The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	**Comparative Assessment of Emerging Technologies for Body Fluid Identification**
Author(s):	**Mirna Ghemrawi, Ph.D., MSFS, Jack Ballantyne, Ph.D., Bruce McCord, Ph.D., Glendon Parker, Ph.D.**
Document Number:	**308841**
Date Received:	**April 2024**
Award Number:	**2020-DQ-BX-0015**

This resource has not been published by the U.S. Department of Justice. This resource is being made publicly available through the Office of Justice Programs' National Criminal Justice Reference Service.

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.

Comparative Assessment of Emerging Technologies for Body Fluid Identification

FREDRIC RIEDERS FAMILY RENAISSANCE FOUNDATION

Mirna Ghemrawi, PhD, MSFS

Associate Director of Forensic Biology
Grant Award Administrator

Co-Principal Investigators:

Jack Ballantyne, PhD
University of Central Florida

Bruce McCord, PhD
Florida International University

Glendon Parker, PhD
University of California, Davis

Award Number 2020-DQ-BX-0015

Award Period: January 1st 2021 to June 30th 2023

Award amount: $501,115.00

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.

PROJECT SUMMARY

1. Goals and objectives

Biological fluid detection and identification provides important contextual information to a forensic investigation. While genetic testing can help establish from whom DNA may have come from, only serological testing can provide an indication of the body fluid or tissue from which a DNA profile may have originated. Given the limitations, current serological techniques lack sufficient sensitivity and specificity. As a result, several novel approaches to identifying biological fluids have been explored in recent years. These include the use of epigenetic modifications (DNA methylation), messenger ribonucleic acid (mRNA) markers, and proteomic identification of protein biomarkers. The goal of this study is to perform a comparative assessment of these emerging "omic" based technologies for body fluid identification (epigenome, transcriptome, proteome). The work will entail a thorough evaluation of error rates, sensitivity, and specificity of several approaches; namely epigenetics, mRNA profiling, and proteomics, as compared to conventional workflows using immunochromatographic assays.

2. Research questions

While the gap in data and research for each individual approach to serological identification is growing smaller, the gap in application is significant. Transition of any new technology to an operational laboratory is challenging and incorporation of any of these novel serological approaches will require a change in culture as much as development of forensically useful technology. The data generated from this study will make this paradigm shift easier by providing practitioners a point-in-time assessment of the advantages of these "omic" based approaches as compared to currently available immunoassay-based methodologies. Overall, we hope that the result of this study can provide forensic investigators with a better understanding of these flexible and informative tools and hope that this study will also aid individual labs interested in considering technology transfer.

3. Summary of project design and methods

The first phase of this project was to standardize sample preparation and analyze a control sample set to permit the participating laboratories to calibrate and ready their respective body fluid identification technologies for phase two. For this, a total of fourteen samples were prepared in triplicate (n=42) from five fluids in high and low volumes. Samples were prepared in five sets according to Table 1. Each set was sent to one lab: The McCord lab (for DNA methylation testing), The Ballantyne lab (for mRNA testing), The Parker lab (for proteomics testing), The DNA Labs International (for serological-immunochromatographic and STR testing) and NIST (for extraction and STR typing). For extraction, NIST employed the AllPrep kit (QIAGEN) following the manufacturer guidelines. STR typing was also conducted at NIST using the GlobalFiler kit, and average peak heights were reported.

Table 1. Summary on control sample set preparation in high and low volume.

Biological Fluid	Collection Method	High	Low
Semen	Fresh Collection	20µL neat	10µL of 1:100 dilution using ultrapure water (0.1µL equivalent)
Venous Blood	Purchased, collected in EDTA	20µL neat	10µL of 1:100 dilution using ultrapure water (0.1µL equivalent)
Saliva	Spit in Cup	20µL neat	10µL of 1:100 dilution using ultrapure water (0.1µL equivalent)
	Swab Slurry 5 oral swabs in 5mL ultrapure water solubilized	50µL	2µL
Vaginal Fluid	Soft Cup Collection	20µL neat	10µL of 1:1 dilution using ultrapure water (5µL equivalent)
	Swab Slurry 5 oral swabs in 5mL ultrapure water solubilized	50µL	2µL
Menstrual Blood	Swab Slurry 5 oral swabs in 5mL ultrapure water solubilized	50µL	2µL

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.

The second phase was to prepare fifty-eight blind mock casework samples (n=58) in four sets. These samples included peripheral blood, menstrual blood, semen, vaginal fluid, saliva. breast milk, urine, non-human blood, and nasal secretions deposited on various substrates including denim, leather, and cotton. A variety of volumes as low as 2.5 µL were applied, and some samples were subjected to degradation and inhibition. The CFSRE completed immunoassay testing of each sample by first extracting samples with universal buffer and allowing for an incubation period, following manufacturer guidelines. After extraction, samples were diluted and tested on the various immunoassays selected (Table 2). The CFSRE completed DNA methylation testing of the samples by first extracting DNA from the set of 58 blind fluid samples. The extracts were then quantified, normalized, and bisulfite converted to allow methylation quantification. Converted samples were then amplified with each body fluid amplification primer set (Table 2) and sequenced via pyrosequencing for analysis. If a sample displayed atypical DNA quantification or methylation results, the sample was then amplified with a species ID primer set [1], sequenced via pyrosequencing, and the species of the sample was determined. The Ballantyne lab completed mRNA testing by first extracting total RNA from the blind samples. Extracts were then reverse transcribed to cDNA. The newly transcribed DNA was then quantified and amplified with a body fluid multiplex. Finally, the amplified product was detected via capillary electrophoresis and analyzed. The Parker lab completed shotgun proteomics of the samples by first rehydrating and incubating the samples to promote protein solubilization. Supernatants were then transferred and filtered. Filtrates were incubated, prepared for digestion, and digested overnight. Peptides were then injected into a mass spectrometer, data was processed, and datasets were searched against the human proteome. The NYC OCME conducted targeted proteomics analysis using nano-HPLC reverse phase chromatography, followed by direct injection of the eluted peptides into a Q-TOF mass spectrometer operated in IDA (Information Dependent Acquisition) mode. Peptides were identified through automated database searching of raw data using ProteinPilot 5.0.2 against a mammalian protein database. The third phase was to evaluate the results, conduct a comprehensive discussion and assess the performance of each technology. Table 2 summarizes the markers targeted per 'omic' assay and Figure 1 illustrates the workflow of each 'omic' method.

Table 2. Summary of the markers targeted for each assay. Note: for the targeted proteomics assay employed by the OCME, check [3] (grey:(no available test or marker for that particular 'omic' method)

Fluid	Immunoassays	DNA Methylation	mRNA	Proteins
Blood	Human Glycophorin A (RSID blood), Human Hemoglobin (Seratec P)	MDFI	ANK1, ALAS2	sp|P01009|A1AT,sp|P68871|HBB,sp|P02790|HEMO,sp|P02647|APOA1,sp|P02787|TRFE
Semen	Human Semenogelin (RSID semen)	ZC3H12D, CG06379435	PRM2, SEMG1	sp|P15309|PPAP,sp|P07288|KLK3,sp|Q02383|SEMG2,sp|P61916|NPC2,sp|P04279|SEMG1
Saliva	Human Salivary alpha-amylase (RSID Saliva)	FAM43A	STATH, HTN3	sp|P04745|AMY1,sp|P02808|STAT,sp|P02814|SMR3B,sp|P09228|CYTT
Vaginal		VE8	CYP2B7P1	sp|Q9UBC9|SPRR3,sp|Q9UBG3|CRNN,sp|P80188|NGAL,sp|O95274|LYPD3,sp|Q6UWP8|SBSN,sp|O60437|PEPL,sp|P07476|INVO
Menstrual Blood	Human D-Dimer (Seratec PMB)	SLC26A10	MMP10, LEFTY2	
Nasal		SOX2OT		

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.

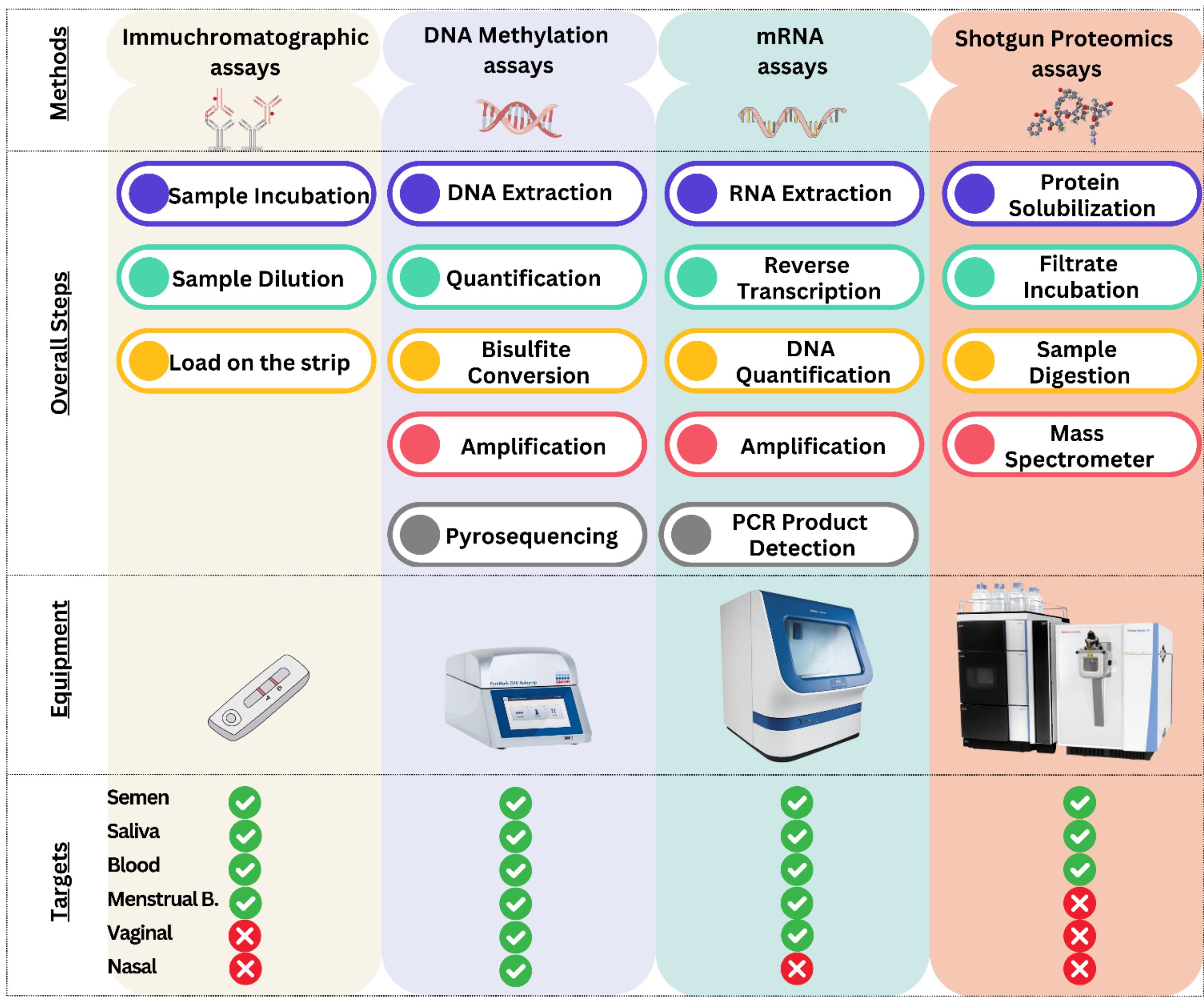

Figure 1. Infographics highlighting the workflow of each method, the main equipment(Pyromark Q48 (methylation), ABI 3500 (mRNA), TripleTOF 6600 (Proteomics), the targets Green check means that assays/markers were developed. Red x means that no marker is yet discovered – or validated.

4. Summary of results

Phase 1 results from each laboratory were compiled. The McCord lab using DNA methylation was able to detect 34/42 of the samples. Due to low levels or no DNA being present in the sample, the semen – low set, the saliva swab slurry – low set, and 2/3 of the saliva spit – low set were unsuccessful. The Ballantyne lab using mRNA testing was able to detect 19/28 samples they were able to test. One sample from each set was unable to produce results due to a bad reagent. Similar to DNA methylation results, all the samples that were unsuccessful in the identification originated from a low valume set, while all the high-volume samples were able to be detected. The Parker lab was able to detect protein biomarkers in 40/42 samples, with the two unsuccessful samples being a high volume of semen and a low volume of menstrual blood. DNA Labs International was able to obtain a full STR profile in 8/14 of the samples tested, while 4/14 samples detected a partial STR profile, and 2/14 samples had no profile (both are from the saliva slurry). DNA Labs International also tested the samples on various immunochromatographic tests. Expected results were seen for the semen tested with RSID Semen, and the blood samples tested with the Seratec PMB (P) test. RSID Blood was able to detect the high volume of blood but did not detect the low volume. RSID Saliva was able to detect both the high and low concentrations of the saliva slurry and the high concentration of the saliva spit, , however it did not detect the low concentration of the saliva spit. The Seratec PMB (M) test did not detect either the high or low concentration of menstrual blood. Finally, the vaginal fluid was not falsely detected on

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.

any of the immunochromatographic tests. NIST conducted STR testing on the 42 samples and reported an average peak height (Table3)

Table 3: Summary of results from Phase 1. Immunoassay results for venous blood include a (*) due to differing results from Seratec PMB (+) and RSID Blood (-) tests. For details about sample preparation refer to Table 1.

Sample	Body Fluid	Volume	STR		Immunoassays (DLI)	DNA Methylation (FIU)	mRNA (UCF)	Shotgun Proteomics (UC Davis)
			DLI(Profile)	NIST (Avg PH)	Succesful / No crossreactivity (Yes/No)	Successful (Yes/No)	Successful (Yes/No)	Successful (Yes/No)
1	Semen	High	Full	11032	Yes	Yes	N/A	No
2				10794		Yes	Yes	Yes
3				12423		Yes	Yes	Yes
4		Low	Partial	578	Yes	No	N/A	Yes
5				5764		No	Yes	Yes
6				1684		No	No	Yes
7	Venous Blood	High	Full	10975	Yes	Yes	N/A	Yes
8				15201		Yes	Yes	Yes
9				10469		Yes	Yes	Yes
10		Low	Partial	0	Yes*	Yes	N/A	Yes
11				0		Yes	Yes	Yes
12				0		Yes	Yes	Yes
13	Saliva Spit in Cup	High	Partial	10598	Yes	Yes	N/A	Yes
14				9900		Yes	Yes	Yes
15				8761		Yes	Yes	Yes
16		Low	Full	2190	No	No	N/A	Yes
17				2889		Yes	Yes	Yes
18				0		No	No	Yes
19	Saliva Swab Slurry	High	NR	6440	Yes	Yes	N/A	Yes
20				10249		Yes	Yes	Yes
21				8579		Yes	Yes	Yes
22		Low	NR	620	Yes	No	N/A	Yes
23				355		No	No	Yes
24				0		No	No	Yes
25	Vaginal Fluid Soft Cup	High	Full	12872	Yes	Yes	N/A	Yes
26				10602		Yes	Yes	Yes
27				8716		Yes	Yes	Yes
28		Low	Full	11162	Yes	Yes	N/A	Yes
29				9061		Yes	Yes	Yes
30				10837		Yes	No	Yes
31	Vaginal Fluid Swab Slurry	High	Full	13775	Yes	Yes	N/A	Yes
32				13759		Yes	Yes	Yes
33				15998		Yes	Yes	Yes
34		Low	Partial	478	Yes	Yes	N/A	Yes
35				0		Yes	No	Yes
36				0		Yes	No	Yes
37	Menstrual Blood	High	Full	6938	No	Yes	N/A	Yes
38				10151		Yes	No	Yes
39				11552		Yes	Yes	Yes
40		Low	Full	11340	No	Yes	N/A	Yes
41				7860		Yes	Yes	No
42				9804		Yes	No	Yes

In light of the findings, it was evident that sample preparation should closely emulate conditions encountered in a forensic exhibit. Therefore, in phase two, no dilution with water was implemented, and swab slurry was abstained from.

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.

Phase 2 results from each laboratory were gathered. Results were reported as True, Partial, False Positive, False Negative, Not Targeted, and Inconclusive (Table 4):

- **True Positive**: Correct identification of all body fluid(s) present in the sample.
- **Partial results**: Identification of at least one correct body fluid from a mixture (Other component of the mixture was not identified, but possible mixture was notated)
- **False Positive:** Identification of at least one body fluid that was not present in the sample.
- **False Negative:** Reported no identification of any body fluid that was present in the sample.
- **Not Targeted:** no body fluid is identified, and the specific body fluid present within the sample is not among the designated targets of the assay(s).
- **Inconclusive:** Differing results between RSID Blood and Seratec peripheral blood target / methylation percentages outside of ranges for all targeted body fluids.

Figure 2 is a bar plot illustrating the overall results of the four methods on the 58 blind samples. Table 4 shows the detailed results per sample.

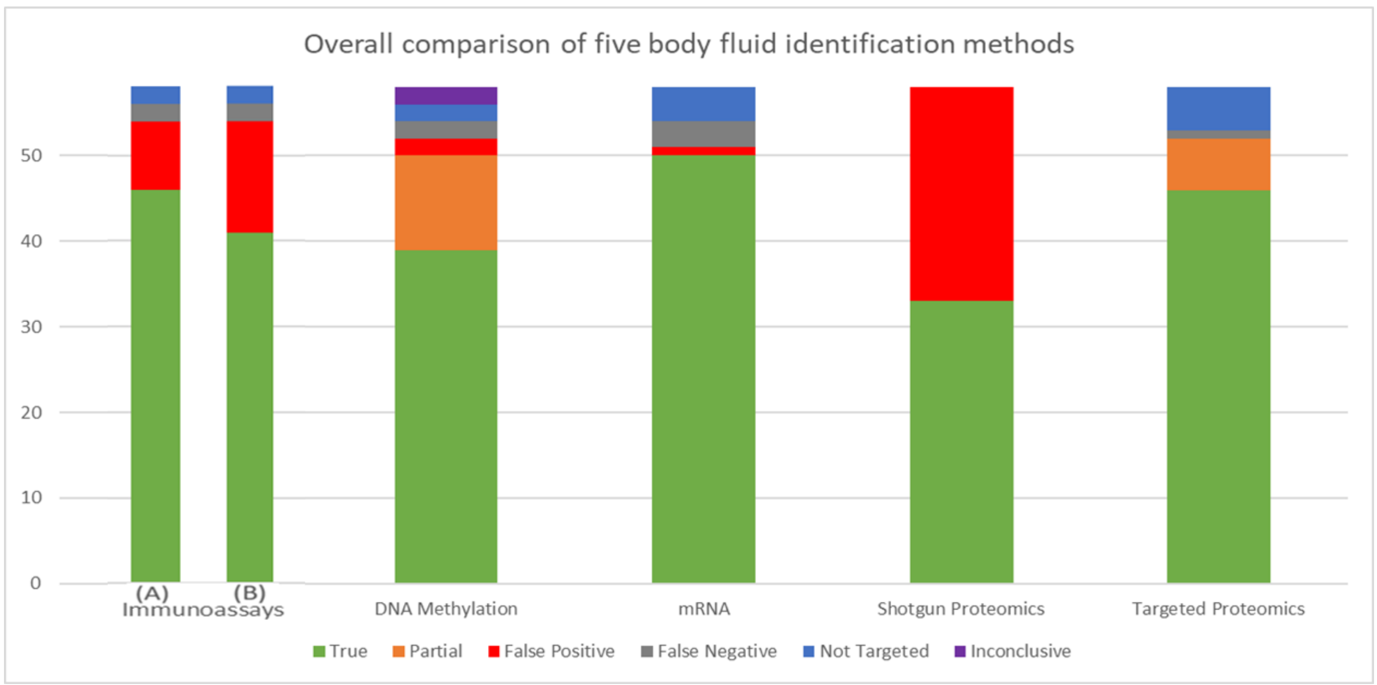

Figure 2: Comparison of five different body fluid identification assays – Immunoassays separated by the use of RSID Blood (A) or Seratec Peripheral Blood target (B) as the blood test. Below in Table 4 are the total number of samples per category per assay.

	Immunoassays (A)	Immunoassays (B)	DNA Methylation	mRNA	Shotgun Proteomics	Targeted Proteomics
True	46	41	39	50	33	46
Partial	0	0	11	0	0	6
False Positive	8	13	2	1	25	0
False Negative	2	2	2	3	0	1
Not Targeted	2	2	2	4	0	5
Inconclusive	0	0	2	0	0	0

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.

Table 5: Overview of 58 blind sample compositions and the corresponding results for each method (means one of the fluids present was not targeted)*

Sample #	Biological Fluid	Preparation	Immunoassays	DNA Methylation	mRNA	Shotgun Proteomics	Targeted Proteomics
1	Semen	25 µL on Swab	True	True	False Negative	False Positive	True
2	Saliva	25 µL on Swab	True	True	True	True	True
3	Blood	5 µL on Swab	True	True	True	False Positive	True
4	Vaginal Fluid	Vaginal Swab	Not Tested	True	True	False Positive	Not Tested
5	Degraded Blood	2.5 µL on Denim, Heat for 1 hour	True	False Negative	True	False Positive	True
6	Urine	50 µL on Bedsheet	Not Tested	Not Tested	Not Tested	False Positive	Not Tested
7	Degraded Blood	5 µL on Swab, Heat for 1 hour	True	True	True	False Positive	True
8	Vaginal Fluid/Semen	5 µL Semen on Vaginal Swab	Inconclusive	Partial	True	False Positive	True *
9	Blank	Blank Swab	True	True	True	True	True
10	Menstrual Blood/Semen	500 µL MB + 25 µL Semen, 50 µL per swab	True	Partial	True	True	Partial
11	Degraded Saliva	50 µL on Swab, Heat for 1 hour	True	True	True	False Positive	True
12	Blood	5 µL on Demin	True	True	True	False Positive	True
13	Degraded Saliva	25 µL on Swab, Heat for 2 hours	True	True	True	True	True
14	Degraded Blood	5 µL on Demin, Heat for 1 hour	False Negative	True	True	True	True
15	Semen	25 µL on Swab	True	True	True	False Positive	True
16	Saliva	5 µL on Cigarette butt	True	True	True	True	True
17	Degraded Blood	10 µL on Swab, Heat for 1 hour	False Positive	True	True	True	True
18	Vaginal Fluid/Saliva	50 µL Saliva on Vaginal Swab	Inconclusive	Partial	True	False Positive	False Negative *
19	Menstrual Blood/Saliva	500 µL MB + 25 µL Saliva, 50 µL per swab	True	False Positive	True	False Positive	Partial
20	Inhibited Blood	100 µL Blood + 100 µL Humic Acid, 25 µL per swab	True	True	False Negative	False Positive	True
21	Degraded Saliva	50 µL on Swab, Heat for 2 hours	True	True	True	True	True
22	Blood	5 µL on Swab	True	True	True	False Positive	True
23	Menstrual Blood/Semen	500 µL MB + 125 µL Semen, 50 µL per swab	False Positive	Partial	True	True	True
24	Breast milk	25 µL on Shirt	False Positive	Not Tested	Not Tested	False Positive	Not Tested
25	Menstrual Blood/Semen/Saliva	500 µL MB + 125 µL Semen + 125 µL Saliva, 25 µL per swab	True	Partial	True	True	Partial
26	Blood	2.5 µL on Demin	True	False Negative	True	True	True
27	Inhibited Blood	100 µL Blood + 200 µL Humic Acid, 25 µL per swab	True	True	False Negative	True	True
28	Saliva	25 µL on Swab	True	True	True	True	True
29	Blank	Blank Demin	True	True	True	True	True
30	Blood	2.5 µL on Leather	True	True	True	True	True
31	Degraded Saliva	5 µL on Swab, Heat for 2 hours	True	True	True	True	True
32	Degraded Saliva	25 µL on Swab, Heat for 1 hour	True	True	True	True	True
33	Vaginal Fluid/Semen	50 µL Semen on Vaginal Swab	False Positive	True	True	False Positive	True *
34	Degraded Blood	25 µL on Swab, Heat for 1 hour	False Positive	True	True	True	True
35	Blank	Blank Leather	True	True	True	True	True
36	Vaginal Fluid/Saliva	25 µL Saliva on Vaginal Swab	Inconclusive	Partial	True	False Positive	True *
37	Monkey Blood	10 µL on Swab	Inconclusive	True	Not Tested	True	Not Tested
38	Blood	5 µL on Swab	True	True	True	True	True
39	Inhibited Saliva	250 µL Saliva + 100 µL Humic Acid, 25 µL per swab	True	True	True	True	True
40	Nasal secretions	Nasal Swab	False Positive	Inconclusive	Not Tested	False Positive	Not Tested
41	Inhibited Blood	250 µL Blood + 100 µL Humic Acid, 25 µL per swab	True	True	True	False Positive	True
42	Menstrual Blood	5 µL on Swab	True	False Positive	True	True	True
43	Blood	10 µL on Demin	True	True	True	True	True
44	Blank	Blank Shirt	True	True	True	False Positive	True
45	Saliva	25 µL on Swab	True	True	True	True	True
46	Menstrual Blood/Saliva	500 µL MB + 125 µL Saliva, 50 µL per swab	True	Partial	True	True	Partial
47	Blood	5 µL on Leather	True	True	True	False Positive	True
48	Inhibited Saliva	100 µL Saliva + 200 µL Humic Acid, 25 µL per swab	True	True	True	True	True
49	Semen	25 µL on Swab	True	True	True	False Positive	True
50	Degraded Blood	10 µL on Demin, Heat for 1 hour	True	True	True	True	True
51	Menstrual Blood/Saliva	500 µL MB + 250 µL Saliva, 50 µL per swab	True	Partial	True	False Positive	Partial
52	Vaginal Fluid/Saliva	5 µL Saliva on Vaginal Swab	Inconclusive	Partial	True	True	True *
53	Menstrual Blood/Semen	500 µL MB + 250 µL Semen, 50 µL per swab	False Positive	Partial	True	True	Partial
54	Degraded Saliva	5 µL on Swab, Heat for 1 hour	False Negative	Inconclusive	True	True	True
55	Vaginal Fluid/Semen	25 µL Semen on Vaginal Swab	False Positive	Partial	True	False Positive	True *
56	Inhibited Saliva	500 µL Saliva + 100 µL Humic Acid, 25 µL per swab	True	True	True	True	True
57	Blank	Blank Bedsheet	True	True	False Positive	True	True
58	Blood	10 µL on Leather	True	True	True	False Positive	True

True	Partial	False Positive	False Negative	Not Tested	Inconclusive

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.

4.1. Phase 2 Immunoassay Results

The CFSRE tested the blind set of 58 samples using four different immunoassays (RSID Blood, RSID Semen, RSID Saliva, and Seratec PMB). Results from the set were divided based on which peripheral blood test was used to report blood results, as there were instances where the results from RSID Blood and the peripheral target of the Seratec PMB test were inconsistent.

Utilizing RSID Blood as the designated peripheral blood test, the immunoassays demonstrated their capacity to accurately identify 46 out of the 58 samples. Eight false positives were reported, six of which were erroneously identified as saliva (nasal secretion, breast milk, mix of semen and vaginal fluid, and mix of semen and menstrual blood), and two as menstrual blood (degraded peripheral blood). Two samples were designated as false negatives, a degraded blood and degraded saliva sample. Furthermore, two samples were categorized as not targeted due to the absence of a corresponding immunoassay designed to detect the specific fluid; these samples included urine and vaginal fluid specimens.

Conversely, when the Seratec peripheral blood target was employed, the number of accurately identified samples decreased to 41 out of the 58 samples, while the occurrence of false positives expanded to 13 samples. It is noteworthy that the introduction of vaginal fluid, when mixed with either semen or saliva, consistently triggered false positives in the Seratec PMB test, on the peripheral blood target line, an observation verified through repeated testing. An anomaly was observed in the case of monkey blood, as it returned a negative result when tested via RSID Blood but was found positive when assessed using the Seratec peripheral blood target. This phenomenon was reported in the Seratec PMB user instructions. Both assays are highlighted as human specific but their degree of cross-reactivity with non-human primates may differ.

4.2. Phase 2 DNA Methylation Results

Using results and methodology developed by the McCord lab, the CFSRE generated a decision tree for DNA methylation testing, using a positive body fluid control set, which allowed for thresholds to be set to determine body fluid presence. Using this conservative decision tree, the body fluid type was determined for each sample which fit within established thresholds. Samples with intermediate levels of methylation were deemed mixtures and those which clearly included at least one body fluid were identified. For the set of 58 samples, body fluid(s) were fully identified for 39 samples and for an additional 11 samples at least one component of a mixture was identified, indicating that 52 of the 58 samples were correctly identified. These samples included single source (on various substrates such as cotton swabs, cloth, cigarette butts, leather, and denim) and mixed body fluid samples (two and three body fluid mixtures), as well as degraded samples. Two samples in the set contained urine and breast milk and were classified as not targeted, since no markers were included for these fluids. No cross-reactivity was seen indicating that the specificity of the assay was strong.

Two samples were reported as inconclusive, a degraded saliva sample and a nasal fluid sample. The nasal fluid and degraded saliva samples did not meet any of the thresholds set for classification, and therefore were reported as inconclusive. This may be due to the previously mentioned conservative interpretation thresholds, and with more testing and refining of these thresholds, these samples may be able to be identified.

Two samples were reported as false negative, a degraded blood sample and a sample with a low volume of blood. These samples were reported as little to no DNA present as the DNA quantification value was below the threshold set for analysis. Bisulfite conversion requires a certain amount of starting DNA to ensure reliable conversion, and these samples did not meet this threshold.

Two of the samples resulted in a false positive, however these false positives were menstrual blood samples that were incorrectly identified as peripheral blood. Menstrual blood's methylation patterns can vary depending on when

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.

the menstrual blood was collected [2], so these results were not surprising. One sample had unusual quantification results, with a large target detected but the small target undetermined, and the pyrograms showed low peak height. These findings prompted a subsequent in-depth examination using a pyrosequencing-based species identification assay, which conclusively established the origin of the contributor as a member of the primate group.

4.3. Phase 2 mRNA Results

The Ballantyne lab using mRNA testing obtained correct results for 54 of the 58 samples in the blind set using the mRNA CE body fluid identification multiplex. These samples included single source (on various substrates such as cotton swabs, cloth, cigarette butts, leather, and denim) and admixed body fluid samples (two and three body fluid mixtures), as well as degraded samples. This included 4 samples containing body fluids not specifically tested for using the mRNA 10-plex system (urine, breast milk, monkey blood and nasal secretions). No cross-reactivity with the mRNA 10-plex system was observed indicating the high degree of specificity of the assay. The blood mRNA targets appear to be human specific as exemplified by the negative mRNA result for the monkey blood sample. Interestingly, for the nasal secretion sample, an inconclusive result was reported but did note that STATH was detected in the absence of HTN3, which is a characteristic finding with nasal secretion samples. A correct presumptive inference of the possible presence of nasal secretions was made, although a confirmatory identification could not be made since the assay was not designed specifically to detect nasal secretions. However, future work with this assay will include a further evaluation of nasal secretion samples to determine if suitable interpretation guidelines can be developed for identification of nasal secretions.

The remaining 4 samples returned three false negative results (rate of 5.2%) in which the assay failed to detect a semen sample and two inhibited blood samples, and one false positive (rate of 1.7%) in which saliva was detected in a sample designated as an extraction blank. Each of these samples will be discussed below:

- Semen sample. This sample was 25 µl semen deposited onto a cotton swab. It is unclear why the assay was not successful in detecting this sample. Interestingly, we were able to detect possible trace levels of semen, but these fell below the calling threshold of the assay.
- The two blood samples that were not detected by the mRNA 10-plex system were both mixed with humic acid and therefore contained known PCR inhibitors. These inhibitors likely impacted the reverse transcription reaction which is why the assay failed to detect the presence of blood in these samples. Future work will include an evaluation of whether RT-PCR/PCR additives such as BSA might improve the recovery of mRNA targets from otherwise inhibited samples.

The one false positive result was the identification of saliva in a negative control sample that should have contained no body fluids. Upon investigation we believe this false positive was nevertheless genuine detection of the presence of saliva in the sample, which likely originated from a contamination event from the immediately adjacent high level saliva sample either during sample preparation or sample extraction, which are both done manually. Thus, while unfortunate, we are not concerned that target cross-reactivity or non-specificity of the assay was to blame.

4.4. Phase 2 Shotgun Proteomics Results

The Parker Lab using their Shotgun Proteomics protocol obtained correct results for 33/58 of the unknown samples, and 25/58 false positive. These samples comprised single-source fluids (on substrates such as cotton swabs, cloth, cigarette butts, leather, and denim), mixed body fluid samples (including two and three fluid mixtures), as well as degraded and inhibited samples. Of the 25 false positives, 6 contained signals for blood proteins where no blood was added during sample preparation, 4 contained signals for salivary proteins where no saliva was added during sample preparation, 10 contained signals for semen proteins where no semen had been added during sample preparation, and 5 contained signals for two or more fluid types not added during sample preparation.

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.

False positive results were due to two factors: 1) More concentrated samples were used in phase 2, resulting in the detection of low levels of blood proteins in non-blood containing samples 2) Carryover (~1%) occurred in the LC/MS instrument resulting in the presence of low levels of semenogelin and amylase proteins. Other factors included problems detecting mixtures containing saliva due to matrix effects. These issues can be corrected in future assays by correcting for carryover and sample concentration.

Clearly the dynamic range of proteomics, of over 6 orders of magnitude, presents a challenge to the categorization of body fluids in forensic science. Over 10,000 genes are expressed in any one cell, and can be regularly detected, resulting in loss of specificity of biomarkers. More research will be required to determine better thresholds and ratios for the reporting of body fluids by shotgun proteomic mass spectrometry. The data from this project suggests that even degradation or inhibition does not eliminate the signal from endogenous blood peptides in non-blood samples, a potential advantage of the method. It should be noted that no markers distinguishing menstrual blood were included in the current study, and since all proteins from vaginal fluid appeared to be present in saliva at low levels, detection of vaginal fluid could only be made when salivary peptides were absent or accounted for.

4.5. ***Phase 2 Targeted Proteomics Results***

The NYC OCME's laboratory utilized a validated proteomic mass spectrometry assay on a SCIEX high-performance liquid chromatography (HPLC) instrument to identify blood, saliva, and semen. The analysis included 26 peptides and 104 transitions. Menstrual and peripheral blood were categorized together as blood since the assay does not differentiate between these sources. Out of 58 samples, body fluids were fully identified in 46 samples, and at least one component of a mixture was identified in an additional 6 samples, resulting in a total of 52 correctly identified samples. These samples comprised single-source fluids (on substrates such as cotton swabs, cloth, cigarette butts, leather, and denim), mixed body fluid samples (including two and three fluid mixtures), as well as degraded and inhibited samples.

Among the remaining 6 samples, 5 contained fluids such as breast milk, urine, vaginal fluid, nasal secretions, and primate blood that are not targeted by the OCME validated method. The other sample produced false negative, failing to detect saliva mixed with vaginal secretions. Importantly, no false positives were observed. These six samples contained quantifiable proteins, but no semen, blood, or saliva was detected. A presumptive test was conducted using the SCIEX 6600 triple TOF instrument, accurately identifying all of these 6 samples, including breast milk, nasal secretion, urine, two vaginal secretions, and primate blood.

4.6. Discussion

To assess the performance of each assay, a comprehensive analysis encompassing specificity, sensitivity, and error rates was undertaken. A micro-average approach was followed to evaluate the overall performance of each method. Specificity was assessed by computing the ratio of true negative results to the summation of true negatives and false positives. Sensitivity was determined by the ratio of true positive identifications to the combined total of true positives and false negatives. Error rates were established by computing the ratio of false positives to the total sum of false positives and true positives. In Table 6, the data reveals that DNA methylation, mRNA and targeted proteomics assays produced the highest levels of specificity, achieving values of 99.5%,99.7%, and 100% respectively. Within the immunoassay category, results were segregated between RSID (pertaining to blood identification based on RSID-blood assay) and Seratec (associated with blood identification based on the Peripheral blood target line in the PMB test). On average, the specificity of immunoassays was 96%, while the shotgun proteomics testing, on average, yielded a specificity rate of 93.4%. Regarding sensitivity metrics, the mRNA assay exhibited the highest degree of sensitivity at 94.1%, closely followed by targeted proteomics assays, which achieved a notable sensitivity level of 88.5.1%, followed by the immunoassay with 87.1% DNA methylation assays displayed a sensitivity rate of 72.5%, while shotgun proteomics assays registered a sensitivity level of 67.7%. Error rates were found to be relatively low for the mRNA assays (1.5%) and DNA methylation assays (3.8%), with 0% error rate for the targeted proteomics assay. However, it is noteworthy that the error rate was high (15.9% on average) for the immunoassays, and higher (32.3%) for the shotgun proteomics method.

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.

Table 6. Summary of the overall performance of the conventional serological technique (immunoassays) in comparison to emerging technologies. Three metrics were computed: sensitivity, specificity and % error rates.

	Specificity	Sensitivity	% error
Assay / Formula	TN/(TN+FP)	TP/ (TP+FN)	FP/(FP+TP)
Immunoassays (RSID)	97.1	89.2	12.1
Immunoassays (Seratec)	95.4	85.1	19.7
DNA Methylation	99.5	72.5	3.8
mRNA	99.7	94.1	1.5
Shotgun Proteomics	93.4	67.7	32.3
Targeted Proteomics	100.0	88.5	0.0

The obtained findings shed light on the limitations of contemporary serologic techniques, with a particular focus on immunoassays, revealing a notable prevalence of elevated error rates and the conspicuous absence of assays targeting the critical examination of vaginal fluid—a pivotal fluid in forensic investigations. Emerging technologies, particularly targeted proteomics-based, mRNA-based and DNA methylation-based assays, were found to be highly specific. Overall, mRNA had the best performance on all three metrics. The main challenge were samples treated with humic acid to mimic inhibition. Future research will investigate the potential benefit of using additives such as BSA in RT-PCR/PCR to improve the recovery of mRNA targets in samples prone to inhibition.

DNA methylation, while specific and producing a low error rate, had lower sensitivity based on the current conservative decision tree. Mitigating this limitation may necessitate the inclusion of additional molecular markers per fluid type to facilitate the effective interpretation of complex mixtures. Furthermore, menstrual blood analysis remains a formidable task, especially during low bleeding days, a concern acknowledged in the literature [2]. In the case of non-human DNA, specifically the sample attributed to monkey blood, the DNA quantification outcomes, coupled with the pyrogram findings, prompted a subsequent in-depth examination involving a pyrosequencing-based species identification assay [1]. This analysis conclusively established the origin of the contributor as a member of the primate group.

In the case of shotgun proteomics, the procedure produced a lower specificity than other "omic" methods, mainly due to the detection of low levels of protein markers previously believed to be indicative for other body fluids. This result is a consequence of the wide dynamic range of the mass spectrometric method and indicates that additional work is needed to establish conservative threshold values, particularly for mixtures and other samples such as menstrual blood which can also contain blood markers. It should be noted that the shotgun proteomics methodology implemented in this investigation is distinct from the targeted mass spectrometry procedure validated and used by the Office of the Chief Medical Examiner (OCME) in New York City. The OCME laboratory, using their validated method for identifying semen, blood, and saliva [3], achieved a notable outcome with no false positives. However, the primary challenge was detecting saliva in mixtures containing two or three body fluids. Additionally, the assay has limitations, such as its inability to identify vaginal fluid or distinguish between peripheral and menstrual blood, both of which are commonly encountered body fluids.

5. Applicability to criminal justice

Forensic practitioners have repeatedly stressed that the identification of biological stains can be a significant challenge for case-working serologists. They often find it difficult to detect and identify low levels of biological stains, particularly those containing vaginal/menstrual fluids. However, such low-level samples routinely produce STR profiles. Furthermore, for certain types of biological fluids, the lack of confirmatory test methods can prevent a positive determination of the serological origin of the DNA profile. It is therefore important that newer, more sensitive and specific technology be carefully examined, especially given the investment in training and

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.

instrumentation that would be required for implementation. Multiple research studies have already demonstrated that the emerging serological techniques described in this proposal exhibit improved detection capabilities and can detect a wider range of potential body fluid types.

The data from this study represents a snapshot in time for the development of these "omic" methods. Extensive research and validation remain for these procedures; however, the technology has been widely heralded and applied in medical diagnostics and its implementation in forensics is long overdue. As a community, we have reached the stage where operational laboratories require distinct and detailed guidelines for these technologies to be effectively transferred to the casework laboratory. The comparative assessment of the strategies discussed in this study provides valuable information to the forensic community which can aid in the development of new research and facilitate technology transfer.

Acknowledgements:

We would like to acknowledge Katherine Gettings from the National Institute of Standards and Technology (NIST) who assisted in analysis the first set of samples and Dr. Donald Siegel and his team at the Office of the Chief Medical Examiner (OCME) in New York City for performing the targeted proteomics analysis.

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.

PRODUCTS:

A manuscript is being prepared.

REFERENCES:

[1] Ghemrawi, M., Fischinger, F., Duncan, G., Dukes, M. J., Guilliano, M., & McCord, B. (2021). *Developmental validation of SpeID: A pyrosequencing-based assay for species identification*. Forensic Science International: Genetics, 55, 102560.

[2] Lee, H. Y., Jung, S. E., Lee, E. H., Yang, W. I., & Shin, K. J. (2016). DNA methylation profiling for a confirmatory test for blood, saliva, semen, vaginal fluid and menstrual blood. *Forensic Science International: Genetics*, *24*, 75-82.

[3] OCME, https://www.nyc.gov/assets/ocme/downloads/pdf/molecular_serology_validation.pdf

FIU published two papers: a review paper on Pyrosequencing and a research article on skin/sweat markers in the Electrophoresis Journal under the "Young and Inspiring Scientists Special Issue" where the NIJ was acknowledged with the award # **2020-DQ-BX-0015.**

i. Ghemrawi, M., Tejero, N. F., Duncan, G., & McCord, B. (2023). Pyrosequencing: Current forensic methodology and future applications—A review. Electrophoresis, 44(1-2), 298-312.

ii. Fernandez-Tejero, N., Gauthier, Q., Cho, S., & McCord, B. R. (2023). High-resolution melt analysis for the detection of skin/sweat via DNA methylation. Electrophoresis, 44(3-4), 371-377.

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.