		FQELESETLK	2	612.3	
					[y8] - 948.4884
					[y7] - 819.4458
					[y6] - 706.3618
	Prostate Specific Antigen (PSA)	LSEPAELTDAVK	2	636.8	
					[y10] - 1072.5521
					[y9] - 943.5095
					[y7] - 775.4196
		IVGGWE[Cys(CAM)]EK	2	539.3	
					[y8] - 964.4193
					[y7] - 865.3509
	Semenogelin-2	DIFTTQDELLVYNK	2	849.9	
					[y11] - 1323.6791
					[y10] - 1222.6314
					[y8] - 993.5251
		DVSQSSISFQIEK	2	734.4	
					[y9] - 1038.5466
					[y8] - 951.5146
					[y6] - 751.3985

Table 3 Saliva Biomarker, Peptide and Transition List

Fluid	BioMarker	Peptide Sequence	Charge State	m/z	Targeted Ions
Saliva	Statherin	FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF	3	1215.2	
					[y13] - 1687.7751
					[y8] - 1074.4891
					[b11] - 1229.5626
	Submaxillary Gland Androgen Regulated Protein	IPPPPPAPYGPGIFPPPPPQP	3	710.7	
					[y7] - 729.3930
					[b10] - 987.5298
					[b12] - 1141.6041
		GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR	3	1034.5	
					[b9] - 850.4458
					[y12] - 1228.6473
					[b12] - 1172.6099
	Amylase	LSGLLDLALGK	2	550.3	
					[y10] - 986.6
					[y9] - 899.6
					[y7] - 729.5
		IAEYMNHLIDIGVAGFR	3	640.3	
					[y9] - 947.5
					[y16] - 903.5
					[y15] - 867.9

Table 4 Vaginal/Menstrual Fluid Biomarker, Peptide and Transition List

Fluid	BioMarker	Peptide Sequence	Charge State	m/z	Targeted Ions
Vaginal/ Menstrual Fluid*	Cornulin	ISPQIQLSGQTEQTQK	2	893.5	
					[y11] - 1247.6226
					[y10] - 1119.5640
					[y9] - 1006.4800
		GQNRPGVQTQGQATGSAWVSSYDR	3	850.7	
					[y11] - 1228.5593
					[y10] - 1127.5116
					[b10] - 1066.5388
	Matrigel-induced Gene C4 protein	DGVTGPGFTLSGSC[Cys(CAM)]C[Cys(CAM)]QGSR	3	971.9	
					[y11] - 1212.5
					[y10] - 1111.5
					[y9] - 998.4
		GC[Cys(CAM)]VQDEFC[Cys(CAM)]TR	3	636.3	
					[y8] - 1054.5
					[y7] - 955.4
					[y6] - 827.3
	Suprabasin	ALDGINSGITHAGR	3	461.2	
					[y9] - 912.5
					[y8] - 798.4
					[y12] - 599.3
					[y11] - 541.8
	Neutrophil Gelatinase-Associated Lipocalin	SYPGLTSYLVR	2	628.3	
					[y9] - 1005.5728
					[y8] - 908.5200
					[y6] - 738.4145
		WYVVGLAGNAILR	2	716.4	
					[y11] - 1082.6681
					[y10] - 983.5996
					[y9] - 884.5312

* Although menstrual and vaginal fluids are distinct body fluids in theory, menstrual fluids markers are always mixed with vaginal fluid markers in practice. Thus the protein biomarkers for these two fluids are combined into a single vaginal/menstrual fluid panel.

Table 5 Peripheral Blood Biomarker, Peptide and Transition List

Fluid	Protein	Peptide Sequence	Charge State	m/z	Targeted Ions
Peripheral Blood	Alpha-1 Antitrypsin	LSITGTYDLK	2	555.8	
					[y9] - 997.5
					[y8] - 910.5
					[y7] - 797.4
					[y6] - 696.4
		SVLGQLGITK	2	508.3	
					[y7] - 875.5098
					[y6] - 761.4668
					[y5] - 662.3984
	Hemopexin	NFPSPVDAAFR	2	610.8	
					[y9] - 959.4945
					[y8] - 862.4417
					[y7] - 775.4097
		GGYTLVSGYPK	2	571.3	
					[y8] - 864.4825
					[y7] - 763.4349
					[y6] - 650.3508
	Hemoglobin subunit beta	GTFATLSELH[Cys(CAM)]DK	2	739.9	
					[y9] - 1102.5197
					[y8] - 1001.4721
					[y7] - 888.3880
		SAVTALWGK	2	466.8	
					[y7] - 774.4509
					[y6] - 675.3824
					[y5] - 574.3348

Performance Optimization of the Multiplex QQQ Assay

From a forensic practitioner's perspective, one of the critical advantages of the QQQ-MRM approach to multiplex assays is that this platform is specifically engineered for rapid and efficient data acquisition. This makes it possible to shorten the assay run times by approximately 50% relative to competing mass spectrometry platforms such as a Q-TOF mass spectrometer. The ability to significantly shorten sample turnaround time has a profound effect on the ability of forensic testing labs to take advantage of the analytical strengths of this assay. It results in a more accessible price point/sample and level of throughput. To achieve the shortest possible run time without sacrificing assay sensitivity or reliability, however, requires a detailed one-time front-end optimization of the multiplex method beginning with the chromatographic separation process.

Chromatographic Separation: Optimizing chromatographic separation makes it possible to shorten the overall QQQ-MRM assay time. While many factors that can impact the chromatographic fractionation, the front-end separation by liquid chromatography is typically the most important factor as the complexity of the sample increases. Chromatographic optimization, therefore, necessitated a thorough performance evaluation of several HPLC Chip options. Table 6 provides a comparison of the features and capabilities of commercially available HPLC-chips for proteomics and bioanalytical applications. Each was evaluated on the basis of reproducibility, separation efficiency and detection sensitivity.

Table 6 Commercially-Available HPLC-Chips for Proteomic/Bioanalytic Applications

		ProtID-Chip-43	ProtID-Chip-150	Large Capacity Chip	Polaris-HR-Chip
	analytical column length (mm)	43	150	150	150
stationary phase	packing material	C18 SB-ZORBAX	C18 SB-ZORBAX	C18 SB-ZORBAX	Polaris C18-A
	particle length (A)	300	300	300	180
	particle diameter (µm)	5	5	5	3
	enrichment column (nl)	40	40	160	360
	loading capacity (µg protein)	1	1	4	4

In general, a longer analytical column allows for better separation of peaks while a larger enrichment column allows for increased loading capacity. The ProtID-Chip-150 was found to achieve faster run times though salivary peptides were ineffectively bound by the smaller 40 nl enrichment column. This resulted in a nearly 20-minute assay with just two minutes worth of unusable data (Figure 2). The performance of the Polaris-HR-Chip (Figure 3) produced the narrowest peaks, however, salivary peptides still overloaded the column. These limitations were resolved by use of the Large Capacity Chip (Figure 4). This chip provided for improved sample partitioning and yielded usable chromatographic data throughout the entire assay run time.

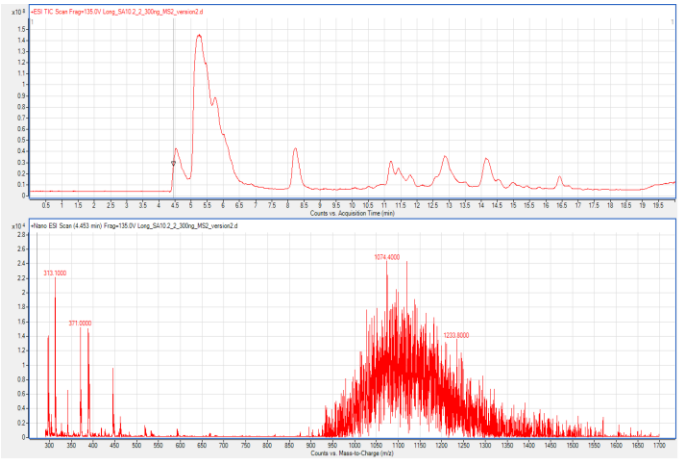


Figure 2 MS2 scan of salivary proteins on the ProtID-Chip-150. **(Top)** Overloaded peaks appear starting at 4.5 minutes. **(Bottom)** Irresolvable mass spectral data resulting from overloaded peaks.

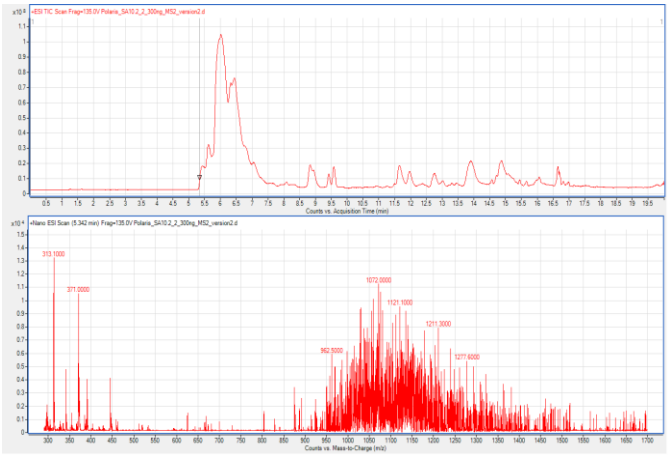


Figure 3 MS2 scan of salivary proteins on the Polaris-HR-Chip. **(Top)** Overloaded peaks appear starting at 5.5 minutes. **(Bottom)** Irresolvable mass spectral data resulting from overloaded peaks.

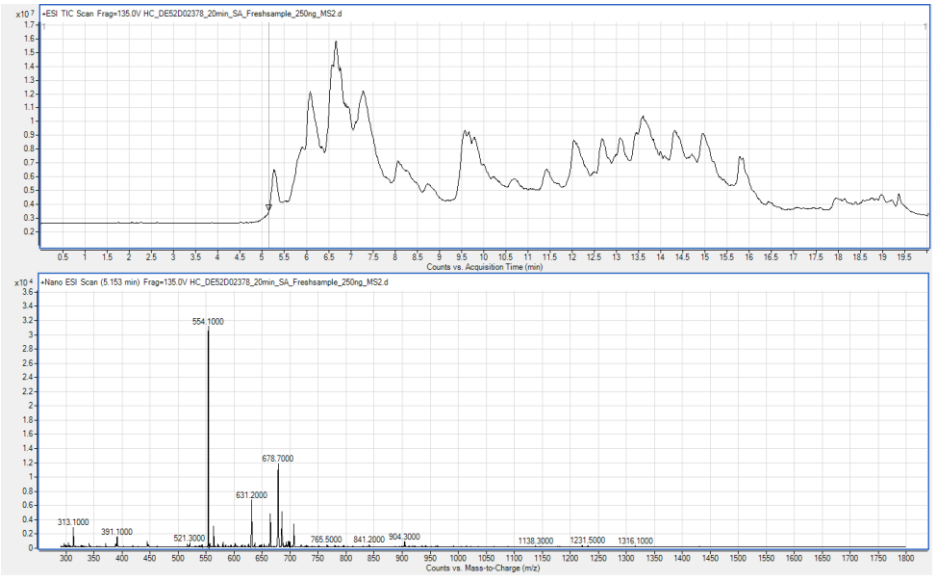


Figure 4

(Top) MS2 scan of salivary proteins on the Large Capacity Chip characterized by the absence of overloaded peaks.

(Bottom) The higher quality chromatography results in a more readily resolvable set of mass spectral data.

The best peak morphology was achieved using the Polaris-HR-Chip. Fluctuation in retention times (illustrated for serum albumin in Figure 5 but observed across all body fluids), however, raised concern with respect to reproducibility. In addition, an inability to detect statherin (a high-specificity protein biomarker for saliva) (Figure 6) suggested the likelihood of chemical incompatibility between the stationary phase and some of the target ions required for the multiplex assay. MRM analyses of the same sample on the Large Capacity Chip, by contrast, yielded a high-quality peak for statherin (Figure 7). The Large Capacity Chip also provided more balanced overall performance in terms of resolution and peak morphology without clogging. The Large Capacity Chip was therefore identified as having the best overall performance and thus was the best option for development of the multiplex body fluid assay.

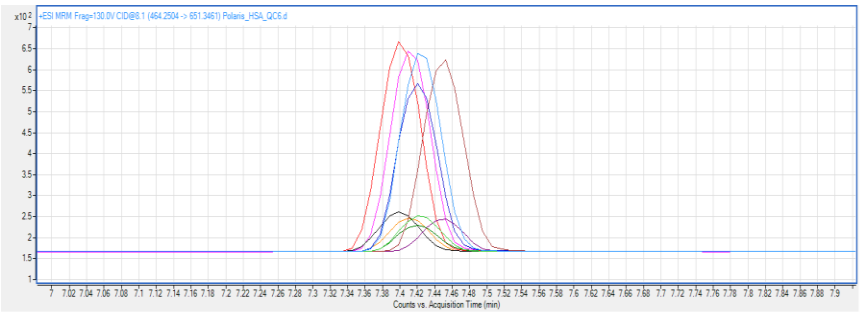


Figure 5

MRM assay of human serum albumin (QC sample) on the Polaris-HR-Chip. Note the fluctuation in retention times of the selected ion fragment over multiple injections.

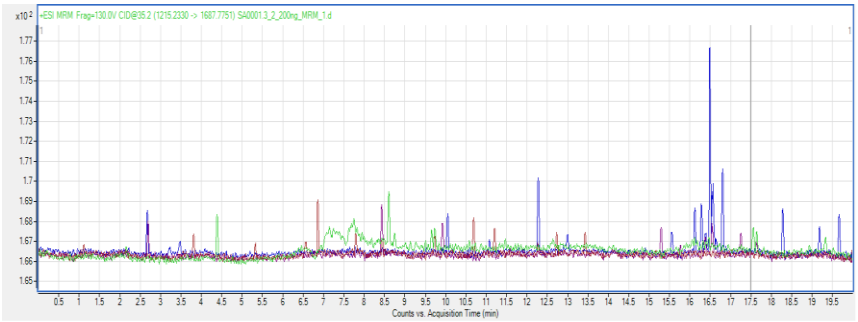


Figure 6

MRM assay of a salivary protein sample on the Polaris-HR-Chip. Note the detection of only base-line noise for targeted statherin ions. This likely reflects chemical incompatibility with the Polaris-HR-Chip.

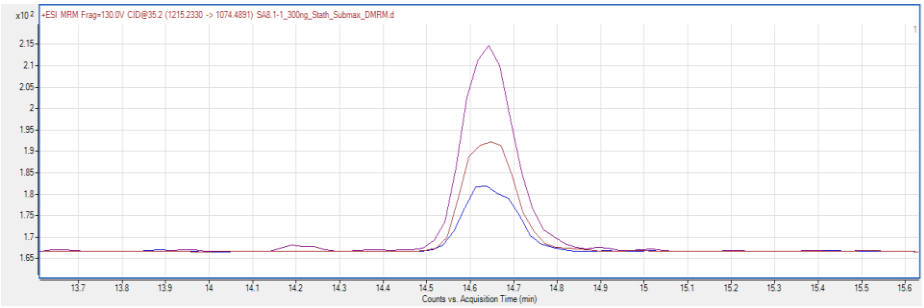


Figure 7

Clear detection by MRM on the Large Capacity Chip of statherin (a high-specificity saliva protein biomarker) in a saliva reference sample.

HPLC Gradient: In order to minimize assay time, the gradient profile for the chromatographic separation was also optimized. Seven variations of a 30-minute method (Figure 8) were evaluated to identify an optimal mobile phase gradient for the detection of statherin. Using extracts of salivary proteins as a model, gradient #7 was found to yield the optimal separation across the run and efficient detection of statherin. All other body fluids were then tested on gradient #7 to ensure compatibility with this gradient. Figure 9 shows an MS scan of a reference saliva sample with the gradient overlay and statherin elution before the final column flush.

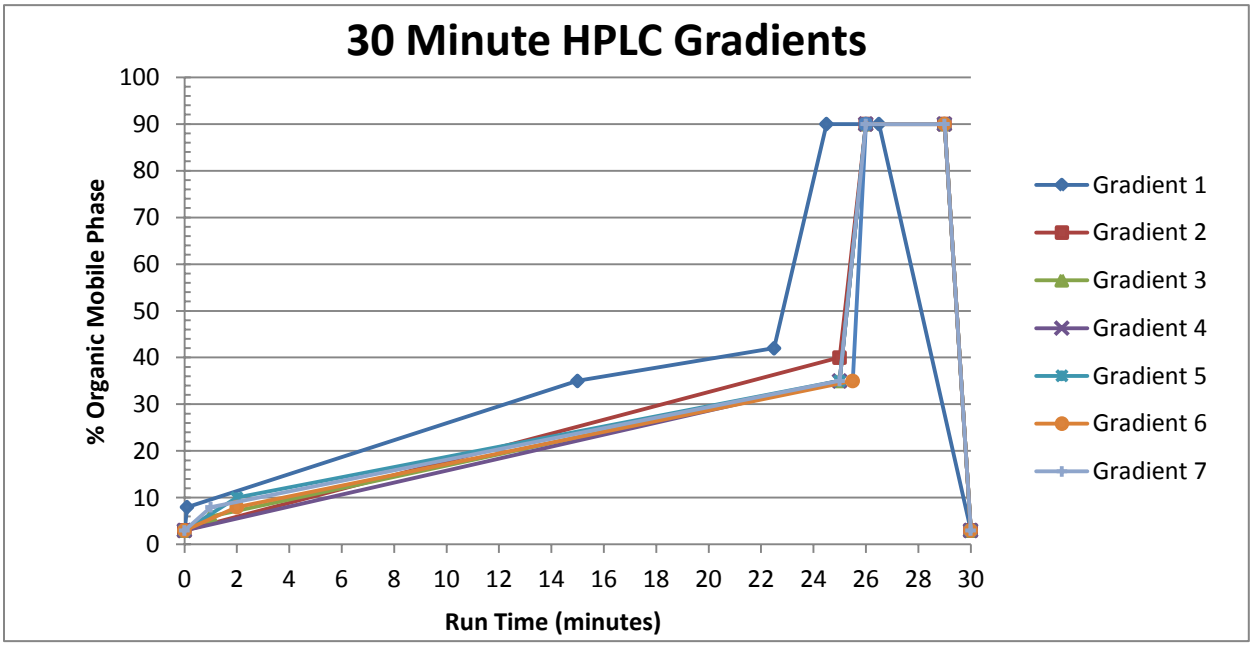


Figure 8 A total of seven variations of a 30-min HPLC gradient profile were assessed to optimize chromatographic separation. The gradient profiles evaluated were as follows:

Gradient 1: 3%-8% organic mobile phase in 0.1 min; 8%-35% in 14.9 min; 35%-42% in 7.5 min; 90% flush.

Gradient 2: 3%-40% organic mobile phase in 25 min; 90% flush.

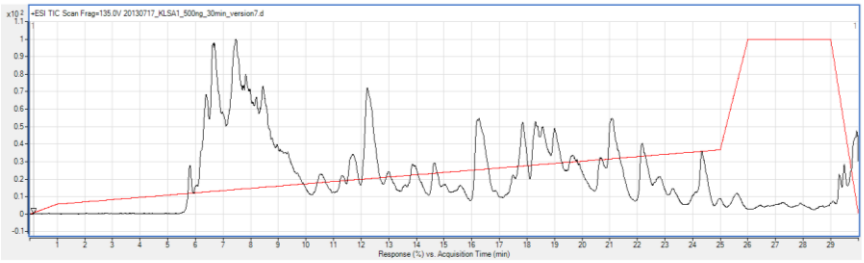
Gradient 3: 3%-6% organic mobile phase in 1 min; 6%-35% in 24 min; 90% flush.

Gradient 4: 3%-35% organic mobile phase in 25 min; 90% flush.

Gradient 5: 3%-10% organic mobile phase in 2 min; 10%-35% in 23 min; 90% flush.

Gradient 6: 3%-8% organic mobile phase in 2 min; 8%-35% in 23.5 min; 90% flush.

Gradient7: 3%-8% organic mobile phase in 1 min; 8%-35% in 24 min; 90% flush



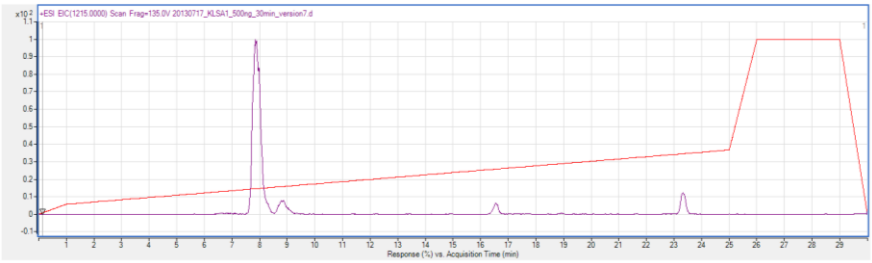


Figure 9

(Top) MS scan of a reference saliva sample analyzed using gradient 7 with the % organic phase overlaid in red.

(Bottom) Extracted ion chromatogram showing the elution of statherin eluting prior to the organic phase flush in red.

Collision Energy: Instrument settings for collision energy and fragmenter voltage were optimized to achieve performance gains by maximizing ion detection. This optimization initiated with *in silico* predictions of the optimal collision energy output by the SKYLINE Proteomics Environment Software. Synthetic peptides were then used to avoid possible matrix effects from the biological fluids of interest. Figure 10 shows replicate results for the saliva peptide, FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF. The empirically determined optimal collision energy is indicated by the highest peak (in light cerulean).

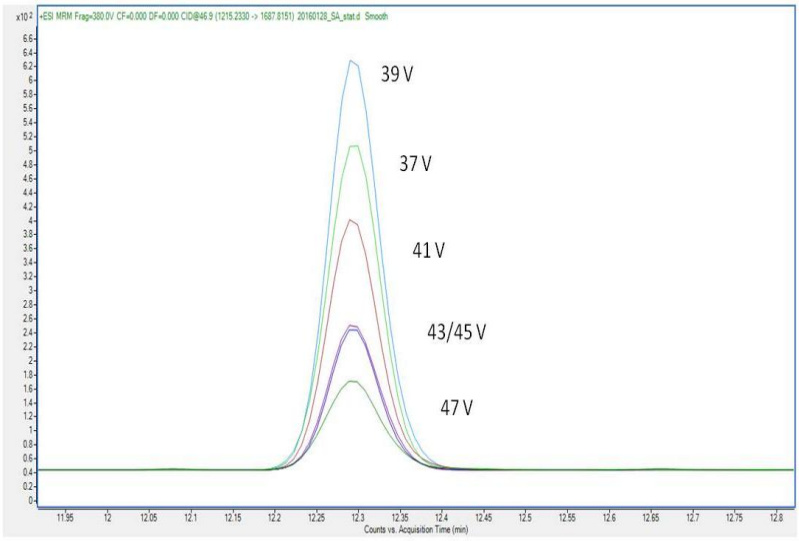


Figure 10

Determination of the optimal collision energy for the Statherin peptide (saliva panel) FGYGYGP YQPVPEQPLYPQPYQPQYQQYTF (Transition 1215.2 m/z → 1687.7 m/z). The optimal collision energy is indicated by peak intensity and percent of regression peak area. The highest peak (light cerulean) represents the optimal energy observed in empirical studies.

Confirmation of Optimized Assay Parameters: To test the specificity of the optimized assay parameters, synthetic peptides for all body fluid biomarkers were combined at equimolar concentration into a target biomarker master mix for each fluid. Each master mix was then analyzed using the optimized instrument/assay parameters. In combination, the assays of the six biomarker master mixes identified sixteen body fluid-specific proteins based on the detection of 26 peptides and 88 transitions. Each of the simplex assays was also tested in duplicate using

single-source body fluids. The results of these injections are shown in Table 7 under the post-optimization column. When compared to the reliability of target peptide detection shown in the pre-optimization column, the significant improvement achieved through empirical optimization is evident. Using the optimized settings, all target protein biomarkers were detected for each sample.

Table 7 Pre-/Post-Optimization Performance of the QQQ-MRM Multiplex Assay with Individual Biological Fluids.

Fluid	Protein	Peptide Sequence	Pre-Optimization	Post-Optimization
Urine	Uromodulin	TLDEYWR		
		STEYGEGYACDTDLR		
Semen	Prostatic Acid Phosphatase	ELSELSLLSLYGIHK		
		FQELESETLK		
	Prostate Specific Antigen (PSA)	LSEPAELTDAVK		
		IVGGWECEK		
	Semenogelin-2	DIFTTQDELLVYNK		
		DVSQSSISFQIEK		
Saliva	Statherin	FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF		
	Submaxillary Gland Androgen Regulated Protein	IPPPPPAPYGPGIFPPPPPQP		
		GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR		
	Amylase	LSGLLDLALGK	*NA: Added later*	
		IAEYMNHLIDIGVAGFR	*NA: Added later*	
Vaginal Fluid	Cornulin	ISPQIQLSGQTEQTQK		
		GQNRPGVQTQGQATGSAWVSSYDR		
	Matrigel-Induced Gene C4 Protein	DGVTGPGFTLSGSC[Cys(CAM)]C[Cys(CAM)]QGSR	*NA: Added later*	
		GC[Cys(CAM)]VQDEFC[Cys(CAM)]TR	*NA: Added later*	
	Suprabasin	ALDGINSGITHAGR	*NA: Added later*	
	Neutrophil Gelatinase-Associated Lipocalin	SYPGLTSYLVR		
		WYVVGLAGNAILR		
Peripheral Blood	Alpha-1 Antitrypsin	LSITGTYDLK	*NA: Added later*	
		SVLGQLGITK	*NA: Added later*	
	Hemopexin	NFPSPVDAAFR		
		GGYTLVSGYPK		
	Hemoglobin Subunit Beta	GTFATLSELHCDK		
		SAVTALWGK		
Menstrual Fluid	Cornulin	ISPQIQLSGQTEQTQK		
		GQNRPGVQTQGQATGSAWVSSYDR		
	Matrigel-Induced Gene C4 Protein	DGVTGPGFTLSGSC[Cys(CAM)]C[Cys(CAM)]QGSR	*NA: Added later*	
		GC[Cys(CAM)]VQDEFC[Cys(CAM)]TR	*NA: Added later*	
	Suprabasin	ALDGINSGITHAGR	*NA: Added later*	
	Neutrophil Gelatinase-Associated Lipocalin	SYPGLTSYLVR		
		WYVVGLAGNAILR		
	Hemopexin	NFPSPVDAAFR		
		GGYTLVSGYPK		
	Alpha-1 Antitrypsin	LSITGTYDLK	*NA: Added later*	
		SVLGQLGITK	*NA: Added later*	
	Hemoglobin Subunit Beta	GTFATLSELHCDK		
		SAVTALWGK		

[**Green**: target peptide detected in 100% of samples; **Yellow**: target peptide detected in <100% of samples; **Red**: target peptide not detected.]

Internal Reference Standards: Biological fluids, analyzed in duplicate, were compared to known synthetic peptide reference standards to ensure consistency in ion ratios and retention time. While ion ratios were consistent between synthetic peptide reference standards and native peptides present in biological fluids, shifts in retention times of up to 1 minute between runs were occasionally noted. Such shifts are not uncommon in nano-flow systems but could compromise data interpretation. To circumvent the potential impact of such variability on assay interpretation, internal reference standards were designed and purchased from New England Peptide to normalize for any shift in retention time. Much like the internal size standards that are employed in STR based DNA profiling systems, these reference standards are identical to the peptides present in a biological fluid of interest except that stable "heavy isotope" labels are used to produce a shift in mass. This allows the standard and the natural peptide from a questioned sample to be monitored simultaneously on the LC-QQQ system (Figure 11). These "heavy" reference standards (a.k.a., Absolute Quantitation or AQUA peptides) allow for the "in matrix" confirmation of a peptide.

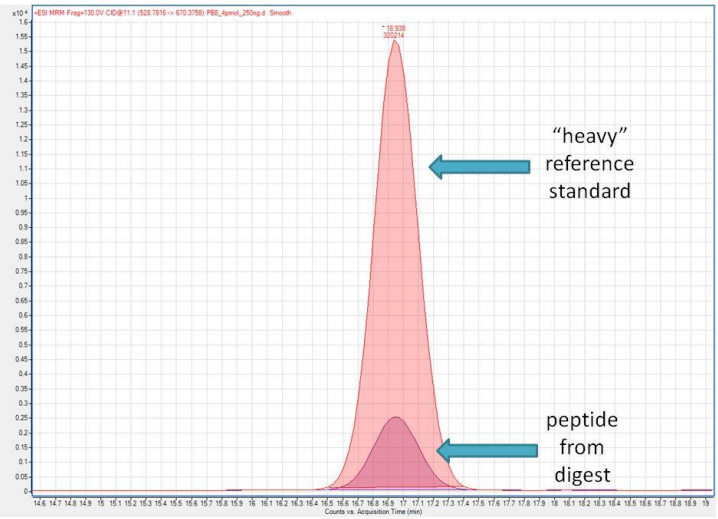


Figure 11
Co-elution of an Absolute Quantitation (*i.e.*, AQUA) peptide reference standard and a natural peptide generated from sample digest.

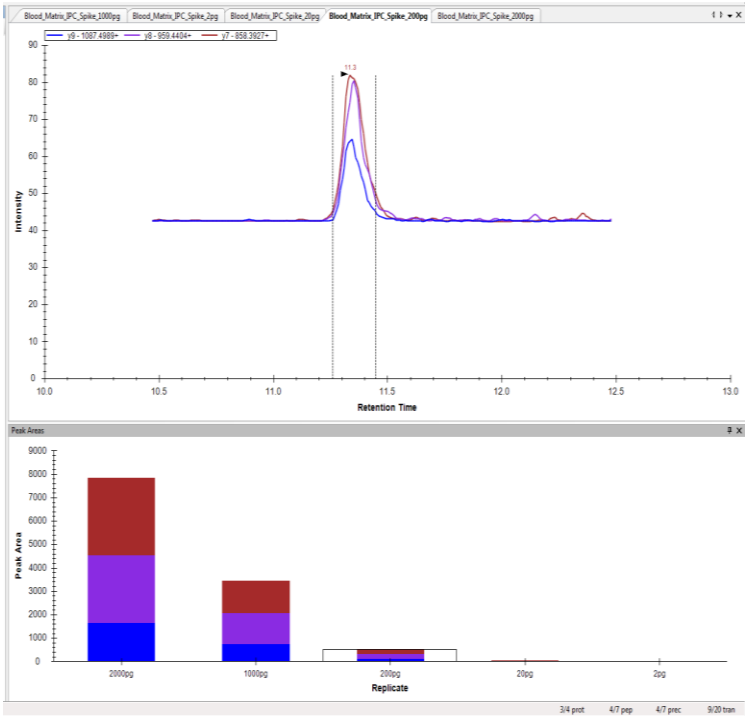
Internal Positive Control: Because forensic casework-type samples are typically of unknown composition and may contain "contaminants" with the potential to inhibit protein digestion, an Internal Protein Control (IPC) was developed to detect possible inhibition and/or matrix-associated ion suppression. The IPC consists of two trypsin-cleavable proteins (bovine aprotinin and bovine myelin basic protein) which are added to each sample and cleaved during sample preparation. A second set of peptides having the same amino acid sequence as the cleaved precursor peptides (*i.e.*, bovine aprotinin peptides YFYNAK; AGLCQTFVYGGCR and myelin peptide DTGILDSLGR) but labeled with a non-radioactive "heavy" carbon isotope allows the signal intensity ratio between the normal and heavy peptides to be monitored during sample injections. These peptide targets were evaluated *in silico* against the UniProt/Swiss-Prot database to ensure that the sequences would not be mistaken for any biomarker of interest or any other protein sequence found in humans. All three peptides were readily detected down to 200 picograms in the presence of 10 µg of peripheral blood protein (Figure 12) Although the response ratios of protein:AQUA peptide did not attain their theoretical maximum, their signal

intensity ratio showed a linear response across the entire range of protein quantities tested (down to 12.5 fmol).

Figure 12

(Top) Representative MRM chromatogram for a 200 pg input of the IPC aprotinin peptide (*i.e.*, AGLCQTFVYGG CR) in the presence of 10 µg of peripheral blood. Detection of all three expected transitions is indicated by the blue, fuchsia and red traces.

(Bottom) Reliable peak shape and consistent ion ratios across three serial dilutions of IPC protein (2000 pg, 1000 pg and 200 pg). The peak area of each of the three expected transitions is indicated by the blue, fuchsia and red bars.



In these studies, run-to-run carryover was also assessed by injecting total protein from body fluids with high-abundance biomarkers (*i.e.*, hemoglobin in blood) and very hydrophobic biomarkers (*i.e.*, amylase in saliva) at a maximum column capacity of approximately 1000 ng of total protein (*i.e.*, "maxim protein capacity"). Each "maximum protein capacity" injection was followed by at least one blank control to monitor for sample carry over. While no detectable sample carryover from the blood digests was observed, low-level (≈0.3%) carryover of amylase was detected. Such sample carryover, was readily eliminated by inclusion of a blank sample (*i.e.*, mobile phase 0.1% formic acid pH 3.0) between runs.

Developmental Validation Studies

In addition to the studies required for base method development and optimization, a series of developmental validation studies were carried out to demonstrate the quality and robustness of the multiplex assay. These studies which encompassed assessments of assay sensitivity, repeatability, reproducibility, species specificity, performance in the context of mixture and performance using a range of casework type sample were designed to meet Standard 8.2 of the FBI's "Quality Assurance Standards for Forensic DNA Testing Laboratories". In designing these studies, the 2012 Scientific Working Group on DNA Analysis Methods (SWGDAM) Validation Guidelines for DNA Analysis Methods[33] were followed. While these guidelines were developed for the validation of DNA associated methods, they are in large part also applicable to the validation of novel serological methods such as those described here.

Sensitivity: The multiplex assay for human body fluids was designed with internal redundancy such that each biomarker protein can be independently identified on the basis of either of two cleavage peptides. The only exception to this is the saliva biomarker statherin which is a small protein for which there is only a single target cleavage product. In order to positively identify a peptide, all daughter ions needed to be detected at the same ion ratio as established in the reproducibility and repeatability studies for both the natural and AQUA peptides. The quality criteria for acceptability of the data included accuracy within ±20% of the average ion ratio. These same criteria also applied to the allowable retention time ratio for the natural and AQUA peptide. Thus, interpretation criteria require at least one peptide of one biomarker to be unambiguously identified in order to positively identify a biological fluid.

While it would be tempting and appear straight forward to determine a limit of detection for each target protein biomarker, this information would have little relevance to assay sensitivity in a real world context. This is because studies under prior NIJ awards to the authors have demonstrated wide ranging interindividual variability in the expression levels of target protein biomarkers. A more useful measure of the sensitivity limit for each of the six biological fluids, therefore was determined based on serial dilutions of body fluid samples pooled from ten individuals so as to obtain an averaged indication of assay sensitivity. Results from these studies are presented in Table 8 where dark green cells indicate detection of all target transitions; light green cells indicate the detection of a minimum of one transition for the target peptide.

It is important to emphasize that "detection limits" for an assay based on the detection of proteins are best expressed in terms of the minimum amount of protein that can be confidently detected. Because, it was demonstrated under a prior NIJ award that protein expression levels can vary by an order of magnitude among individual humans, protein detection levels can only be related to body fluid detection quantities in terms of body fluid averages. If the body fluids of individual donors were used as the basis for determining a detection limit then that limit would appear to be higher for people who expressed low levels of target proteins and lower for those that expressed high levels of proteins. To avoid such potentially misleading results, mass spectrometry assay detection limits were determined in a manner that was similar to other serological assays such as P30. With traditional P30 assays detection limits are expressed in terms of ng of P30 protein rather than in terms of a given volume of semen. Any detection limit based on a volume of semen would be different for different males.

The limit of sensitivity for menstrual fluid was 1:64 based on the NGAL peptide SYPGLTSYLVR which can be used to indicate that a stain has a vaginal origin. The limit of sensitivity for the blood component of menstrual fluid was 1:32,768 based on the hemoglobin peptide SAVTALWGK. For peripheral blood, the sensitivity limit exceeded a 1:131,072 dilution also based on the hemoglobin peptide SAVTALWGK. For vaginal fluid, the limit of sensitivity was 1:1,024 based on the cornulin peptide ISPQIQLSGQTEQTQK. Saliva was also detected down to a 1:1,024 dilution as indicated by the IPPPPPAPYGPGIFPPPPPQP peptide derived from submaxially protein. Semen was detected at a dilution of 1:16,384 based on the detection

of the Semenogelin 2 peptide DVSQSSISFQIEK. Urine was detected at a dilution of 1:128 based on the uromodulin marker STEYGEGYACDTDLR.

Table 8 Sensitivity Limits for Each of the Six Biological Matrices

µl of Target Body Fluid Detected at Each Dilution Level

		Dilution Series																		
		1:1	1:2	1:4	1:8	1:16	1:32	1:64	1:128	1:256	1:512	1:1,024	1:2,048	1:4,096	1:8,192	1:16,384	1:32,768	1:65,536	1:131,072	1:262,144
Menstrual Fluid	Alpha 1 Antitrypsin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031					
	Hemopexin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031					
	Hemoglobin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031	0.0015	0.0008	0.0004		
	Cornulin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906												
	LY6	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953											
	NGAL	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953											
	Suprabasin	25.0	12.5	6.25	3.125	1.5625	0.7813													
Peripheral Blood	Alpha 1 Antitrypsin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031	0.0015	0.0008			
	Hemopexin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031	0.0015				
	Hemoglobin	25.0	12.5	6.25	3.125	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031	0.0015	0.0008	0.0004	0.0002	0.0001
Vaginal Fluid	Cornulin	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488								
	LY6	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906											
	NGAL	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977									
	Suprabasin	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906											
Saliva	Statherin	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813												
	SubMax	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488								
	Amylase	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244							
Semen	Acid Phosphatase	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061					
	Prostate Specific Antigen	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061					
	Semenogelin 2	50.0	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906	0.1953	0.0977	0.0488	0.0244	0.0122	0.0061	0.0031				
Urine	Uromodulin	50	25.0	12.50	6.250	3.1250	1.5625	0.7813	0.3906											

*Dark green cells indicate detection of all target transitions; light green cells indicate the detection of a minimum of one transition for the target peptide.

The potential impact of matrix effects in the context of any sensitivity study employing serial dilution is an important concern because such matrix effects could be diluted out. It was demonstrated in the mixture study, however, that no matrix effects were observed for any of the target fluids alone or in combination with another biological matrix. In addition, the original evaluation and selection process for the target peptides had as a major objective the selection of peptides that did not suffer from matrix effects. To further confirm this, the dilution series can be plotted against the inverse of the response ratio of a target peptide to its internal standard. In the absence of any matrix effect, there will be a solid linear relationship between the dilution series and the response ratio. Figure 13 provides representative illustrations of the linear relationship for a hemoglobin peptide in peripheral blood and a semenogelin-2 peptide in semen. The coefficients of determination for the two peptides shown in these examples are 0.98 and 0.99 respectively.

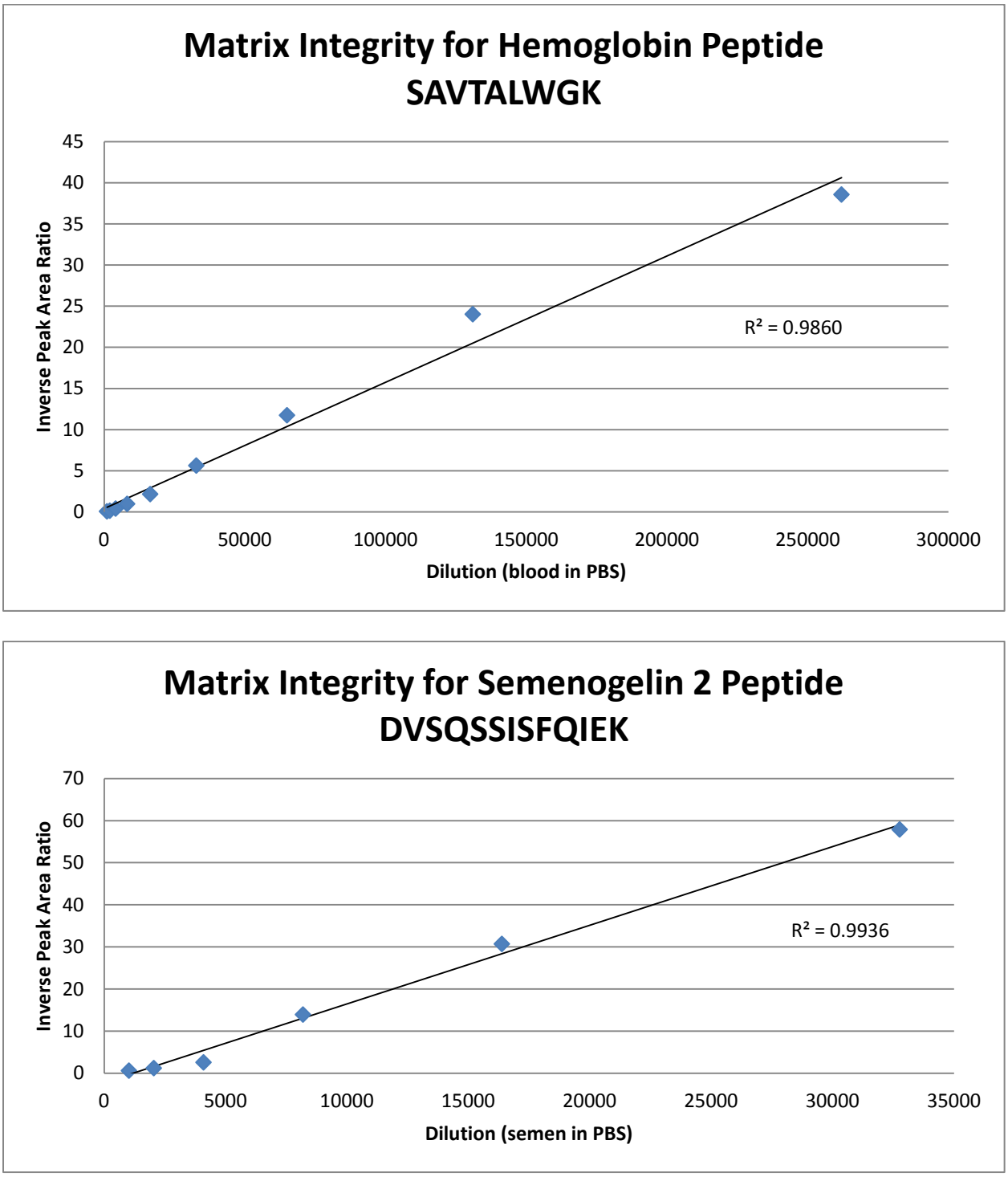


Figure 13

Representative plots of the sensitivity dilution series against the inverse of the response ratio of a target hemoglobin peptide in peripheral blood **(Top)** and a semenogelin-2 peptide in semen **(Bottom)** to their internal standards. The linear relationship demonstrates the absence of a detectible matrix effect.

Repeatability and Reproducibility: To assess repeatability and reproducibility, a pooled sample of each body fluid from 10 individuals was aliquoted into 2.0 ml microcentrifuge tubes. From these, a total of 18 replicates were prepared as follows. Three aliquots of each single body fluid were extracted by each of two analysts per day over a period of three days for a total of 9 samples per analyst. Samples were quantified using a BCA assay and then analyzed using the QQQ-MRM multiplex assay. All injections were performed in triplicate. From these data, reproducibility was indicated by the percent coefficient of variation (%CV) values. The maximum allowable coefficient of variation was 25% (i.e., a common industry practice).

Reproducibility is a function of the ability to consistently extract a sample as well as detect target peptides by mass spectrometry. Protein extraction reproducibility was assessed on the basis of variation in average BCA protein quantitation values as a function of two or more people performing the same extraction using the same procedure. The calculated %CV values for each body fluid extract are indicated in Tables 9-14 where each sample name consists of four letters and two numerals. The first two letters indicate the body fluid; the second two letters indicate the analyst; the first numeral indicates the day and the last numeral indicates the replicate number.

This study was repeated with two analysts and thus the overall %CV for each body fluid reflects an average across both analysts. As such, it provides a measure of the reproducibility for the extraction step of the overall method. Aside from those exceptions discusses above, all peptides met the allowable coefficient of variation. The only exception was the reproducibility associated with urine extraction. Greater variation in extracted protein concentrations were seen with urine due to the specific preparation method required for urine. Due to the presence of urea, urine samples require a precipitation step using 80% acetone. This introduces an unavoidably greater amount of variability in terms due to artificial inflation of protein content values by non-sepcific interaction of urea with the BCA reagent. In turn, this results in a higher %CV value.

Table 9 Peripheral Blood Extraction Reproducibility

Day	Sample Name	Prep Date	Concentration (ug/mL)
1	PB HM 1.1	5/20/2014	14956.8
	PB HM 1.2	5/20/2014	14597.7
	PB HM 1.3	5/20/2014	13045.2
	PB KL 1.1	5/20/2014	19776.3
	PB KL 1.2	5/20/2014	19970.5
	PB KL 1.3	5/20/2014	18028.0
2	PB HM 2.1	5/21/2014	19962.2
	PB HM 2.2	5/21/2014	17290.2
	PB HM 2.3	5/21/2014	17785.2
	PB KL 2.1	5/21/2014	17567.9
	PB KL 2.2	5/21/2014	19646.1
	PB KL 2.3	5/21/2014	19994.3
3	PB HM 3.1	5/22/2014	17493.9
	PB HM 3.2	5/22/2014	18880.3
	PB HM 3.3	5/22/2014	20135.6
	PB KL 3.1	5/22/2014	18196.7
	PB KL 3.2	5/22/2014	20626.5
	PB KL 3.3	5/22/2014	19180.4
statistics		average	18174.10
		min	13045.20
		max	20626.50
		stdev	2129.29
		%CV	11.72

Table 10 Urine (Precipitated) Extraction Reproducibility

Day	Sample Name	Prep Date	Concentration (ug/mL)
1	UR HM 1.1	5/20/2014	1468.45
	UR HM 1.2	5/20/2014	1352.54
	UR HM 1.3	5/20/2014	1574.76
	UR KL 1.1	5/20/2014	1542.11
	UR KL 1.2	5/20/2014	1533.78
	UR KL 1.3	5/20/2014	1380.86
2	UR HM 2.1	5/21/2014	2339.26
	UR HM 2.2	5/21/2014	2443.35
	UR HM 2.3	5/21/2014	2280.23
	UR KL 2.1	5/21/2014	2458.87
	UR KL 2.2	5/21/2014	2336.19
	UR KL 2.3	5/21/2014	2458.21
3	UR HM 3.1	5/22/2014	2785.34
	UR HM 3.2	5/22/2014	2541.96
	UR HM 3.3	5/22/2014	2168.36
	UR KL 3.1	5/22/2014	2389.34
	UR KL 3.2	5/22/2014	2724.81
	UR KL 3.3	5/22/2014	1735.90
statistics		average	2084.129
		min	1352.540
		max	2785.340
		stdev	495.555
		%CV	23.778

Table 11 Saliva Assay Extraction Reproducibility

Day	Sample Name	Prep Date	Concentration (ug/mL)
1	SA HM 1.1	6/17/2014	841.486
	SA HM 1.2	6/17/2014	814.049
	SA HM 1.3	6/17/2014	821.049
	SA KL 1.1	6/17/2014	856.588
	SA KL 1.2	6/17/2014	847.261
	SA KL 1.3	6/17/2014	833.773
2	SA HM 2.1	6/18/2014	843.897
	SA HM 2.2	6/18/2014	815.05
	SA HM 2.3	6/18/2014	808.575
	SA KL 2.1	6/18/2014	820.63
	SA KL 2.2	6/18/2014	844.34
	SA KL 2.3	6/18/2014	869.063
3	SA HM 3.1	6/19/2014	853.969
	SA HM 3.2	6/19/2014	832.153
	SA HM 3.3	6/19/2014	826.215
	SA KL 3.1	6/19/2014	837.357
	SA KL 3.2	6/19/2014	861.67
	SA KL 3.3	6/19/2014	859.033
statistics		average	838.120
		min	808.575
		max	869.063
		stdev	17.966
		%CV	2.144

Table 12 Seminal Fluid Extraction Reproducibility

Day	Sample Name	Prep Date	Concentration (ug/mL)
1	SE HM 1.1	6/30/2014	29071
	SE HM 1.2	6/30/2014	32419.9
	SE HM 1.3	6/30/2014	30793.9
	SE KL 1.1	6/30/2014	28036.1
	SE KL 1.2	6/30/2014	32299.6
	SE KL 1.3	6/30/2014	27349.3
2	SE HM 2.1	7/1/2014	35244.3
	SE HM 2.2	7/1/2014	29658.7
	SE HM 2.3	7/1/2014	28308
	SE KL 2.1	7/1/2014	26520.3
	SE KL 2.2	7/1/2014	23027.4
	SE KL 2.3	7/1/2014	32292.7
3	SE HM 3.1	7/2/2014	29543.8
	SE HM 3.2	7/2/2014	28485
	SE HM 3.3	7/2/2014	24971
	SE KL 3.1	7/2/2014	27410.4
	SE KL 3.2	7/2/2014	35218.8
	SE KL 3.3	7/2/2014	25495.4
statistics		average	29230.31111
		min	23027.4
		max	35244.3
		stdev	3354.964766
		%CV	11.47769093

Table 13 Vaginal Fluid Extraction Reproducibility

Day	Sample Name	Prep Date	Concentration (ug/mL)
1	VF HM 1.1	6/24/2014	2836.31
	VF HM 1.2	6/24/2014	2906.03
	VF HM 1.3	6/24/2014	2882.94
	VF KL 1.1	6/24/2014	2825.14
	VF KL 1.2	6/24/2014	2849.03
	VF KL 1.3	6/24/2014	2817.3
2	VF HM 1.1	6/25/2014	3045.83
	VF HM 1.2	6/25/2014	2994.42
	VF HM 1.3	6/25/2014	2901.85
	VF KL 1.1	6/25/2014	2871.75
	VF KL 1.2	6/25/2014	2886.37
	VF KL 1.3	6/25/2014	2902.25
3	VF HM 1.1	6/26/2014	2825.77
	VF HM 1.2	6/26/2014	2802.32
	VF HM 1.3	6/26/2014	2969.51
	VF KL 1.1	6/26/2014	2876.38
	VF KL 1.2	6/26/2014	2972.95
	VF KL 1.3	6/26/2014	2905.92
statistics		average	2892.893
		min	2802.320
		max	3045.830
		stdev	66.637
		%CV	2.303

Table 14 Menstrual Fluid Extraction Reproducibility

Day	Sample Name	Prep Date	Concentration (ug/mL)
1	MB HM 1.1	6/30/2014	1013.83
	MB HM 1.2	6/30/2014	1143.63
	MB HM 1.3	6/30/2014	1161.83
	MB KL 1.1	6/30/2014	1212.78
	MB KL 1.2	6/30/2014	1347.44
	MB KL 1.3	6/30/2014	1395.02
2	MB HM 1.1	7/1/2014	966.13
	MB HM 1.2	7/1/2014	997.87
	MB HM 1.3	7/1/2014	875.64
	MB KL 1.1	7/1/2014	1181.03
	MB KL 1.2	7/1/2014	1360.06
	MB KL 1.3	7/1/2014	1198.12
3	MB HM 1.1	7/2/2014	1354.83
	MB HM 1.2	7/2/2014	1385.63
	MB HM 1.3	7/2/2014	1325.55
	MB KL 1.1	7/2/2014	1062.13
	MB KL 1.2	7/2/2014	917.4
	MB KL 1.3	7/2/2014	1492.08
statistics		average	1188.388889
		min	875.64
		max	1492.08
		stdev	185.0152186
		%CV	15.56857526

The %CV values for each fluid for the overall analytical method in terms of the peak area ratios and peak retention times for the native versus the AQUA peptides and the ion response ratios for the native and AQUA peptides are indicated in Table 15. The %CV values for analyst

#1 and for analyst #2 provides two measures of repeatability for the analytical method. Here again, the maximum allowable coefficient of variation percentage was 25%. As was observed with the reproducibility studies on protein extraction, the majority of assay results showed excellent reproducibility. The only notable exceptions were again associated with the assay of urine samples. Both uromodulin peptides demonstrate greater variability and, as was discussed previously, this appears to be due to the requirement for an 80% acetone precipitation step when preparing urine samples. The Hemoglobin peptides (SAVTALWGK and GTFATLSELHCDK) and the semenogelin-2 peptide (DIFTTQDELLVYNK) were associated with suboptimal peak morphologies which resulted in poor peak integration and a greater %CVs for overall peak area ratios. This, however, was due to the high abundance of these targets. This can be readily ameliorated by sample dilution. Conversely, the ALDGINSGITHAGR peptide for suprabasin showed an elevated ion ratio %CV value for the natural peptide. This is due to the fact that this peptide was present at only very low quantities. The difficulty of detection at the lower limit of the assay unavoidably leads to a greater %CV as it does with any type of analytical assay.

Table 15 Repeatability (Analyst 1 and Analyst 2) and Reproducibility (Overall) of the Analytical Method

			Area Ratio Natural v AQUA			Retention Time Natural v AQUA			Ion Response Ratio Natural			Ion Response Ratio AQUA		
			Analyst 1	Analyst 2	Overall	Analyst 1	Analyst 2	Overall	Analyst 1	Analyst 2	Overall	Analyst 1	Analyst 2	Overall
			%CV	%CV	%CV	%CV	%CV	%CV	%CV	%CV	%CV	%CV	%CV	%CV
Menstrual Blood	Alpha 1 Antitrypsin	LSITGTYDLK	8.90	7.70	8.30	0.09	0.08	0.09	3.10	4.00	3.80	29.50	19.10	26.90
		SVLGQLGITK	10.10	5.70	8.30	0.06	0.08	0.07	1.60	1.80	1.80	2.80	3.20	3.00
	Hemopexin	NFPSPVDAAFR	9.00	7.90	8.40	0.12	0.14	0.13	13.00	9.10	11.10	3.20	3.50	3.40
		GGYTLVSGYPK	23.50	20.50	21.80	0.56	0.44	0.51	NR	NR	NR	6.50	5.90	6.40
	Hemoglobin	SAVTALWGK	9.30	4.20	7.20	0.08	0.05	0.07	2.60	2.70	2.70	5.10	6.10	5.50
	Cornulin	ISPQIQLSGQTEQTQK	21.90	20.80	21.30	0.35	0.32	0.33	NR	NR	NR	5.50	6.50	6.00
	LY6	GCVQDEFCTR	13.50	12.50	13.00	0.20	0.24	0.22	15.00	16.90	15.80	3.50	3.30	3.40
	NGAL	WYVVGLAGNAILR	26.50	24.40	26.70	0.18	0.21	0.21	NR	NR	NR	7.10	7.30	7.10
		SYPGLTSYLVR	7.00	6.80	7.10	0.14	0.13	0.13	15.10	16.90	16.00	2.80	2.90	2.90
	Suprabasin	ALDGINSGITHAGR	28.60	21.00	25.70	0.24	0.23	0.24	NR	NR	NR	4.70	3.40	4.10
Peripheral Blood	Alpha 1 Antitrypsin	LSITGTYDLK	16.40	17.30	16.70	0.23	0.17	0.20	3.90	3.80	3.80	21.60	14.60	18.20
		SVLGQLGITK	15.30	17.20	16.10	0.08	0.08	0.08	1.40	3.20	2.40	3.80	4.00	3.90
	Hemopexin	NFPSPVDAAFR	13.10	10.70	11.80	0.10	0.09	0.10	3.50	3.70	3.70	2.70	3.70	3.20
		GGYTLVSGYPK	10.50	8.80	9.70	0.24	0.29	0.26	9.00	12.15	10.60	5.50	6.80	6.10
	Hemoglobin	SAVTALWGK	46.20	27.10	39.90	0.09	0.09	0.09	3.50	3.30	3.30	6.60	4.50	5.60
		GTFATLSELHCDK	44.20	29.20	29.20	0.82	0.17	0.17	12.30	11.30	11.30	11.00	20.70	16.40
Vaginal Fluid	Cornulin	GQNRPGVQTQGQATGSAWVSSYDR	7.80	9.20	8.60	0.19	0.22	0.20	17.10	19.40	18.10	6.40	8.20	7.30
		ISPQIQLSGQTEQTQK	7.40	7.50	7.40	0.16	0.18	0.17	13.50	8.40	7.60	17.30	5.20	5.60
	LY6	GCVQDEFCTR	7.50	9.50	8.50	0.14	0.22	0.18	6.10	7.30	6.70	2.90	2.80	2.90
	NGAL	WYVVGLAGNAILR	6.20	4.30	5.30	0.11	0.12	0.11	7.20	4.00	5.80	3.80	3.80	3.90
		SYPGLTSYLVR	4.30	3.00	3.70	0.06	0.08	0.07	5.50	5.30	5.50	3.60	3.00	3.30
	Suprabasin	ALDGINSGITHAGR	12.20	12.10	12.20	0.16	0.20	0.18	26.60	27.10	26.70	3.40	4.00	3.70
Saliva	Statherine	FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF	6.10	5.00	5.60	0.09	0.10	0.10	8.10	7.10	8.00	15.00	12.00	14.10
	SubMax	GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR	3.70	2.60	3.40	0.13	0.14	0.14	1.90	1.50	1.70	9.40	7.40	8.80
		IPPPPPAPYGPGIFPPPPPQP	6.70	7.50	7.10	0.25	0.33	0.29	2.60	2.50	2.50	3.60	7.60	6.00
	Amylase	LSGLLDLALGK	6.80	7.80	7.20	0.11	0.05	0.08	1.80	1.80	1.80	27.00	21.40	24.60
		IAEYMNHLIDIGVAGFR	9.10	6.70	8.00	0.03	0.04	0.03	1.80	1.60	1.70	5.10	4.70	4.80
Semen	Prostatic AP	ELSELSLLSLYGIHK	25.40	20.20	22.70	0.03	0.03	0.03	2.80	1.70	2.40	4.60	2.20	3.60
	Prostate Specific Antigen	IVGGWECEK	24.10	17.30	20.70	0.09	0.00	0.06	1.90	1.90	2.00	4.40	3.80	4.20
		LSEPAELTDAVK	21.10	16.30	18.60	0.06	0.09	0.08	3.90	2.70	3.30	3.90	4.10	4.10
	Semenogelin 2	DIFTTQDELLVYNK	34.50	31.30	32.70	0.06	0.04	0.05	1.30	1.20	1.30	7.50	5.60	6.50
		DVSQSSISFQIEK	24.30	18.70	21.50	0.16	0.16	0.15	2.50	1.60	2.10	5.30	4.90	5.10
Urine	Uromodulin	TLDEYWR	46.90	45.30	46.50	0.14	0.14	0.14	14.90	21.10	18.10	49.90	48.70	48.90
		STEYGEGYACDTDLR	69.80	81.80	79.10	0.23	0.19	0.22	6.70	7.50	7.20	5.60	4.70	4.70

* Measured as a coefficient of variation for area ratios of the natural peptide as compared to the Aqua peptide, retention time for the natural peptide as compared to the Aqua peptide, ion response ratios for the natural peptide and ion response ratios for the Aqua peptide.

Species Specificity: To demonstrate the human-specificity of the QQQ-MRM multiplex assay, both *in silico* and empirical methods were employed. It is recognized and appreciated that species specificity studies have traditionally involved empirical testing of a variety of non-human specimens. The justification for this versus *in silico* based studies is well grounded when validating DNA or RNA based assay systems and especially those employing PCR-based amplification technologies. Originally, the mere lack of comprehensive genomic/transcriptomic sequences made *in silico* analysis impossible. Even as high-quality databases encompassing a wide range of common non-human genomes/transcriptomes became available, there was still good reason to carry out empirical studies. The fact that it is not possible to accurately predict the exact melting point of PCR primers nor is it possible to accurately predict the annealing behavior and thus all amplification products continues to necessitate empirical testing. This is because primers can and do anneal to less than perfect target sequences resulting in spurious amplicons. This is a fundamental difference between nucleic acid based tests and mass spectrometry based tests.

Mass spectrometry, however, looks at mass with an accuracy of a fraction of a single atomic mass unit. The principles of mass spectrometry in general and triple quadrupole mass spectrometry in particular are well established. Thus, unlike STR-based DNA analyses, where amplified allele designation is based on assessment of amplicon length rather than actual nucleotide sequence data, protein biomarker identification by mass spectrometry is based on the precise amino acid sequence and the associated mass of the target peptide and its product ions (transitions). Thus non-human peptides would have to possess not only the same total mass but also the same amino acid primary sequence as a target peptide in the multiplex assay to be mistaken for a human target.

Species specificity studies therefore initiated with and relied primarily on a rigorous *in silico* approach. The amino acid sequences of the target peptides for each candidate biomarker were first screened against the SWISS-PROT databases of 550,116 distinct proteins from 13,257 species. In addition, the target peptide sequences were also screened against the conceptual amino acid translations of all DNA sequences in major genomic databases (*i.e.* GenBank and NCBI RefSeq). This was done to rigorously evaluate the possibility of obtaining false positive results from proteins of non-human origin even if those proteins have not yet been isolated and directly sequenced. The results of these database searches which are provided in Tables 16 through 20 demonstrate the degree of specificity associated with each of the target biomarker peptides. As anticipated, and as has been seen with species specificity studies of DNA typing systems, nearly all of the non-human peptides that were identified as being identical in amino acid sequence to target human biomarker peptides were found in higher order primates. This is expected given the close evolutionary relatedness of these species to modern humans. It should be pointed out that in most cases only one of the two target peptides for a specific biomarker is a perfect match for a non-human protein. This fact has been taken into consideration in the design of the interpretation guidelines for the multiplex assay.

Table 16 Species Specificity of Urine Biomarker Peptides

Fluid	BioMarker (Accession #)	Peptide Sequence	Non-Human Organisms with Shared Sequence
Urine	Uromodulin (P07911)	TLDEYWR	Higher order primate (Sumatran orangutan)
		STEYGEGYACDTDLR	Higher order primate (Sumatran orangutan)

Table 17 Species Specificity of Seminal Fluid Biomarker Peptides

Fluid	BioMarker (Accession #)	Peptide Sequence	Non-Human Organisms with Shared Sequence
Seminal Fluid	Prostatic Acid Phosphatase (P15309)	ELSELSLLSLYGIHK	Mouse and Rat
	Prostate Specific Antigen (PSA) (P07288)	LSEPAELTDAVK	none
		IVGGWECEK	Primate (Rhesus Monkey and Crab-Eating Macaque)
	Semenogelin-2 (Q02383)	DIFTTQDELLVYNK	Primate
		DVSQSSISFQIEK	none

Table 18 Species Specificity of Saliva Biomarker Peptides

Fluid	Biomarker (Accession #)	Peptide Sequence	Non-Human Organisms with Shared Sequence
Saliva	Statherin (P02808)	FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF	none
	Submaxillary Gland Androgen Regulated Protein 3B (P02814)	IPPPPPAPYGPGIFPPPPPQP	none
		GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR	none
	Amylase (P04745)	LSGLLDLALGK	none
		IAEYMNHLIDIGVAGFR	none

Table 19 Species Specificity of Vaginal/Menstrual Fluid Biomarker Peptides

Fluid	BioMarker (Accession #)	Peptide Sequence	Non-Human Organisms with Shared Sequence
Vaginal / Menstrual Fluids	Cornulin (Q9UBG3)	ISPQIQLSGQTEQTQK	none
		GQNRPGVQTQGQATGSAWVSSYDR	none
		GCVQDEFCTR	none
	Suprabasin (Q6UWP8)	ALDGINSGITHAGR	none
	Neutrophil Gelatinase-Associated Lipocalin (P80188)	SYPGLTSYLVR	none
		WYVVGLAGNAILR	none

Table 20 **Species Specificity of Peripheral Blood Biomarker Peptides**

Fluid	Biomarker (Accession #)	Peptide Sequence	Non-Human Organisms with Shared Sequence
Peripheral Blood	Alpha-1 Antitrypsin (P01009)	LSITGTYDLK	Primate
		SVLGQLGITK	none
	Hemopexin (P02790)	NFPSPVDAAFR	Higher order primate (Sumatran orangutan)
		GGYTLVSGYPK	Higher order primate (Sumatran orangutan)
	Hemoglobin Subunit Beta (P68871)	GTFATLSELHCDK	Some Mammal Genera
		SAVTALWGK	Some Mammal Genera

While database searches represent a near exhaustive approach to assessing species specificity, there is certainly some added value to be gained by empirically demonstrating the human specificity of target biomarker peptides as well. If for no other reason, empirical testing serves to confirm the fundamental reliability of mass spectrometry. This was achieved by subjecting blood samples of non-human animal origin to the multiplex assay for human body fluid identification. Blood was chosen for two reasons. First, the hemoglobin biomarker peptides are more widely conserved across non-human mammalian species than any of the other target peptides. Second, the blood proteome is exceptionally complex in terms of the total number of expressed proteins – being more challenging in this regard than other body fluids.

Several blood samples were collected from domestic pets (dogs, cats) and species commonly hunted in Pennsylvania (bear, turkey, coyote). Data obtained from QQQ-MRM multiplex analyses of these non-human blood samples have been carefully reviewed for the presence of any detectible peaks that could be mistaken for a positive detection of a target human protein biomarker. No peaks corresponding to human proteins were detected in any tested samples. Table 21 shows a summary of the empirical results obtained from a series of non-human blood samples. It should be emphasized that in each assay addition, the bovine myelin basic protein which serves as the internal positive control consistently indicated successful digestion and sample processing. This ensures that these negative results are not due to digestion or processing failures but that these non-human samples do not contain detectible levels of any proteins that could be mistaken as biomarker peptides targeted by the QQQ-MRM multiplex assay (Figure 14).

Table 21 **Human Specific QQQ-MRM Assay Results when Using Non-Human Blood**

			Dog 1	Dog 2	Cat 1	Cat 2	Deer 1	Deer 2	Bear 1	Bear 2	Otter	Turkey	Coyote
Peripheral Blood	Alpha 1 Antitrypsin	LSITGTYDLK	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
		SVLGQLGITK	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	Hemopexin	NFPSPVDAAFR	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
		GGYTLVSGYPK	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	Hemoglobin	SAVTALWGK	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
		GTFATLSELHCDK	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

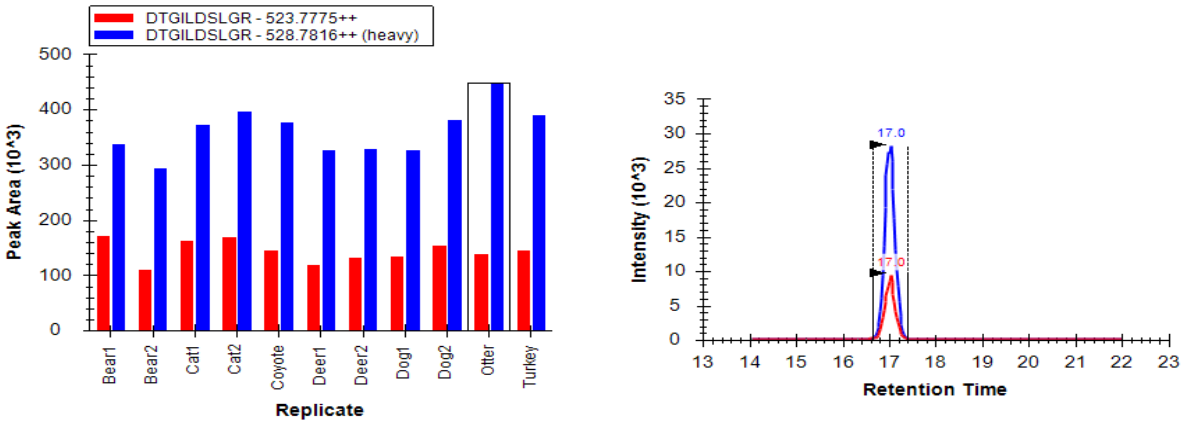


Figure 14 Results from the myelin basic protein internal positive control used with non-human blood samples. Successful digestion of extracted proteins is indicated by the presence of digested myelin basic protein (red) and corresponding AQUA peptide standard (blue).

Mixtures: With traditional PCR-based DNA typing systems mixture studies serve to assess the impact of competition for amplification between alleles from different contributors. With mass spectrometry based body fluid identification there is no competition for peptide detection between different contributors. Rather, mixtures of different body fluids are conducted to assess assay performance in the context of potential matrix effects. For these studies, all possible pair-wise combinations of the six body fluids targeted by the QQQ-MRM multiplex assay were evaluated using equal volumes of corresponding body fluids. As described previously, pooled samples were used to ameliorate the impact of inter-individual variability in protein expression levels. The data from these studies was used to determine the biomarker detection limits in the context of mixed body fluids. This information has been compared to the results of previously described sensitivity studies to determine the extent to which differences in sensitivity might exist between single-fluid and mixed-fluid stains.

Mixtures (1:1 ratio) prepared in triplicate were generated using 25 µl of each of two body fluids. These were placed into 2.0 ml microcentrifuge tubes. Samples were quantified, digested, and analyzed using the QQQ-MRM multiplex assay. In total, 45 mixed samples were tested. The results of these tests are presented in Tables 22-26. In all, 42 of the 45 samples that were tested, accurately characterized each fluid in the mixture. As with the sensitivity studies, successful identification requires that at least one (or more) proteins for a given body fluid be confidently identified. Urine was the only fluid which yielded mixed results. Uromodulin was successfully identified across all replicates when mixed with saliva or vaginal fluid. Uromodulin peptides, however were not detected when mixed with semen, menstrual or peripheral blood. This most likely reflects the very large difference in protein content associated with equal volumes of urine (typically a low protein content fluid) versus semen, menstrual or peripheral blood (average protein concentrations of ≈30 mg/ml in semen and > 200 mg/ml in menstrual or peripheral blood). Since mixed stains are analyzed on the basis of the total protein concentration of initial extract, the inability to detect a body fluid with a low amount of protein when mixed with a body fluid of high protein content is expected.

Table 22 Saliva Mixtures

			Menstrual Blood	Peripheral Blood	Vaginal Fluid	Semen	Urine
Saliva	**Biomarkers**	Statherin	ND	ND	ND	ND	
		SubMax					
		Amylase					
Sufficient Saliva Markers Detected for Identification			**YES**	**YES**	**YES**	**YES**	**YES**

Table 23 Semen Mixtures

			Menstrual Blood	Peripheral Blood	Vaginal Fluid	Saliva	Urine
Semen	**Biomarkers**	Acid Phosphatase					
		Prostate Specific Antigen					
		Semenogelin 2					
Sufficient Semen Markers Detected for Identification			**YES**	**YES**	**YES**	**YES**	**YES**

Table 24 Vaginal Fluid Mixtures

			Menstrual Blood	Peripheral Blood	Saliva	Semen	Urine
Vaginal Fluid	**Biomarkers**	Cornulin				ND	
		LY6		ND			
		NGAL					
		Suprabasin		ND		ND	
Sufficient Vaginal Fluid Markers Detected for Identification			**YES**	**YES**	**YES**	**YES**	**YES**

Table 25 Menstrual Fluid Mixtures

			Peripheral Blood	Vaginal Fluid	Saliva	Semen	Urine
Menstrual Fluid	Biomarkers	Alpha 1 Antitrypsin					
		Hemopexin					
		Hemoglobin					
		Cornulin		ND	ND	ND	
		LY6					
		NGAL	ND				
		Suprabasin					
		Sufficient Menstrual Blood Markers Detected for Identification	YES	YES	YES	YES	YES

Table 26 Urine Mixtures

			Menstrual Blood	Peripheral Blood	Vaginal Fluid	Saliva	Semen
Urine	Biomarkers	Uromodulin	ND	ND			ND
		Sufficient Urine Markers Detected for Identification	NO	NO	YES	YES	NO

Casework-Type Samples: Even though an assay may perform flawlessly in a research and development environment, the demands of an operational environment may create challenges that were unforeseen during development. Among the most important of the developmental validation studies, therefore are those that seek to replicate the types of challenging samples encountered by forensic practitioners. By simulating the characteristics of authentic forensic samples, it has often been possible to identify previously unrecognized areas for improvement not only in the assay itself but also in the Standard Operating Procedures (SOPs) and interpretation guidelines that are essential to an operational environment.

For the current project over 100 casework-type samples were tested to assess the potential impact of a broad range of sample parameters. Casework-type samples (Appendix C) included single-source and sexual assault type stains recovered from a variety of substrates (*e.g.*, cotton, leather, skin, latex, styrofoam, denim, sanitary pad). The potential impact of environmental contaminants (*e.g.*, spermicides, personal lubricants, detergent, fecal matter) and including

several of those that have been previously identified as having inhibitory impacts on DNA profiling technologies (*e.g.*, soil, leather, indigo dyes in denim, bleach and tobacco juice) were assessed by spotting samples of all body fluids onto varied substrates contaminated with these compounds. The results of these tests are presented in Table 27 where dark green cells indicate detection of all target peptides; light green cells indicate the detection of a minimum of one target peptide and red cells indicate a failure to detect any of the targeted peptides. All casework-type type samples prepared with urine, vaginal fluid, semen or saliva were unambiguously and confidently identified. From the menstrual fluid samples, the blood component was always successfully identified. The identification of the non-blood components of menstrual fluid necessary to link the sample to the female reproductive system, however, proved to be more variable by comparison.

Among those casework samples that contained potential inhibitors or that were spotted onto inhibitory substrates, blood biomarkers in peripheral blood and menstrual fluids were successfully extracted and identified when mixed with soil, bleach or when extracted from leather or denim. However, mixing blood with household laundry detergent or 10% bleach resulted in a failure to detect any blood-specific proteins. It is likely that the detergent or bleach in these samples denatured the trypsin during the digestion leading to a failed assay. This is clearly supported by the results from the IPC for these samples. As can be seen in Figure 15, the myelin basic protein which is added to each sample and serves as the basis for the IPC was undetectable in this sample. As is the case for DNA analysis of samples containing *Taq* polymerase inhibitors, the development of additional front-end sample preparation protocols may enable the successful processing of these samples. Microscale solid phase extraction and/or single use size exclusion chromatography cartridges which are known to successfully remove small molecule contaminants, detergents, salts etc. represent a promising option for further evaluation for sample clean up.

Table 27 QQQ-MRM Detection of Body Fluid Biomarkers in Forensic Casework-Type Samples

		Fluid Confirmation					
		Menstrual Blood	Peripheral Blood	Vaginal Fluid	Saliva	Semen	Urine
Menstrual Blood	Cotton		ND	ND	ND	ND	ND
	Denim		ND	ND	ND	ND	ND
	Pad		ND	ND	ND	ND	ND
	Rectal Swab		ND	ND	ND	ND	ND
	Spermicide		ND	ND	ND	ND	ND
	Lubricant		ND	ND	ND	ND	ND
	Vaginal Swab+Lube		ND	ND	ND	ND	ND
Peripheral Blood	Soil	ND		ND	ND	ND	ND
	10% Bleach	ND		ND	ND	ND	ND
	Leather	ND		ND	ND	ND	ND
	Detergent	ND		ND	ND	ND	ND
	Cotton	ND		ND	ND	ND	ND
Vaginal Fluid Case Samples	Cotton	ND	ND		ND	ND	ND
	Finger Swab	ND	ND		ND	ND	ND
	Penile Swab	ND	ND		ND	ND	ND
Saliva Case Samples	Styrofoam	ND	ND	ND		ND	ND
	Gum	ND	ND	ND		ND	ND
	Cup	ND	ND	ND		ND	ND
	Cotton	ND	ND	ND		ND	ND
	Condom	ND	ND	ND		ND	ND
	Tobacco	ND	ND	ND		ND	ND
Semen Case Samples	Condom	ND	ND	ND	ND		ND
	Cotton	ND	ND	ND	ND		ND
	Denim	ND	ND	ND	ND		ND
	Oral Swab	ND	ND	ND	ND		ND
	Rectal Swab	ND	ND	ND	ND		ND
	Spermicide	ND	ND	ND	ND		ND
	Lubricant	ND	ND	ND	ND		ND
Urine Case Samples	Cotton	ND	ND	ND	ND	ND	
	Soda	ND	ND	ND	ND	ND	
	Ceramic Cup	ND	ND	ND	ND	ND	
	Styrofoam	ND	ND	ND	ND	ND	

*Dark green cells indicate detection of all target peptides; light green cells indicate the detection of a minimum of one target peptide and red cells indicate a failure to detect any of the targeted peptides.

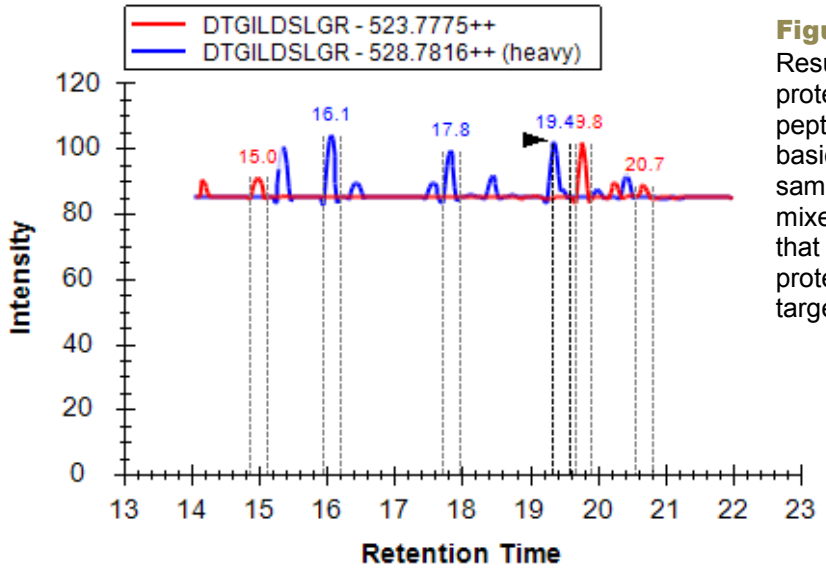


Figure 15
Results obtained for the myelin basic protein internal positive control. Target peptides representing bovine myelin basic protein were undetectable in this sample consisting of peripheral blood mixed with detergent. This indicates that the digestion of peripheral blood proteins that is required to produce the target peptides had failed.

The frequency with which degraded samples are encountered by forensic DNA testing laboratories raises the possibility that protein degradation may also be a concern. There is abundant evidence, however, that protein survives longer than DNA. The oldest recognized protein sequence is that of a collagen protein isolated from a 3.4 million year old high arctic camel from the Middle-Pliocene[35]. By comparison, the oldest sequenced DNA sample comes from the remains of a 700,000-year-old Middle-Pleistocene horse recovered from the Yukon Territory in Canada's northwest[34]. While there is evidence that proteins are generally more stable over time than nucleic acids, extensive degradation may adversely impact the ability to obtain interpretable data from aged or weathered materials. To explore the impact of degradation in aged body fluid samples, a series of saliva, peripheral blood, semen, and urine samples which had been aged at room temperature from 2 to 7 years were analyzed. The results of these assays are presented in Tables 28-31.

Table 28 QQQ-MRM Detection of Peripheral Blood Biomarkers in Aged Blood Stains

			Blood				
			7 years	5 years	4 years	2 years	2 years
Peripheral Blood	Biomarkers	Alpha 1 Antitrypsin					
		Hemopexin					
		Hemoglobin					

Table 29 QQQ-MRM Detection of Saliva Biomarkers in Aged Stains

			Saliva		
			3 Years	3 Years	3 Years
Saliva	Biomarkers	Statherin	ND	ND	ND
		SubMax			
		Amylase			ND

Table 30 QQQ-MRM Detection of Seminal Fluid Biomarkers in Aged Stains

			Semen						
			5 years	5 years	4 years	4 years	2 years	2 years	2 years
Semen	Biomarkers	Acid Phosphatase		ND	ND		ND	ND	ND
		Prostate Specific Antigen		ND	ND		ND		ND
		Semenogelin 2							

Table 31 QQQ-MRM Detection of Urine Biomarkers in Aged Stains

			Urine		
			3 years	3 years	3 years
Urine	Biomarkers	Uromodulin			

As indicated by the results obtained, the QQQ-MRM multiplex assay is able to obtain confident body fluid identification even from aged samples. At least one, and often multiple, target protein biomarkers were confidently identified by mass spectrometry in all samples tested. This result was not unexpected given that the target peptides are only 7 to 31 residues in length. Thus, extensive protein degradation can take place before trypsin cleavage recognition sequences or the peptides themselves are lost. This is in sharp contrast to antibody or enzyme activity based assays where even mild degradation can be expected to result in a loss of functional activity or loss of the three-dimensional integrity required for antibody binding.

Standard Operating Procedures and Interpretation Guidelines

Based on the results of the method development and optimization work and on the results from the developmental validation studies, Standard Operating Procedures (SOPs) have been drafted using document templates from an ISO17025 and ASCLD/LAB-International accredited caseworking laboratory (see Appendices A and B). These SOPs include instructions and information on sample extraction/preparation, instrumental method parameters and data interpretation guidelines.

Because of the superior sensitivity and resolution of the QQQ-MRM assay relative to non-target ion approaches, it would be reasonable to anticipate that some of the high-specificity protein biomarkers selected for the multiplex assay may still be expressed at low but potentially detectible levels in non target tissues. This has necessitated that the interpretation guidelines take into consideration the relative expression patterns of biomarkers across tissues and the overall profile of detected protein biomarkers in a questioned sample. As a result, limited protein identification data from highly degraded or trace samples may in some cases still limit analysts to reporting only a presumptive result. For the majority of samples, however, the sensitivity and accuracy of the QQQ-MRM assay will provide analysts with true confirmatory identifications of questioned biological stains. The SOPs and interpretation guidelines have taken these possibilities into consideration. These documents were developed in close consultation with practicing forensic analysts and the co-Principle Investigator who is the director of the Criminalistics Unit of NMS Labs. While these documents will provide valuable guidance to analysts in an operational caseworking environment, it is recognized that these are “living documents” which may and should evolve based on the experiential input of practitioners.

IV. Conclusions and Implications

Based on the results of the assay development and optimization work and on the data generated in the course of developmental validation studies, a targeted-ion mass spectrometry-based workflow may offer significant advantages compared to the existing serological methods employed by case-working forensic laboratories. The use of mass spectrometry for human body fluid identification is based on the unique chemical composition and mass spectra of multiple peptide targets derived from multiple protein biomarkers. Accordingly, assay results are not easily compromised by interfering chemicals; dependence on a single protein target; potential

antibody cross reactivity, weak test lines in low abundance samples or false negative results from degraded samples or undetected hook effects that are intrinsic to widely used immunochromatographic methods.

In toto, assays based on the use of mass spectrometry offer a more reliable approach to the confirmatory identification of those human body fluids that are most frequently encountered by serologists in forensic caseworking laboratories.

Implications for Policy and Practice

Excellent working relationships with forensic practitioners in the US and abroad have been essential in productively guiding the current R&D efforts. Their advice has played an important role in shaping our experimental design. They have repeatedly stressed that the identification of biological stains can still be a significant challenge for forensic serologists. Practitioners have indicated a need for a reliably means of achieving the confirmatory identification of stains such as saliva and semen and vaginal secretions and a means of better differentiating between peripheral vs. menstrual fluid.

Commercial kits that have been developed for the identification of blood, semen and saliva, use proteins as diagnostic markers of these forensically important substances. While these protein markers have proven useful, they were selected at a time when the field of proteomics was in its infancy. Funding to the principle investigator through NIJ award 2012-DN-BX-K035 have made it possible apply cutting edge proteomic technologies to the detection of previously validated high-specificity protein biomarkers for human body fluids commonly encountered in a forensic context. The immediate benefit of this research is the development and availability of a sensitive, reliable, human-specific assay for the confirmatory identification of questioned biological stains. This will directly address the inherent limitations of the approaches currently employed in case-working laboratories. The use of Multiple Reaction Monitoring has allowed for an unparalleled degree of multiplexing. Not only is it now possible to readily test for different body fluids in a single pass, but also, multiple assays for different diagnostic fragments of individual biomarkers can be performed simultaneously.

This method of peptide identification for body fluid characterization does not eliminate the utility of visual enhancement aids in instances where no visible stains are present on an item of evidence. What the assay is capable of doing is replacing routine presumptive serological testing. The standard workflow for processing evidentiary items, therefore, does not change. For items of evidence with visual stains present, collections are made and the stain would be processed for peptide identification via mass spectrometry (vs. presumptive characterization via colorimetric and/or lateral flow immunochromatographic assays). The cellular component and remaining supernatant, after an aliquot is removed for protein digestion, remains available for DNA analysis. If however, an item of evidence has no visible staining, enhancement techniques commonly employed such as Luminol/BlueStar®, alternate light source (ALS) or acid phosphatase mapping will be useful for identifying areas for collection of biological material prior to peptide identification.

Implementing QQQ-MRM assays in a caseworking laboratory will involve a significant up-front investment. QQQ mass spectrometers retail for $200K - $400K, depending on configuration. However, the cost-to-benefit ratio of the instrument is substantial and QQQ-MRM assays are unequaled for detecting trace quantities of compounds in a complex background. In fact, Agilent produces a Forensic Toxicology MRM kit that enables labs to test for a wide variety of target compounds in blood/urine. Combined with automated sample preparation, this approach can greatly increase sample throughput while lowering the cost of identifying an unknown stain. By multiplexing assays, the cost will drop below that of many existing immunoassays. Moreover, the multiple levels of internal verification that are intrinsic to QQQ-MRM assays will greatly increase the level of confidence that analysts have in the accuracy of the results obtained. Finally the development of a QQQ-MRM approach to body fluid identification will provide forensic practitioners with an opportunity to explore the potential value of using a common instrument platform for serology, toxicology/drug testing and even some types of trace analyses.

The availability of a confirmatory assay for biological stains of forensic interest has significant potential to assist forensic serologist linking DNA profiles to specific biological fluids. Nonetheless, some casework samples can and will still present challenges that may not necessarily be anticipated or that can complicate interpretations. The release of small quantities of blood into the oral cavity as a result of using dental floss or a minor injury to the inside of the mouth may be detected as a mixed stain – which it is. In such cases, it will fall to the experienced judgment of the serologist to make an interpretation with regard to the potential significance of the mixture. Similarly, even with the most accurate of protein biomarkers markers, it may still not be possible in some cases (*e.g.,* mixtures) to definitively say that a DNA profile came from a specific type of epithelial cell. This would be true even if the biomarker were a cell surface molecule. In such challenging cases, however, high specificity biomarkers used in combination with technologies such as Laser Capture Microdissection may enable an analyst to make a definitive statement on the source of a DNA profile.

Finally, it is recognized that Daubert *v.* Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 593-94 (1993), Frye's "general acceptance" test, Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923) and federal rules of evidence specifically Rules CRE 403 and CRE 702 provide the standard for admitting scientific evidence in the federal courts. The study results reported here coupled with publication in peer-reviewed journals, interlaboratory validation studies and adoption by serologists (even as an investigational tool), will help to place the findings of this research on sound legal footing.

Implications for Further Research

This proposal continues to build on the principle investigator's successful completion of previous NIJ funded projects. These projects have charted a course from basic research to laboratory-ready application. This work began with the comparative proteomic mapping of thousands of proteins from six forensically relevant human body fluids. This made it possible to identify a panel of candidate high-specificity protein biomarkers for each stain. Following biomarker discovery, the specificity of each candidate protein, the consistency with which it can

be detected and the degree of interindividual variability in its expression was evaluated across a larger population of human subjects. Using these verified protein targets, the current project successfully completed the developmental validation of a QQQ-MRM based serological assay, a method that can now be utilized by a caseworking laboratory.

As with any new analytical technique, additional improvements in cost, sensitivity, and throughput should be explored. For example, implementation of multi-capillary genetic analyzers for coupled with sample automation systems have led to tremendous improvements in casework throughput over the years. Likewise, a promising direction for future proteomics based serology research would be to modify the current mass spectrometry-based workflow for high throughput analysis. To do this, manual sample preparation could be automated and the frontend nanoflow liquid chromatography system could be replaced with a high flow ultra-performance liquid chromatography interface.

A multitude of vendors offer automation solutions for the trypsin which is core to this application. For example, the Beckman Biomek® NXP Laboratory Automation Workstation, the Agilent Technologies Bravo platform, or Tecan Freedom EVO® all offer "off the shelf" proteomics automation solutions. All of these platforms would be able to digest several 96 well plates/day. For the improved liquid chromatography system, replacing the 1200 series nanoflow chip cube employed during this project with a high flow 1290 series UHPLC, run times of 30 minutes/sample could be reduced to 3-10 minutes/sample – ideal for the high-throughput demands of forensic laboratories. Taking these improvements together, it would likely be possible to process upwards of 300 serological stains a day.

V. References Cited

1. *Biology Methods Manual*. 1978: Metropolitan Police Forensic Science Laboratory.
2. *Protocol Manual*. 1989: FBI Laboratory Serology Unit.
3. Hochmeister, M.N., et al., *Evaluation of prostate-specific antigen (PSA) membrane test assays for the forensic identification of seminal fluid.* J Forensic Sci, 1999. **44**(5): p. 1057-60.
4. Rees, B. and T.J. Rothwell, *The identification of phosphoglucomutase isoenzymes in semen stains and its use in forensic casework investigation.* Med Sci Law, 1975. **15**(4): p. 284-93.
5. Thomas, F.a.v.H., W., *The demonstration of recent sexual intercourse in the male by the Lugol method.* . Medicine, Science and the Law, 1963. **3**: p. 169-171.
6. Rothwell, T.J. and K.J. Harvey, *The limitations of the Lugol's iodine staining technique for the identification of vaginal epithelial cells.* J Forensic Sci Soc, 1978. **18**(3-4): p. 181-4.
7. Hausmann, R., C. Pregler, and B. Schellmann, *The value of the Lugol's iodine staining technique for the identification of vaginal epithelial cells.* Int J Legal Med, 1994. **106**(6): p. 298-301.
8. Jones, E.L., Jr. and J.A. Leon, *Lugol's test reexamined again: buccal cells.* J Forensic Sci, 2004. **49**(1): p. 64-7.

9. Laux, D.L., A.J. Tambasco, and E.A. Benzinger. *Forensic Detection of Semen II. Comparison of the Abacus Diagnostics OneStep ABAcard p30 Test and the Seratec PSA Semiquant Kit for the Determination of the Presence of Semen in Forensic Cases*. Available from: http://mafs.net/pdf/laux2.pdf.
10. Schweers, B.A., et al., *Developmental validation of a novel lateral flow strip test for rapid identification of human blood (Rapid Stain Identification--Blood).* Forensic Sci Int Genet, 2008. **2**(3): p. 243-7.
11. Laffan, A., et al., *Evaluation of semen presumptive tests for use at crime scenes.* Med Sci Law, 2011. **51**(1): p. 11-7.
12. Hobbs, M.M., et al., *Vaginal swab specimen processing methods influence performance of rapid semen detection tests: a cautionary tale.* Contraception, 2010. **82**(3): p. 291-5.
13. Juusola, J. and J. Ballantyne, *Multiplex mRNA profiling for the identification of body fluids.* Forensic Sci Int, 2005. **152**(1): p. 1-12.
14. Hanson, E.K., H. Lubenow, and J. Ballantyne, *Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs.* Anal Biochem, 2009. **387**(2): p. 303-14.
15. Danaher, P., et al., *Facile semi-automated forensic body fluid identification by multiplex solution hybridization of NanoString(R) barcode probes to specific mRNA targets.* Forensic Sci Int Genet, 2015. **14**: p. 18-30.
16. Erin Hanson, S.I., Cordula Haas, Jack Ballantyne, *Targeted multiplexed next generation RNA sequencing assay for tissue source determination of forensic samples.* Forensic Science International: Genetics Supplement Series, 2015. **5**: p. 441-443.
17. Luo, X.Y., et al., *MicroRNA markers for forensic body fluid identification obtained from miRCURY™ LNA array.* Forensic Science International: Genetics Supplement Series, 2015. **5**: p. e630-e632.
18. Silva, S.S., et al., *Forensic miRNA: potential biomarker for body fluids?* Forensic Sci Int Genet, 2015. **14**: p. 1-10.
19. Frumkin, D., et al., *DNA methylation-based forensic tissue identification.* Forensic science international Genetics, 2011. **5**(5): p. 517-24.
20. Lee, H.Y., et al., *Potential forensic application of DNA methylation profiling to body fluid identification.* International journal of legal medicine, 2012. **126**(1): p. 55-62.
21. Antunes, J., et al., *Tissue-Specific DNA Methylation Patterns in Forensic Samples Detected by Pyrosequencing(R).* Methods Mol Biol, 2015. **1315**: p. 397-409.
22. Antunes, J., et al., *High-resolution melt analysis of DNA methylation to discriminate semen in biological stains.* Anal Biochem, 2015. **494**: p. 40-45.
23. Sikirzhytski, V., A. Sikirzhytskaya, and I.K. Lednev, *Multidimensional Raman spectroscopic signature of sweat and its potential application to forensic body fluid identification.* Analytica chimica acta, 2012. **718**: p. 78-83.
24. McLaughlin, G. and I.K. Lednev, *In Situ Identification of Semen Stains on Common Substrates via Raman Spectroscopy.* J Forensic Sci, 2015. **60**(3): p. 595-604.
25. Zapata, F.l., M.Ã.n. FernÃ¡ndez de la Ossa, and C. GarcÃa-Ruiz, *Emerging spectrometric techniques for the forensic analysis of body fluids.* TrAC Trends in Analytical Chemistry, 2015. **64**: p. 53-63.
26. Mechthild Prinz., Y.T., Donald Siegel, Heyi Yang, Bo Zhou, Haiteng Deng. *Establishment of a Fast and Accurate Proteomic Method for Body Fluid/Cell Type Identification*. 2011 [cited 2012; Available from: https://www.ncjrs.gov/pdffiles1/nij/grants/236538.pdf.

27. Legg, K.M., et al., *Discovery of highly specific protein markers for the identification of biological stains.* Electrophoresis, 2014. **35**(21-22): p. 3069-78.
28. Cappellini, E., et al., *Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins.* Journal of proteome research, 2012. **11**(2): p. 917-26.
29. Lindgren, J., et al., *Microspectroscopic evidence of cretaceous bone proteins.* PLoS One, 2011. **6**(4): p. e19445.
30. Johnston, N.L., et al., *Multivariate analysis of RNA levels from postmortem human brains as measured by three different methods of RT-PCR.* Journal of Neuroscience Methods, 1997. **77**(1): p. 83-92.
31. *Recommended minimum criteria for the validation of various aspects of the DNA profiling process.* 2010, ENFSI DNA Working Group
32. Boyd, R., Basic, C, & Bethem, R, *Trace quantitative analysis by mass spectrometry*. 20008, West Sussex, England: Wiley.
33. *SWGDAM Validation Guidelines for DNA Analysis Methods*. 2012, Scientific Working Group on DNA Analysis Methods.
34. Orlando, L., et al., *Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse*. Nature, 2013. **499**(7456): p. 74-78.
35. Rybczynski, N., et al., *Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution.* Nature communications, 2013. **4**: p. 1550.
36. Keating, S.M., *Oral Sex--a review of it's prevalence and proof.* Journal of the Forensic Science Society, 1988. **28**: p. 341-355.
37. Balsells, D., et al., *Reference values for alpha-amylase in human serum and urine using 2-chloro-4-nitrophenyl-alpha-D-maltotrioside as substrate.* Clin Chim Acta, 1998. **274**(2): p. 213-7.
38. Quarino, L., et al., *An ELISA method for the identification of salivary amylase.* J Forensic Sci, 2005. **50**(4): p. 873-6.
39. Singh, V.N., *Human uterine amylase in relation to infertility.* Horm Metab Res, 1995. **27**(1): p. 35-6.
40. Allery, J.P., et al., *Cytological detection of spermatozoa: comparison of three staining methods.* J Forensic Sci, 2001. **46**(2): p. 349-51.
41. French, C.E., et al., *A novel histological technique for distinguishing between epithelial cells in forensic casework.* Forensic Sci Int, 2008. **178**(1): p. 1-6.
42. Sensabaugh, G.F., *Isolation and characterization of a semen-specific protein from human seminal plasma: a potential new marker for semen identification.* J Forensic Sci, 1978. **23**(1): p. 106-15.
43. Wimpissinger, F., et al., *The female prostate revisited: perineal ultrasound and biochemical studies of female ejaculate.* J Sex Med, 2007. **4**(5): p. 1388-93; discussion 1393.
44. Yu, H. and E.P. Diamandis, *Prostate-specific antigen in milk of lactating women.* Clin Chem, 1995. **41**(1): p. 54-8.
45. Iwakiri, J., et al., *An analysis of urinary prostate specific antigen before and after radical prostatectomy: evidence for secretion of prostate specific antigen by the periurethral glands.* J Urol, 1993. **149**(4): p. 783-6.
46. Old, J., et al., *Developmental validation of RSID-Semen: a lateral flow immunochromatographic strip test for the forensic detection of human semen.* J Forensic Sci, 2012. **57**(2): p. 489-99.

47. *PSA in body fluids – an overview for users of the SERATEC PSA SEMIQUANT Tests* 2006, SERATEC Ges.f. Biotechnologie mbH.
48. Miller, K.W., et al., *Developmental validation of the SPERM HY-LITER kit for the identification of human spermatozoa in forensic samples.* J Forensic Sci, 2011. **56**(4): p. 853-65.
49. Griffin, T.J., et al., *Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae.* Mol Cell Proteomics, 2002. **1**(4): p. 323-33.
50. Bauer, M. and D. Patzelt, *Evaluation of mRNA markers for the identification of menstrual blood.* J Forensic Sci, 2002. **47**(6): p. 1278-82.
51. Juusola, J. and J. Ballantyne, *mRNA profiling for body fluid identification by multiplex quantitative RT-PCR.* J Forensic Sci, 2007. **52**(6): p. 1252-62.
52. Buel, E., Noreault-Conti, T., *Bringing tissue identification into the 21st century: mRNA analysis as the next molecular biology revolution in forensic science?* Forensic Genetics Research Progress, ed. F. Gonzalez-Andrade. Vol. 1. 2009: Nova Publishers.
53. Setzer, M., J. Juusola, and J. Ballantyne, *Recovery and stability of RNA in vaginal swabs and blood, semen, and saliva stains.* J Forensic Sci, 2008. **53**(2): p. 296-305.
54. Zubakov, D., et al., *MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation.* Int J Legal Med, 2010. **124**(3): p. 217-26.
55. Wang, Z., et al., *A model for data analysis of microRNA expression in forensic body fluid identification.* Forensic Sci Int Genet, 2012. **6**(3): p. 419-23.
56. Fazzari, M.J. and J.M. Greally, *Introduction to epigenomics and epigenome-wide analysis.* Methods Mol Biol, 2010. **620**: p. 243-65.
57. DeAngelis, J.T., W.J. Farrington, and T.O. Tollefsbol, *An overview of epigenetic assays.* Mol Biotechnol, 2008. **38**(2): p. 179-83.
58. Bell, C.G. and S. Beck, *The epigenomic interface between genome and environment in common complex diseases.* Brief Funct Genomics, 2010. **9**(5-6): p. 477-85.
59. Jones, P.A. and S.B. Baylin, *The epigenomics of cancer.* Cell, 2007. **128**(4): p. 683-92.
60. Zhao, M., et al., *Abnormal epigenetic modifications in peripheral blood mononuclear cells from patients with alopecia areata.* The British journal of dermatology, 2012. **166**(2): p. 226-73.
61. Turner, B.M., *Epigenetic responses to environmental change and their evolutionary implications.* Philosophical transactions of the Royal Society of London Series B, Biological sciences, 2009. **364**(1534): p. 3403-18.
62. Baker, D.J., E.A. Grimes, and A.J. Hopwood, *D-dimer assays for the identification of menstrual blood.* Forensic science international, 2011. **212**(1-3): p. 210-4.

VI. Dissemination of Research Findings

The central deliverable of this project was the development and validation of a QQQ-MRM approach to the identification of six human body fluids with direct forensic utility. This approach to body fluid identification provides forensic practitioners with the ability to use the same instrument platform for serology as is currently used for toxicology/drug testing and even some types of trace analyses. With the completion of this project, laboratories now have a validated approach to confirmatory body fluid identification that is adaptable to the specific case-working needs of each practitioner's laboratory. During the course of the project, the principle investigator strived to provide the professional forensic and the broader scientific communities along with the general public with information on the progress and potential benefits this research. This was achieved through ongoing forensic science workshops and conferences. During the period of the award, the following presentations were made:

Presentations and Publications

2016 Legg, KM and Danielson, PB, "Forensic Serology: Current and Emerging Technologies. In: Sexual Assault: Victimization Across the Life Span 2nd Ed." (In Press for 2016 release) (Managing Editor Liz Fergus) STM Learning, Inc.

2015 Advanced Topics for Human Identification and Data Interpretation, "Body Fluid Identification by Proteomics" (2015) The Center for Forensic Science Research and Education, Philadelphia, PA

2015 The International Association of Forensic Toxicologists, "Confirmatory Blood Identification by Mass Spectrometry: Adapting Proteomics to Forensic Biology", (2015) TIAFT annual meeting, Florence, Italy

2015 Green Mountain DNA Conference, "Body Fluid Identification by Mass Spectrometry: Ready for Primetime", (2015) 8th Green Mountain DNA Conference. Burlington, VT.

2015 American Society of Crime Laboratory Directors. "Body Fluid Identification by Mass Spectrometry from Sexual Assault Evidence" (2015) ASCLD symposium, Washington, D.C.

2015 American Academy of Forensic Sciences, "Confirmatory Body Fluid Identification by Mass Spectrometry from Sexual Assault Evidence", (2015) 67th annual scientific meeting, Orlando, FL

2014 Golden Helix Foundation, "Application of Mass Spectrometry in Forensic Serology," (2014) Golden Helix Summer School: Pharmacogenomics and Genomic Medicine - Bridging Research and the Clinic, Aegina, Greece.

2014 National Meeting of Forensic Chemistry, "Development of a Mass Spectrometry Based Assay for Biological Fluid Identification", (2014) 4th annual ENQFOR meeting, Riberiao Preto, Sao Palo, Brazil

2014 Rape, Abuse, Incest National Network (RAINN) and the National Center for Victims of Crime (NCVC), "Sexual Assault Backlog Reduction Through Protein Based Mass Spectrometry" (2014) Forensic DNA Day on Capitol Hill, Washington, DC

2014 American Society of Crime Laboratory Directors. "Confirmatory Body Fluid Identification by Mass Spectrometry from Sexual Assault Evidence" (2014) ASCLD symposium, Scottsdale, AZ

2014 American Academy of Forensic Sciences, "Development and Testing of a Mass Spectrometry-Based Assay for the Identification of Biological Stains", (2014) 66th annual scientific meeting, Seattle, WA

2014 Legg, KM, Powell, R, Reisdorph, N, Reisdorph, R, Danielson, PB, Discovery of highly specific protein markers for the identification of biological stains. (2014)*Electrophoresis*. 35:21-22 pp.3069-3078 Special Issue: SI

- Additional manuscripts will be submitted to peer-reviewed forensic science and proteomics journals in 2016.

Webinars:

2015 Agilent technologies hosted Webinar, "Confirmatory Blood Identification by Mass Spectrometry". https://www.forensiced.org

Other:

2014 The dataset of validation studies produced under award 2012-DN-BX-K165 is used as the foundation for mock Schrek and Daubert hearings as part of an experiential learning capstone project for law students at the University of Denver's Sturm College of Law.

It was not the objective of this research to commercialize a mass spectrometry assay system for human stain identification. Rather, it was our goal to provided information that would facilitate adoption by commercial and other interested partners to facilitate the adoption and dissemination of such systems. This includes information on all sample preparation protocols, analytical mass spectrometry parameters (*e.g.*, optimal HPLC separation gradients, target transitions, collision energies, internal controls) and a detailed set of interpretation guidelines. Two important initiatives have already arisen as a result of the availability of this technical information.

Initiative 1: Application to Sexual Assault Evidence

NMS Labs, which is a private forensic testing laboratory, has used information from this project to conduct feasibility studies on the implementation of a mass spectrometry-based serological assay as a means of processing backlog sexual assault kits. Currently, seminal fluid screening at NMS Labs for the prioritization of sexual samples is carried out via two later flow immunoassay strips - RSID Semen and Abacus Diagnostics ABA p30. Using the seminal fluid specific biomarkers from the QQQ-MRM developmental validation, a semen specific method has been developed on a Sciex 6500 Qtrap mass spectrometer. This platform had been able to shorten the assay turnaround time to only 15 minutes/sample (Figure 16).

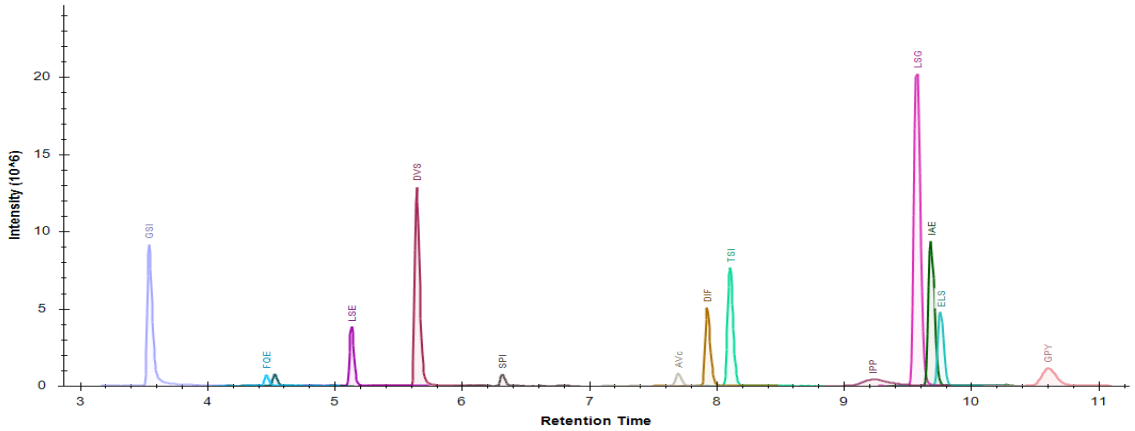


Figure 16 Scan from a 15-minute semen specific assay. Each individual peak represents a unique seminal fluid peptide target. The assay targets human semenogelin I/II, Prostate Specific Antigen, and Prostatic Acid Phosphatase.

This rapid 15 minute allows for the unambiguous identification of human seminal fluid. As with the developmental validation of the QQQ-MRM assay, this approach is also less consumptive of potentially precious evidentiary material and provides true confirmatory stain characterization. Regardless of its technical superiority in terms of stain specificity, however, the practical applicability of a targeted mass spectrometry-based assay for use with casework samples requires a detection sensitivity that is at least equivalent to that of the widely used immunochromatographic test systems. One of the outcomes of an ongoing collaboration with NMS Labs, therefore, has been to evaluate the sensitivity of a targeted mass spectrometry assay for seminal fluid relative to that of commercial immunochromatographic tests currently employed for the screening of sexual assault evidence.

To assess the relative sensitivities of these two approaches, a dilution series of seminal fluid samples (prepared in triplicate on cotton swabs) as well as samples representing a range of post-coital intervals (12 hours to 8 days after intercourse) were tested. All swabs were processed in tandem following the SOPs that were an important deliverable under award 2012-DN-BX-K035. For the commercial immunochromatographic kits, the manufacturer's recommended protocol for the RSID Semen and Abacus Diagnostics ABA p30 assays were followed.

In total, 84 samples were assayed and the results are summarized in Table 32. All seminal fluid dilutions were successfully identified using the Sciex 6500 Qtrap method with all seminal fluid biomarkers being clearly detected. The antibody based tests, however, were unable to reliably identify seminal fluid even at the 1:4096 dilution.

Table 32 **Sensitivity of Immunochromatographic vs. Mass Spectrometry Assays for Seminal Fluid**

Dilution Factor	Antibody Tests	6500 MRM/Trap
2	+	+
4	+	+
8	+	+
16	+	+
32	+	+
64	+	+
128	+	+
256	+	+
512	+	+
1024	+	+
2048	+	+
4096	WP+	+
8192	-	+
16384	-	+

"**+**" indicates a clear positive result; "**WP+**" indicates a weak positive result; "**-**" indicates a negative result.

Eight post-coital samples were assayed using immunochromatographic mass spectrometry-based assays (Table 33). Only the 12 hour post-coital samples could be reliably identified by all three methods. Using the antibody-based P30 assay, only one of the Day-2 post-coital samples yielded a positive result whereas the RSID and mass spectrometry-based approaches both yielded clear positive results. Past day two, however, only the Sciex 6500 Qtrap was able to identify the presence of trace levels of seminal fluid.

Table 33 **Semen identification in post-coital samples**

Interval	P30/Abacus	Semenogelin/RSID	6500
12 hour	+	+	+
Day 2	wp	+	+
Day 2	-	+	+
Day 3	-	-	-
Day 3	-	-	+
Day 4	-	-	+
Day 5	-	-	-
Day 8	-	-	+

"**+**" indicates a clear positive result; "**WP+**" indicates a weak positive result; "**-**" indicates a negative result.

These results clearly indicate that a mass spectrometry-based assay for screening sexual assault evidence will yield superior results compared to traditional immunochromatographic methods. In addition, even with post-coital of 8 days, semenogelin could be reliably detected with very strong signal intensity (Figure 17), this implies that serological identification may be obtainable beyond the time frame encompassed by these experiments. Based in part on these encouraging results, a Rape Kit Advisory board has been set up by NMS Labs to bring together stakeholders from the practitioner, legal, and victim's rights communities. The overall goal of this advisory board is to explore the implementation of a casework assay for use by the NMS Labs criminalistics laboratory.

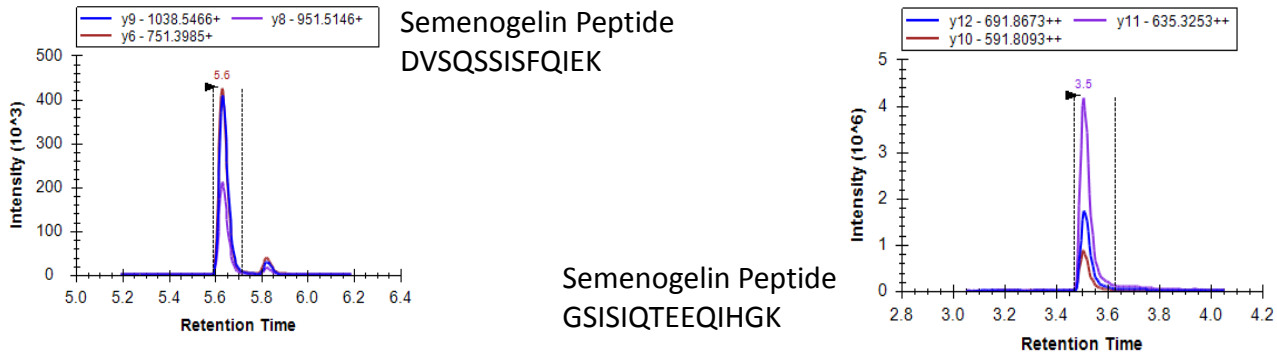


Figure 17 Confirmatory detection of Semenogelin (seminal fluid marker) at 8 days post-intercourse. chromatograms show robust detection of two targeted semenogelin peptides.

Initiative 2: Brasilia Civil Police and Agilent Technologies Peripheral Blood Assay

Peripheral blood is the biological fluid most frequently encountered by forensic practitioners in association with violent crimes against persons. Because of the challenges associated with the analysis of aged blood stain evidence in Brazil, a collaboration with Agilent Technologies and the Brasilia Federal Police was set up to implement a mass spectrometry method for the analysis of aged and/or severely degraded blood stain evidence. Using protein targets from the QQQ-MRM developmental validation studies, a rapid, ultra performance liquid chromatography quadrupole time-of-flight (UHPLC Q-TOF) assay was developed and evaluated in collaboration with the Brasilia Civil Police. This method has now been successfully applied to both mock and challenging casework samples collected by the Brasilia Civil Police.

Thirty casework swabs originating from presumed blood stains were collected from various crime scenes in Brasilia, Brazil. A portion of these samples had been previously analyzed using by the laboratory using contemporary chemical reaction-based (*i.e.*, Kastle-Meyer) and antibody-based (*i.e.*, Hexagon OBTI and Fecal Occult Blood kits) methodologies for blood identification. All samples were reanalyzed using the UHPLC Q-TOF assay. The results obtained from the samples tested by each method are provided in Table 34. Of the samples processed by the immunochromatographic- and chemical reaction- based tests, all but two samples (18 and 27) yielded positive test results. However, because of the known limitations of these tests, even a positive result is not in itself evidence of the presence of human blood. Rather, it must be

reported as only a presumptive indication of blood. In contrast, all samples analyzed by the UHPLC Q-TOF method yielded a true confirmatory identification for human blood. Of particular note were samples 18 and 27. Although all three historical tests had previously yielded negative results, both of these samples were confidently identified as containing human blood by the mass spectrometry-based assay.

Table 34 Comparison of Four Blood Identification Methods in Aged and Challenging Casework Samples

Sample Number	Hexagon OBTI	Fecal Occult Blood	Kastle-Meyer	Mass Spectrometry
Sample 1	NT	NT	NT	+
Sample 2	NT	NT	NT	+
Sample 3	NT	NT	NT	+
Sample 4	NT	NT	NT	+
Sample 5	NT	NT	NT	+
Sample 6	NT	NT	NT	+
Sample 7	NT	NT	NT	+
Sample 8	NT	NT	NT	+
Sample 9	NT	NT	NT	+
Sample 10	NT	NT	NT	+
Sample 11	+	+	+	+
Sample 12	+	+	+	+
Sample 13	+	+	+	+
Sample 14	WP+	WP+	WP+	+
Sample 15	+	+	+	+
Sample 16	+	+	+	+
Sample 17	+	+	+	+
Sample 18	-	-	-	+
Sample 19	NT	+	+	+
Sample 20	NT	+	+	+
Sample 21	NT	+	+	+
Sample 22	+	+	+	+
Sample 23	+	+	+	+
Sample 24	NT	NT	+	+
Sample 25	NT	NT	+	+
Sample 26	+	+	+	+
Sample 27	-	-	+	+
Sample 28	+	+	+	+
Sample 29	+	+	+	+
Sample 30	+	+	+	+

"**+**" indicates a clear positive result; "**WP+**" indicates a weak positive result; "**-**" indicates a negative result; "NT" indicates a sample that was not tested.

With the successful completion of this study, additional collaborations are planned between the Center for Forensic Science Research & Education and the Brasilia Police. A scientist from The Center will be sent to Brazil in early 2016 to complete a thorough study comparing the performance of the UHPLC Q-TOF method relative to more traditional methodologies currently being employed. This study will include additional casework samples as well as further evaluations of sensitivity and reproducibility/repeatability prior to full implementation of the UHPLC Q-TOF workflow for routine casework. Finally, with the successfully completion of this collaborative project, Agilent Technologies is enthusiastic about the possibility of deploying the workflow at other sites. This includes forensic laboratories in Europe as well as Asia.

VII. Appendices

Appendix A: Sample Preparation for Protein Analysis

Appendix B: Protein Interpretation Guidelines and Policies

Appendix C: Preparation of Casework-Type Samples

NAME OF TEST:	**SAMPLE PREPARATION FOR PROTEIN ANANLSIS**
METHOD PRINCIPLE:	To separate protein from cellular and genetic components for serological analysis using liquid chromatography tandem mass spectrometry.
ALTERNATIVE METHODS:	None.
ACCEPTABLE SPECIMENS:	Biological specimens of known or unknown identity.
SPECIAL HANDLING:	Specimens for forensic analysis should be air-dried, kept refrigerated or frozen whenever possible, and placed in paper bags (never plastic). Liquid specimens should be refrigerated and never frozen. Exposure to heat, rain, and other inclement weather should be minimized as best as possible. Proper chain of custody should always be maintained for forensic specimens.
REPORTING LIMIT:	Not applicable
LIMITATIONS OF METHOD:	Not applicable
PHARMACOTOXICOLOGIC DATA:	Not applicable

REFERENCES:

FEDERAL BUREAU OF INVESTIGATION. FBI Laboratory Serology Unit Protocol Manual. 1989. U.S. Government Printing Office, Washington, DC.

GAENSSLEN, R.E. Sourcebook in Forensic Serology, Immunology, and Biochemistry. 1983. U.S. Department of Justice, Washington, DC.

SAFERSTEIN, R. (Editor). Forensic Science Handbook, Volume II. 1988. Prentice-Hall, Englewood Cliffs, NJ.

SEROLOGICAL RESEARCH INSTITUTE. SERI Laboratory Protocol Manual. 1991. SERI, Richmond, CA.

Appendix A: Sample Preparation for Protein Analysis

A. MATERIALS:

1. Bench paper
2. Disposable scalpels
3. Spin baskets
4. 2.0 mL microcentrifuge tubes (preferably low-binding proteomics grade)
5. 1.5 mL microcentrifuge tubes (preferably low-binding proteomics grade)
6. Labels
7. Racks
8. Sonicator
9. Pipettors and aerosol resistant tips
10. Refrigerated microcentrifuge and vortex
11. RotoMixer
12. ThermoMixer
13. Forceps
14. Weigh paper
15. Sterilizing agents

B. REAGENTS:

Note: NMS Labs' procedures for handling chemicals and potentially infectious materials can be found in the Chemical Hygiene Plan and the Bloodborne Pathogen Exposure Control Plan. The analyst must be familiar with and follow the policies and procedures in these plans.

For specific information on the source, preparation, and pertinent safety precautions please refer to the Forensic Biology Unit Master QA/QC Book for the following reagents:

1. DI water
2. HemogloBind (for blood samples)
3. 50 mM Ammonium Bicarbonate (for urine samples)
   Dilute 39.53 mg ABC in 10 mL LC grade water.

C. STANDARDS:

Not applicable to this procedure.

D. CONTROLS:

1. Positive Control: internal positive control (bovine myelin and aprotinin) in each sample

2. Reagent (negative) Control (RC) - a blank tube, which undergoes the same extraction process as the samples to test the reagents for contaminants. The RC should be placed as the last sample in the batch.

E. INSTRUMENT PARAMETERS:

Instrument: Refrigerated microcentrifuge: for 2.0 mL tubes set for 12,000 RCF and 4°C
ThermoMixer set at 23°C and 800 RPM

F. PROCEDURE:

CAUTION: During this procedure the analyst will be working with potentially infectious materials (blood, *etc.*) and potentially hazardous chemicals. The analyst must follow the safety procedures for handling these materials as detailed in NMS Labs' Bloodborne Pathogen Exposure Control Plan and Chemical Hygiene Plan.

1. Initial preparations

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

a. Clean off the workspace and wipe the surface using 20% bleach followed by dH_2O.

b. Cover the counter surface with a clean piece of paper.

c. Clean scissors and forceps by soaking in 20% bleach followed by a diH_2O rinse. Wipe with clean paper towel before use.

d. Check that all reagents have not expired. Replace any reagents that have expired.

2. Extracting/preparing the samples

a. Stains suspected of containing biological fluid must be extracted from the substrate prior to mass spectrometric analysis.

i. Place the sample material (for validation studies add the following amounts of biological material to ½ precut fiber-tipped swabs and allow to dry: 150 µL semen, saliva, urine or vaginal secretions; 50 µL blood) in a clean 2.0 mL low-retention microcentrifuge tube and label accordingly. Extract cellular material by soaking in 300 µL of DI water for 15 minutes in a thermomixer set to 23ºC and 800 RPM. Vortex every 5 minutes. Pulse spin for 10 seconds to remove droplets from the lid and transfer the sample substrate into a clean spin basket and centrifuge at 4°C and 12,000 RCF for 10 minutes.

ii. Following centrifugation, transfer the extracted sample material into a clean 2.0 mL microcentrifuge tube and label accordingly (NOTE: Extracted sample materials should be labeled accordingly and marked as Extracted Sample Material for identification purposes) OR discard.

iii. Transfer the supernatant to a 1.5 mL microcentrifuge tube

iv. For samples possibly containing blood:
   a. Shake/Mix HemogloBind and add 200 µl to sample tube with a 1mL trimmed pipette tip
   b. Vortex samples for 30 seconds
   c. Mix by inversion for 15 minutes
   d. Centrifuge at 23°C and 7,000 RCF for 2 minutes
   e. Transfer supernatant to a 1.5 mL microcentrifuge tube

   For samples being analyzed for urine:
   a. Add 1.2 mL acetone. (NOTE: Acetone should be stored at -20°C and kept on ice while out of refrigeration).
   b. Vortex 2.0 mL microcentrifuge tube containing acetone and supernatant then store at -20°C for 30 minutes.
   c. Centrifuge samples at 4°C and 12,000 RCF for 10 minutes. Discard supernatant.
   d. Resuspend cellular material in 300µL acetone. Store at -20°C for 15 minutes.
   e. Centrifuge samples at 4°C and 12,000 RCF for 10 minutes. Discard supernatant.
   f. Resuspend samples in 150 µL 50 mM ABC. Place the samples in a shaking thermomixer set for 800 RPM at 30°C for a 15 minute incubation.
   g. Centrifuge at 12,000 RCF for 10 minutes. Transfer the supernatant to a 1.5 mL microcentrifuge tubes.

v. Estimate the quantity of protein in the samples by an appropriate quantification method. After quantification, the samples may be digested.

vi. Store the samples at 4°C (short term – same day) or at -80°C (long term >1 day). Prior to using stored samples, samples should be thawed, vortexed, and spun in a microcentrifuge for 5 seconds.

G. QUALITY CONTROL:

Controls should behave appropriately as described under the test method used. The results of the tested controls should be documented in the appropriate case folder.

H. NOTES:

1. The methods described in this procedure are designed to cover the normal circumstances encountered during casework. Due to the numerous factors involved in forensic specimens, it is not possible to encompass them all. Ultimately, the final course of action may be dictated by the Technical Leader.

K. REPORTING RESULTS:

Not applicable

L. CHANGE CONTROL IDENTIFIER:

Original Effective Date:

Original Author:

Methods Manual:

SOP Number:

REVISED BY: _____ Date: _____

REVIEWED BY: _____ Date: _____

ACCEPTED BY: _____ Date: _____

EFFECTIVE DATE: _____

REVIEWED (no changes to current version):

Date: _____ By: _____

Date: _____ By: _____

Date: _____ By: _____

Date: _____ By: _____

Date Retired: _____

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

NAME OF TEST:	**PROTEIN INTERPRETATION GUIDELINES AND POLICIES**
METHOD PRINCIPLE:	To provide guidelines for the interpretation of peptide data for biological fluid identification
ALTERNATIVE METHODS:	None
ACCEPTABLE SPECIMENS:	Agilent 6430 LC-MS/MS instrument data
SPECIAL HANDLING:	Not applicable
REPORTING LIMIT:	See individual guidelines for specific information
LIMITATIONS OF METHOD:	See individual guidelines for specific information
PHARMACOTOXICOLOGIC DATA:	Not applicable
REFERENCES:	See individual manufacturer user manuals and internal validation data

A. MATERIALS:

1. Agilent 6430 LC-MS/MS
2. Pens
3. Worksheets

B. REAGENTS:

Not applicable

C. STANDARDS:

Not applicable

D. CONTROLS:

See “Protein Sample Preparation and Digest” SOP and “LC-MS Analysis of Tryptic Digest” SOP

E. INSTRUMENT PARAMETERS:

See “LC-MS Analysis of Tryptic Digest” SOP

F. PROCEDURE:

See “LC-MS Analysis of Tryptic Digest” SOP

G. INTRODUCTION

The interpretation of data from genetic analysis is a matter of professional training and expertise. The following objective criteria are to be used by analysts to guide most routine data interpretation scenarios. Not every situation, however, may be fully covered by these interpretation standards. Samples that fall outside these standards must be addressed through discussion with the Technical Reviewer in order to reach agreement on a reportable opinion. In the event that agreement on a reportable opinion cannot be reached, the Technical Leader shall be consulted to issue a final decision on a reportable opinion. These criteria are based upon validation studies, literature and professional training and expertise.

These interpretation standards establish a solid framework of quality standards to ensure that:

a. conclusions in the casework report are scientifically supported by the analytical data, including that obtained from appropriate standards and controls;

b. interpretations are made objectively; and

c. interpretations are consistent and accurate from analyst to analyst and case to case.

H. INTERPRETATION CONSIDERATIONS:

1. LC-MS Data Preprocessing

In order to interpret data across runs, LC-MS data must undergo the following pre-processing step(s) as outlined below:

Normalizing: in order to normalize LC-MS data, a constant amount of internal standard is added to all samples. Systematic variation in peak intensity can then be normalized by observing ions’ intensities in relation to the intensities of the ions in the internal standards. Additionally, the retention time of any ion may drift across different samples and this drift cannot be controlled for. However, normalization via the addition of internal standard allows for a comparison of ion peaks to internal standard ion peaks which would demonstrate identical shift within the same sample.

2. Evaluation of Blank Runs

The laboratory has established quality criteria to prevent the occurrence off carryover between sample runs. A Blank (neat methanol) is run after every sample in order to wash the column and prevent sample carryover. As a preventative wash step, these blank samples do not need to be evaluated for data interpretation purposes. The analyst does however have to confirm that a Bank was analyzed between

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

every sample. In the event that a Bank was missed, the preceding runs meet the laboratory's quality standard, however any subsequent samples in the injection list up until another Blank sample is analyzed should be deemed "inconclusive" and must be re-injected.

3. Evaluation of Controls

The laboratory has established quality criteria for evaluation of experimental control data collected as part of the laboratory's DNA testing activities. Controls include, but are not limited to positive extraction controls, reagent negative controls, internal controls and internal positive controls. The laboratory has also established steps to be taken that may allow for the interpretation and documentation of results in the event that the controls do not perform as expected.

Positive Extraction Control: This control (*e.g.*, human semen) serves to demonstrate that the sample preparation, digest and instrument analysis processes performed successfully. The positive extraction control must be evaluated and meet the laboratory's quality standard. Specifically, all 5 targeted peptide peaks for semen should be identified based on the appearance of two MRM transitions per peptide. above 10,000 counts. Additionally, peaks should appear above 10,000 counts and fall within one minute of the acceptable retention times outlined in the table below.

a. If there appears to be an injection or other chromatographic problem, the control should be re-injected.

b. If the positive extraction control fails to meet the laboratory's quality standard but question samples with similar protein concentrations yield positive results, the failure shall be deemed tube specific and the sample preparation and digest shall be considered to have met the quality standard.

Protein	Peptide Sequence	MRM Transitions	Retention time (min)
PSA	IVGGWECEK	539.2 → 964.4, 865.3	10.87
	LSEPAELTDAVK	636.8 → 943.5, 846.4	16.44
PAP	FQELESETLK	612.3 → 948.4, 819.4	13.89
	ELSELSLLSLYGIHK	567.9 → 730.4, 622.3	29.07
SMG	DIFTTQDELLVYNK	849.9 → 1323.6, 993.5	24.47
	DVSQSSISFQIEK	734.3 → 1038.5, 751.3	17.98

Reagent (Negative) Control: This control (*e.g.*, a sample processed in parallel with the casework samples of a batch but to which no protein source material was added) serves to demonstrate that the protein extraction and processing reagents do not contain targeted protein. The negative reagent control must be evaluated and meet the laboratory's quality standard. Specifically, the negative reagent control should be free of contaminating protein upon analysis. The occurrence of more than one targeted peptide peak shall be considered an indication of protein contamination. Such findings shall be necessary but not sufficient for failing the negative reagent control. The official designation of a failure of the negative reagent control must be reviewed and signed off by the supervisor or Technical Leader.

a. If protein contamination is observed in a negative reagent control, acknowledgment of the contaminant and subsequent actions must be included in the case files. In addition, the analyst should endeavor to determine the point at which the contamination was introduced and the scope of the samples affected by the contamination.

b. If the contaminating source affects all samples in the extraction set, the analysis must be redone from the point at which the contamination was introduced.

c. If it is unclear at what point the contamination was introduced, the analysis must be repeated from the protein sample preparation step forward.

d. If additional negative reagent controls were prepared with the batch show no sign of contamination AND the associated samples show no sign of contamination, the incident may be considered tube-specific. The data already derived from these samples can be used for fluid identification purposes.

Internal Standard: This control (*e.g.*, heavy labeled peptides added to every sample and control) serves to normalize the data (See LC-MS Data Preprocessing above) and aid in peak identification. The internal

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

standard must be evaluated and meet the laboratory's quality standard. Specifically, all peptide peaks which are expected to be present must be properly called (with a signal to noise ratio greater than 3) in the negative reagent control. In question samples, internal standards are used only to confirm the presence of detected peptide markers.

Protein	Peptide Sequence	MRM Transitions	Retention Time (min)
PSA	IVGGWECEK	539.2 → 964.4, 865.3	10.87
	LSEPAELTDAVK	636.8 → 943.5, 846.4	16.44
PAP	FQELESETLK	612.3 → 948.4, 819.4	13.89
	ELSELSLLSLYGIHK	567.9 → 730.4, 622.3	29.07
SMG	DIFTTQDELLVYNK	849.9 → 1323.6, 993.5	24.47
	DVSQSSISFQIEK	734.3 → 1038.5, 751.3	17.98
A1AT	LSITGTYDLK	559.8 → 837.5, 805.4	15.3
	SVLGQLGITK	512.3 → 724.4, 837.5	17.9
HEMOPEX	NFPSPVDAAFR	615.8 → 969.5, 485.2	18.7
	GGYTLVSGYPK	575.3 → 872.4, 658.3	13.1
HEMOGLO	SAVTALWGK	470.7 → 782.4, 683.3	15.2
	GTFATLSELHCDK	743.8 → 1110.5, 896.4	13.7
STATH	FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF	1221.9 → 1239.5, 1593.7	30
SMAX	GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR	1039.0 → 1238.6, 856.4	28.2
	IPPPPPAPYGPGIFPPPPPQP	715.0 → 735.4, 1148.6	14.5
AMY1	LSGLLDLALGK	554.3 → 907.5, 737.4	25.9
	IAEYMNHLIDIGVAGFR	643.6 → 908.4, 872.9	28
UROMOD	TLDEYWR	496.7 → 778.3, 663.3	13.1
	STEYGEGYACDTDLR	873.8 → 1080.4, 860.3	11.1
NGAL	WYVVGLAGNAILR	721.4 → 993.6, 894.5	26.4
	SYPGLTSYLVR	633.3 →918.5, 508.2	20.4
CRNN	GQNRPGVQTQGQATGSAWVSSYDR	854 → 993.4, 962.4	12.8
	ISPQIQLSGQTEQTQK	897.4 → 1127.5, 1014.4	11.6
LY6	GCVQDEFCTR	641.2 → 965.4, 837.3	8.6
	DGVTGPGFTLSGSCCQGSR	976.9 → 790.8, 762.3	14.9
SBSN	ALDGINSGITHAGR	464.5 → 604.3, 546.7	9.9

Internal Positive Control: This control (*e.g.*, bovine myelin basic protein) serves to demonstrate that the sample digest (including digestion, denaturation, reduction and alkylation) performed successfully for each sample in the batch. The internal positive control must be evaluated for each sample and meet the laboratory's quality standard. Specifically, both the natural targeted peptide peak as well as the heavy labeled peak should be identified based on the appearance of two MRM transitions per peptide. Additionally, peaks should fall within one minute of the acceptable retention times outlined in the table below. The area ratio of the heavy labeled peptide to the natural peptide should be 3.0 ± 20%. Area ratios falling outside of this range may indicate the presence of digestion inhibition.

c. If there appears to be an injection or other chromatographic problem, the sample should be re-injected.

d. If the internal positive control fails to generate a peak for the natural peptide, or if the ratio falls to meet the response criteria indicated above, the sample should be considered for re-extraction using the alcohol precipitation protocol (See "Protein Sample Preparation and Digest" SOP).

Protein	Peptide Sequence	MRM transitions	Retention time (min)
Myelin	DTGILDSLGR (Light)	523.7 → 660.3, 547.2	18.39

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

	DTGILDSLGR (Heavy)	528.7 → 670.3, 557.2	18.39

4. **Peak Designation**

Proteins are composed of a sequence of amino acids arranged in a linear order. This allows for the prediction, to a certain degree of confidence, the fragmentation pattern and MS/MS spectra that will be produced. To enhance the specificity of the method, two MRM transitions for each peptide are employed. Detectable peptide peaks are those that meet the following criteria:

a. The peak has a signal to noise ratio greater than 3.

b. The peak for the natural peptide should have the same retention time as the corresponding heavy labeled peptide ± 1%.

c. The ion response ratio for the natural peptide should be equal to that of the ion response ratio of the corresponding heavy labeled internal standard peptide ± 20%.

5. General Categories of Testing Conclusions and Reporting Language

Confirmatory Identification: The presence of at least one confirmatory peptide (see Table below) for a body fluid of interest provides a confirmatory indication of the presence of the corresponding targeted biological fluid. This will be reported as "A confirmatory identification of (blood/vaginal fluid/saliva/semen/urine) was obtained for item…"

Presumptive Detection: The presence of at least one presumptive peptide (see Table below) for a body fluid of interest provides a presumptive indication of the presence of the corresponding targeted biological fluid. This will be reported as "A presumptive indication of (blood/vaginal fluid/saliva/semen/urine) was obtained for item…"

Not Detected: In all cases, a failure to detect a minimum of one targeted peptide for any body fluid represents a negative results. This will be reported as **"**No targeted biological fluids were detected".

Statements Regarding Human Specificity: Within the context of a confirmatory result, reporting of the result as human specific requires the detection of a peptide target unique to humans (see Table below). This will be reported as "the confirmatory identification of human (blood/vaginal fluid/saliva/semen/urine) was obtained"

	Human Specific Confirmatory Peptides	Confirmatory Peptides	Presumptive Peptides
Blood	SVLGQLGITK	LSITGTYDLK	
	NFPSPVDAAFR	SAVTALWGK	
	GGYTLVSGYPK	GTFATLSELHCDK	
Vaginal Fluid	GQNRPGVQTQGQATGSAWVSSYDR		WYVVGLAGNAILR
	ISPQIQLSGQTEQTQK		SYPGLTSYLVR
	GCVQDEFCTR		
	ALDGINSGITHAGR		
Saliva	FGYGYGPYQPVPEQPLYPQPYQPQYQQYTF		LSGLLDLALGK
	GPYPPGPLAPPQPFGPGFVPPPPPPPYGPGR		IAEYMNHLIDIGVAGFR
	IPPPPPAPYGPGIFPPPPPQP		
Semen	DVSQSSISFQIEK	DIFTTQDELLVYNK	ELSELSLLSLYGIHK
			LSEPAELTDAVK
			IVGGWECEK
Urine	TLDEYWR		
	STEYGEGYADTDLR		

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

I. QUALITY CONTROL:

Controls should behave appropriately as described under the test method used. The results of the tested controls should be documented in the appropriate case folder.

J. NOTES:

Body fluid identification is confirmed through the mass spectral identification of multiple protein markers. Those protein markers, in turn, are confirmed through the detection of multiple tryptic peptides per protein. Below is a brief description of each proteins function and necessary statements on protein specificity, if nessesary.

Peripheral Blood is identified through the detection of α-1-antitrypsin, hemopexin, and hemoglobin subunit beta. α-1-antitrypsin is a non-specific serine protease inhibitor found in human plasma. This protein's primary role is as an inhibitor of neutrophil elastase thus protecting tissues from proteolytic damage[1, 2]. Hemopexin is produced in the liver and found in plasma. This protein is responsible for trapping free heme in plasma as well as iron recycling in the liver[3, 4]. Hemoglobin subunit beta - The metalloprotein hemoglobin is responsible for oxygen transport and is the major protein contained within erythrocytes. Hemoglobin exists as a tetramer containing two beta chains and two alpha chains[5].

Vaginal Fluid is confirmed through the detection of neutrophil gelatinase-associated lipocalin, cornulin, ly6/PLAUR domain-containing protein 3, and suprabasin. Neutrophil gelatinase-associated lipocalin belongs to the lipocalin family of transport proteins which have been associated with innate immunity though iron sequestration[6]. As such, this protein can be found in tissues prone to exposure to bacterial and other microorganisms including the respiratory tract, salivary glands, uterus, and prostate[6, 7]. Cornulin is also expressed in squamous where it plays a role in epithelial cell differentiation. It may also play a role in mucosal-epithelial immune response. The protein has been characterized in the cervix and in esophageal tissues[8, 9]. Ly6/PLAUR domain-containing protein 3 is involved in the regulation between extracellular structural support scaffolding and epithelial cell layers [10]. Suprabasin is expressed in keratinocytes and plays a role in epidermal differentiation. It has been reported to be expressed in the uterus as well as the esophagus[11].

Saliva is confirmed through the detection of statherin, submaxillary gland androgen-regulated protein 3B, and alpha amylase. Statherin, as well as submaxillary gland androgen-regulated protein 3B, assist in inhibiting potentially harmful calcium phosphate precipitation in saliva[12, 13]. Alpha Amylase is the most abundant protein found in saliva where it digests starches into glucose & maltose [14]. While highly abundant in saliva, this protein can be found in a number of alternate body fluids including vaginal fluid, breast milk, fecal matter, urine, blood and semen.

Seminal Fluid identification is based off the detection of prostatic acid phosphatase, semenogelin-I/II, and prostate-specific antigen. Prostatic Acid Phosphatase (also known as Seminal Acid Phosphatase or SAP) is a glycoprotein secreted by the epithelial cells of the prostate gland which is capable of hydrolyzing phosphate groups from substrate molecules [15]. SAP is another seminal fluid protein which has seen utility as a clinical marker for prostate cancer[16, 17]. While largely replaced by PSA/p30 for screening purposes, the combination of low expression in non-target tissues and assay detection limits makes this protein useful as a potential marker of seminal fluid. Semenogelin-I/II are the most abundant proteins in seminal plasma and are responsible for the gel-like matrix of human semen. Both isoforms act as substrates for prostate specific antigen (p30), where upon lysis, sperm are able to move freely through the seminal matrix[18, 19]. Prostate-Specific Antigen (also known as PSA or p30) is a serine protease produced by epithelial cells located in the prostate[20]. The primary function of prostate-specific antigen is to cleave semenogelin-I/II thus creating a soluble, liquid medium, for spermatozoa movement[21]. Prostate-specific antigen has been well studied as an indicator for prostate cancer when serum levels reach approximately 4-10 ng/mL[22, 23]. While this protein is not absolutely seminal fluid-specific, the detection limits of most assays make it difficult to detect it in whole blood[24]. As a result, this protein may have utility, in combination with other biomarkers, for the detection of seminal fluid.

Urine identification is based on the detection of Uromodulin, also known as Tamm-Horsfall urinary glycoprotein. Uromodulin is the most abundant protein found in human urine[25]. It has been proposed that uromodulin is involved in preservation of water and electrolyte levels as well as being linked to infection prevention [25, 26]. Uromodulin is also involved as an extracellular protein in bone matrix formation[27].

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

K. REPORTING RESULTS:

See above

L. REFERENCES:

1. Kolarich, D., et al., *Comprehensive glyco-proteomic analysis of human α1-antitrypsin and its charge isoforms.* PROTEOMICS, 2006. **6**(11): p. 3369-3380.
2. Parfrey, H., R. Mahadeva, and D.A. Lomas, *Alpha(1)-antitrypsin deficiency, liver disease and emphysema.* Int J Biochem Cell Biol, 2003. **35**(7): p. 1009-14.
3. Tolosano, E. and F. Altruda, *Hemopexin: structure, function, and regulation.* DNA and cell biology, 2002. **21**(4): p. 297-306.
4. Ascenzi, P., et al., *Hemoglobin and heme scavenging.* IUBMB Life, 2005. **57**(11): p. 749-759.
5. Berg JM, T.J., Stryer L., *Biochemistry*, in *Hemoglobin Transports Oxygen Efficiently by Binding Oxygen Cooperatively*. 2002, W H Freeman: New York.
6. Goetz, D.H., et al., *The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition.* Mol Cell, 2002. **10**(5): p. 1033-43.
7. Cowland, J.B. and N. Borregaard, *Molecular characterization and pattern of tissue expression of the gene for neutrophil gelatinase-associated lipocalin from humans.* Genomics, 1997. **45**(1): p. 17-23.
8. Contzler, R., et al., *Cornulin, a new member of the "fused gene" family, is expressed during epidermal differentiation.* J Invest Dermatol, 2005. **124**(5): p. 990-7.
9. Arnouk, H., et al., *Characterization of Molecular Markers Indicative of Cervical Cancer Progression.* Proteomics Clin Appl, 2009. **3**(5): p. 516-527.
10. Smith, B.A., et al., *Identification of genes involved in human urothelial cell-matrix interactions: implications for the progression pathways of malignant urothelium.* Cancer Res, 2001. **61**(4): p. 1678-85.
11. Park, G.T., et al., *Suprabasin, a novel epidermal differentiation marker and potential cornified envelope precursor.* J Biol Chem, 2002. **277**(47): p. 45195-202.
12. Schlesinger, D.H., D.I. Hay, and M.J. Levine, *Complete primary structure of statherin, a potent inhibitor of calcium phosphate precipitation, from the saliva of the monkey, Macaca arctoides.* Int J Pept Protein Res, 1989. **34**(5): p. 374-80.
13. Hay, D.I., et al., *Relationship between concentration of human salivary statherin and inhibition of calcium phosphate precipitation in stimulated human parotid saliva.* J Dent Res, 1984. **63**(6): p. 857-63.
14. Ramasubbu, N., et al., *Structure of human salivary alpha-amylase at 1.6 A resolution: implications for its role in the oral cavity.* Acta Crystallogr D Biol Crystallogr, 1996. **52**(Pt 3): p. 435-46.
15. Yam, L.T., *Clinical significance of the human acid phosphatases: A review.* The American journal of medicine, 1974. **56**(5): p. 604-616.
16. Kontturi, M., *Is acid phosphatase (PAP) still justified in the management of prostatic cancer?* Acta Oncol, 1991. **30**(2): p. 169-70.
17. Taira, A., et al., *Reviving the acid phosphatase test for prostate cancer.* Oncology (Williston Park), 2007. **21**(8): p. 1003-10.
18. Kise, H., et al., *Characterization of semenogelin II and its molecular interaction with prostate-specific antigen and protein C inhibitor.* Eur J Biochem, 1996. **238**(1): p. 88-96.
19. Malm, J., et al., *Isolation and characterization of the major gel proteins in human semen, semenogelin I and semenogelin II.* Eur J Biochem, 1996. **238**(1): p. 48-53.
20. Ward, A.M., J.W. Catto, and F.C. Hamdy, *Prostate specific antigen: biology, biochemistry and available commercial assays.* Ann Clin Biochem, 2001. **38**(Pt 6): p. 633-51.
21. Suzuki, K., et al., *The interaction among protein C inhibitor, prostate-specific antigen, and the semenogelin system.* Semin Thromb Hemost, 2007. **33**(1): p. 46-52.
22. Thompson, I.M., et al., *Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter.* N Engl J Med, 2004. **350**(22): p. 2239-46.
23. Basch, E., et al., *Screening for prostate cancer with prostate-specific antigen testing: American Society of Clinical Oncology Provisional Clinical Opinion.* J Clin Oncol, 2012. **30**(24): p. 3020-5.
24. Keshishian, H., et al., *Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution.* Mol Cell Proteomics, 2007. **6**(12): p. 2212-29.
25. Rampoldi, L., et al., *The rediscovery of uromodulin (Tamm-Horsfall protein): from tubulointerstitial nephropathy to chronic kidney disease.* Kidney Int, 2011. **80**(4): p. 338-47.
26. Serafini-Cessi, F., A. Monti, and D. Cavallone, *N-Glycans carried by Tamm-Horsfall glycoprotein have a crucial role in the defense against urinary tract diseases.* Glycoconj J, 2005. **22**(7-9): p. 383-94.

27. Reinholt, F.P., et al., *Osteopontin--a possible anchor of osteoclasts to bone.* Proceedings of the National Academy of Sciences, 1990. **87**(12): p. 4473-4475.

M. CHANGE CONTROL IDENTIFIER:

Original Effective Date:

Original Author:

Methods Manual:

SOP Number:

REVISED BY: _____ Date:_____

REVIEWED BY: _____ Date:_____

ACCEPTED BY: _____ Date:_____

EFFECTIVE DATE: _____

REVIEWED (no changes to current version):

Date: _____ By:_____

Date: _____ By:_____

Date: _____ By:_____

Date: _____ By:_____

Date Retired: _____

All mock casework samples were prepared as stains on fabric or swabs and allowed to dry prior to resolubilization, quantification, digestion and analysis.

- **Blood:**
  - Neat Bleach: 50µl bleach and 10µl blood mix prepared and then swab dipped in and allowed to dry prior to analysis as bleach caused blood to coagulate
  - 10% Bleach: 50 µl of 10%bleach added to swab followed by 10µl blood
  - Soil: mud slurry prepared – swab rolled in this and then 10µl blood added
  - Detergent: 10µl detergent on swab then 10µl blood

- **Semen**:
  - Cotton and denim: 10µl neat semen
  - Spermicidal condom: 50µl semen added to condom waited 10min then swabbed with cotton swab
  - Non spermicidal condom: 50µl semen added to condom waited 10min then swabbed with cotton swab
  - Oral swab: 5µl semen added to oral swab
  - Rectal swab: 5µl semen added to rectal swab
  - Vaginal swab: 5µl lubricant and 5µl semen added

- **Saliva:**
  - Cotton: 10µl saliva added
  - Glass bottle, Aluminum can, Plastic bottle: 50µl saliva deposited onto bottles that had contents poured out of them (not used) – waited 10 minutes, swabbed with cotton swab moistened with 2%SDS
  - Chewing tobacco spit: 10µl added to swab

- **Vaginal swab:**
  - 10µl saliva added to vaginal swab

- **Urine:**
  - Detergent: 50µl detergent and 10µl urine applied to swab
  - Styrofoam cup: 100µl urine added, let dry completely, swabbed the inside of the cup with a swab moistened with 2%SDS
  - Cotton substrate: 10µl urine added
  - Soda: 50µl soda and 10µl urine added to swab
  - Lotion: swab dipped into lotion and 10µl urine added

- **Menstrual Fluid:**
  - Vaginal swab: 5µl menstrual blood added to swab
  - Rectal swab: 5µl menstrual blood added to swab
  - Cotton, denim, feminine hygiene pad: 10µl menstrual blood added
  - Vaginal swab: 5µl lubricant and 5µl menstrual blood added
  - Spermicidal condom: 50µl menstrual blood added to condom – waited 10 minutes and then swabbed with cotton swab

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

- **Vaginal Fluid:**
  - Glass bottle: 50µl vaginal fluid added to rim – waited 10 minutes – swabbed with moist swab with 2%SDS
  - Cotton substrate: 10µl vaginal fluid
  - Finger swab: following vaginal penetration, wait 10 minutes swab with moist swab with 2%SDS
  - Penile swab: following vaginal penetration, wait 10 minutes swab with moist swab with 2%SDS

- **Aged samples all came from past proficiency tests from CTS or from mock casework that had been prepared internally for competency testing – all samples had been stored at room temperature**
  - Blood: 2 years (1x), 4 years (2x), 7 years (3x) ALL CTS
  - Blood/semen mix: 2 years (3x), 4 years (2x), 5 years (2x) ALL CTS
  - Urine: 3 years (3x) on khaki pants – competency test
  - Semen/Saliva mix: (3 years (3x) on khaki pants – competency test

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.