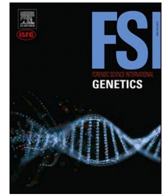
Forensic Science International: Genetics 74 (2025) 103168

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Mass spectrometry-based proteomics for source-level attribution after DNA extraction

Layal Zaarour [a,*], Matthew Padula [b], Roland A.H. van Oorschot [c,d], Dennis McNevin [a]

[a] *Centre for Forensic Science, School of Mathematical and Physical Sciences, Faculty of Science, University of Technology Sydney, Sydney, NSW, Australia*
[b] *School of Life Sciences, University of Technology Sydney, Sydney, NSW, Australia*
[c] *Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Macleod, Victoria, Australia*
[d] *School of Agriculture, Biomedicine and Environment, La Trobe University, Bundoora, Victoria, Australia*

ARTICLE INFO

Keywords:
Forensic
Biomarkers
Body fluid identification
Proteomics
Source level attribution

ABSTRACT

Biological traces recovered from crime scenes serve as vital evidence in forensic investigations. While DNA evidence is frequently used to address the sub-source level of the hierarchy of propositions, the biological source of the DNA can be highly probative at the source level. Current body fluid detection methods pose certain limitations, such as reports of false positive results from some of the presumptive and/or confirmatory tests in current use. These tests are also individual tests for the detection of one body fluid, meaning that if the sample is suspected to be a mixture of multiple body fluids, then different tests would need to be conducted to confirm the body fluid(s) present, which may exhaust small amounts of available biological trace. Proteomics applications for the identification of body fluids have been previously explored, and potential biomarkers indicative of body fluids discovered from liquid-chromatography tandem mass spectrometry (LC-MS/MS) methods have been reported. This work focuses on developing a mass spectrometry-based proteomics approach for the identification of body fluids by targeting discriminating peptide biomarkers from the non-DNA component left over after DNA extraction of samples. The non-DNA component is typically a waste product but with unappreciated evidential value. Our methodology for the purification of proteins from the post-DNA extraction waste includes an acetone precipitation and single-pot solid-phase-enhanced sample preparation (SP3) technique, microwave-assisted trypsin digestion, and LC-MS/MS analysis of the resultant peptides. Preliminary results from this proof-of-concept study include a list of potentially discriminating proteins and peptides for blood, saliva, and semen developed from the analysis of post-DNA extraction waste. Our method allows for multiple analytes to be targeted simultaneously from a DNA profiling waste stream and we anticipate that it could eventually be incorporated into standard forensic laboratory workflows.

1. Introduction

Biological traces found at crime scenes can provide crucial evidence for a forensic investigation. DNA evidence is commonly used for sub-source level attributions [1] but the identification of body fluids can assist forensic scientists make source level attributions tendered as expert opinion in court [2]. The ability to report the nature of the body fluid recovered can help to reconstruct events at the crime scene [3].

Current methods being used for the identification of body fluids include the use of presumptive tests followed by confirmatory tests [2]. Common presumptive tests include chemical [4] and spectroscopic methods [5]. Standard confirmatory tests include chemical, microscopic, immunological, or molecular genetic-based methods [2,6,7]. Some of these presumptive and/or confirmatory methods have been reported to return false positive results, such as vaginal fluid giving positive results for the RSID™- Semen test [8] and saliva showing a false positive result for the ABACard® Hematrace® confirmatory test for blood [9]. Another challenge these current tests pose is exhausting small amounts of available body fluids. Identifying body fluids is not always straightforward, as many may not be visible, present as mixtures, or present in very small amounts [2]. As each of the current presumptive and confirmatory tests require separate tests for each body fluid, there may be instances where a sample needs to be tested for multiple biological fluids. In such cases, insufficient sample may remain to allow the

* Correspondence to: Centre for Forensic Science, School of Mathematical and Physical Sciences, Faculty of Science, University of Technology Sydney, Broadway, NSW 2007, Australia.
E-mail address: layal.zaarour@uts.edu.au (L. Zaarour).

https://doi.org/10.1016/j.fsigen.2024.103168
Received 11 April 2024; Received in revised form 22 October 2024; Accepted 23 October 2024
Available online 26 October 2024
1872-4973/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

generation of full DNA profiles, or no sample at all would be left for DNA testing. As a result, DNA testing is often prioritised, and DNA profiles are readily obtained from body fluids that cannot be identified [10]. There is a need for a multiplex testing approach which minimises sample depletion and allows for the targeting of multiple fluids simultaneously as well as leaving sufficient biological material for DNA profiling.

To address the issues mentioned above, methods for conducting DNA profiling directly from the immunochromatographic tests, and their buffers, used in confirmatory testing have been developed. These remain as single-plex approaches, however, testing for one body fluid at a time [11–13]. Another approach to source level attribution is messenger-RNA (mRNA) profiling, where mRNA is co-extracted with DNA and hence does not consume more starting material, however, the instability of RNA *ex vivo* has posed some limitations for this method of testing [14–16]. MicroRNA (miRNA) has also emerged as a biomarker for forensic body fluid identification and has proven to be stable under various environmental changes [17]. More recently, the identification of body fluids has been achieved by proteomics techniques, specifically by targeting protein-derived biomarkers using liquid-chromatography tandem mass spectrometry (LC-MS/MS) methods [18]. Different sample preparation techniques and instruments have been used to identify protein biomarkers for a range of body fluid types and have been successful in identifying discriminating protein biomarkers [19–37].

This research also focuses on the development of a mass spectrometry-based approach for the forensic identification of body fluids by targeting discriminating peptide biomarkers. Rather than starting with the body fluid prior to DNA profiling, however, the starting material used is the non-DNA component left over after DNA extraction. This method addresses the limitations of the current tests, allowing for multiple analytes to be targeted simultaneously from a DNA profiling waste stream. Peptide-based proteome profiles were generated from the non-DNA component of blood, saliva, and semen samples post-DNA extraction. The overall goal is to extract valuable information from body fluids without compromising their state, amount, or condition in an approach that aims to enhance the efficiency and maximise the use of biological traces found at crime scenes. Van Steendam et al. [21] introduced a mass spectrometry-based method that analyses the protein in the sample but still preserves the DNA by retaining the DNA pellet after the first centrifugation step and pointed out the possibility of using this pellet for further DNA typing, which could address some of the limitations presented in current confirmatory tests. The method developed in this study, however, prioritises DNA extraction and utilises the waste stream after extraction to develop an approach that could potentially be used for the forensic identification of body fluids.

2. Methods

2.1. Laboratory decontamination procedures

Standard practices for decreasing the possibility of sample contamination were implemented. Particularly, for DNA extraction, lysis was conducted in a Biological Safety Cabinet (Class II Type A2: 1.2; EuroClone TopSafe) after wiping the internal surface with 70 % ethanol and subjected to UV irradiation for 30 minutes. All consumables and materials required for DNA extraction were sterilised prior to use by wiping with 10 % bleach, followed by 70 % ethanol, and then UV irradiated for 30 minutes.

2.2. Sample collection

Ethics approval for the collection and storage of human body fluids was granted by the University of Technology Sydney Human Research Ethics Committee (Project ETH18–2521). All body fluids were donated by participants, over the age of 18, who had previously provided informed consent and were subsequently de-identified.

For biomarker identification, four samples of saliva, blood, and semen were collected from different individuals, however not all participants provided samples for all three body fluids. For the collection of saliva, participants were provided with a sterile rayon swab in a plastic sheath (Tubed Sterile Dryswab™ Rayon MW1021, Medical Wire & Equipment, UK) which had been punctured to allow air ingress for drying. They were instructed to rub and rotate the swab on their inner cheeks for 5–10 seconds, and carefully return the swab to the sheath which was then stored at −80°C. For blood samples, participants were first instructed to wipe the tip of one of the fingers on their non-dominant hand with an alcohol swab. Then, using their dominant hand, they were instructed to use an individual sterile lancet (Accu-Chek® Softclix: Roche Diabetes Care) to produce a blood droplet. Blood was deposited on a sterile rayon swab which was stored in its punctured sheath at −80°C. For semen samples, participants were given a sterile specimen jar and instructed to directly deposit semen into it. The jar was stored at 4–8 °C (at the participant's home, for up to 8 hours) and then at −20°C until, at the time of analysis, an aliquot of 10 μL of semen was deposited on a sterile rayon swab. Four mixtures were also prepared, where different body fluids were directly deposited on the same swab. Participants providing the mixture samples were different from those who provided the samples for biomarker identification. The collection of blood for mixture samples was achieved by directly transferring the produced blood droplet into a sterile collection tube and immediately pipetting the required volume onto a swab. The collection of saliva for mixture samples was achieved by instructing the participants to directly spit into a sterile specimen jar and immediately pipetting the required volume onto a swab. The collection of semen for mixture samples was the same as that described above, where the required volume was then immediately pipetted onto the swab. Mixture samples were prepared by transferring 2 μL of the body fluid of interest onto the swabs and consisted of mixture one, where saliva and semen were deposited on the same swab; mixture two, where blood and semen were deposited on the same swab; mixture three, where blood and saliva were deposited on the same swab and mixture four, where blood, saliva, and semen were all deposited on the same swab.

2.3. DNA extraction

DNA extraction of body fluid and mixture samples was performed using the PrepFiler Express™ Forensic DNA Extraction Kit (Applied Biosystems®) and following the 'body fluids' protocol provided by the manufacturer. Each swab head was cut off its sheath with sterile scissors and placed into a LySep Column fitted into a hinge-less PrepFiler™ sample tube. A 500 μL volume of lysis buffer and 5 μL of 1 M DTT was added to each sample. Each Lysep Column / sample tube assembly was incubated in a thermal shaker at 70°C and 750 rpm for 40 minutes, then centrifuged for 2 minutes at 10,000 × g. The LySep columns were removed, and the sample lysate tubes (containing lysed DNA) were then transferred to an AutoMate Express™ Forensic DNA Extraction System (Applied Biosystems®). DNA extraction was performed according to the manufacturer's recommended protocol with an elution volume of 40 μL. At the end of each run, the remaining non-DNA component from each sample (approximately 750 μL) was collected from the second well of the PrepFiler Express™ Cartridge (Applied Biosystems®) and transferred to a sterile 1.5 mL microcentrifuge tube. These tubes were placed in a magnetic rack for one minute in order to capture any residual magnetic particles from the automated DNA extraction. The liquid from each tube was then transferred to another sterile 1.5 mL microcentrifuge tube. The purified DNA samples and non-DNA component samples were then stored at −20°C until further analysis.

2.4. DNA quantitation

For DNA Quantitation, the Quantifiler™ Trio Quantification Kit (Applied Biosystems®) was used. A standard dilution series was prepared by following the protocol provided by the manufacturer. The

required volume of Quantifilier™ Trio Primer Mix (8 μL) and Quantifiler™ THP PCR Reaction Mix (10 μL) needed for each standard and sample were added to create a PCR mix, which was dispensed in each reaction well of a 96-well plate. The standards (2 μL), DNA samples (2 μL), and controls (2 μL) were then added to the applicable wells. Real time PCR was performed in a QuantStudio™ 6 Flex Real-Time PCR System (Applied Biosystems®) and data were analysed on the QuantStudio™ Real-Time PCR Software.

2.5. Sample preparation

For body fluid samples used for biomarker identification, the non-DNA component was collected after DNA extraction from four blood, four saliva, and four semen samples collected from participants. The non-DNA component of mixture samples was also collected after DNA extraction from mixture swabs. All non-DNA component fractions of body fluids and mixtures were prepared as described below.

2.5.1. Protein precipitation and resuspension

The non-DNA component samples were placed in 15 mL tubes and a 5 × volume of cold acetone (-20°C) was added to each one. The tubes were shaken to homogenise the solution and then incubated at −20°C for 30 minutes to precipitate protein. The tubes were then centrifuged at 2588 × g for 5 minutes. The acetone was carefully decanted, and excess acetone in the tube was removed by wiping with a Kimteck wipe inside the tube. The protein pellet at the bottom of the tube was resuspended in 48 μL of buffer (1 % SDS, 50 mM Tris-HCl, pH 8.8) and heated at 95°C for 10 minutes to resolubilise the protein into solution.

2.5.2. Single-pot, solid-phase-enhanced sample preparation (SP3)

Resuspended protein was prepared for protein purification by the single-pot solid-phase-enhanced sample-preparation (SP3) technique [38]. An aliquot of 2 μL of suspended magnetic beads (Sera-Mag SpeedBeads, NY, USA) was added to the samples in buffer solution and the mixture was then gently pipette-mixed to homogenise. A 50 μL aliquot of 100 % ethanol was added to induce the binding of the proteins to the beads.

The tubes for each of the samples were vortexed at low speed and room temperate for five minutes, after which they were placed on a magnetic rack for one minute or until the beads had migrated to the tube wall. The supernatant was then carefully discarded, and the tube removed from the magnetic rack. A volume of 180 μL of 80 % ethanol was added to rinse the beads. Samples were incubated for approximately 5 minutes, with periodic pipette mixing, to enhance surfactant removal. Tubes were then placed again on a magnetic rack, the supernatant was removed, and 180 μL of 80 % ethanol was re-added to rinse the beads. These ethanol wash steps were repeated 3 times. After removing the supernatant during the last wash step, 100 μL of 200 mM ammonium bicarbonate was added to resuspend the beads.

2.5.3. Microwave-assisted trypsin digestion

Trypsin (1 μg, Promega; Madison, WI) was added to the samples after SP3 clean-up. Samples were sonicated for 30 seconds in a water bath to disaggregate the beads fully. To accelerate the trypsin digestion (normally requiring an 18-hour incubation at 37°C) [39], samples were placed in a float inside a glass beaker containing 500 mL of cold water. The beaker was placed in a domestic microwave (LG Electronics: i-wave model) at 170 W for 6 minutes. After digestion, the tubes were left to cool for 5 minutes and then centrifuged for 1 minute at 16,000 × g to ensure that there will be no carryover of SP3 beads during the recovery of peptides. Tubes were placed on a magnetic rack to capture the beads, and digested supernatant (100 μL) was transferred to new Eppendorf tubes.

2.6. Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

2.6.1. Data dependent acquisition

An aliquot of 5 μL of each liquid sample (body fluids and mixture samples) obtained after SP3 clean-up was placed in individual vials for LC-MS/MS analysis. Using an Acquity M-class nanoLC system (Waters, USA), the sample was loaded at 15 μL/min for 3 minutes onto a nanoEase Symmetry C18 trapping column (180 mm×20 mm) before being washed onto a home-made column (75 μmID × 350 mm) with integrated emitter packed with SP-120–1.7-ODS-BIO resin (1.7 μm, Osaka Soda Co, Japan) and heated to 45°C. Peptides were eluted from the column and into the source of a Q Exactive Plus mass spectrometer (Thermo Scientific) using the following program: 5–30 % MS buffer B (98 % ACN + 0.2 % Formic Acid) over 90 minutes, 30–80 % MS buffer B over 3 minutes, 80 % MS buffer B for 2 minutes, 80–5 % for 3 minutes. The eluting peptides were ionised at 3000 V.

A Data Dependant MS/MS (dd-MS2) experiment was performed, with a survey scan (MS1 scan) of 350–1500 Da at 70,000 resolution for peptides of charge state 2+ or higher with an AGC target of 3×10^6 and maximum Injection Time of 50 ms. The Top 12 peptides in abundance were selected and fragmented in the Higher-energy Collisional Dissociation (HCD) cell to produce fragment ions using an isolation window of 1.4 *m/z*, an AGC target of 1×10^5 and maximum injection time of 100 ms. Fragments were scanned in the Orbitrap analyser (MS2 scan) at 17,500 resolution and the product ion fragment masses measured over a mass range of 120–2000 Da. The mass of the precursor peptide was then excluded for 30 seconds.

2.6.2. Data analysis

The identification of proteins in samples requires the confident identification of peptides associated with the discriminating proteins, achieved by the search of raw data against a library of expected peptide fragments derived from expected protein products of protein coding regions in the genome. The raw data files obtained after the instrument run were searched using the software PEAKS® Studio 11 (Bioinformatics Solutions Inc.) against the Human Proteome database (download date: 23/3/23). A common contaminant database was also added to the search. Search parameters were set as follows: parent mass error tolerance: 10.0 pp; fragment mass error tolerance: 0.02 Da; precursor mass search type: monoisotopic; maximum missed cleavages: 3; enzyme: trypsin; enzyme digest mode: semi-specific; peptide length range: 6–45; Variable Modifications: Carbamidomethylation (+57.02), Deamidation (NQ) (+0.98), Oxidation (M) (+15.99); Max Variable Post-translational modification (PTM) Per Peptide: 3; Database: Human; Taxon: all species; Searched Entries: 103789; Contaminant Database: Contaminants; Deep Learning Boost: Yes; False Discovery Rate (FDR) Estimation: Enabled.

Protein and peptide biomarker identification were performed manually in Microsoft Excel (Microsoft Corporation) based on specific criteria detailed in Sections 3.2–3.6. Mixture classification was also analysed based on protein and peptide biomarker selection and described in Section 3.7.

2.6.2.1. Peptide/protein identification threshold. Peptide matches obtained from data searches underwent statistical validation to prevent false positive identifications by calculation of an FDR [18,40,41]. The FDR was set as 1 % for identifications, which ensures reporting accurate and confident identifications of peptides that are potential biomarkers of body fluids.

A −10logP score is a statistical measure of the quality of a peptide-spectrum match (PSM) [31]. The −10logP score of a protein is the sum of −10logP values of peptides associated with that protein—the higher the −10logP score, the more confident the identification. All samples will have a different −10logP value defined by the 1 % FDR.

3. Results

3.1. DNA quantitation

A summary of the quantity of DNA present within DNA extracts of the blood, saliva and semen samples analysed in this study is present in Table 1. Recovered DNA was sufficient for achieving optimal DNA template input for most modern STR profiling assays (~ 1 ng) with highest recoveries for saliva, followed by semen and then blood. There was no DNA detected in extraction and PCR negative controls.

It is important to note that the quantity of DNA reported in each sample is related to the amount of body fluid recovered on the sample swab, which differs between the samples due to the collection procedure detailed in Section 2.2.

3.2. Peptide/protein identification

The number of PSMs and peptides identified in the samples after the 1 % FDR threshold was applied are reported in Table 2. The number of protein groups identified in the samples is also reported in Table 2, where a protein group is a collection of proteins that share one or more peptides in common, offering a broad view of protein composition within a sample.

As shown in Table 2, the number of total protein groups identified was generally highest in saliva and lowest in semen. These differences were also observed for the number of peptides and PSMs identified, with saliva having mostly higher values and semen mostly lower.

3.3. Identification of peptides by MS1 and MS2 Scans

MS2 scans obtained by the mass spectrometer are analysed by PEAKS® Studio 11 to identify the corresponding peptide sequence. The mass differences between adjacent fragment ions indicate the sequence of amino acid residues. Fig. 1 shows an MS1 and MS2 scan of the peptide SSYVLQTEELVVNK derived from Semenogelin-2.

The MS1 scan (Fig. 1A) shows the m/z ratios of precursor ions separated by the LC at a retention time of 109.85 minutes. The peak circled in red, with m/z 804.93, is a precursor ion for peptide SSYVLQTEELVVNK and has been selected for fragmentation in the HCD cell, resulting in different fragment ions shown in Fig. 1B.

The presence and relative intensities of the y and b ions identify a peptide, and the confident identification of a peptide leads to the identification of a protein from which it was derived. In this case, peptide SSYVLQTEELVVNK was identified.

3.4. Protein coverage

Proteins are only identified in a database search if a peptide(s) defining the protein is detected. These peptides are displayed as blue bars below the associated protein sequence, as shown in the protein coverage map in Fig. 2. Peptide SSYVLQTEELVVNK, referred to in Section 3.3, is circled in red and is one of many peptides identified in Semenogelin-2.

Table 1
Concentration (ng/μL) and quantity (ng) of DNA recovered from DNA extraction of blood, saliva and semen samples analysed in this study.

Body Fluid	Sample ID	Concentration of DNA (ng/μL)	Total amount of DNA (ng) in Eluent (40 μL)
Blood	BLD1	0.18	7.2
	BLD2	0.21	8.4
	BLD3	0.34	13.6
	BLD5	0.71	28.4
Saliva	SAL1	12.77	510.8
	SAL2	6.16	246.4
	SAL3	10.26	410.4
	SAL4	3.93	157.2
Semen	SMN1	2.58	103.2
	SMN2	2.56	102.4
	SMN3	3.16	126.4
	SMN4	2.39	95.6

Table 2
Summary of the number of PSMs, peptides and protein groups in the four blood, saliva and semen samples used for biomarker identification.

Body Fluid	Sample ID	#PSMs	#Peptides	#Protein groups
Blood	BLD1	3482	800	108
	BLD2	4036	882	118
	BLD3	4027	1036	140
	BLD4	4948	1215	166
Saliva	SAL1	6536	2211	372
	SAL2	8248	2782	450
	SAL3	7176	2382	375
	SAL4	6009	1913	320
Semen	SMN1	1510	396	58
	SMN2	7160	1356	186
	SMN3	3140	644	101
	SMN4	964	250	32

3.5. Identification of biomarkers for body fluids

After the peptide and protein threshold was set (Section 3.2), the identification of proteins and peptides that can act as biomarkers indicative of body fluids of interest was performed manually. The selection of proteins that were discriminating to a particular fluid was the first step, followed by the identification of peptide biomarkers. The first criterion for selecting discriminating proteins to certain body fluids was specificity, meaning that the proteins only present in the body fluid of interest and absent in all other body fluid samples were identified. The next criterion was sensitivity: proteins must be present at sufficient abundance within all samples analysed of the same body fluid. The ten consistently most abundant proteins across the samples that met the above specificity criteria were selected to make up a list of discriminating biomarkers for each body fluid. Keratin proteins and proteins identified from the common contaminants database were excluded from the list. In addition, sequences annotated as fragments and isoforms from the same protein were also excluded from the list, unless only a specific isoform met the above specificity and sensitivity criteria, in which case it was included.

The final list of selected protein biomarkers is presented in Table 3, ranked from most to least abundant. These identified proteins were present in all samples of the body fluid for which they were indicative and absent in all other body fluid samples. The average −10logP score and average percentage coverage of the protein is provided for the four blood, four saliva, and four semen samples analysed.

The threshold for the confident identification of peptides and proteins was set at 1 % FDR, as mentioned in Section 3.2. All proteins in Table 3 had −10logP values for proteins identified in blood, saliva, and semen that were higher than the −10logP threshold for 1 % FDR, indicating a confident identification of each protein in each body fluid.

In Fig. 2, the peptides identified for each protein are indicated as annotations below the protein sequence. A percentage coverage of the sequence was calculated as the number of amino acids in all identified peptides divided by the number of amino acids in the protein sequence, multiplied by 100. This value is shown in Table 3 as an average for all samples in which the protein was identified. The higher the percentage coverage value, the more peptide sequences were identified across the protein.

3.6. Identification of proteotypic peptides

As the detection of a protein relies on the detection of a peptide indicative of the protein, called a proteotypic peptide (PTP), targeted proteomic experiments target those discriminating peptides for

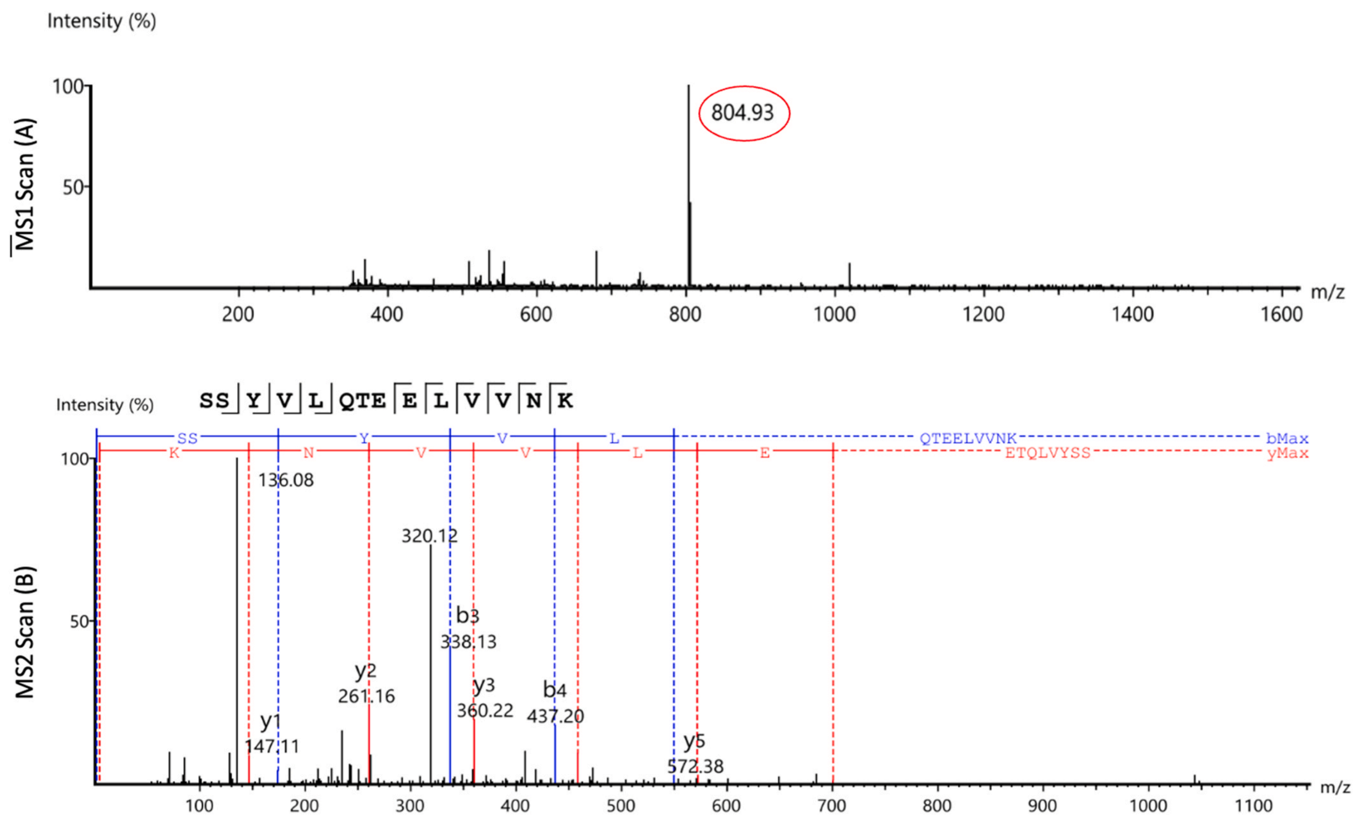


Fig. 1. MS1 and MS2 scans used in the sequencing of peptide SSYVLQTEELVVNK in Semenogelin-2. (A) MS1 scan from an LC retention time of 109.85 minutes. The x-axis shows the *m/z* ratios of all precursor ions present in that fraction at that retention time, and the y-axis shows the respective relative intensities of the ions (%). (B) MS2 scan of the selected precursor ion 804.93. The x-axis shows the fragment ions formed after HCD cell fragmentation, and the y-axis shows the relative intensities of the ions (%).

Fig. 2. Protein Coverage diagram for Semenogelin-2 showing amino acid residues 81 to 240. The blue lines represent the peptide sequences identified by software. PTMs are also shown on the peptides.

Table 3
Summary of discriminating proteins identified for blood, saliva, and semen (ranked from most to least abundant) with protein name, unique protein accession, abbreviation, average −10logP score, and average percentage coverage of samples analysed for each body fluid. The presence of these proteins in body fluids reported by others [19–21,24–26] is noted in the last column. Upon further testing and analysis of other fluid types, the list of potentially discriminating biomarkers for each body fluid may change, as there is evidence of Cornulin and Involucrin being associated with vaginal fluids, for example.

Body Fluid	Accession	Description	Abbreviation	Average −10logP	Average Coverage (%)	Proteins reported in other tissues
Blood	P00915	Carbonic anhydrase 1	CAH1	315.8	54.12	
	P02679	Fibrinogen gamma chain	FIBG	326.5	31.63	
	D6REL8	Fibrinogen beta chain	FIBB	288.11	31.34	
	P02790	Hemopexin	HEMO	278.64	19.59	Peripheral Blood [26]
	C9JVG0	Transferrin (Fragment)	TRFE	171.85	15.67	
	P02549−2	Isoform 2 of Spectrin alpha chain, erythrocytic 1	SPTA1	273.67	4.63	Peripheral blood [19]
	A0A140TA29	Complement C4-B	CO4B	268	5.96	
	P00918	Carbonic anhydrase 2	CAH2	259.53	30	
	C9JRG0	Hemoglobin subunit delta (Fragment)	HBD	153.07	20	
	P01008	Antithrombin-III	ANT3	308.48	23.76	
Saliva	A0A0G2JNB4	Basic salivary proline-rich protein 3	PRB3	319.08	25.08	Saliva [24]
	Q9UBC9	Small proline-rich protein 3	SPRR3	342.64	61.54	Vaginal fluid [24,25]
	A0A0G2JR74	Basic salivary proline-rich protein 1	PRB1	472.45	31.81	
	P02812	Basic salivary proline-rich protein 2	PRB2	472.45	43.51	
	E9PAL0	Basic salivary proline-rich protein 4	PRB4	412.92	55.87	
	Q9UBG3	Cornulin	CRNN	447.36	79.65	Vaginal fluid [20,21,24,26], menstrual blood[21]
	P07476	Involucrin	INVO	419.39	59.4	Vaginal fluid [20,21,24,26]
	P05109	Protein S100-A8	S10A8	259.45	44.09	Vaginal fluid [24]
	P04080	Cystatin-B	CYTB	316.32	79.34	
	A0A140T8X8	Mucin−21	MUC21	184.84	5.67	
Semen	Q02383	Semenogelin−2	SEMG2	489.36	60.91	Semen [19–21,24,26]
	P04279	Semenogelin−1	SEMG1	478.2	62.29	Semen [19–21,24,26]
	P07288	Prostate-specific antigen	PSA	194.11	14.85	Semen [19,21,24,26]
	M0R1F0	Kallikrein related peptidase 3 (Fragment)	KLK3	186.37	15.81	
	Q9NY87	Sperm protein associated with the nucleus on the X chromosome C	SPANXC	110.42	14.43	
	Q9BXN6	Sperm protein associated with the nucleus on the X chromosome D	SPANXD	110.42	14.43	Semen [24]
	Q9NS25	Sperm protein associated with the nucleus on the X chromosome B1	SPANXB	97.99	10.68	
	Q9NS26	Sperm protein associated with the nucleus on the X chromosome A	SPANXA	97.99	11.34	Semen [24]
	G3V5I3	Thioesterase (Fragment)	TE	95.11	24.44	
	Q5JQC9	A-kinase anchor protein 4	AKAP4	169.21	5.04	Semen [24]

identification purposes. Moreover, the coverage of a protein in a sample may not always be high, however, the identification of one PTP indicative of that body fluid can confidently identify the presence of the body fluid.

For each protein biomarker identified in Table 3, at least one PTP was identified and included in Table 4. The selection criteria were:

- The peptide length was between 8–25 amino acids.
- The peptide did not contain any trypsin cleavage sites within the sequence.
- The peptide charge state was a doubly or triply charged ion.
- The peptide was discriminating to the body fluid of interest and present in all samples of that body fluid.
- Where the peptide was not specific to one protein or proteoform, it was shared by proteins or proteoforms that were only found in one body fluid.

Table 4 includes a summary of the peptides identified based on the above criteria. Every peptide is presented including the protein it identifies, mass (m), charge (z), mass-to-charge ratio (*m/z*), length of the peptide, average −10logP score, precursor mass error (ppm), and retention time (RT) in the four samples of each body fluid analysed. The precursor mass error was calculated as 10^6 x (detected precursor mass – theoretical peptide mass)/theoretical peptide mass. This value reflects the error between the detected mass and the theoretical mass of the peptide. The mass error range threshold set for the identification of a peptide was 10 ppm. All peptides identified in the table have a ppm value less than 10, indicating a low mass error between the detected precursor mass and the theoretical mass of the peptide.

3.7. *Mixture classification*

To test the proposition that the detection of the discriminating peptides outlined in Section 3.6 is indicative of the presence of the body fluid of interest, mixed samples were analysed. These samples were collected as outlined in Section 2.2, where 2 μL of each body fluid was transferred directly to a swab before analysis, resulting in a 1:1 vol ratio for mixtures of two body fluids and a 1:1:1 vol ratio for the three-body fluid mixture. Mixture samples consisted of mixture one, where saliva and semen were deposited on the same swab; mixture two, where blood and semen were deposited on the same swab; mixture three, where blood and saliva were deposited on the same swab and mixture four, where blood, saliva, and semen were all deposited on the same swab. The non-DNA component left over after DNA extraction of mixture samples were prepared and analysed as outlined in Sections 2.5–2.6.

The LC-MS/MS data of the non-DNA component from the four mixed samples were searched for the presence of any of the peptides listed in the fourth column of Table 5 and noted as ‘Y’ for the presence of the peptide and ‘N’ for the absence of the peptide. An assessment of each mixture was performed with the presence of any PTP being sufficient for the inclusion of the relevant body fluid and the absence of all PTPs being required for the exclusion of the relevant body fluid. Predicted mixture components were then compared with the true mixture components. This serves as a further investigation of the performance of the selected peptides in identifying body fluids. However, extensive validation of the specificity and detectability of these peptides is required before adoption

Table 4
Summary of proteotypic peptides selected for blood, saliva, and semen protein biomarkers, not ranked according to any particular criteria, with protein name, unique protein accession, and peptide sequence and details including the mass (m), charge (z), mass-to-charge ratio (*m/z),* length, average −10logP score, precursor mass error (ppm), and retention time (RT) in the four samples of the body fluids analysed.

Body Fluid	Accession	Protein Name	Peptide	m	z	Length	-10logP	ppm	*m/z*	RT
Blood	P00915	Carbonic anhydrase 1	LYPIANGNNQSPVDIK	1741.9	2	16	104.83	−0.75	871.96	84.86
			ESISVSSEQLAQFR	1579.78	2	14	99.47	0.3	790.9	91.71
	P02679	Fibrinogen gamma chain	IHLISTQSAIPYALR	1681.95	3	15	108.9	0.03	561.66	101.04
			QSGLYFIKPLK	1292.75	3	11	108.71	0.95	431.92	92.24
	D6REL8	Fibrinogen beta chain	AHYGGFTVQNEANK	1534.72	3	14	109.89	1.23	512.58	55.68
			QGFGNVATNTDGK	1307.61	2	13	107.17	0.48	654.81	51.45
	P02790	Hemopexin	SGAQATWTELPWPHEK	1836.88	3	16	104.47	−0.75	613.3	103.31
			GGYTLVSGYPK	1140.58	2	11	107.7	0.93	571.3	75.14
	C9JVG0	Transferrin (Fragment)	DGAGDVAFVK	977.48	2	10	90.8	0.9	489.75	64.35
	P02549−2	Isoform 2 of Spectrin alpha chain, erythrocytic 1	ALSNAANLQR	1056.57	2	10	104.88	1.6	529.29	47.47
			FEALKEPLATR	1273.7	3	11	101.28	−0.28	425.57	71.74
	A0A140TA29	Complement C4-B	ASAGLLGAHAAAITAYALTLTK	2084.16	3	22	110.32	1.25	695.73	130.21
			LTVAAPPSGGPGFLSIERPDSRPPR	2573.37	4	25	109.08	−1.53	644.35	96.52
	P00918	Carbonic anhydrase 2	SADFTNFDPR	1168.51	2	10	107.38	0.45	585.26	76.99
			GGPLDGTYR	934.45	2	9	105.94	−0.03	468.23	54.82
	C9JRG0	Hemoglobin subunit delta (Fragment)	TAVNALWGK	958.52	2	9	104.86	1.78	480.27	77.12
	P01008	Antithrombin-III	SKLPGIVAEGR	1125.65	3	11	107.99	0.33	376.22	60.82
			TSDQIHFFFAK	1339.66	3	11	102.43	−0.35	447.56	99.65
Saliva	A0A0G2JNB4	Basic salivary proline-rich protein 3	TPPPPGKPEGR	1131.6	3	11	97.96	1.3	378.21	46.36
	Q9UBC9	Small proline-rich protein 3	VPVPGYTK	859.48	2	8	106.02	0.15	430.75	68.32
			VPDQGFIK	902.49	2	8	104.27	0.43	452.25	76.8
	A0A0G2JR74	Basic salivary proline-rich protein 1	GPPPPGKPQGPPPQGDNK	1763.9	3	18	88.8	−1.35	588.97	51.32
			PQGPPPQGDNK	1133.55	2	11	71.31	0.98	567.78	42.82
	P02812	Basic salivary proline-rich protein 2	GPPPPGKPQGPPPQGDNK	1763.9	3	18	88.8	−1.35	588.97	51.32
			PQGPPPQGDNK	1133.55	2	11	71.31	0.98	567.78	42.82
	E9PAL0	Basic salivary proline-rich protein 4	GRPPRPAQGQQPPQ	1512.79	3	14	110.84	0.15	505.27	48.11
	Q9UBG3	Cornulin	NQTTEMRPER	1260.59	3	10	100.53	0.13	421.2	47.89
	P07476	Involucrin	LLDQQLDQELVK	1440.78	2	12	108.53	0.78	721.4	104.56
			HLEHPEQQDGQLK	1557.75	3	13	110.06	1.2	520.26	50.68
	P05109	Protein S100-A8	GNFHAVYR	962.47	2	8	97.8	0.7	482.24	58.92
			KGADVWFK	949.5	2	8	66.81	0.55	475.76	81.28
	P04080	Cystatin-B	AKHDELTYF	1122.53	2	9	105.62	0.93	562.28	87.29
			VHVGDEDFVHLR	1421.71	3	12	110.27	1.58	474.91	88.92
	A0A140T8X8	Mucin−21	NTFNTAVYHPH	1299.6	3	11	90.21	0.55	434.21	71.19
Semen	Q02383	Semenogelin−2	DIFTTQDELLVYNK	1697.85	2	14	100.33	1.45	849.93	125.21
			SSYVLQTEELVVNK	1607.84	2	14	100.61	0.38	804.93	110.77
	P04279	Semenogelin−1	DIFSTQDELLVYNK	1683.84	2	14	100.58	0.53	842.93	124.56
			GISSQYSNTEER	1369.61	2	12	98.93	0.4	685.81	56.93
	P07288	Prostate-specific antigen	LSEPAELTDAVK	1271.66	2	12	97.2	2.23	636.84	90.94
	M0R1F0	Kallikrein related peptidase 3 (Fragment)	LSEPAELTDAVK	1271.66	2	12	97.2	2.23	636.84	90.94
	Q9NY87	Sperm protein associated with the nucleus on the X chromosome C	TSESSTILVVR	1190.65	2	11	97.99	0.7	596.33	86.46
	Q9BXN6	Sperm protein associated with the nucleus on the X chromosome D	TSESSTILVVR	1190.65	2	11	97.99	0.7	596.33	86.46
	Q9NS25	Sperm protein associated with the nucleus on the X chromosome B1	TSESSTILVVR	1190.65	2	11	97.99	0.7	596.33	86.46
	Q9NS26	Sperm protein associated with the nucleus on the X chromosome A	TSESSTILVVR	1190.65	2	11	97.99	0.7	596.33	86.46
	G3V5I3	Thioesterase (Fragment)	AVVEVDESGTR	1160.57	2	11	95.11	0.78	581.29	62.35
	Q5JQC9	A-kinase anchor protein 4	SQSLSYASLK	1082.56	2	10	96.81	0.63	542.29	79.08

as biomarkers for identification.

The ‘mixture classification’ row in Table 5 is the prediction for each mixture based solely on the presence and absence of any of the selected biomarker peptides as detected by LC-MS/MS. For example, the presence of peptides LYPIANGNNQSPVDIK and ESISVSSEQLAQFR in Mixture 3 indicates the presence of protein carbonic anhydrase 1. As these peptides have been identified as indicative of blood, then this mixture would be predicted to contain blood. This process was performed for the four mixtures, and the results were consistent with the true components for Mixtures 2 and 4. For Mixture 1, a false positive was observed, where peptide DGAGDVAFVK from TRFE was detected in the mixture, inferring the presence of blood, however, the true mixture components were saliva and semen only. For Mixture 3, a false negative was observed, i.e., no saliva peptides were detected, inferring that the mixture only contained blood, whereas the true mixture also contained saliva.

4. Discussion

The results of this proof-of-concept study show that it is possible to distinguish body fluids from each other by targeting peptide biomarkers in the non-DNA component of DNA extraction waste products. Peptide-based proteome profiles allowed the identification of potentially discriminating biomarkers in blood, saliva, and semen in these samples. However, extensive validation would need to be conducted on the DNA profiling waste from other body fluid types before determining a final list of potential biomarkers. Source- indicative proteins and peptide biomarkers that can be used to identify body fluids have been previously reported [19–37] (Table 3). Unlike these previous studies, this current study uses the non-DNA component after DNA extraction, which differs

Table 5
Summary of peptides present/absent in four prepared mixture samples, with the mixture classification based on PTPs identified in this study. 'Y' indicates the presence of the peptide and 'N' indicates the absence of the peptide in the sample.

				Mix 1	Mix 2	Mix 3	Mix 4
True Mixture components				saliva, semen	blood, semen	blood, saliva	blood, saliva, semen
Mixture Classification				blood, saliva, semen	blood, semen	blood	blood, saliva, semen
Body Fluid	Accession	Protein Name	Peptide	Mix1	Mix 2	Mix 3	Mix 4
Blood	P00915	Carbonic anhydrase 1	LYPIANGNNQSPVDIK	N	N	Y	N
			ESISVSSEQLAQFR	N	N	Y	Y
	P02679	Fibrinogen gamma chain	IHLISTQSAIPYALR	N	Y	Y	N
			QSGLYFIKPLK	N	Y	Y	N
	D6REL8	Fibrinogen beta chain	AHYGGFTVQNEANK	N	Y	Y	Y
			QGFGNVATNTDGK	N	N	Y	N
	P02790	Hemopexin	SGAQATWTELPWPHEK	N	N	N	N
			GGYTLVSGYPK	N	N	Y	N
	C9JVG0	Transferrin (Fragment)	DGAGDVAFVK	Y	N	N	Y
	P02549-2	Isoform 2 of Spectrin alpha chain, erythrocytic 1	ALSNAANLQR	N	N	Y	N
			FEALKEPLATR	N	N	N	N
	A0A140TA29	Complement C4-B	ASAGLLGAHAAAITAYALTLTK	N	N	Y	N
			LTVAAPPSGGPGFLSIERPDSRPPR	N	N	Y	N
	P00918	Carbonic anhydrase 2	SADFTNFDPR	N	Y	Y	Y
			GGPLDGTYR	N	N	Y	N
	C9JRG0	Hemoglobin subunit delta	TAVNALWGK	N	Y	Y	Y
	P01008	Antithrombin-III	SKLPGIVAEGR	N	N	Y	Y
			TSDQIHFFFAK	N	Y	Y	Y
Saliva	A0A0G2JNB4	Basic salivary proline-rich protein 3	TPPPPGKPEGR	N	N	N	N
	Q9UBC9	Small proline-rich protein 3	VPVPGYTK	Y	N	N	N
			VPDQGFIK	N	N	N	Y
	A0A0G2JR74	Basic salivary proline-rich protein 1	GPPPPGKPQGPPPQGDNK	Y	N	N	N
			PQGPPPQGDNK	N	N	N	N
	P02812	Basic salivary proline-rich protein 2	GPPPPGKPQGPPPQGDNK	Y	N	N	N
			PQGPPPQGDNK	N	N	N	N
	E9PAL0	Basic salivary proline-rich protein 4	GRPPRPAQGQQPPQ	N	N	N	N
	Q9UBG3	Cornulin	NQTTEMRPER	N	N	N	N
	P07476	Involucrin	LLDQQLDQELVK	N	N	N	N
			HLEHPEQQDGQLK	N	N	N	N
	P05109	Protein S100-A8	GNFHAVYR	N	N	N	N
			KGADVWFK	N	N	N	N
	P04080	Cystatin-B	AKHDELTYF	N	N	N	N
			VHVGDEDFVHLR	N	N	N	N
	A0A140T8X8	Mucin-21	NTFNTAVYHPH	N	N	N	N
Semen	Q02383	Semenogelin-2	DIFTTQDELLVYNK	Y	Y	N	Y
			SSYVLQTEELVVNK	Y	Y	N	Y
	P04279	Semenogelin-1	DIFSTQDELLVYNK	Y	Y	N	Y
			GISSQYSNTEER	Y	Y	N	Y
	P07288	Prostate-specific antigen	LSEPAELTDAVK	Y	Y	N	N
	M0R1F0	Kallikrein related peptidase 3	LSEPAELTDAVK	Y	Y	N	N
	Q9NY87	Sperm protein associated with the nucleus on the X chromosome C	TSESSTILVVR	Y	N	N	N
	Q9BXN6	Sperm protein associated with the nucleus on the X chromosome D	TSESSTILVVR	Y	N	N	N
	Q9NS25	Sperm protein associated with the nucleus on the X chromosome B1	TSESSTILVVR	Y	N	N	N
	G3V5I3	Thioesterase	AVVEVDESGTR	Y	N	N	N
	Q9NS26	Sperm protein associated with the nucleus on the X chromosome A	TSESSTILVVR	Y	N	N	N
	Q5JQC9	A-kinase anchor protein 4	SQSLSYASLK	N	N	N	N

significantly from the starting material used for other mass spectrometry-based body fluid identification techniques that typically start with the body fluid of interest and undergo standard sample preparation for mass spectrometry analysis. For this reason, it was important to conduct a new identification process for proteins and peptides from this new sample type rather than adopt those discovered and developed by others in order to observe whether similar results would be achieved. It is important to emphasise that the proteins selected in this study as potentially discriminating are only based on three body fluids and the small number of samples of each tested, and further refinement of the protein and peptide list will need to be conducted by conducting further validation studies with other body fluids.

The recovered DNA was sufficient to achieve optimal DNA template input for most modern STR profiling assays, as shown in Table 1. The differences in the quantity of DNA reported in each sample (Table 1) could be due to the difference in the amount of body fluid collected from each individual, as detailed in Section 2.2, and the difference in DNA content based on the fluid type [42]. The ability to prioritise DNA recovery from samples while also giving source-level information from the non-DNA component after DNA extraction is potentially valuable for forensic investigations.

The total number of peptides identified in samples prepared by the method described in Section 2 is generally highest in saliva and lowest in semen, as seen in Table 2. There is also a difference observed in the number of peptides recovered amongst the four samples for each body fluid. A similar trend is observed for the number of PSMs and protein groups. This could be due to several reasons, including the difference in the starting volume of body fluid recovered from the volunteers, described in Section 2.2. Since the same volume of semen was deposited onto swabs for the analysis of all four semen samples (10 μL), the differences in the number of peptides could be attributed to variations in sperm count in the semen of individuals [43,44]. Although the same volume of semen is deposited onto swabs for analysis, the volume of blood and saliva deposited on swabs by volunteers was variable. It has been estimated that 1 μg of protein can be derived from approximately 3 nL of blood, 50 nL of semen, and 400 nL of saliva [19]. This means that for the same volumes of blood, saliva, and semen, we would expect to see more protein in blood, less in semen and least in saliva. The proportion of protein recovered in the non-DNA component of DNA extraction waste is also unknown, which may account for the difference in protein levels across samples.

It is also important to note that Proteinase K was not used in the DNA extraction procedures of any samples. Proteinase K is a broad-spectrum serine protease that hydrolyses peptide bonds to degrade proteins, and its use in DNA extraction procedures is mainly to aid in isolating DNA by removing proteins that could degrade DNA and RNA in the sample [45]. Broad specificity proteases, such as proteinase K, are not widely used when preparing samples for proteomic analysis due to the high complexity and random nature of the peptide mixtures that they could generate, reducing peptide-based identifications [46]. In comparison to the well-established reproducibility of peptides with trypsin digestion, it is more difficult to reproducibly identify the same peptides in replicates with proteinase K digestion [47]. Whilst the usage of proteinase K is standard in some DNA extraction silica-based kits used in forensic biology, such as QIAamp® DNA Investigator Kit (Qiagen) and QIAamp® DNA Mini Kit (Qiagen), the PrepFiler Express™ Forensic DNA Extraction Kit (Applied Biosystems®) used in this study does not require its addition to the extraction procedure when analysing body fluids [48], hence preserving the proteomic content in the waste product after DNA extraction, and allowing for further proteomic analysis to be conducted. QIAamp® DNA Blood Mini Kit (Qiagen) also does not use proteinase K, but a QIAGEN® Protease is added instead, and this extraction method can be used to prepare DNA from blood, cells, and body fluids [48]. The impact of the QIAGEN protease on proteome analyses would need to be studied. Another kit not requiring the use of Proteinase K is the DNA IQ™ System (Promega) for DNA extraction from body fluids [48], where Proteinase K is only added for specific samples such as hair and bones [49]. Moreover, DNA kits used for the preparation of samples for direct PCR processing, such as Casework Direct® Kit (Promega) and Investigator Casework GO!® Kit (Qiagen) do not require the addition of Proteinase K during the extraction process [50]. When extracting DNA from semen samples, it is a routine practice to add Proteinase K [51], together with dithiothreitol (DTT), to degrade proteins in the acrosome of spermatozoa, but this is not necessary for other types of cells, including blood and buccal cells, and interferes with downstream peptide analysis [52]. It is possible to extract DNA from semen without the addition of Proteinase K, however [53]. Kranes et al. conducted a study where proteinase K was replaced with trypsin for a standard DNA extraction method using Millipore Microcon MW100 filter units, and the results show that DNA yields with this method were similar to those where proteinase K was employed, but the PCR-STR results were better [54]. Further experimentation would need to be conducted to validate the efficacy of DNA extraction with and without the addition of Proteinase K, and its impact on proteome analyses, when using workflows where this is routinely added. This is essential before the potential adoption of this method in routine analysis.

Processes described in Sections 3.3–3.6 lead to the identification of potentially discriminating peptide biomarkers for blood, saliva, and semen, with a list of proteins reported in Table 3, and a list of PTPs reported in Table 4. The selected biomarkers were limited to those in the database that appeared in only four blood, four saliva, and four semen samples. As only three body fluids were analysed in this study, the results only reflect the analysis of these samples, which are not exhaustive, and the selected proteins and peptides will require further testing and validation with a larger sample size from each body fluid and more body fluid types. Although this is a limitation in terms of sample size, the reported results are only preliminary and indicative of the potential of employing such a method for forensic body fluid identification, therefore, this direct comparison approach was fit for the purpose of identifying potentially discriminating biomarkers to distinguish between body fluids for forensic purposes. Moreover, it was important to use the non-DNA component of DNA extraction waste specifically. The selection of ten discriminating proteins for each body fluid as a first step was conducted to refine and limit the list of discriminating proteins and to increase confidence in the results by selecting the discriminating proteins that are of consistently high abundance within the samples of the same body fluid. Upon further testing, other possible proteins could be considered if those in the current list were found in other fluid types.

For blood, the proteins CAH1, FIBG, FIBB, HEMO, TRFE, SPTA1, CO4B, CAH2, HBD, and ANT3 (Table 3) were identified as potentially discriminating biomarkers to be used for fluid identification purposes. Hemoglobin subunit alpha (HBA) and Hemoglobin subunit beta (HBB) were present in all four blood samples, but they were also found in some saliva samples, so were excluded. Other studies with a similar purpose have reported one or both as potential biomarkers for blood [19,21,55]. HBA and HBB have also been found in shotgun proteomic studies of saliva samples, however [56]. Moreover, the ABACard® Hematrace® confirmatory test for the presence of hemoglobin has been shown to give a high false positive rate with saliva samples [9]. Hemoglobin is known to be an abundant protein in plasma, but not present in glandular salivary secretion [57]. For this reason, hemoglobin has been used as a marker to measure the level of blood contamination in saliva [57]. Also, an increase in hemoglobin levels in saliva has been detected in obese patients [58]. For forensic body fluid identification purposes, it is important to select biomarkers that are not only highly abundant in certain body fluids but also proteins that are not readily found in other body fluids. The sensitivity of nano-LC coupled with hybrid quadrupole-orbitrap MS/MS used in this study was sufficient to detect low hemoglobin levels in saliva, which would impact the use of HBA and/or HBB as biomarkers for blood. HBD was only found in the blood samples analysed in this study. HBD is far less abundant in blood than HBA and HBB since a mutation in the promoter region of the δ-globin gene causes the production of the δ-chain (HBD) to be low [59]. SPTA1 has previously been reported in blood and not in the proteomes of 15 other body fluids [19], in agreement with the findings of this study. HEMO has also been previously selected as a potential biomarker for blood [26].

For saliva, proteins PRB3, SPRR3, PRB1, PRB2, PRB4, CRNN, INVO, S10A8, CYTB, and MUC21 (Table 3) were identified as discriminating biomarkers for saliva in this study. However, as seen in Table 3, many of the proteins identified as potentially discriminating for saliva have been reported in other body fluids by other groups [20,21,24–26], which means they may be unsuitable for identifying saliva. For example, CRNN and INVO have been identified elsewhere as potential biomarkers for vaginal fluid [20,21,24,26]. Until the non-DNA component of more body fluids are analysed, including urine, menstrual blood and vaginal fluid, proteins SPRR3, CRNN, INVO and S10A8 cannot be confidently assigned as discriminating biomarkers for saliva, given their reported presence in other body fluids by other groups (Table 3). For semen, discriminating proteins included SEMG2, SEMG1, PSA, KLK3, SPANXC, SPANXD, SPANXB, SPANXA, TE and AKAP4 (Table 3). SEMG2, SEMG1, and PSA have also been reported elsewhere as potential biomarkers for the identification of semen [19,21,26,55]. Upon further testing and

analysis of other fluid types, the list of potential discriminating biomarkers for each body fluid will likely change. Further testing with other body fluids, such as menstrual blood, vaginal fluid, and urine, will allow further validation of the specificity of the identified biomarkers to their respective body fluids.

Discovery proteomics involves analysing peptides that are the result of trypsin digestion of proteins using MS/MS instruments to predict their amino acid sequence. This infers the presence of proteins in the sample [60]. The peptides are matched against a protein sequence database to determine the open-reading frame from which the protein originated (Sections 3.3–3.4). The method used in this study leverages the advantages that come with analysing peptides rather than intact proteins, which include the increase in sensitivity derived from small molecules that are more easily ionised and fragmented [60]. In this study, the biomarkers that were used to identify body fluids by mass spectrometry-based proteomics were PTPs (Table 4) derived from proteins. While a specific peptide sequence may be present in multiple open-reading frame products, a single peptide indicative of body fluid is adequate for identification purposes. This means that a PTP that is discriminating to a body fluid may originate from multiple proteins which are also discriminating to the body fluid of interest.

Due to the complexity of bottom-up proteomics, it is common for database identifications to assign protein clusters, rather than specific molecular forms of a protein, or proteoforms [60,61]. Proteoforms have a range of structures and molecular forms derived from a parent protein coded by a single specific gene, including genetic variations, alternative splicing, and PTMs [62–64]. A single peptide can also identify multiple proteoforms of the same protein. For example, a peptide like IHLISTQSAIPYALR (Table 4) may identify multiple forms of Fibrinogen gamma chain (P02679 and P02679–2). However, this peptide is indicative of blood, so it can serve as a biomarker for forensic body fluid identification regardless of its protein origin, as long as the proteins themselves are also indicative of that body fluid. In fact, when a single peptide identifies multiple proteoforms of a protein, it increases the sensitivity for detection.

We identified the ten consistently most abundant proteins for each body fluid type, but it is important to note that while proteins are initially selected during the biomarker identification steps, the study is fundamentally peptide-derived. One benefit of identifying peptide biomarkers is that they can be used to develop targeted approaches on instruments more readily available in forensic laboratories [65]. Table 4 lists at least one PTP for each protein in the list, that satisfies the selection criteria described in Section 3.6. Where possible, a PTP that is completely specific to the protein of interest is listed, e.g., PTP DIFTTQDELLVYNK from SEMG2 and PTP TAVNALWGK from HBD. However, some of the PTPs presented in Table 4 are present in multiple proteins, such as GPPPPGKPQGPPPQGDNK, shared by PRB1 and PRB2, and PQGPPPQGDNK, shared by PRB2 and PRB4. Because PRB1, PRB2 and PRB4 are all indicative of saliva, these PTPs are also discriminating to saliva.

Mixture assessment was based solely on the list of PTPs, and proteins reported from this study. It could change if other PTPs from other proteins were selected as biomarkers for identification. This small mixture assessment has shed light on the performance of the selected peptides in identifying body fluids. However, extensive validation of these peptides is required before they can be used as biomarkers for all body fluids. It is important to test the specificity of these peptides for their respective body fluids when analysing other body fluid types that have not been included in this study, such as vaginal fluid, urine, menstrual blood, etc. Further studies could include determining the limit of detection in single body fluid samples to which they are discriminating. Moreover, observing the stability of these peptides when exposed to different environmental factors would also be valuable to validate these peptides as body fluid biomarkers.

The classification according to the presence of prospective biomarker peptides was accurate for mixtures two and four (Table 5). For Mixture 1, a false positive result was reported, where the presence of one PTP from TRFE falsely indicated the presence of blood in the mixture, which only contained saliva and semen. Transferrin (TRFE) is a protein synthesised by the liver and released into the blood [66], however, it has also been found in cerebrospinal fluid (CSF) and semen [67]. This could suggest that PTP DGAGDVAFVK and protein TRFE may not be specific to blood [68], and further validation studies are required with other forensically relevant body fluids, a larger sample size, and more complex mixtures. A false negative was reported for Mixture 3 (blood/saliva), where no PTPs for saliva were detected. This could suggest that the selected PTPs for saliva are not universally present in saliva or not present in all samples at detectable levels. It may also reflect the fact that the saliva used to produce the mixture was derived from spitting, which may have included more non-cell components and may not have included many buccal cells that were used in the discovery phase of the project. The results for Mixtures 1 and 3 could also be due to the amount of protein present in the body fluids of each of the mixtures. Mixtures were prepared in a 1:1 vol ratio for Mixtures 1, 2 and 3 and a 1:1:1 vol ratio for Mixture 4 before being subjected to DNA extraction and sample preparation for proteomic analysis. It is known that the protein content in body fluids differs, where 1 μg of protein is derived from approximately 3 nL of blood, 50 nL of semen, and 400 nL of saliva [19]. The total protein content of saliva is low, with studies showing a concentration of 0.5–2 mg/mL of proteins in saliva from healthy individuals [69]; this may have had an impact on the detection of salivary peptides when saliva was added to mixtures at a similar ratio with other body fluids having a higher density of proteins. Whilst this mixture assessment does provide some information on the performance of the peptides, it is essential to conduct extensive validation studies of these biomarkers, including mixtures of different body fluids with a variety of compositions. These recommended studies include different volumes of single body fluid source samples to determine the detectability of each peptide at different starting volumes of body fluid. This is particularly important as the PTPs listed in Table 4 have been detected in all four of the respective single-source body fluid samples analysed; however, when the relevant body fluid was present in the mixtures analysed, not all discriminating PTPs identified for that fluid were present. For example, peptide TPPPPGKPEGR from PRB3 was found to be indicative for saliva and present in all four samples analysed; however, it was not detected in mixtures 1, 3, and 4, which contained saliva. The selection of appropriate candidate PTPs may require the consideration of further selection criteria for inclusion [70], in addition to those described in Section 3.6, as the physicochemical properties of a PTP may affect detectability in LC-MS/MS analyses [71], preventing it from being a suitable biomarker for body fluid identification. As seen in Table 5, many peptides were not detected in any of the mixtures where the presence of a body fluid suggests they should be present, resulting in a high false negative rate. The absence of those peptides in the mixtures could be attributed to the different concentrations of proteins in the body fluids analysed. Further testing and experimentation are required to assess the performance of these peptides as biomarkers for their respective body fluids, and those that are not readily detectable by LC-MS/MS should be screened out. Moreover, for preliminary peptide screening, an assessment of the specificity of the peptide to the respective fluid should be conducted by analysing more samples of the same body fluid and comparing them to other body fluids not analysed in this study. Sensitivity studies should also be conducted on various dilutions of body fluid. Then, the prevalence of true and false positives for each peptide could be assessed. Ultimately, once validation is conducted, allowing for a refinement of a suitable candidate PTP list, it may be reasonable to assign a threshold of a certain number of PTPs required to confidently identify a body fluid. In practice, to meet the standard required for a forensic application, many peptide biomarkers are likely to be required for a body fluid classification. This work demonstrates the potential peptide markers that could be used but is not an exhaustive list.

The Scientific Working Group on DNA Analysis Methods (SWGDAM)

suggests guidelines for validation studies involving DNA analysis [72], which could be applied to proteomic methods and the validation of peptides used for identification. Biomarker identification in proteomics is complex [73], and requires validation on different levels, including the specificity of proteins/peptides to a body fluid, the selection of the appropriate proteins, and then the selection of the relevant peptides.

5. Conclusion

Preliminary results from this proof-of-concept study have shown that it is possible to identify body fluid biomarkers from the non-DNA component of DNA extraction waste as a starting material. However, results are limited to three body fluids: blood, semen and saliva. This study is only the beginning of the discovery aspect of body fluid identification, which requires further development and validation prior to any implementation. The next phase of our research is to validate the specificity of the selected protein biomarkers by analysing other body fluid types not included in this study. More samples from other individuals and different body fluid types (such as menstrual blood, vaginal fluid, and urine) need to be analysed. Further experiments are required to determine the detectability of discriminating makers in forensically relevant samples, including samples of different volumes, a more comprehensive range of mixture ratios and biological fluid combinations, and sample stains exposed to different environmental insults over various time periods. Extensive validation studies, perhaps using the SWGDAM guidelines [72], would also need to be conducted before this method can be adopted by the forensic community. We have demonstrated, however, the feasibility of a mass spectrometry-based method that allows for multiple analytes to be targeted simultaneously from a DNA profiling waste stream. Results from this discovery method could potentially be translated into a targeted approach on an instrument readily available in forensic laboratories, such as a triple-quadrupole mass spectrometer [74]. The overall goal is to extract valuable information from body fluids without compromising their state, amount, or condition in an approach that maximises the use of biological traces found at crime scenes. Whilst this study aims to address limitations presented by current confirmatory tests by developing a multiplex testing approach, a goal shared with other laboratories developing a mass spectrometry-based method for body fluid identification [19–37], it also prioritises and preserves the DNA from the biological trace.

CRediT authorship contribution statement

Layal Zaarour: Writing - review & editing, Writing - original draft, Validation, Methodology, Formal analysis, Data curation. **Matthew Padula:** Writing - review & editing, Supervision, Resources, Methodology. **Roland van Oorschot:** Writing - review & editing, Supervision, Conceptualization. **Dennis McNevin:** Writing - review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is supported by an Australian Government Research Training Program Scholarship.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [75] partner repository with the dataset identifier PXD057374.

References

[1] I.W. Evett, P.D. Gill, G. Jackson, J. Whitaker, C. Champod, Interpreting small quantities of DNA: the hierarchy of propositions and the use of Bayesian networks, J. Forensic Sci. 47 (3) (2002) 520–530.
[2] S.A. Harbison, R.I. Fleming, Forensic body fluid identification: state of the art, Res. Rep. Forensic Med. Sci. 6 (2016) 11.
[3] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: deciding which level to address in casework, Sci. Justice 38 (4) (1998) 231–239.
[4] K. Virkler, I.K. Lednev, Analysis of body fluids for forensic purposes: From laboratory testing to non-destructive rapid confirmatory identification at a crime scene, Forensic Sci. Int. 188 (1) (2009) 1–17.
[5] F. Zapata, I. Gregório, C. García-Ruiz, Body fluids and spectroscopic techniques in forensics: a perfect match, J. Forensic Med 1 (1) (2015) 1–7.
[6] J.-H. An, K.-J. Shin, W.-I. Yang, H.-Y. Lee, Body fluid identification in forensics, BMB Rep. 45 (10) (2012) 545–553.
[7] D. McNevin, M. Padula, Serology: Blood, in: M.M. Houck (Ed.), Encyclopedia of Forensic Sciences, Third Edition, Elsevier, Oxford, 2023, pp. 470–483.
[8] M.M. Hobbs, M.J. Steiner, K.D. Rich, M.F. Gallo, L. Warner, M. Macaluso, Vaginal swab specimen processing methods influence performance of rapid semen detection tests: a cautionary tale, Contraception 82 (3) (2010) 291–295.
[9] I. Horjan, L. Barbaric, G. Mrsic, Applicability of three commercially available kits for forensic identification of blood stains, J. Forensic Leg. Med. 38 (2016) 101–105.
[10] G.A. Williams, Body fluid identification: A case for more research and innovation, Forensic Sci. Int. Rep. 2 (2020) 100096.
[11] L. Wallis, C. Hitchcock, D. McNevin, J. Raymond, Source Level Attribution: DNA Profiling from the ABAcard® HemaTrace® Kit, Forensic Sci. (2021) 116–129.
[12] T.-T. Ho, R. Roy, Generating DNA profiles from immunochromatographic cards using LCN methodology, Forensic Sci. Int.: Genet. 5 (3) (2011) 210–215.
[13] S. Turrina, G. Filippini, R. Atzei, E. Zaglia, D. De Leo, Validation studies of rapid stain identification-blood (RSID-blood) kit in forensic caseworks, Forensic Sci. Int. Genet. Suppl. Ser. 1 (1) (2008) 74–75.
[14] A.D. Roeder, C. Haas, Body Fluid Identification Using mRNA Profiling, Methods Mol. Biol. 1420 (2016) 13–31.
[15] Y. Xu, J. Xie, Y. Cao, H. Zhou, Y. Ping, L. Chen, L. Gu, W. Hu, G. Bi, J. Ge, X. Chen, Z. Zhao, Development of highly sensitive and specific mRNA multiplex system (XCYR1) for forensic human body fluids and tissues identification, PloS One 9 (7) (2014) e100123-e100123.
[16] T. Sijen, S. Harbison, On the Identification of Body Fluids and Tissues: A Crucial Link in the Investigation and Solution of Crime, Genes 12 (11) (2021).
[17] M. Hamza, D. Sankhyan, S. Shukla, P. Pandey, Advances in body fluid identification: MiRNA markers as powerful tool, Int. J. Leg. Med. 138 (4) (2024) 1223–1232.
[18] G.J. Parker, H.E. McKiernan, K.M. Legg, Z.C. Goecker, Forensic Proteom., Forensic Sci. Int.: Genet. 54 (2021), 102529-102529.
[19] H. Yang, B. Zhou, H. Deng, M. Prinz, D. Siegel, Body fluid identification by mass spectrometry, Int. J. Leg. Med. 127 (6) (2013) 1065–1077.
[20] M. Zhao, Y. Yang, Z. Guo, C. Shao, H. Sun, Y. Zhang, Y. Sun, Y. Liu, Y. Song, L. Zhang, Q. Li, J. Liu, M. Li, Y. Gao, W. Sun, A Comparative Proteomics Analysis of Five Body Fluids: Plasma, Urine, Cerebrospinal Fluid, Amniotic Fluid, and Saliva, Proteom. Clin. Appl. 12 (6) (2018) e1800008.
[21] K. Van Steendam, M. De Ceuleneer, M. Dhaenens, D. Van Hoofstat, D. Deforce, Mass spectrometry-based proteomics as a tool to identify biological matrices in forensic science, Int. J. Leg. Med. 127 (2) (2013) 287–298.
[22] K.M. Legg, R. Powell, N. Reisdorph, R. Reisdorph, P.B. Danielson, Verification of protein biomarker specificity for the identification of biological stains by quadrupole time-of-flight mass spectrometry, Electrophoresis 38 (6) (2017) 833–845.
[23] FLN-TWG, Implementation Strategies: Proteomic Mass Spectrometry for Biology Fluid Identification, in: N.I.o.J. Forensic Technology Center of Excellence (Ed.) Forensic Technology Center of Excellence, National Institute of Justice, Washington, DC, 2021.
[24] J. Zhang, M. Yan, A. Ji, Q. Sun, W. Ying, Mass spectrometry-based proteomic analysis of biological stains identifies body fluids specific markers, Forensic Sci. Int. 357 (2024) 112008.
[25] A. Igoh, Y. Doi, K. Sakurada, Identification and evaluation of potential forensic marker proteins in vaginal fluid by liquid chromatography/mass spectrometry, Anal. Bioanal. Chem. 407 (23) (2015) 7135–7144.
[26] K.M. Legg, R. Powell, N. Reisdorph, R. Reisdorph, P.B. Danielson, Discovery of highly specific protein markers for the identification of biological stains, ELECTROPHORESIS 35 (21-22) (2014) 3069–3078.
[27] M. Pieri, A. Silvestre, M. De Cicco, G. Mamone, E. Capasso, F. Addeo, G. Picariello, Mass spectrometry-based proteomics for the forensic identification of vomit traces, J. Proteom. 209 (2019) 103524.
[28] A. Illiano, V. Arpino, G. Pinto, A. Berti, V. Verdoliva, G. Peluso, P. Pucci, A. Amoresano, Multiple Reaction Monitoring Tandem Mass Spectrometry Approach for the Identification of Biological Fluids at Crime Scene Investigations, Anal. Chem. 90 (9) (2018) 5627–5636.
[29] H.E. McKiernan, C.O. Brown, L.C. Arantes, P.B. Danielson, K.M. Legg, NextGen Serology: Leveraging Mass Spectrometry for Protein-Based Human Body Fluid Identification. Applications in Forensic Proteomics: Protein Identification and Profiling, American Chemical Society, 2019, pp. 47–80.
[30] D. Siegel, Molecular Serology Validation: Proteomic Body Fluid Assay Validation, New York City Office of Chief Medical Examiner, New York, 2023, pp. 1–44.

[31] K. Legg, Validation of a Mass Spectrometry Based Serological Assay, United States Department of Defense, Research Triangle Park, North Carolina, 2016.
[32] H. McKiernan, Targeted-Ion Mass Spectrometry for the Identification of Forensically Relevant Biological Fluids and Samples from Sexual Assault Evidence. Natural Sciences and Mathematics, University of Denver, Electronic Theses and Dissertations, 2019.
[33] Y.T. Mechthild Prinz, Donald Siegel, Heyi Yang, Bo Zhou, Haiteng Deng, Establishment of a Fast and Accurate Proteomic Method for Body Fluid/Cell Type Identification, in: U.S.D.o. Justice (Ed.) U.S. Department of Justice, 2011, pp. 1-62.
[34] D. Siegel, Development of a Proteomic Assay for Menstrual Blood, Vaginal Fluid and Species Identification in: U.S.D.o. Justice (Ed.) Office of Chief Medical Examiner, New York, 2018, pp. 1-95.
[35] H. Yang, B. Zhou, M. Prinz, D. Siegel, Proteomic Analysis of Menstrual Blood, Mol. Cell. Proteom. 11 (10) (2012) 1024–1035.
[36] R. Davidovics, Y.L. Saw, C.O. Brown, M. Prinz, H.E. McKiernan, P.B. Danielson, K. M. Legg, High-throughput seminal fluid identification by automated immunoaffinity mass spectrometry, J. Forensic Sci. 67 (3) (2022) 1184–1190.
[37] H.E. McKiernan, P.B. Danielson, C.O. Brown, M. Signaevsky, C.G. Westring, K. M. Legg, Developmental validation of a multiplex proteomic assay for the identification of forensically relevant biological fluids, Forensic Sci. Int. 326 (2021) 110908.
[38] C.S. Hughes, S. Moggridge, T. Müller, P.H. Sorensen, G.B. Morin, J. Krijgsveld, Single-pot, solid-phase-enhanced sample preparation for proteomics experiments, Nat. Protoc. 14 (1) (2019) 68–85.
[39] A.R. Vaezzadeh, J.M. Deshusses, P. Waridel, P. François, C.G. Zimmermann-Ivol, P. Lescuyer, J. Schrenzel, D.F. Hochstrasser, Accelerated digestion for high-throughput proteomics analysis of whole bacterial proteomes, J. Microbiol Methods 80 (1) (2010) 56–62.
[40] J. Zhang, L. Xin, B. Shan, W. Chen, M. Xie, D. Yuen, W. Zhang, Z. Zhang, G. A. Lajoie, B. Ma, PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification, Mol. Cell. Proteom. 11 (4) (2012). M111.010587-M111.010587.
[41] K. Jung, Statistical Analysis in Proteomics, 1st ed. 2016. ed., Springer New York, New York, NY, 2016.
[42] J. Siuta, A. Dobosz, J. Kawecki, T. Dobosz, DNA Content of Various Fluids and Tissues of the Human Body, Genes 15 (1) (2024) 17.
[43] E. Carlsen, J.H. Petersen, A.-M. Andersson, N.E. Skakkebaek, Effects of ejaculatory frequency and season on variations in semen quality, Fertil. Steril. 82 (2) (2004) 358–366.
[44] R. Rahban, S. Nef, Regional difference in semen quality of young men: a review on the implication of environmental and lifestyle factors during fetal life and adulthood, Basic Clin. Androl. 30 (2020) 16.
[45] W. Qamar, M.R. Khan, A. Arafah, Optimization of conditions to extract high quality DNA for PCR analysis from whole blood using SDS-proteinase K method, Saudi J. Biol. Sci. 24 (7) (2017) 1465–1469.
[46] T. Dau, G. Bartolomucci, J. Rappsilber, Proteomics Using Protease Alternatives to Trypsin Benefits from Sequential Digestion with Trypsin, Anal. Chem. 92 (14) (2020) 9523–9527.
[47] J.C. Rogers, R.D. Bomgarden, Sample Preparation for Mass Spectrometry-Based Proteomics; from Proteomes to Peptides, in: H. Mirzaei, M. Carrasco (Eds.), Modern Proteomics – Sample Preparation, Analysis and Practical Applications, Springer International Publishing, Cham, 2016, pp. 43–62.
[48] M.M. Foley, Applied Biosystems™ PrepFiler™ Forensic DNA Extraction Kit (Manual and Semi-automated via AutoMate Express™), in: C. Cupples Connon (Ed.), Forensic DNA Analysis: Methods and Protocols, Springer US, New York, NY, 2023, pp. 53–81.
[49] Promega, DNA IQ System - Small Sample Casework Protocol, Tech. Bull. (2021).
[50] L. Dierig, M. Schwender, P. Wiegand, Looking for the pinpoint: Optimizing identification, recovery and DNA extraction of micro traces in forensic casework, Forensic Sci. Int.: Genet. 44 (2020) 102191.
[51] A. Weyrich, Preparation of Genomic DNA from Mammalian Sperm, Curr. Protoc. Mol. Biol. 98 (1) (2012) 2.13.1–2.13.3.
[52] H.E. McKiernan, P.B. Danielson, Chapter 21 - Molecular Diagnostic Applications in Forensic Science, in: G.P. Patrinos (Ed.), Molecular Diagnostics (Third Edition), Academic Press, 2017, pp. 371–394.
[53] M.G. Brevnov, H.S. Pawar, J. Mundt, L.M. Calandro, M.R. Furtado, J.G. Shewale, Developmental Validation of the PrepFiler™ Forensic DNA Extraction Kit for Extraction of Genomic DNA from Biological Samples*, J. Forensic Sci. 54 (3) (2009) 599–607.
[54] S. Kranes, S.A. Sterling, K. Mason, D. Anex, B. Hart, G. Parker, M. Prinz, Simultaneous DNA and protein extraction using trypsin, Forensic Sci. Int.: Genet. Suppl. Ser. 6 (2017) e203–e204.
[55] S. Kamanna, J. Henry, N.H. Voelcker, A. Linacre, K.P. Kirkbride, Direct identification of forensic body fluids using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, Int. J. Mass Spectrom. 397-398 (2016) 18–26.
[56] T.Md.S. Ventura, N.R. Ribeiro, A.S. Dionizio, I.T. Sabino, M.A.R. Buzalaf, Standardization of a protocol for shotgun proteomic analysis of saliva, J. Appl. Oral. Sci. 26 (2018) e20170561-e20170561.
[57] J.-H. Kang, H.-S. Kho, Blood Contam. Sali Diagn.: Curr. Methods their Limit. 57 (8) (2019) 1115–1124.
[58] H. Rangé, T. Léger, C. Huchon, C. Ciangura, D. Diallo, C. Poitou, O. Meilhac, P. Bouchard, C. Chaussain, Salivary proteome modifications associated with periodontitis in obese patients, J. Clin. Periodontol. 39 (9) (2012) 799–806.
[59] C.N. Otto, 7 - Hemoglobin metabolism, in: E.M. Keohane, C.N. Otto, J.M. Walenga (Eds.), Rodak's Hematology (Sixth Edition), Elsevier, St. Louis (MO), 2020, pp. 91–103.
[60] E.J. Dupree, M. Jayathirtha, H. Yorkey, M. Mihasan, B.A. Petre, C.C. Darie, A Critical Review of Bottom-Up Proteomics: The Good, the Bad, and the Future of this Field, Proteomes 8 (3) (2020) 14.
[61] L.M. Smith, N.L. Kelleher, Proteoforms as the next proteomics currency, Science 359 (6380) (2018) 1106–1107.
[62] L.M. Forgrave, M. Wang, D. Yang, M.L. DeMarco, Proteoforms and their expanding role in laboratory medicine, Pr. Lab Med 28 (2022) e00260.
[63] L.M. Smith, J.N. Agar, J. Chamot-Rooke, P.O. Danis, Y. Ge, J.A. Loo, L. Paša-Tolić, Y.O. Tsybin, N.L. Kelleher, P. The Consortium for Top-Down, The Human Proteoform Project: Defining the human proteome, Science Advances 7(46) eabk0734.
[64] L.M. Smith, N.L. Kelleher, M. Linial, D. Goodlett, P. Langridge-Smith, Y. Ah Goo, G. Safford, L. Bonilla, G. Kruppa, R. Zubarev, J. Rontree, J. Chamot-Rooke, J. Garavelli, A. Heck, J. Loo, D. Penque, M. Hornshaw, C. Hendrickson, L. Pasa-Tolic, C. Borchers, D. Chan, N. Young, J. Agar, C. Masselon, M. Gross, F. McLafferty, Y. Tsybin, Y. Ge, I. Sanders, J. Langridge, J. Whitelegge, A. Marshall, P, The Consortium for Top Down, Proteoform: a single term describing protein complexity, Nat. Methods 10 (3) (2013) 186–187.
[65] C. Wu, T. Shi, J.N. Brown, J. He, Y. Gao, T.L. Fillmore, A.K. Shukla, R.J. Moore, D. G. Camp, II, K.D. Rodland, W.-J. Qian, T. Liu, R.D. Smith, Expediting SRM Assay Development for Large-Scale Targeted Proteomics Experiments, J. Proteome Res. 13 (10) (2014) 4479–4487.
[66] S. Kundrapu, J. Noguez, Chapter Six - Laboratory Assessment of Anemia, in: G. S. Makowski (Ed.), Advances in Clinical Chemistry, Elsevier, 2018, pp. 197–225.
[67] A.S. Ogun, A. Adeyinka, Biochemistry, Transferrin, StatPearls, StatPearls Publishing Copyright © 2024, StatPearls Publishing LLC. Treasure Island (FL) ineligible companies. Disclosure: Adebayo Adeyinka declares no relevant financial relationships with ineligible companies, 2024.
[68] J.H. Kang, Y.H. Lee, H.S. Kho, Clinical factors affecting salivary transferrin level, a marker of blood contamination in salivary analysis, BMC Oral. Health 18 (1) (2018) 49.
[69] R. Mohamed, J.-L. Campbell, J. Cooper-White, G. Dimeski, C. Punyadeera, The impact of saliva collection and processing methods on CRP, IgE, and Myoglobin immunoassays, Clin. Transl. Med. 1 (1) (2012) 19.
[70] J.G. Bollinger, A.B. Stergachis, R.S. Johnson, J.D. Egertson, M.J. MacCoss, Selecting Optimal Peptides for Targeted Proteomic Experiments in Human Plasma Using In Vitro Synthesized Proteins as Analytical Standards, Methods Mol. Biol. 1410 (2016) 207–221.
[71] K. Demeure, E. Duriez, B. Domon, S.P. Niclou, PeptideManager: a peptide selection tool for targeted proteomic studies involving mixed samples from different species, Front. Genet. 5 (305) () (2014).
[72] SWGDAM, SWGDAM Validation Guidelines for Forensic DNA Analysis Methods, in: Do Justice (Ed.) Federal Bureau of Investigation, Department of Justice, Washington DC, 2016.
[73] E.S. Nakayasu, M. Gritsenko, P.D. Piehowski, Y. Gao, D.J. Orton, A.A. Schepmoes, T.L. Fillmore, B.I. Frohnert, M. Rewers, J.P. Krischer, C. Ansong, A.M. Suchy-Dicey, C. Evans-Molina, W.-J. Qian, B.-J.M. Webb-Robertson, T.O. Metz, Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation, Nat. Protoc. 16 (8) (2021) 3737–3760.
[74] P. Picotti, R. Aebersold, Selected reaction monitoring-based proteomics: Workflows, potential, pitfalls and future directions, Nat. Methods 9 (6) (2012) 555–566.
[75] Y. Perez-Riverol, J. Bai, C. Bandla, S. Hewapathirana, D. García-Seisdedos, S. Kamatchinathan, D. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, M. Walzer, S. Wang, A. Brazma, J.A. Vizcaíno, The PRIDE database resources in 2022: A Hub for mass spectrometry-based proteomics evidences, Nucleic Acids Res. 50 (D1) (2022) D543–D552. PubMed ID: 34723319.