Forensic Science International: Genetics 79 (2025) 103303

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

mRNA profiling and donor association of mock casework samples: Results of a 3rd and 4th EDNAP collaborative exercise

Nadescha Viviane Hänggi [a], António Amorim [b,c], Heloísa Afonso Costa [b,d], Jeppe Dyrberg Andersen [e], Marie-Louise Kampmann [e], Cornelius Courts [f], Maximilian Neis [f], Denise Syndercombe-Court [g], Federica Giangasparo [g], Ane Elida Fonneløp [h], Helen Johannessen [h], Thorsten Hadrys [i], Angelika Fürst [i], Walther Parson [j,k], Harald Niederstätter [j], Maja Sidstedt [l], Siri Aili Fagerholm [l], Titia Sijen [m], Margreet van den Berge [m], Erin Hanson [n,o], Jack Ballantyne [n,o], Cordula Haas [a,*]

[a] Zurich Institute of Forensic Medicine, University of Zurich, Zurich, Switzerland
[b] Instituto Nacional de Medicina Legal e Ciências Forenses, Serviço de Genética e Biologia Forenses, Portugal
[c] Faculdade de Ciências da Universidade de Lisboa, Portugal
[d] Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Portugal
[e] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
[f] Institute of Legal Medicine, University Hospital Cologne, Cologne, Germany
[g] Department of Pharmacy and Forensic Science, King's College London, London, UK
[h] Department of Forensic Sciences, Oslo University Hospital, Oslo, Norway
[i] Bavarian State Criminal Police Office, Forensic Science Institute, Munich, Germany
[j] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[k] Forensic Science Program, The Pennsylvania State University, University Park, PA, USA
[l] National Forensic Centre, Swedish Police Authority, Linköping, Sweden
[m] Division Biological Traces, Netherlands Forensic Institute, The Hague, the Netherlands
[n] National Center for Forensic Science, Orlando, FL, USA
[o] Department of Chemistry, University of Central Florida, Orlando, FL, USA

ARTICLE INFO

Keywords:
Forensic science
Body fluid identification
mRNA profiling
Coding region SNPs (cSNPs)
Massively parallel sequencing (MPS)

ABSTRACT

Simultaneous identification and association of body fluids to donors can serve as a powerful tool in the criminal investigation of mixed traces. Massively parallel sequencing of mRNA targets not only identifies the origin of the body fluids but may also provide additional contextual information about the body fluid donors of a (binary) mixture using coding region SNPs (cSNPs). Within the European DNA Profiling Group (EDNAP), two consecutive collaborative exercises (3rd and 4th EDNAP exercise) were organized, with the objective to evaluate the performance of two previously published high-resolution mRNA sequencing assays. In the 3rd EDNAP exercise, the BFID-cSNP-BSS assay (cSNPs for blood, saliva, and semen) was evaluated, while in the 4th EDNAP exercise, the BFID-cSNP-6F assay (cSNPs for six fluids/tissues, including blood, saliva, semen, vaginal secretion, menstrual blood, and skin) was tested. Each RNA cSNP assay was accompanied by a genomic DNA assay for the genotyping of the cSNPs in the individual(s) or body fluid donor(s) of interest. A total of 11 laboratories participated in one or both collaborative exercises. In each exercise, the participating laboratories received a set of 16 standardized mock case stains for analysis and were encouraged to analyze additional, self-prepared stains and reference samples. Laboratories could participate using either the Ion Torrent S5™ or the Illumina MiSeq™ sequencing system. The results of the 16 mock case stains were very encouraging in both exercises, as body fluid components could be reliably identified for most of the stains. Since successful donor association depends on the number of body fluid markers covered in the sequencing results, we found that for stains consisting of blood, menstrual blood, vaginal secretion or a mixture thereof, the cSNPs provided substantial genetic discriminatory information for successful association of the respective body fluid to its donor. In mixtures, the difficulty in interpreting the

* Correspondence to: University of Zurich, Zurich Institute of Forensic Medicine, Winterthurerstrasse 190/Y52, CH-8057 Zurich, Switzerland.
E-mail address: cordula.haas@irm.uzh.ch (C. Haas).

https://doi.org/10.1016/j.fsigen.2025.103303
Received 13 December 2024; Received in revised form 17 April 2025; Accepted 12 May 2025
Available online 23 May 2025

cSNP genotypes might be attributed to the masking effect of the other body fluid(s) present. Body fluid identification and donor association of skin samples proved to be a significant challenge. In conclusion, body fluid identification and donor association using the BFID-cSNP-BSS and -6 F assays is a promising and effective method across laboratories and sequencing platforms.

1. Introduction

Valuable information about the circumstances of a crime can be obtained from biological evidence, in part by identifying the type of body fluid or tissue present. Body fluid identification (BFID) is a fundamental part of forensic genetics [1], as the origin of a tissue can be a crucial factor in a case. Currently, the first analytical step involves screening to ascertain the presence of body fluids in biological evidence. However, stain consumption is a limitation of current presumptive BFID tests, their specificity is limited and separate analyses are required for each body fluid tested [2]. BFID techniques developed in recent years include RNA profiling [1,3–12], tissue-specific methylation analysis [13, 14], microbial forensics [15–17], and proteomics [18,19].

RNA analysis [1,3–12] has emerged as a promising method for BFID, using the tissue-specific composition of the transcriptome of body fluids. BFID becomes feasible through the detection of specific mRNA transcripts and a number of useful mRNA markers have been identified for the forensically most relevant body fluids/tissues, such as blood, menstrual blood, nasal and vaginal secretion, saliva, semen, skin, and rectal tissue [20–30]. The European DNA Profiling Group (EDNAP) has evaluated the performance, robustness, and reproducibility of different mRNA markers for the identification of blood, saliva, semen, vaginal secretion, menstrual blood, and skin [31–35]. mRNA profiling has been implemented and accredited according to ISO17025 in several laboratories, where it is regularly used in casework and accepted in court. A classical mRNA profiling workflow starts with RNA extraction, followed by DNase treatment, reverse transcription, marker-specific endpoint PCR, and separation of the amplicons by capillary electrophoresis, or detection by qPCR [5]. Moreover, the classical workflow can be improved by performing an RNA/DNA co-extraction, since STR profiling and tissue/body fluid identification can be performed on the same sample, minimizing sample consumption and facilitating the integration of RNA- and DNA-derived information for sub-source and source-level evaluations [34,35].

The advent and development of massively parallel sequencing (MPS) had a significant impact on the forensic field through the introduction of DNA and RNA sequencing into stain analysis. Targeted RNA sequencing methods became available that allowed to simultaneously analyze multiple mRNA markers with higher analytical sensitivity compared to whole transcriptome sequencing (RNA-Seq) [10,30,36]. In 2018, a BFID set of markers suitable for MPS was introduced, including 33 tissue-specific mRNA markers, designed for the Illumina MiSeq™ platform [10]. A collaborative exercise was then organized within the EUROFORGEN-NoE and EDNAP groups to evaluate the MPS assay, which provided robust and reliable results with no or low signals in non-target body fluids [30]. Partial Least Squares (PLS) analysis supported the specificity of the assay by showing well-clustered body fluid markers and samples. A second exercise was organized within the Forensic RNA Profiling (FoRNAP) group [37], which allowed the participants to analyze 16 provided stains using their respective in-house analysis methods (PCR/CE or MPS) and markers, demonstrating that mRNA profiling including different extraction, analysis methods, and markers performs well on samples resembling forensic casework.

The availability of sequence information within the mRNA targets not only allows the identification of body fluids, but can also provide information on who contributed which body fluid component to a (binary) mixture, thereby adding important contextual insights. Coding region single nucleotide polymorphisms (cSNPs) in body fluid-specific transcripts enable the association of an individual's DNA cSNP profile with the body fluid-specific RNA cSNP genotype. In a proof-of-concept study, a targeted MPS assay for the Illumina MiSeq™ platform was developed, including 35 cSNPs in body fluid-specific mRNA transcripts, to assign body fluids to their respective donors in binary mixtures of same-type and different types of body fluids [12]. However, some problems emerged, i.e. some of the markers were not cDNA-specific in the presence of gDNA contamination, and low-expressed cSNP markers were prone to dropouts. Therefore, further development was required. Two improved, dual-function (BFID and donor association), high-resolution mRNA sequencing assays were developed, optimized for the Ion Torrent S5™ platform [38]. Each RNA assay includes markers for the identification of six body fluids/tissues, namely blood, semen, saliva, vaginal secretion, menstrual blood and skin. In the original blood, semen and saliva (BSS) assay, some of these BFID markers contain cSNPs (21 in total) for the donor association of blood, semen and saliva. The expanded 6 F (all six fluids/tissues) assay includes the BSS assay and an additional 23 cSNPs located in the BFID markers for vaginal secretion, menstrual blood and skin. Each RNA assay is accompanied by a genomic DNA assay for the DNA genotyping of the cSNPs in the person(s)/ body fluid donor(s) of interest.

For this study, the Zurich Institute of Forensic Medicine (ZIFM) in Switzerland organized two consecutive collaborative exercises, with the original (BSS) and the extended (6 F) BFID/cSNP assays. The 3rd Exercise explored the BSS assay using the Ion Torrent S5™. The 4th Exercise focused on the 6 F assay and included both the Ion Torrent S5™ and the Illumina MiSeq™ sequencing platforms, allowing more laboratories to participate. For each collaborative exercise, the participating laboratories received a set of 16 standardized mock case stains, the respective targeted RNA assay (BSS or 6 F) and its corresponding DNA assay, as well as detailed instructions on how to analyze the stains. Additionally, participating laboratories were encouraged to prepare some additional *bona fide* or mock case stains (referred to as self-prepared stains) for BFID/cSNP analysis. Participation in the exercises provided laboratories with an opportunity to become familiar and gain experience with the MPS-based method. Here, we report the compiled results of the 3rd and 4th EDNAP collaborative exercises and the performances of the BSS and the 6 F assays. Furthermore, the BFID-cSNP-6F assay was tested on two sequencing platforms. In addition, the organizing laboratory compared the BFID-cSNP-6F performance with another recently published alternative cSNP assay [39], by re-analyzing the cDNA of the 16 standardized stains of the 4th EDNAP collaborative exercise.

2. Materials and methods

2.1. Participants and samples

In the 3rd collaborative exercise, six laboratories used the Ion Torrent S5™ (S5) platform, while the 4th collaborative exercise had 11 participating laboratories using either the S5 or the Illumina MiSeq™ (MiSeq) sequencing platforms. Two laboratories performed the 4th exercise on both sequencing platforms. The results for each sequencing platform of the two laboratories were given their own label; thus, we present the results of thirteen "laboratories" in the 4th exercise. For each exercise, 16 mock case stains were prepared by the organizing laboratory (ZIFM). For the 3rd exercise, the stains consisted of nine single-donor body fluid/tissue samples and seven mixtures (Table 1A). The 4th exercise included six single-donor body fluid/tissue stains, and ten mixed body fluid samples were provided (Table 1B). Throughout this study, the following abbreviations will be used: BL for blood, MB for

menstrual blood, SA for saliva, SE for semen, VAG for vaginal secretion, and SK for skin. Moreover, blood markers will be colored in red, semen markers in yellow, saliva markers in blue, vaginal secretion in green, menstrual blood markers in pink, and skin markers in apricot.

The study protocols were reviewed and approved by the CEBES Review Board at the University of Zurich (case number 2021–11e). Finger-prick blood was used for blood samples. Fresh ejaculate was collected in 50 ml tubes and frozen in aliquots within a few hours of collection, and thereafter thawed only once for stain preparation. Menstrual blood and vaginal secretion were collected on cotton swabs (Deltalab, Barcelona, Spain) from female volunteers. For menstrual blood and vaginal secretion stains on other substrates (pad, underwear), the carrier material was in contact with the genital area for several hours (pad) or up to 11 hours (underwear). Saliva was collected in tubes before being applied directly to the respective carrier material. For skin samples, 50 µL of nuclease-free H_2O was applied to a swab before the whole inside of the underarm was swabbed. Two swabs were taken per arm per day (one swab in the morning, one in the evening).

The stains were shipped at ambient temperature with the corresponding DNA and RNA assays for each exercise, i.e. the BFID-cSNP-BSS RNA and respective DNA assay for the 3rd exercise (Table 2), and the BFID-cSNP-6F RNA and respective DNA assay for the 4th exercise (Table 2). All packages arrived within one week. Participants were instructed to store the stains in the dark at room temperature and the primer mixes at −20°C until processing. Participants analyzed the stains without knowledge of their composition (except for the organizing laboratory, which also participated in the exercises).

Participating laboratories were also encouraged to prepare and analyze additional/own mock case stains. The composition of the self-prepared stains was only disclosed to the organizing laboratory after the analysis and interpretation of the results were completed. The laboratories collected reference DNA samples from the body fluid donors for their self-prepared stains.

2.2. Analysis of the standardized mock case stains

Prior to shipment, the organizing laboratory tested each set of 16 stains using an in-house RNA 12plex CE assay [67] in order to check the quality of the samples. The 12plex PCR products were analyzed on a Genetic Analyzer 3130xl for the 3rd exercise and on a SeqStudio Genetic Analyzer (both Thermo Fisher Scientific) for the 4th exercise, and the results were evaluated using GeneMapper ID-X Software v1.4 (Thermo Fisher Scientific).

For both collaborative exercises, the organizing laboratory provided recommendations and instructions for each of the steps required to analyze the stains (Supplementary Material, Table S1). For some of the steps, the instructions consisted of recommendations, while for other steps strict adherence to the protocol was required. The process for analyzing the stains included extracting RNA (and DNA) from the stains, treatment with DNase, conversion of the RNA into cDNA by reverse transcription, followed by library preparation and sequencing. The co-extracted DNA was used to assess the quality and quantity and to determine the number of contributors to the stain.

2.2.1. RNA Analysis of the standardized mock case stains

In the 3rd exercise, four out of the six participating laboratories used the recommended mirVana™ miRNA Isolation kit, while the remaining two laboratories used other RNA extraction methods (Supplementary Material, Table S2, column B). All laboratories performed DNase treatment using the TURBO DNA-*free*™ kit and quantified the RNA extracts. It was mandatory to use the SuperScript™ IV VILO Master Mix for reverse transcription. An RNA input of 50 ng was recommended for a manual workflow, whereas 25 ng were recommended for an automated workflow. Two laboratories used an automated library preparation workflow with the Ion Chef™, while three laboratories prepared the libraries manually. Target amplification was performed using either 23 (four laboratories) or 24 (two laboratories) PCR cycles. All libraries were quantified using the Ion Library TaqMan™ Quantification Kit and sequenced on the S5 platform. The pooled libraries were diluted to a concentration between 25 pM and 50 pM. The HID Ion S5™ Chef Templating 02 protocol was used with 500 flows.

In the 4th exercise, five out of eight S5 laboratories and four out of five MiSeq laboratories used the recommended mirVana™ miRNA isolation Kit, while the remaining four laboratories used other RNA extraction methods (Supplementary Material, Table S2, columns C + D). All laboratories performed a DNase treatment using the TURBO DNA-*free*™ kit. Seven of the eight S5 laboratories and four of the five MiSeq laboratories quantified their RNA extracts using fluorometric methods. All laboratories used the SuperScript™ IV VILO Master Mix for reverse transcription. An RNA input of 50 ng was recommended for the manual S5 and for the MiSeq workflow, whereas 25 ng were recommended for an automated S5 workflow. Three S5 laboratories used an automated library preparation protocol, while the remaining five S5 laboratories prepared the libraries manually. Target amplification was performed

Table 1
Details on the 16 standardized mock case stains for the 3rd Exercise (A) and 4th Exercise (B). The color code indicates high-input stains (dark blue), low-input stains (light blue), and mixtures (orange). BF/T stands for body fluid/tissue. Stains were classified as low-input if they contained less than 20 µL of a single body fluid, skin swabs, and carrier materials that had brief contact (<1 min) with a bodily opening. High-input stains were swabs containing 50 µL of a single body fluid, and carrier materials that were in prolonged contact (several hours) with a bodily opening.

A)

Stain N°	BF/T	Amount	Stain Provided
1	SE	10 µl	piece of fabric (boxer shorts)
2	BL-MB	1/2 Swab + 25 µl	1/2 swab
3	SE	50 µl	artificial cotton
4	SA-SE	50 µl + 25 µl	part of a T-shirt
5	BL	50 µl	1 swab
6	SK	1 swab	1 swab
7	BL-BL	25 µl + 25 µl	part of a T-shirt
8	SA	Licked plastic spoon	spoon
9	SA-SA	25 µl + 25 µl	1 swab
10	BL-SA	25 µl +25 µl	1 swab
11	SA	50 µl	part of a T-shirt
12	VAG	1/2 swab	1/2 swab
13	BL	Nose bleed on tissue	part of a tissue
14	SA-SE	25 µl + 25 µl	piece of fabric (boxer shorts)
15	MB	1/2 swab	1/2 swab
16	SE-VAG	½ Swab+ 25 µl SE	1/2 swab

B)

Stain N°	BF/T	Amount	Stain Provided
1	SK	1 swab	1 swab
2	MB-BL	1 swab + 25 µl	1/4 swab
3	VAG-SA	1 swab + 25 µl	1/4 swab
4	MB-SA	1 swab + 25 µl	1/4 swab
5	BL-SE	25 µl + 25 µl	part of T-Shirt
6	SE-SE	25 µl + 25 µl	1 swab
7	MB-SA	1 swab + 50 µl	1/4 swab
8	SK-SA	1 swab + 25 µl	1 swab
9	VAG	cotton part of a slip	a piece of slip
10	MB	menstrual pad	a piece of pad
11	SE	50 µl	part of a glove (latex)
12	BL	20 µl	part of a T-Shirt
13	SA-SE	50 µl + 10 µl	artificial cotton
14	VAG-BL	1 swab + 25 µl	1/4 swab
15	SA	50ul	piece of stockings (nylon)
16	VAG-SE	1 swab + 25 µl	1/4 swab

Table 2
Markers included in the BFID-cSNP-BSS assay (3rd exercise) and the BFID-cSNP-6F assay (4th exercise) [38]. The BFID-cSNP-BSS assay contains BFID markers for the detection of 19 body fluid-specific transcripts, and 21 cSNPs in 11 mRNA targets. These markers can be used for BFID of all six body fluids/tissues and for the association of donors with saliva, semen, and blood. The BFID-cSNP-6F assay contains 23 BFID markers and 46 cSNPs in 20 mRNA markers (all six body fluids/tissues).

		3rd Exercise BFID-cSNP-BSS assay		4th Exercise BFID-cSNP-6F assay	
Body Fluid	BFID Marker	included	# cSNPs	included	# cSNPs
Blood (BL)	ANK1		2		2
	CD3G		1		1
	SPTB		4		4
Semen (SE)	PRM1		1		1
	SEMG2		1		1
	KLK3		2		2
	TGM4		4		4
Saliva (SA)	HTN3		3		3
	PRB4		1		1
	PRH2		1		1
	MUC7		1		1
	STATH				
Vaginal Secretion (VAG)	CYP2A6				1
	MUC22				7
	CYP2B7P1				
Menstrual Blood (MB)	MMP10				2
	MMP3				1
	COL6A3				5
	COL12A1				3
	LEFTY2				
Skin (SK)	LCE1C				3
	COL17A1				1
	IL37				2

with 23 or 24 cycles. All S5 laboratories used the Ion S5™ Precision ID Chef & Sequencing Kit and the HID Ion S5™ Chef templating 02 protocol (500 flows). All MiSeq laboratories conducted the library preparation with the AmpliSeq library PLUS for Illumina kit (with 23 PCR cycles for target amplification), quantified with different kits, and sequenced with the MiSeq™ FGx Reagent Kit on a micro flow cell, including PhiX control, but at different concentrations.

2.2.2. DNA Analysis of the standardized mock case stains

In the 3rd exercise, only the organizing laboratory used a co-extraction method to generate STR profiles of the standardized stains using the NGM Select kit (Supplementary Material, Table S3).

In the 4th exercise, the participating laboratories were asked to co-extract DNA and RNA to facilitate the interpretation of the results by generating STR profiles to determine the number of contributors. Among the S5 and MiSeq participants, nine laboratories extracted DNA from the stains using the QIAamp DNA Mini Kit (Qiagen) which is part of the recommended co-extraction protocol. The other four laboratories co-extracted DNA using different methods (Supplementary Material, Table S2, columns C + D). Ten laboratories further quantified and analyzed the co-extracted DNA using some of the different established DNA quantification and STR amplification kits (NGM Detect™ and NGM SElect™ PCR Amplification Kits (Thermo Fisher Scientific), PowerPlex® ESX 17 and PowerPlex® Fusion 6 C Systems (Promega)) to determine the number of contributors and their sex.

2.2.3. Analysis of DNA Reference Samples (contributors to the standardized mock case stains)

For both exercises, the organizing laboratory collected and analyzed DNA reference samples from the body fluid donors of the 16 standardized mock case stains in each exercise. The DNA reference samples were analyzed through DNA extraction, DNA quantification, library preparation, and sequencing. Biological samples were extracted using the QIAamp® DNA Investigator Kit following the manufacturer's protocol for *Isolation of Total DNA from Surface and Buccal Swabs*. NAO™ Baskets were used to separate the lysate from the swab. Target amplification was performed using the respective DNA cSNP assay with a manual library preparation and 21 cycles for target amplification.

2.3. Analysis of the self-prepared stains

Participating laboratories had the option of preparing and analyzing up to eight additional single-body-fluid or mixed stains for both the 3rd and 4th exercise, consisting of the respective body fluids from up to eight donors. In the 3rd exercise, four out of six laboratories prepared additional stains. In the 4th exercise, five S5 laboratories and four MiSeq laboratories prepared additional stains. Each participating laboratory was responsible for obtaining the approval from its local ethics committee for the sample collection.

2.3.1. RNA analysis of the self-prepared stains

RNA analysis of the self-prepared stains was performed in the same way as for the standardized mock case stains (2.2.1).

2.3.2. DNA analysis of the self-prepared stains

In the 4th exercise, all laboratories were asked to analyze the co-extracted DNA from the stains. Three out of five S5 laboratories and three out of four MiSeq laboratories used the QIAamp DNA Mini Kit, while the remaining three laboratories used the Qiagen AllPrep DNA/RNA/miRNA Universal Kit, the EZ1&2 Virus Mini Kit v2.0 and the DNA IQ System (DC6700) for co-extraction. The DNA extracts were then quantified and STR profiles were generated using some of the established STR kits (NGM Detect™, NGM SElect™and GlobalFiler™ PCR Amplification Kits (Thermo Fisher Scientific), PowerPlex® ESX 17 and PowerPlex® Fusion 6 C Systems (Promega)).

2.3.3. Analysis of DNA reference samples (contributors to the self-prepared stains)

Laboratories were free to use any method of their choice to extract DNA from the reference samples (contributors to the self-prepared stains).

In the 3rd exercise, four laboratories prepared additional stains and therefore also collected DNA reference samples. One laboratory used buccal swabs, while the other three laboratories used a DNA/RNA co-extraction method [7] to generate the DNA reference from their self-prepared stains. This approach is only effective for single donor stains, as the DNA genotype will be a combination of all contributors, making it unsuitable for deconvolution purposes. Three out of four laboratories used a manual library preparation protocol and performed target amplification with 21 PCR cycles using 2 ng (or 3 µL max.) of DNA input. One laboratory used an automated workflow with 24 cycles for target amplification on the same amount of input DNA.

In the 4th exercise, laboratories mainly collected buccal swabs from their body fluid donors as a reference. One laboratory used co-extracted DNA from their self-prepared single donor stains. DNA extraction was performed with one of the various established DNA extraction kits (Supplementary Material, Table S2, columns 2 and 3) or a Chelex-based method [40]. One laboratory used the EZ1&2 Virus Mini Kit v2.0, which is not intended for the extraction of human DNA. Target amplification was performed with 21, 23, or 24 cycles (manual or automated workflows).

2.4. Data analysis

For both exercises, S5 laboratories were provided with an RNA reference file containing the target genes and a target file to configure the run plan correctly in the Ion Torrent suite before sequencing. This enabled the alignment using the integrated Torrent Mapping Alignment Program (TMAP) run with default settings. Each participating S5 laboratory then sent the BAM/BAI files and the run report to the ZIFM. Following sequencing, the MiSeq participants sent the FASTQ files to ZIFM, where the reads were merged using FLASH [41], and aligned with the BWA aligner (v. 0.7.17) [42]. The resulting BAM files were then processed in the same way as the S5 files.

A multiple sequence alignment algorithm (developed by Thermo Fisher Scientific, [38]) re-aligned the TMAP-identified reference gene, requiring all SNP positions of the targeted microhaplotype to be present in the sequence of interest. Two mechanisms reliably eliminated genomic DNA contamination and contaminating sequence products: First, the alignment gap parameters do not allow for multiple base insertions, thus excluding reads containing intronic sequences. Second, the careful selection of the SNP positions for the microhaplotypes ensures that only sequences containing the targeted exon coding regions are aligned. Sequences are phased based on their microhaplotype content, and microhaplotype genotypes are identified and tallied. A Perl script summed the sequence coverage and cSNP genotypes for each of the body fluids present.

A threshold of 0.5 % of the total number of reads was introduced for noise reduction previous to body fluid identification, with read counts below the threshold set to zero [38]. In addition, read counts corresponding to 0.15 % of the total reads were calculated, thereby introducing a further point of reference. The 0.15 % threshold was introduced to address the issue of potential dominance of one body fluid component in mixtures or one body fluid-specific marker, which can result in read counts of markers with good coverage being set to zero when using a 0.5 % threshold because BFID markers of the dominant body fluid/marker received a much higher coverage. Such markers/read counts, although set to zero, were marked in the results tables (Supplementray Material, Tables S7 and S8). The results tables (Supplementary Material, Tables S7 and S8) only contain read counts above the 0.5 % threshold. All BFID markers with remaining positive coverage were considered for interpretation and no additional (potentially absolute) threshold was imposed for BFID. Stains with relatively low coverage (i.e., <500 reads per marker) across multiple markers for different body fluids were considered unsuitable for BFID, as the BFID marker profile was not clear or ambiguous. For instance, if BFID markers specific to three or more body fluids were detected, the marker expression profile was considered ambiguous and no BFID prediction was made. The stains from both exercises were also analyzed in terms of which body fluid components were predicted, calculated as percentages by summing up the read counts of all markers specific to one body fluid and relating it to the total number of reads in that stain.

Once BFID was completed, the consistency across laboratories was evaluated, calculating the average performance score per laboratory and the number of true positives (TP) (i.e., the number of instances where the body fluid(s)/tissue was correctly identified), false negatives (FN) (i.e., the number of instances where the body fluid(s)/tissue was not identified), and false positives (FP) (i.e., the number of instances where the body fluid(s)/tissue was incorrectly identified). It should be noted that the set of standardized mock case stains did not include any blanks or body fluid(s)/tissues that could not be identified with either of the two cSNP assays, hence there are no true negatives.

The average performance per laboratory was computed as the average weighted performance score based on the assigned scores for each sample. The performance score was defined as follows considering the results of the BFID: True positive (TP) BFID received a score of 1. Partial correctness (PC) received a score of 0.5. The category PC consists of both, BFID results of mixtures for which only one body fluid component was identified correctly, and correctly identified stains with low certainty (relatively low number of reads in one or two BF markers for one body fluid). Both FN and FP received a score of 0. The average score was calculated by taking the mean of the score values for all stains analyzed by each laboratory. To assess the distribution of identification outcomes across laboratories, we calculated the proportion of each category (TP, FN, PC, FP) for each laboratory. The total number of occurrences of each category was determined, and the proportion was calculated as the ratio of each category's occurrences to the total number of stains analyzed by the laboratory.

Once the body fluid(s) present have been identified, the possibility of associating the body fluid(s) with the donor(s) can be investigated. For the following analyses, all reported sequencing results were examined. For donor association, the RNA cSNP genotypes were compared with the genotypes obtained from the reference samples using the respective DNA assay. cSNPs in all body fluid markers that have been identified were considered and examined. In addition, the number of missing, correctly and incorrectly (i.e. those showing discrepancies with the DNA genotype) typed cSNP genotypes were calculated in percentages. Failed stain analysis (grayed out in Supplementary Material, Tables S12 and S13) resulted in the base being adjusted when calculating the percentages in the respective categories (correct/missing/incorrect). Dropouts for heterozygous alleles were considered as incorrect. No absolute threshold was set for the coverage of correct calls. Moreover, potential imbalances in allele expression for certain heterozygous cSNPs were not taken into account when evaluating whether the observed cSNP genotype was correct. For further calculations and visualizations, only table entries with a single corresponding cSNP were considered.

2.5. Analysis with an alternative cSNP Panel

The BFID-cSNP-6F assay was compared with an alternative cSNP panel [39], developed by a participating laboratory (Cologne). The alternative cSNP panel contains 30 markers including 70 cSNPs, allowing the identification of five body fluids and skin (Supplementary Material, Table S5). The ZIFM re-analyzed their cDNA of the 16 stains from the 4th exercise with the alternative RNA cSNP panel. For target amplification, 23 cycles were applied. Library preparation was performed using the Precision ID Library Kit and the 16 RNA libraries were sequenced on an Ion 520™ chip. The pooled RNA libraries were diluted to a concentration of 30 pM.

The 14 reference DNA samples were analyzed with the corresponding alternative DNA cSNP panel [39]. The corresponding DNA libraries were sequenced on an Ion 520™ chip along with 20 other samples (not related to the exercise). Templating was performed using the Ion 510™ & Ion 520™ & Ion 530™ Kit-Chef Kit and the HID Ion S5™ Chef templating 02 protocol (500 flows).

Following sequencing, automatic alignment was performed using TMAP with default settings in the Torrent Suite Software (Thermo Fisher Scientific). The BAM/BAI files were further analyzed using several of the Torrent Suite Software plugins. The *ampliSeqRNA*, the *coverageAnalysis*, the *HID_SNP_Genotyper_v5_2_2,* and the *variantCaller* plugin were run. The *ampliSeqRNA* plugin was used with the following settings: Library Type – AmpliSeqRNA, filter barcodes – true, ERCC tracking – false. The *coverageAnalysis* was run with the following settings: Library Type - AmpliSeq DNA, Read filters - uniquely mapped, minimum aligned length= 30. The *HID_SNP_Genotyper_v5_2_2* was run with default settings and the respective hotspot and target files. For the *variantCaller*, the following settings were employed: Library Type – AmpliSeq, Configuration - Generic – S5/S5XL (510/520/530) – Germ Line – Low Stringency. The output of the *coverageAnalysis* was further analyzed by applying the same thresholds (0.5 % and 0.15 %) as for the data analysis of the BFID-cSNP-6F assay, setting read counts below these thresholds to zero, or setting them to zero but marking them if they contribute to the interpretation.

3. Results

3.1. Results of the standardized mock case stains

3.1.1. Nucleic acid extraction and quantification

In the 3rd exercise, RNA yields per stain were generally comparable across laboratories (Supplementary Material, Fig. S1), with average, minimum, and maximum RNA yields for each stain reported in the Supplementary Material (Table S3). In contrast, RNA yields in the 4th exercise exhibited greater variability (Supplementary Material, Fig. S1), with the corresponding averages, minimum, and maximum values also provided in the Supplementary Material (Table S3).

In the 4th exercise, laboratories were additionally encouraged to co-extract DNA from the stains for STR profiling. DNA quantification values showed considerable variation between laboratories. However, stains that yielded low or high amounts of DNA were relatively consistent across laboratories, despite inter-laboratory differences (Supplementary Material, Fig. S1). Three laboratories did not report DNA quantification values.

3.1.2. STR Profiling

The laboratories participating in the 4th exercise were instructed to perform STR profiling of the co-extracted DNA from the stains to determine the number of contributors, the major and minor contributors (if applicable), and the sex of the contributor(s). For stain 8 (a saliva-skin mixture), most laboratories detected only the female contributor (saliva donor). Similarly, for stain 14 (a mixture of vaginal secretion and blood), the male contributor (blood donor) was missing in 8 out 12 cases (Supplementary Material, Table S4).

3.1.3. Body fluid identification

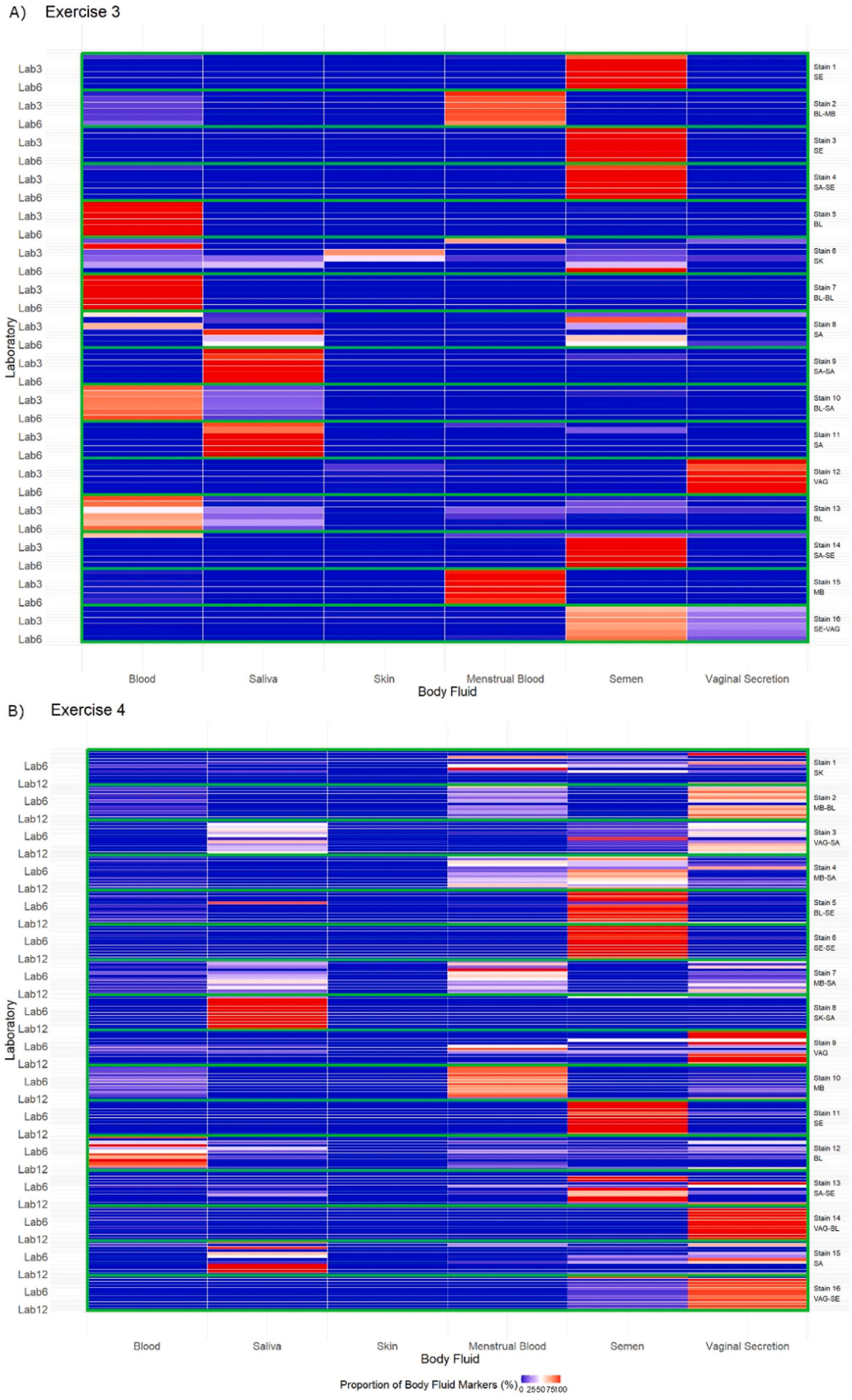
Before shipment, each set of 16 stains (3rd and 4th exercise) was tested by the organizing laboratory using an in-house RNA 12plex CE assay [43]. This assay contains 12 body fluid-specific markers for blood (2), saliva (2), semen (2), vaginal secretions (4), and menstrual blood (2). For both sets of stains, the CE analysis of the 12plex PCR product revealed many of the expected peaks corresponding to the correct body fluids (Supplementary Material, Table S6). For some low-input stains, no results were obtained or one body fluid result in a mixture was not detected. No results were obtained for skin samples because the multiplex assay does not include skin-specific markers. Sporadic not-expected markers were observed, mostly in high input samples, usually as small peaks. Some stains involving vaginal fluid (VAG or MB) also showed semen markers, presumably from recent sexual intercourse, as the body fluid donors were neither asked about their sexual behavior, nor instructed to remain abstinent prior to sample donation. Semen markers (*PRM1, SEMG1*) were prominently observed in stain 3 analyzed in the 4th exercise. In stains 7, 10, and 14 from the 4th exercise low peaks (<360 RFU) for *PRM1* were detected.

In the 3rd exercise, the cSNP-BSS-assay enabled the successful identification of the body fluid components in 13 out of 16 stains across laboratories, meaning that at least one, but usually several of the laboratories correctly identified all body fluid components in the stain (Supplementary Material, Table S7). In the 4th exercise, the BFID-cSNP-6F assay allowed for correct body fluid identification for 14 out of 16 stains across laboratories (Supplementary Material, Table S8). For BFID predictions, only mRNA markers with positive coverage after the 0.5 % noise removal were considered. However, stain 14 (VAG-BL) in the 4th exercise was an exception: vaginal secretion markers dominated the sequencing results to such an extent that blood markers, when detected, were consistently below the threshold in all laboratories. Thus, the 0.15 % threshold introduced to flag informative mRNA markers below the noise-correcting threshold were used for interpretation (Supplementary Material, Table S8). In both exercises, mRNA markers included in the assays to identify body fluids achieved high coverage overall, although for each body fluid one or two markers appeared to be more dominant, i.e. showed a higher coverage compared to the other body fluid-specific markers. This was especially evident for semen where *PRM1* showed the highest coverage. Blood is an exception as each marker was equally well covered (Supplementary Material, Tables S7 + S8). The mRNA markers for skin obtained the lowest coverage overall. The sequencing results of each participating laboratory in both exercises were visualized in heat maps (Fig. 1, A + B), showing the proportion of the total read counts in each of the body fluid categories, including on-target and off-target reads. The heat maps demonstrate the extent to which the sequencing results are consistent across laboratories. Especially in the 3rd exercise (Fig. 1, A), high proportions of the total read counts are in the expected body fluid categorie(s). However, in the 4th exercise, particularly evident in the mixtures, a greater degree of variation between the laboratories is seen in the distribution of read counts across the six body fluid categories (Fig. 1, B). For instance, in the case of stains 2 and 7, the proportions of each body fluid component differed between the laboratories that identified the correct body fluids. The majority of laboratories did not correctly identify stains 1 and 15, resulting in the heat maps not showing a congruent distribution of read counts across the six body fluid categories.

Typically, in high input samples, multiple markers of a specific body fluid are amplified, thereby enhancing a correct identification. Conversely, low input stains and mixed stains pose greater challenges, in particular the detection and identification of skin, either as a single tissue stain or as a component of a mixture. In both exercises, the percentage of non-body-fluid-specific targets being amplified in correctly predicted stains is negligible. Considering all sequencing results of the standardized stains in the 3rd and 4th exercise, an average of 87.65 % and 73.3 % of the total reads were attributable to the body fluids expected based on each stains' composition, respectively. A table including average, minimum and maximum on- and off-target amplification percentages per stain and exercise can be found in the Supplementary Material (Table S9). Fig. 2 shows the percentages of body fluid specific reads for each stain, averaged across all participating laboratories in each exercise. This shows nicely how well the body fluid components were predicted for most of the stains. For the stains whose body fluid components could not be determined, a certain percentage is allocated to several body fluids, all with very low read counts. Stacked bar plots showing the proportion of each body fluid detected in the standardized stains for all laboratories across both exercises can be found in the Supplementary Material (Fig. S2 and Fig. S3).

The average weighted performance score was calculated for each laboratory in both exercises as an overall measure for each laboratory's BFID accuracy (Supplementary Material, Fig. S4). The average scores varied between laboratories, with values ranging from 0.75 to 0.88 (3rd Exercise) and 0.59–0.81 (4th Exercise). Laboratory 4 (3rd Exercise) and laboratory 8 (4th Exercise) demonstrated the highest overall accuracy with an average score of 0.88, and 0.81, respectively, indicating a high proportion of correct identifications (True Positives). In contrast, laboratories 9 and 12 (4th Exercise) had the lowest average score of 0.59, reflecting a lower performance, with a higher occurrence of incorrect or partial identifications. The average score was slightly higher in the 3rd exercise (0.81) compared to the 4th exercise (0.69).

The distribution of identification outcomes—True Positive (TP), False Negative (FN), Partially Correct (PC), and False Positive (FP)—were analyzed for each laboratory to further characterize the potential BFID errors made (Fig. 3). In the 3rd exercise, all six participating laboratories showed a high proportion of TP (≥0.69), while the proportion of PC was marginally higher than that of FN. In the 3rd exercise, stain 8 was twice predicted as a saliva-semen mixture based on the sequencing results of Laboratories 2 and 6, resulting in two false positive BFID predictions. In the 4th exercise, the distribution of identification outcomes between the different laboratories showed a bigger variation. The proportion of TP ranged from 0.375 to 0.75. While there was a certain proportion of FP BFID predictions, these results show that the main source of error in most laboratories was either identifying only one



(caption on next page)

Fig. 1. Body fluid identification results across laboratories in Exercises 3 and 4 visualized as heat maps. Each heat map displays the proportions of sequencing reads assigned to six body fluid marker categories (Blood, Saliva, Skin, Menstrual Blood, Semen, and Vaginal Secretion) for each stain analyzed by the participating laboratories. Each row represents one laboratory's result for a given stain; each column corresponds to one body fluid category. Green frames group the sequencing results by stain, showing the inter-laboratory agreement within each set, with the true body fluid composition of the stain indicated next to the frame. The color intensity, ranging from blue to white to red, reflects the proportion of reads assigned to each fluid, normalized per stain.

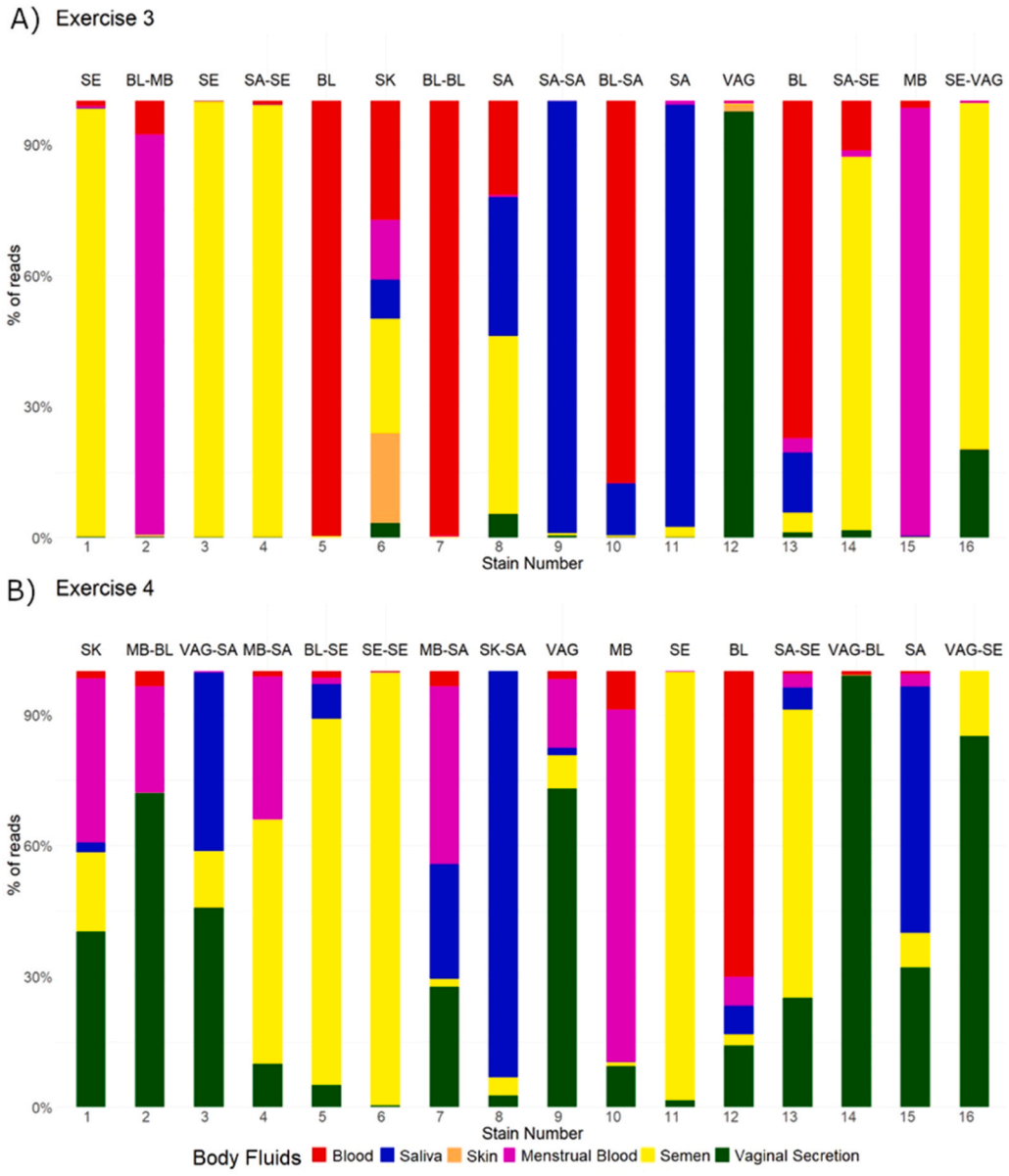


Fig. 2. Percentages of body fluid specific reads for each of the 16 stains (averaged across all participating laboratories per exercise).

component in a mixture (PC) or failing to identify the body fluids (FN). The FP category included three misclassifications of stain 5 (SE-BL) where the second component, blood, was incorrectly predicted as vaginal secretion, menstrual blood, or saliva. In stain 13, composed of saliva and semen, the saliva component was incorrectly predicted as vaginal secretion, and in another case only weak signals for vaginal secretion markers were detected.

3.1.4. Donor Association

While the BSS cSNP assay examined in the 3rd exercise allowed the matching of body fluids and their respective donor for blood, saliva, and semen, the 6 F cSNP assay additionally contained body fluid-specific markers with cSNPs for vaginal secretion, menstrual blood, and skin. The higher the number of body fluid-specific cSNP markers detected and the higher the coverage of each cSNP marker, the more accurately and completely the RNA genotypes reflect the DNA genotypes.

In the 3rd exercise, we observed a relatively high number of reads for the single-source stains 1, 3, and 5 in the body fluid-specific markers, and the donor DNA genotypes were mostly well reflected in the RNA genotypes (138 cSNPs considered, 8.7 % missing and 2.17 % discrepant genotypes) (Supplementary Material, Table S12). For stains 11 and 13, most of the body fluid-specific markers had good coverage and therefore discrepancies between RNA and DNA genotypes were rarely observed (78 cSNPs considered, 1.3 % missing and 1.3 % discrepant genotypes, Supplementary Material, Table S12). For single donor stains 6, 12, and 15, no donor associations could be made because the BFID-cSNP-BSS

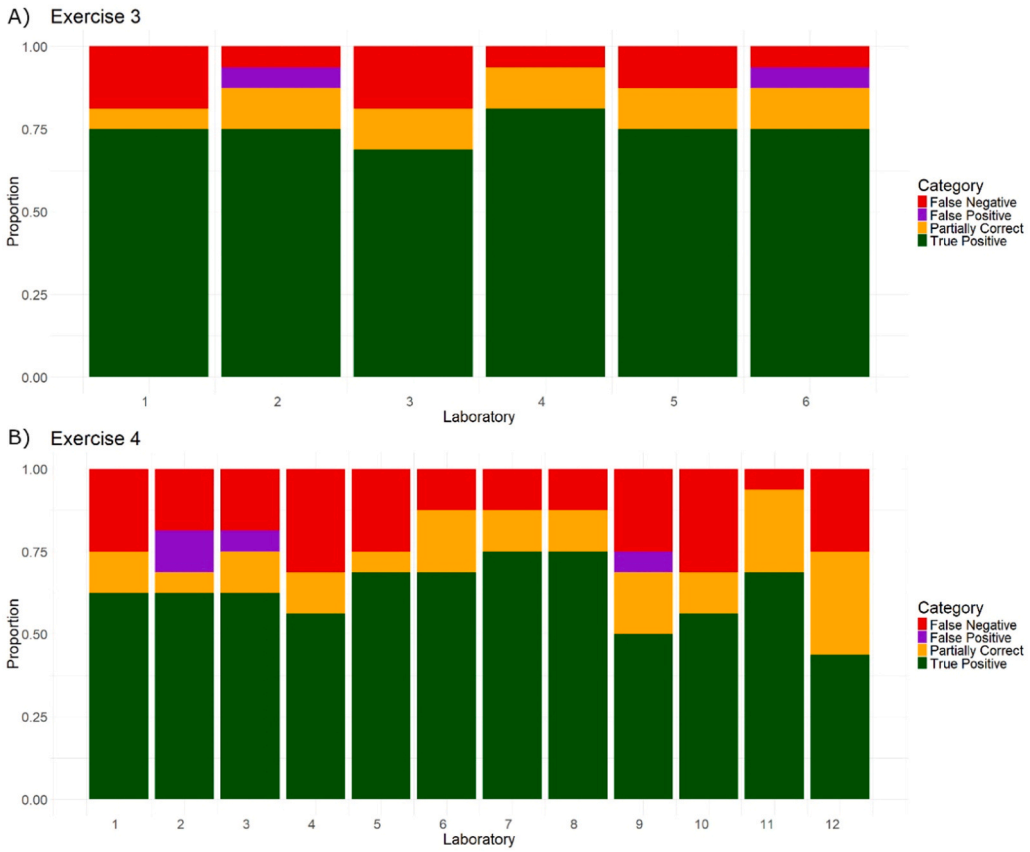


Fig. 3. Proportions of False Negatives, False Positives, Partially Correct and True Positives.

assay used in the 3rd exercise does not contain cSNPs for skin, vaginal secretion, and menstrual blood. Stain 8 (saliva on a plastic spoon) was challenging, but one laboratory obtained high-coverage RNA genotypes that reflected the DNA genotypes well (Supplementary Material, Table S12).

Mixed stains composed of the same body fluid from different donors (stains 7 and 9) showed high numbers of reads and the allele ratios of the sequenced RNA genotypes were close to the theoretical expectations (Supplementary Material, Table S12). Stains 2, 10, and 16 yielded high numbers of reads, resulting in few observations of discrepant RNA and DNA genotypes (168 cSNPs considered, 2.98 % missing and 7.14 % discrepant cSNP genotypes between all three stains) (Supplementary Material, Table S12). Table 3 shows the cSNP read counts of all laboratories for stain 10, a blood-saliva mixture. Stains 4 and 14 (saliva-semen mixtures) showed few reads overall, with the saliva portion of each stain being particularly problematic. For the semen portion of the stains, a few markers were not well covered, resulting in a high number of dropouts and some discrepancies between RNA and DNA genotypes (96 cSNPs considered, 45.12 % of missing and 15.4 % discrepant cSNP genotypes between the two stains, semen portion only).

In the 4th exercise, six of the 16 stains were single donor stains. Stains 9, 10, 11, 12, and 15 were prepared on realistic carrier materials and the remaining were on cotton swabs. Stains 9 and 10 worked well overall, meaning that body fluid-specific markers gave a high number of reads and cSNP genotypes matched between RNA and DNA (400 cSNPs considered, 16.75 % missing, 2 % discrepant cSNP genotypes between the two stains) (Supplementary Material, Table S13). In stain 11, the semen marker *PRM1* obtained a high coverage, but some of the other semen markers were less well covered, therefore making donor association more difficult (Supplementary Material, Table S13).

The set of stains provided in the 4th exercise included ten mixtures. From a forensic perspective, mixtures, especially semen and vaginal secretion, are highly relevant in casework. Overall, components like vaginal secretion and menstrual blood were consistently detected, and donor association was successful since several markers for each of these body fluids showed a high coverage. For instance, stain 16, a vaginal secretion and semen mixture (Table 3), obtained high coverage in most of the expected body fluid markers, and the RNA cSNP genotypes mostly reflected the DNA donor genotypes (192 cSNPs considered, 11.45 % missing, 3.12 % discrepant cSNP genotypes). Similar results were obtained for stain 4 (a mixture of menstrual blood and semen), however, not all laboratories detected the blood component of the menstrual blood to the same extent, which is reflected in the percentage of dropouts (408 cSNPs considered, 23.77 % missing and 3.2 % discrepant cSNP genotypes) (Supplementary Material, Table S13). For some of the mixtures, one component was missing in several of the laboratories' sequencing results (e.g., stains 13 and 14) or was entirely missing (skin in stain 8) (Supplementary Material, Table S13).

For both exercises, we assessed the number of correctly typed, missing and incorrectly typed (i.e., those discrepant from the DNA genotypes) cSNPs, in percentages. The corresponding numbers for each marker (except for the skin markers because of data scarcity), examined per stain and exercise can be found in the Supplementary Material (Tables S12 + S13). The percentages of cSNPs in each category (correct/missing/incorrect) are presented per body fluid (Fig. 4). For blood, semen and saliva, the data from both exercises were pooled, since the same cSNPs were analyzed.

For the blood cSNPs analyzed in the 3rd exercise set of stains, we found that most cSNP genotypes were correctly called, regardless of whether the stain was a low input, high input or mixed stain (Fig. 4). In the 4th exercise, a rather high percentage of missing cSNP genotypes was observed, especially in the low input stain 12. For saliva cSNP genotypes, we found a higher percentage of missing genotypes, especially in low-input stains, and mixtures (depending on the other body fluid

Table 3
The number of full cSNP profiles identified by the participating laboratories in each exercise per stain, and per body fluid component (if applicable). BF/T stands for body fluid/tissue, BL for blood, MB for menstrual blood, SA for saliva, SE for semen, VAG for vaginal secretion, and SK for skin. BF/T 1 corresponds to the first body fluid, BF/T 2 to the second body fluid listed for each stain (2nd column). The last column (Complement) lists additional body fluids that may co-occur with menstrual blood (VAG, BL) and whose cSNP profiles may also be relevant for interpretation. * indicates the systematically missing cSNP genotype for *CYP2A6* in VAG cSNP profiles. The color light green is used to indicate that at least 50 % of the laboratories participating in the respective exercise reported a full cSNP profile, i.e. the correct cSNP genotype for each cSNP BFID marker of interest. The light yellow color is used to indicate that between 25 % and 50 % of the full cSNP profiles were reported, whereas the color light red is used to indicate that less than 25 % of the full cSNP profiles were reported.

		Full cSNP Profile(s)		
Stain N°	BF/T	BF/T 1	BF/T 2	Complement
Ex.3-1	SE	3/6	NA	
Ex.3-2	BL-MB	5/6	NA	
Ex.3-3	SE	2/6	NA	
Ex.3-4	SA-SE	0/6	0/6	
Ex.3-5	BL	6/6	NA	
Ex.3-6	SK	NA	NA	
Ex.3-7	BL-BL	6/6	6/6	
Ex.3-8	SA	0/6	NA	
Ex.3-9	SA-SA	3/6	3/6	
Ex.3-10	BL-SA	6/6	0/6	
Ex.3-11	SA	5/6	NA	
Ex.3-12	VAG	NA	NA	
Ex.3-13	BL	3/6	NA	
Ex.3-14	SA-SE	0/6	0/6	
Ex.3-15	MB	NA	NA	
Ex.3-16	SE-VAG	4/6	NA	
Ex.4-1	SK	0/12	NA	
Ex.4-2	BL-MB	9/12	9/12	
Ex.4-3	SA-VAG	5/12	0/12	
Ex.4-4	MB-SE	6/12	6/12	BL: 1
Ex.4-5	BL-SE	5/12	2/12	
Ex.4-6	SE-SE	1/12	1/12	
Ex.4-7	MB-SA	6/12	6/12	VAG:8*, BL: 4
Ex.4-8	SA-SK	7/12	0/12	
Ex.4-9	VAG	9*/12	NA	
Ex.4-10	MB	7/12	NA	VAG:11*, BL:8
Ex.4-11	SE	0/12	NA	
Ex.4-12	BL	0/12	NA	
Ex.4-13	SA-SE	0/12	0/12	
Ex.4-14	VAG-BL	11/12	3/12	
Ex.4-15	SA	0/12	NA	
Ex.4-16	VAG-SE	11/12	6/12	

present) (Fig. 4). Nevertheless, a high percentage of correctly called cSNPs could also be found, and the percentage of incorrectly typed cSNPs was rather low. Semen showed the highest number of incorrectly typed cSNP genotypes, based on the observed percentages in the respective categories (correct/missing/incorrect) (Fig. 4). Such mistyped cSNP genotypes often correspond to the loss of one allele in a heterozygous cSNP. At the BFID level, we had already observed that PRM1 had a high coverage, which was also reflected in the high percentage of correctly called cSNP genotypes. For the other markers, the low coverage greatly affected the ability to correctly type both alleles for heterozygous cSNPs. However, we also observed instances where homozygous alleles turned out to be heterozygous, with the correct allele accounting for the majority of calls and a minor proportion attributed to an incorrectly called allele (Supplementary Material, Tables S12 + S13). For vaginal secretion, the highest percentage of correctly typed cSNPs was found, regardless of whether the stain was in the high-input or mixed stain category (Fig. 4). Moreover, the cSNP genotype for *CYP2A6* was systematically not called. Similar to vaginal secretion, menstrual blood cSNP profiles were found to be highly congruent with the DNA reference profiles, resulting in a low number of incorrectly called cSNP genotypes (Fig. 4). In contrast to saliva and semen, the percentage of missing cSNP genotypes was rather low.

We compiled the number of complete cSNP profiles in the standardized stains, considering each detected body fluid component separately (Table 3). For vaginal secretion, a cSNP profile was considered complete even though *CYP2A6* was always missing.

3.2. Self-prepared stains of the Laboratories

3.2.1. Nucleic Acid Extraction and Quantification

For self-prepared stains, RNA yields after extraction varied widely between stains and laboratories (Supplementary Material, Fig. S5). In the 3rd exercise, the RNA yield was highest for the self-prepared stains from laboratory 6. In the 4th exercise, the RNA yield of the self-prepared stains was overall lower than in the 3rd exercise. Moreover, the co-extracted DNA from the self-prepared stains was quantified, showing that RNA and DNA yields do not necessarily correlate.

3.2.2. STR Profiling

Participants in the 3rd exercise did not conduct STR analysis, hence stains were interpreted based solely on the body fluid components detected (Supplementary Material, Table S10, column H). Most laboratories participating in the 4th exercise provided STR profiles generated from the co-extracted DNA from the self-prepared stains (Supplementary Material, Table S10, column I and J). The organizing laboratory considered this additional information to facilitate the interpretation of the sequencing results of the self-prepared stains.

3.2.3. Body Fluid Identification and Donor Association

During the 3rd exercise, four out of the six participating laboratories analyzed self-prepared stains (eight stains each). Laboratories were free to include body fluids and/or tissues that the BFID-cSNP-BSS assay was not designed to identify, and/or the assay did not include markers with cSNPs for donor association. Hence, the inability to predict the body fluids present in a stain may be due to the absence of body fluids, the lack of markers for BFID, or a failed amplification. Predictions on BFID level could be made for 21 out of 32 stains (65 %) (Supplementary Material, Table S10, column H). Presumable donors could be identified for ten of these stains.

In the 3rd exercise, interpretations were also categorized as correct in case there was no body fluid predicted for stains whose composition could not be identified using the BFID-cSNP-BSS assay or for stains containing no body fluids (extraction blanks, Supplementary Material, Table S10, column M). Laboratory 4 included an ear-wax stain (*Ex3_Lab4_own7*) that could not be identified, while a mixture of nasal secretion and epithelial cells (*Ex3_Lab4_own8*) as well as a saliva stain (*Ex3_Lab5_own4*), were mistakenly classified as a mixture of menstrual blood and saliva.

Additionally, mRNA analysis of self-prepared stains required the extraction of DNA reference samples from body fluid donors. Laboratories could also submit potential DNA references from non-body fluid donors. While the concept of the cSNP analysis was new to several laboratories, one of the participating laboratories of the 3rd exercise inadvertently used co-extracted DNA from its self-prepared stains for the DNA reference target amplification. In case of mixtures, donor matching was then impossible. Another laboratory intentionally used the DNA co-extraction method to prepare their DNA reference samples, as all of their

Fig. 4. Percentage of cSNP genotypes found in the respective categories (correct/missing/incorrect) per body fluid and stain. The code above each individual stacked bar plot is composed of the following elements: The number indicates the exercise the stain was part of (3 = 3rd exercise, 4 = 4th exercise), low input/high input/mixture denotes the stain type, and the last number corresponds to the stain number.

stains were single body fluid stains.

Within the 4th exercise, sequencing results were obtained for 68 self-prepared stains, but body fluid predictions were possible for only 42 stains (62 %). The reasons for this include inadequate coverage of BFID markers, the presence of body fluids that the assay does not cover, and the use of extraction blanks/non-template controls (NTCs). Each laboratory correctly analyzed the reference DNA samples from the body fluid donors. Donor association was rather difficult as the read counts for some markers were low and therefore did not provide the full genetic information possibly available. Nevertheless, donor association was possible for 21 stains.

Furthermore, BFID predictions were interpreted by the organizing laboratory, additionally considering the STR analysis results of the self-prepared stains (Supplementary Materials, Table S10, column M). In some cases, STR profiles may reveal a second contributor to a stain that is not immediately apparent from a BFID-perspective. This could be due to the absence or concealment of a second body fluid. For example, it is challenging to discriminate between menstrual blood and blood in a mixture (Sample *Ex4_Lab6_own8*) based on mRNA profiling results alone. Blood markers are usually detected in menstrual blood, as blood is one of its components, in addition to vaginal secretion. The interpretation of body fluid and tissue samples for which the cSNP assays do not include any BFID markers is equally challenging, not least because of potential cross-reactivity. For instance, stain *Ex4_Lab7_own1* (a rectal mucosa sample) was predicted to be a menstrual blood-skin mixture. However, after considering the STR results, which indicated that the stain originated from a single male donor, it is highly unlikely that menstrual blood is present. Another rectal mucosa sample (*Ex4_Lab7_own4)* appeared to be a mixture of menstrual blood and skin based on BFID analysis. This observation was consistent with the STR analysis

result. A third rectal mucosa sample (*Ex4_Lab2_own7*) was misclassified as vaginal secretion based on BFID results and could not be further characterized since an STR profile was not provided. Among the self-prepared stains were three nasal mucosa samples (*Ex4_Lab5_own8, Ex4_Lab7_own5, Ex4_Lab7_own6*). Based on the sequencing results, Ex4_Lab5_own8 was not classified, while the latter two stains were classified as saliva, and a saliva-vaginal secretion mixture, respectively. For *Ex4_Lab12_own5*, the STR profiling result refined the interpretation, but the body fluid corresponding to the male component in the stain was not detected (Supplementary Material, Table S10, column M, highlighted in yellow). Furthermore, a menstrual blood sample (*Ex4_Lab2_own2*) was correctly characterized as such, yet an incorrect donor from the pool of potential donors was identified from the RNA cSNP genotypes (Supplementary Material, Table S10, column M, highlighted in orange). All sequencing results of the donor association of the self-prepared stains from both exercises can be found in the Supplementary Material (Tables S15-S26). Both BFID and donor association proved more complex for the self-prepared stains in both exercises, as these samples showed more variability and also lacked a reference frame (e.g. a defined hypothesis) to facilitate the interpretation.

3.3. Comparison between Laboratories and Sequencing Platforms

Within both exercises, six laboratories used only the S5 platform, while three laboratories used only the MiSeq sequencing system. Two laboratories performed the experiment on both sequencing platforms.

In the 3rd exercise, we directly compared the relative read counts obtained per marker for the six participating laboratories. After correcting for the total number of reads (including off-target reads), we found that the percentages per stain per marker showed a high degree of comparability between laboratories, for single source and mixed stains (Supplementary Material, Fig. S6). However, we found a higher variation in the percentages per body fluid for the more challenging stains 6 and 8 (a skin swab and a licked plastic spoon, respectively). While laboratory 4 achieved satisfactory results with high coverage for several of the expected saliva markers, the other laboratories reported overall low coverage or off-target markers, potentially leading to misinterpretations regarding the body fluids present. This greater variation in the distribution of body fluids detected across laboratories can be seen in a stacked column chart (Supplementary Material, Fig. S2).

In the 4th exercise, high input single-source stains (MB, SE, and VAG) gave relatively consistent results between laboratories using the same sequencing platform (Supplementary Material, Fig. S7, A) + B)). The remaining single-source stains yielded more variable results, as was already evident from the body fluid identification. In general, MiSeq laboratories detected fewer off-target markers at low percentages. For the mixed stains analyzed in the 4th exercise the results are more homogeneous between laboratories using the same sequencing platform (Supplementary Material, Fig. S8, A) + B)).

Two laboratories performed library preparation and sequencing on both platforms starting from the same cDNA (results are shown in Supplementary Material, Fig. S9). For the first laboratory, the results were considered sufficiently comparable, since the BFID markers representing the largest proportion were amplified to the same level in both workflows, with minor off-target amplification, which was slightly more prevalent in the S5 workflow. Target amplification failed for stains 12 (blood on T-shirt) and 13 (saliva-semen mixture) in the MiSeq and the S5 workflows. The sequencing results for the second laboratory (Supplementary Material, Fig. S10), were consistent on both platforms. Target amplification failed for stain 1 (skin) in both workflows.

3.4. Comparison with an alternative cSNP Panel

In Laboratory 3, the 16 stains analyzed in the 4th exercise were subjected to an additional analysis using an alternative cSNP panel [39] on the same cDNA. This panel includes markers containing cSNPs for body fluid identification and donor association of blood, menstrual blood, saliva, semen, vaginal secretion and skin. A comparison of the sequencing results of the two cSNP panels at the BFID level revealed that the BFID-cSNP-6F assay accurately characterized the nature of the body fluids in 10 of the 16 stains, while the alternative panel determined the body fluid components in 8 of 16 stains (Table 4, and Supplementary Material, Table S11). The BFID-cSNP-6F assay was only able to identify one of the stain's components for two stains. Furthermore, for stain 15, the read counts for saliva were relatively low, which affected the accuracy of the prediction. The alternative cSNP panel [39] incorrectly identified stain 8 as a mixture of saliva and blood, when in fact it was a mixture of saliva and skin. Stain 12 was also misidentified as menstrual blood instead of blood.

The two panels showed comparable results in terms of assigning percentages to specific body fluids, based on the markers that were amplified (Fig. 5). The alternative panel [39] demonstrated greater variability in the percentages of detected body fluids for stains 1, 13, and 15, likely due to a low total read count. Similar observations were made with the BFID-cSNP-6F assay if total read counts were low for a stain. Although the majority of reads for stain 12 are associated with blood markers, some menstrual blood markers also obtained a coverage that exceeded the 0.5 % threshold. This ultimately leads to misinterpretation of the nature of the stain. It is noteworthy that the BFID-cSNP-6F assay identified vaginal secretion markers in greater proportions across all stains.

A more detailed analysis of the markers that received coverage within the body fluid categories revealed that across all stains, markers found in both panels (Fig. 6, green boxes) were amplified to a similar extent. Although the alternative cSNP panel [39] contains a greater number of markers, the distribution of percentages among body fluid-specific markers may vary, resulting in cases where a marker shared between panels may receive more or fewer reads.

Of the 16 analyzed stains, the alternative panel [39] generated several full cSNP profiles for the menstrual blood component (stains 2, 3, 4, 7, 9, and 10). Moreover, the cSNP profile for the vaginal secretion component of stain 3, the semen component of stain 4, and the blood component of stain 10 yielded a complete cSNP profile. However, discrepancies were found between the RNA and DNA cSNP genotypes, with several cases of discordant genotypes observed (Supplementary Material, Table S14).

4. Discussion

The objective of this study was to evaluate the BFID-cSNP-BSS and BFID-cSNP-6F assays on a variety of mock case stains in different laboratories and on two sequencing platforms. The laboratories participating in the study were provided with mock case stains and comprehensive instructions regarding the analysis of the standardized stains. Despite the harmonization of the experiment, variations in methods persisted, reflecting a realistic scenario in which each laboratory prefers its established protocols for certain analysis steps. In addition to the standardized stains, the laboratories also analyzed self-prepared samples. The sequencing data were collected and analyzed at the ZIFM, with comparisons made at the BFID level, regarding donor association and across MPS platforms. The self-prepared stains were interpreted by the ZIFM without prior knowledge of the sample's identity, and were therefore challenging, but were considered to reflect actual casework scenarios where sample composition in terms of body fluids present is unknown.

We first discuss some aspects of the RNA analysis methods. The RNA/DNA co-extraction is a critical factor, while most participants followed the recommended protocol (mirVana™ miRNA Isolation Kit). Furthermore, as shown in a previous study [37], our findings confirm that higher RNA concentrations do not necessarily lead to better results. This is because the sensitivity and specificity of different RNA quantification methods can vary considerably. Additionally, the quantification of

human or non-human (microbial) RNA content is a recognized problem, as there is no human-specific or sensitive method for the forensic low template RNA samples [44,45]. Consequently, total RNA is quantified and not only the target mRNA [46]. Therefore, it is difficult to draw conclusions about the efficacy of different extraction methods solely based on quantification values. Greater RNA input into reverse transcription (RT) does not automatically lead to better results [47]. Although the protocol required an input of 50 ng or 25 ng RNA for SuperScript™ IV VILO RT for a manual/MiSeq or automated (Chef DL8) RT, respectively, inaccurate or lack of quantification can lead to variation in cDNA yield. In addition, even when using the maximum input volume, the recommended 50 ng input could not always be achieved. Significant variability in the RT step has been reported in the literature [48,49], with even greater variation observed at low template concentrations [50]. The reported results reflect the chosen primer design strategy, reaction conditions and initial RNA concentration [51], as well as the choice of RT enzyme [52]. Moreover, an excess of template RNA has been shown to result in marker cross-reactivity, whereas an insufficient amount of RNA can lead to false negative results [53–56]. While the stain analysis encompasses several steps beyond RNA extraction and subsequent RT, target amplification represents the most crucial part of the protocol. In forensic DNA analysis, the presence of low levels of starting template promotes stochastic sampling effects, which become more pronounced when traditional methods are employed to enhance sensitivity, such as the use of an increased number of PCR cycles [57]. These caveats in the workflow likely contribute to the variability in results observed between laboratories. Similarly, inconsistent RNA input amounts due to the lack of or inaccurate RNA quantification may have resulted in reduced cDNA yields and inconsistent marker detection. Discrepancies in marker detection or sequencing depth between laboratories could be mitigated by using the maximum RNA input for RT instead. Nonetheless, it remains difficult to determine which specific step in the workflow is primarily responsible for the observed variation.

Regardless of the different levels of experience in RNA analysis among the participants in both exercises, our results indicate that both laboratories with little experience and others with more experience were able to complete the analysis successfully. In the 3rd exercise, laboratory 3 co-extracted DNA of the stains instead of DNA of the reference persons and donor association was not possible for mixtures. In the 4th exercise, the sequencing results from laboratory 13 were excluded from further analysis. We assume that the stain analysis failed because the laboratory used the EZ1&2 Virus Mini Kit v2.0 to co-extract RNA and DNA from the stains, which may not be optimal for the extraction of human genetic material. Aside from these shortcomings, which clearly affected the success of the stain analysis, there is no indication that the experience of the participating laboratories played a significant role in achieving reliable results.

For body fluid identification, most of the high-input samples were correctly classified (stains 3, 5, 11, 12, and 15 in the 3rd exercise, and stains 9, 10, and 11 in the 4th exercise). Still, the identification of stain 15 in the 4th exercise, consisting of 50 µL of saliva on thick nylon tights (thread size of 120 Denier), proved to be a challenging task. A significant number of laboratories reported low read counts for markers of several body fluids, including saliva, vaginal secretion and semen. The low absorbency of nylon may contribute to the negative results. Additionally, the stain was not visible to the naked eye, hence laboratories were instructed to cut a circular piece of the nylon from the center of the stain. For the low-input samples (stains 1, 6, 8, and 13 in the 3rd exercise, stains 1 and 12 in the 4th exercise), the results showed greater variability. Some stains were consistently typed correctly by all laboratories (stains 1 and 13 in the 3rd exercise), while others were more challenging, with results varying between laboratories. In the 4th exercise, no participating laboratory identified skin markers in stain 1, only sporadic low-coverage peaks for other body fluid-specific markers were observed. The analysis of gene expression in skin samples is a significant challenge, not only in the context of forensic science [58]. Nevertheless, some laboratories were able to successfully detect skin markers in the stains composed of skin only, in addition to detecting skin markers in other stains (e.g. in vaginal secretion stains). Skin samples were found to be equally challenging to analyze in another collaborative exercise [37]. In contrast, van den Berge et al. [59] found that RNA profiling was more sensitive for skin than DNA profiling (the opposite was true for blood, menstrual blood, saliva and semen). Markers for semen were still detected in stains containing relatively small amounts of semen (about a quarter of 25 µl), although semen is one of the body fluids with relatively low levels of RNA [60]. In samples containing a mixture of body fluids, one fluid typically dominates [12,30,36]. The greater amount of RNA present in the more prominent body fluid will then predominate over the other body fluid component, particularly if the other body fluid in the mixture is generally low in RNA [30,61]. This is consistent with our findings: In cases where a mixture was analyzed with one body fluid naturally containing less RNA, such as skin, the more RNA-rich body fluid predominated (Fig. 2B).

Interpretation of RNA results can be complicated by the detection of additional markers from body fluids not expected to be present in the sample, but we did not frequently encounter the presence of such markers. In the 3rd exercise, for stain 13, a blood stain, all laboratories obtained high coverages for the majority of blood-specific markers, but several laboratories detected additional markers for saliva, semen and menstrual blood. However, the coverage of these additional markers was low. The occasional cases in which additional markers were consistently observed in the 4th exercise may be attributed to the previous sexual intercourse of the vaginal secretion donor. This observation is exemplified by stain 3 (a saliva-vaginal secretion mixture). Additional vaginal secretion and/or menstrual blood markers were sporadically observed in other mixed stains where one component was semen (stains 5 and 13) and in a semen-only stain (stain 11). Conversely, stain 13 showed low coverage for most body-fluid specific markers in several laboratories (Supplementary Material, Table S8), indicating a rather low template sample. Analysis of the same stain using the alternative cSNP panel resulted in the amplification of several markers that appear to be less specific (Supplementary Material, Table S11), making interpretation impossible.

Cross-reactivity of body fluid markers can potentially lead to complications in mRNA profiling. This has been observed in other studies for certain markers between saliva and nasal mucosa [9], and between

Table 4
Summarized BFID results using the BFID-cSNP-6F and the alternative cSNP panel on the 16 standardized stains. The presence of a question mark indicates that no interpretation in terms of which body fluid(s) are present could be made.

Stain Nr.	BF/T	BFID-6F-cSNP Assay	Alternative cSNP Panel
1	SK	?	?
2	BL-MB	BL-MB	BL-MB
3	SA-VAG	VAG-SA (SE in VAG?)	?
4	MB-SE	MB-SE	MB-SE
5	BL-SE	SE-BL	?
6	SE-SE	SE-SE	SE-SE
7	MB-SA	MB-SA	MB-SA
8	SA-SK	SA	SA-BL
9	VAG	VAG	?
10	MB	MB	MB
11	SE	SE	SE
12	BL	?	MB
13	SA-SE	?	?
14	VAG-BL	VAG	VAG-BL
15	SA	SA(?)	?
16	VAG-SE	VAG-SE	VAG-SE

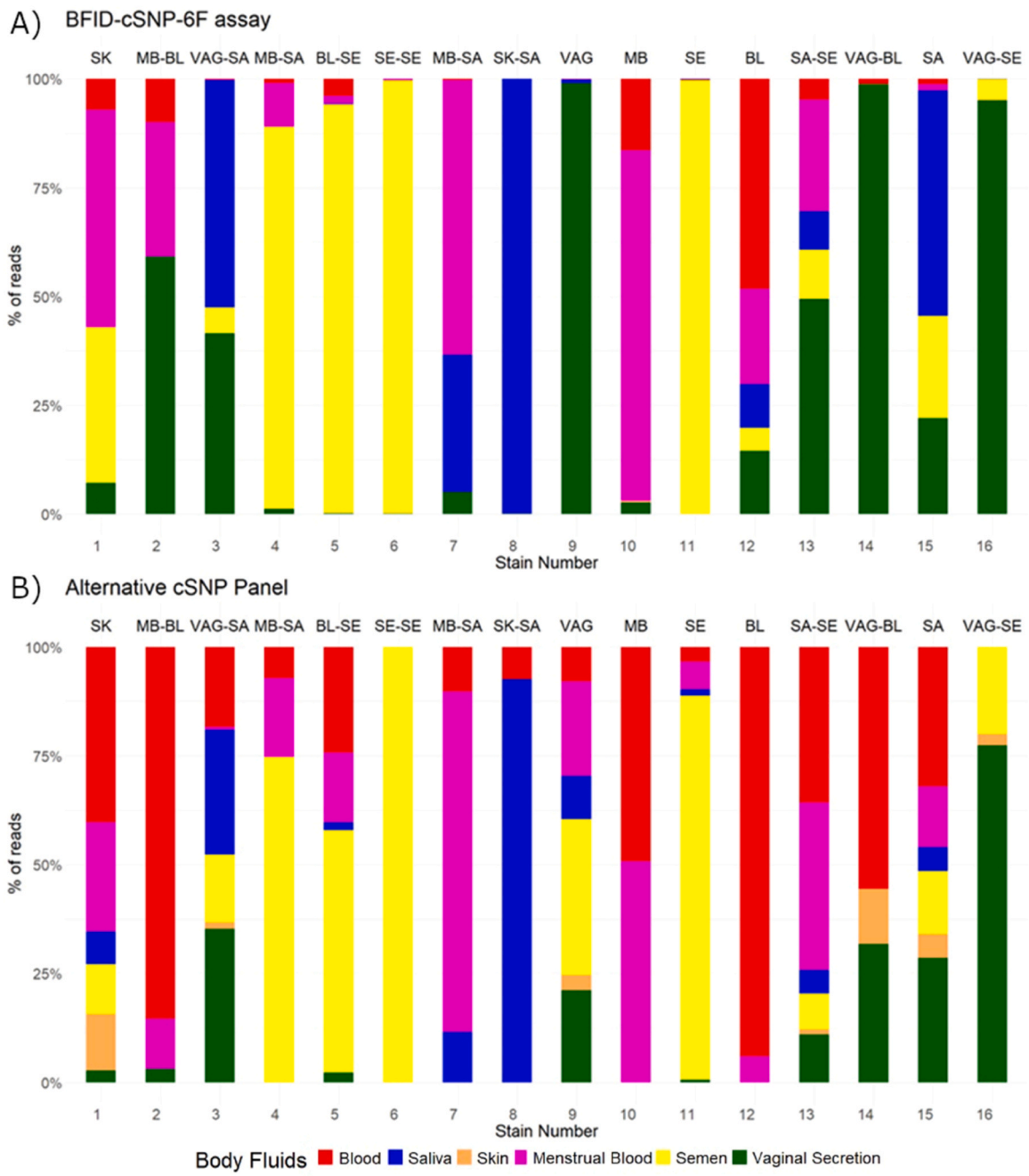


Fig. 5. Composition analysis of stains by body fluid percentages. All body fluid-specific mRNAs are summed up per body fluid.

saliva and vaginal secretion [31]. Similar observations were made for some of the stains prepared by the participating laboratories in the 4th exercise. The occurrence of spurious amplification is more prevalent when the target body fluid markers are underrepresented, and the total number of reads sequenced is low. Another issue that arose was the amplification of body fluid-specific markers from nearby sites. For instance, laboratory 7 prepared a rectal mucosa stain that was interpreted as a menstrual blood–skin mixture based on the mRNA markers. While skin markers would be expected, menstrual blood markers could have come from the vaginal tract of a female donor and therefore represent a sampling artefact. Despite the absence of notable cross-reactivity between the body fluids derived from mucous membranes in the 3rd and 4th exercise, cross-reactivity for such body fluids has been reported previously [44]. False-positive signals were obtained for vaginal secretion (*MUC4* was detected), and in nasal mucosa samples the saliva markers *HTN3* and *STATH* were detected [62,63]. Cross-reactivity seemed to be more pronounced in the alternative cSNP panel [39] than in the BFID-cSNP-6F assay. This is particularly evident between saliva markers and vaginal secretion, where saliva markers are observed at relatively high levels in several mixtures including vaginal secretion. In single body fluid stains (e.g., stains 9 and 10 from the 4th exercise), high read counts are observed not only for markers of the target body fluids, but also for markers of other body fluids. These results may be attributed to cross-reactivity of some of the markers for several body fluids or a lack of body fluid specificity. Consequently, high input samples will obtain high read counts for the majority of the markers analyzed. In addition, several studies have observed cross-reactivity of vaginal secretion and/ or menstrual blood mRNA markers when analyzing rectal mucosa samples [39,64,65]. Such cross-reactivity was also observed in the 4th exercise for the rectal

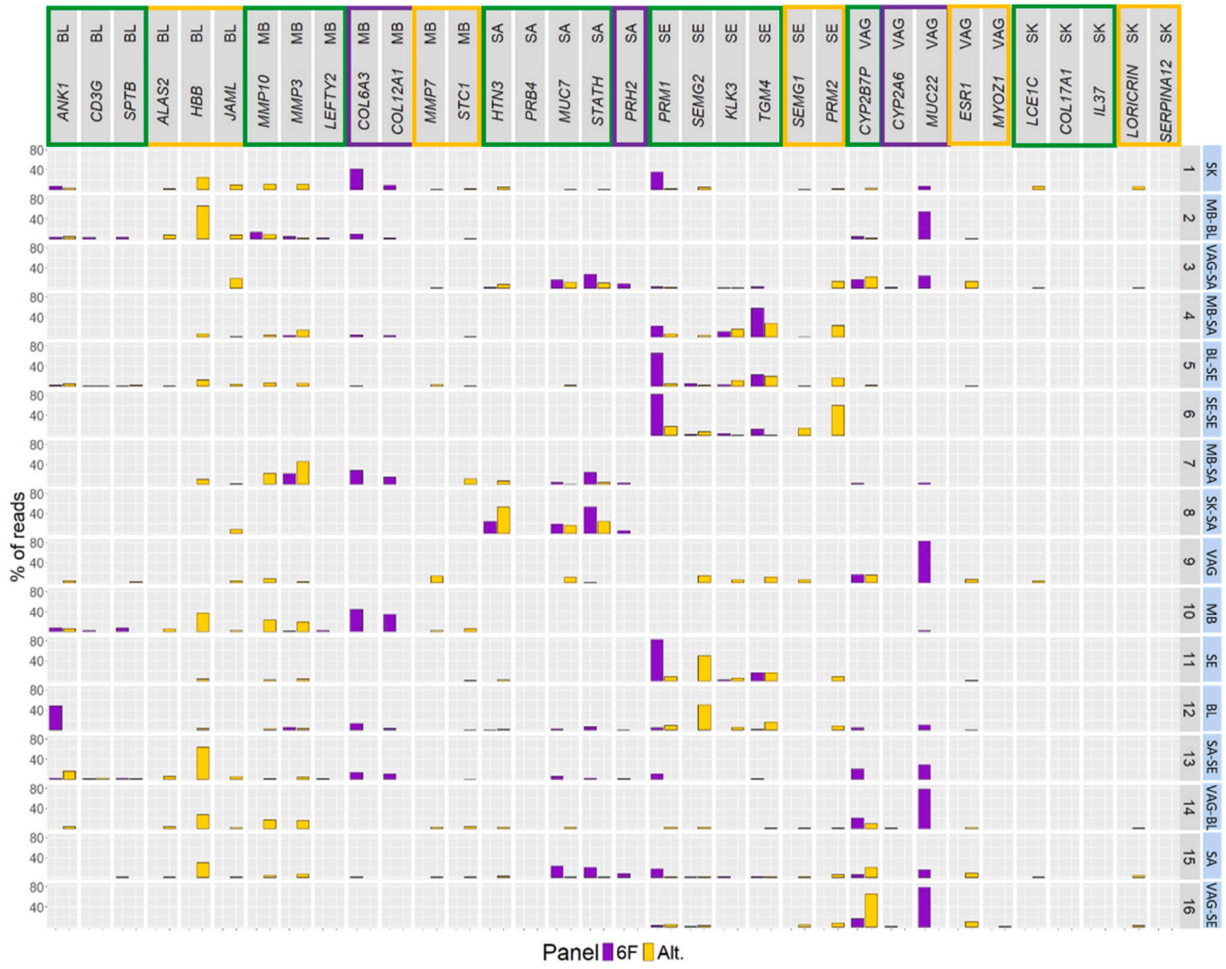


Fig. 6. Percentage of reads per marker per stain for both panels: BFID-cSNP-6F panel (6 F in purple) and alternative panel (Alt. in yellow) in the 16 standardized stains. The green boxes represent markers that are shared between the two panels, the purple boxes indicate markers specific to the 6 F panel, and the yellow boxes indicate markers specific to the alternative panel. The actual stain compositions are indicated in the light blue boxes.

mucosa samples that were included in the self-prepared stains of participating laboratories.

As previously stated, some of the standardized stains had a variety of characteristics which presented a challenge not only in predicting the body fluid but also in subsequently associating the donor. In both exercises, skin could not be identified or matched to the respective donor. The donor matching was very successful in both exercises for stains containing blood, and/or vaginal secretions, and/or menstrual blood using the BFID-cSNP-6F panel. However, the success of donor association is less pronounced for saliva and semen as single markers are sequenced at high coverage (e.g. *PRM1* in semen), which is sufficient for BFID, but not for donor association. The association of body fluids and donors is largely dependent on the number of markers with body-fluid specific cSNPs amplified in the sample. The more information on the RNA cSNP genotype of the target body fluid is available, the more accurate and potentially discriminatory the reflection of the DNA genotype will be.

We decided for both BFID and donor association to abstain from absolute thresholds. Although an absolute threshold may be a reasonable approach in the context of BFID, as demonstrated by Johannessen et al. [66] (coverage $>$ 100 reads), it is nevertheless an arbitrary measure. However, such a threshold may facilitate the interpretation of sequencing results by eliminating a greater number of artefacts. In CE-based BFID, the question of when to conclude that a body fluid is present or absent is a well-known issue. Approaches to addressing this issue include the use of a scoring method, such as the X = n/2 scoring method [67]. While this approach relies on repeated amplification to determine whether the presence of certain BFID markers is consistently observed [67], an alternative scoring system weights the markers according to their body fluid specificity [68]. Subsequently, a threshold score is established for each body fluid, which must be exceeded by summing the BFID marker-specific values. Furthermore, probabilistic models that employ quantitative data for BFID have been developed [36, 69]. These models consider the co-expression of all markers for the determination of the probability of the presence of a specific body fluid, as opposed to relying on the presence or absence of individual markers. In the future, such a probabilistic approach could potentially be used in conjunction with the BFID-cSNP-BSS or BFID-cSNP-6F assays. Establishing an absolute threshold for cSNP coverage for donor association represents a demanding task, not least because of the observation of dropouts for heterozygous cSNP positions even at high coverage levels. To ensure reliable SNP genotyping, a 30X coverage is typically recommended for DNA sequencing [70], however, considering the RNA sequencing data from both exercises, most cSNPs received a much higher coverage.

We compared the performance of the BFID-cSNP-6F assay on two different sequencing platforms. We had seven S5 participants and three MiSeq laboratories, and two laboratories did the analysis on both sequencing platforms. When comparing the results from laboratories using either platform, we found that the markers that proportionally accounted for most of the amplified targets in each stain were the same for both platforms. However, on average, fewer markers were amplified

per body fluid on the MiSeq platform. This observation was confirmed by the results of the two laboratories that analyzed the experiment on both platforms. These discrepancies may be due to differences in the sequencing chemistry, read depth, or library preparation efficiency, since each platform relies on its own kits for the respective steps of the workflow.

The subsequent retrieval of information on the stains prepared by the laboratories further demonstrated the ability of the assay to correctly identify the body fluids present (in a mixture). This highlights the importance of a frame of reference. For example, in the absence of additional information, it is a somewhat arbitrary decision whether an unknown stain showing menstrual blood and blood markers should be considered as a single body fluid stain. The number of contributors from STR profiling may be informative in such a case, but an assumption or hypothesis may guide the interpretation of the results more purposefully. Such a frame of reference is usually provided when analyzing casework samples. Moreover, the potential for false positives, particularly due to cross-reactivity or background amplification, has important legal implications. Misinterpretation of such results could have consequences, reinforcing the need for confirmatory testing and cautious reporting of results.

Mechanisms that may be responsible for RNA/DNA differences include RNA editing [71,72], transcriptional errors [73,74], alternative splicing, transcriptional noise, contamination, allele-specific expression [75], somatic mutations, copy number variations (CNVs), and low-quality or degraded RNA, to name some. RNA editing is a post-transcriptional process involving nucleotide modification, substitution, deletion, and addition in RNA molecules [71,72]. Two families of deaminases are capable of performing RNA editing, i.e. ADARs (adenosine deaminases that act on RNA) and APOBECs (apolipoprotein B mRNA editing enzymes, catalytic polypeptide-like), which produce A-to-G and C-to-U changes, respectively [76]. RNA/DNA differences, however, are not well covered in the forensic literature [77]. In a study on individual identification from low template blood samples using whole transcriptome shotgun sequencing, the authors initially compiled a set of candidate blood cSNPs for individual identification from fresh whole blood samples. The degree of cSNP concordance was then investigated in degraded blood samples, reflecting those relevant from a forensic perspective. Despite the observation of a consistent cSNP concordance, they noted a significant presence of no-call rates, mostly in the form of locus drop-outs [78]. Previously, allelic drop-out in the cSNP BFID markers *TGM4* and *CYP2A7* and sporadic allelic drop-in were observed for some RNA cSNPs [12]. While differential or unequal (potentially monoallelic) expression may be responsible for a drop-out [75,79], a drop-in could result from stochastic PCR amplification, which produces pseudo-heterozygous genotypes [12]. Significant heterozygous allele imbalances were again observed for the RNA cSNPs in the *TGM4* gene in both the BFID-cSNP-BSS and the BFID-cSNP-6F-assays [38], and the alternative panel [39] tested in this study. In the 3rd and the 4th exercise, we have observed few discordant RNA and DNA genotypes with the BFID-cSNP-BSS or BFID-cSNP-6F assays that are most likely due to low-quality and degraded RNA or the result of technical artefacts. For the alternative cSNP panel, we found occasional discordance between the RNA and DNA cSNP genotypes, i.e. RNA/DNA differences that can be traced back to gDNA contamination. This, in turn, highlights the necessity for an RNA-specific primer design, spanning exon-exon junctions, which is not the case for 36 out of 76 markers in the alternative cSNP panel.

5. Conclusion & Outlook

In this study, we have critically evaluated two mRNA cSNP assays (the BFID-cSNP-BSS and -6F assays) for their potential to identify forensically relevant body fluids and the association to their donors on mock case stains in two consecutive exercises. While body fluid identification worked well for all body fluids, with minimal cross-reactivity observed, the ability to associate these body fluids to their donors is highly dependent on the number of markers, and therefore the number of cSNPs covered. However, identifying and associating skin to its donor proved to be a significant challenge, due to the low RNA amounts. The BFID-cSNP-6F assay performed equally well on the S5™ and MiSeq™ sequencing platforms. Careful interpretation of both BFID and donor association results is needed, ideally including a frame of reference for optimal outcomes. In the 4th EDNAP exercise, the 16 mock case stains were additionally analyzed with a recently published alternative cSNP MPS assay, which also gave adequate results.

Given the relatively low RNA content in skin and generally in low template samples, it may be beneficial to explore methods to increase the sensitivity of the method. Moreover, the assay could be expanded to include new markers for other body fluids/ tissues of forensic relevance. For instance, the inclusion of rectal mucosa markers could prove valuable as they would reduce misidentification of stains due to cross-reactivity with vaginal/menstrual markers.

In summary, while the assays demonstrated a solid ability to identify body fluid(s) and in most cases donor association, particularly for body fluids with a high RNA content, certain challenges remain, especially for low-template samples like skin. Therefore, future improvements in assay sensitivity and marker coverage are warranted. This should include streamlining the interpretation of results, if the potential of these assays for broader forensic applicability is to be realized.

CRediT authorship contribution statement

Nadescha Viviane Hänggi: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis. **António Amorim:** Writing – review & editing. **Heloísa Afonso Costa:** Investigation. **Jeppe Dyrberg Andersen:** Writing – review & editing. **Marie-Louise Kampmann:** Writing – review & editing, Investigation. **Cornelius Courts:** Writing – review & editing. **Maximilian Neis:** Writing – review & editing, Investigation. **Denise Syndercombe-Court:** Writing – review & editing. **Federica Giangasparo:** Writing – review & editing, Investigation. **Ane Elida Fonneløp:** Writing – review & editing. **Helen Johannessen:** Writing – review & editing, Investigation. **Thorsten Hadrys:** Writing – review & editing. **Angelika Fürst:** Investigation. **Walther Parson:** Writing – review & editing. **Harald Niederstätter:** Writing – review & editing, Investigation. **Maja Sidstedt:** Writing – review & editing, Investigation. **Siri Aili Fagerholm:** Writing – review & editing. **Titia Sijen:** Writing – review & editing. **Margreet van den Berge:** Writing – review & editing, Investigation. **Erin Hanson:** Writing – review & editing, Methodology, Investigation, Data curation. **Jack Ballantyne:** Writing – review & editing. **Cordula Haas:** Writing – review & editing, Writing – original draft, Supervision, Project administration, 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

The authors thank Anja Ladegaard Jørgensen and Annica Gosch for laboratory assistance and the Forensic Genetics and Casework section of the Institute of Legal Medicine, Medical University of Innsbruck for technical assistance, and the anonymous volunteers who donated body fluid samples and/or skin.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fsigen.2025.103303.

References

[1] C. Haas, N. Hänggi, E. Hanson, J. Ballantyne, Body Fluid Identification by mRNA and MicroRNA. Encyclopedia of Forensic Sciences, Third Edition, Elsevier, 2023, pp. 390–401.

[2] E.K. Hanson, J. Ballantyne, RNA Profiling for the Identification of the Tissue Origin of Dried Stains, Forensic DNA Anal.: Curr. Pract. Emerg. Technol. 81 (2013).

[3] M. Bauer, D. Patzelt, Evaluation of mRNA markers for the identification of menstrual blood, J. Forensic Sci. 47 (2002) 1278–1282.

[4] J. Juusola, J. Ballantyne, Multiplex mRNA profiling for the identification of body fluids, Forensic Sci. Int. 152 (2005) 1–12, https://doi.org/10.1016/j.forsciint.2005.02.020.

[5] C. Haas, B. Klesser, C. Maake, W. Bär, A. Kratzer, mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR, Forensic Sci. Int. Genet 3 (2009) 80–88, https://doi.org/10.1016/j.fsigen.2008.11.003.

[6] R.I. Fleming, S. Harbison, The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids, Forensic Sci. Int. Genet 4 (2010) 244–256, https://doi.org/10.1016/j.fsigen.2009.10.006.

[7] A. Lindenbergh, M. de Pagter, G. Ramdayal, M. Visser, D. Zubakov, M. Kayser, T. Sijen, A multiplex (m)RNA-profiling system for the forensic identification of body fluids and contact traces, Forensic Sci. Int. Genet 6 (2012) 565–577, https://doi.org/10.1016/j.fsigen.2012.01.009.

[8] M.-H. Lin, D.F. Jones, R. Fleming, Transcriptomic analysis of degraded forensic body fluids, Forensic Sci. Int. Genet 17 (2015) 35–42, https://doi.org/10.1016/j.fsigen.2015.03.005.

[9] M. van den Berge, B. Bhoelai, J. Harteveld, A. Matai, T. Sijen, Advancing forensic RNA typing: On non-target secretions, a nasal mucosa marker, a differential co-extraction protocol and the sensitivity of DNA and RNA profiling, Forensic Sci. Int. Genet 20 (2016) 119–129, https://doi.org/10.1016/j.fsigen.2015.10.011.

[10] E. Hanson, S. Ingold, C. Haas, J. Ballantyne, Messenger RNA biomarker signatures for forensic body fluid identification revealed by targeted RNA sequencing, Forensic Sci. Int. Genet 34 (2018) 206–221, https://doi.org/10.1016/j.fsigen.2018.02.020.

[11] E. Hanson, S. Ingold, G. Dorum, C. Haas, R. Lagace, J. Ballantyne, Assigning forensic body fluids to DNA donors in mixed samples by targeted RNA/DNA deep seqeuncing of coding region SNPs using ion torrent technology, Forensic Sci. Int.: Genet. Suppl. Ser. 7 (2019) 23–24.

[12] S. Ingold, G. Dørum, E. Hanson, J. Ballantyne, C. Haas, Assigning forensic body fluids to donors in mixed body fluids by targeted RNA/DNA deep sequencing of coding region SNPs, Int. J. Leg. Med. 134 (2020) 473–485, https://doi.org/10.1007/s00414-020-02252-w.

[13] D. Frumkin, A. Wasserstrom, B. Budowle, A. Davidson, DNA methylation-based forensic tissue identification, Forensic Sci. Int. Genet 5 (2011) 517–524, https://doi.org/10.1016/j.fsigen.2010.12.001.

[14] A. Vidaki, F. Giangasparo, D. Syndercombe Court, Discovery of potential DNA methylation markers for forensic tissue identification using bisulphite pyrosequencing, Electrophoresis 37 (2016) 2767–2779, https://doi.org/10.1002/elps.201600261.

[15] E.N. Hanssen, E. Avershina, K. Rudi, P. Gill, L. Snipen, Body fluid prediction from microbial patterns for forensic application, Forensic Sci. Int. Genet 30 (2017) 10–17, https://doi.org/10.1016/j.fsigen.2017.05.009.

[16] A. Dobay, C. Haas, G. Fucile, N. Downey, H.G. Morrison, A. Kratzer, N. Arora, Microbiome-based body fluid identification of samples exposed to indoor conditions, Forensic Sci. Int. Genet 40 (2019) 105–113, https://doi.org/10.1016/j.fsigen.2019.02.010.

[17] C. Díez López, D. Montiel González, C. Haas, A. Vidaki, M. Kayser, Microbiome-based body site of origin classification of forensically relevant blood traces, Forensic Sci. Int. Genet 47 (2020) 102280, https://doi.org/10.1016/j.fsigen.2020.102280.

[18] H. Yang, B. Zhou, H. Deng, M. Prinz, D. Siegel, Body fluid identification by mass spectrometry, Int J. Leg. Med 127 (2013) 1065–1077, https://doi.org/10.1007/s00414-013-0848-1.

[19] T.H. Clarke, A. Gomez, H. Singh, K.E. Nelson, L.M. Brinkac, Integrating the microbiome as a resource in the forensics toolkit, Forensic Sci. Int. Genet 30 (2017) 141–147, https://doi.org/10.1016/j.fsigen.2017.06.008.

[20] R. Stark, M. Grzelak, J. Hadfield, RNA sequencing: the teenage years, Nat. Rev. Genet 20 (2019) 631–656, https://doi.org/10.1038/s41576-019-0150-2.

[21] R. Hrdlickova, M. Toloue, B. Tian, RNA-Seq methods for transcriptome analysis, Wiley Interdiscip. Rev. RNA 8 (2017) e1364, https://doi.org/10.1002/wrna.1364.

[22] S. Zhao, Y. Zhang, R. Gamini, B. Zhang, D. von Schack, Evaluation of two main RNA-seq approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus rRNA depletion, Sci. Rep. 8 (2018) 4781, https://doi.org/10.1038/s41598-018-23226-4.

[23] M. Vennemann, A. Koppelkamm, mRNA profiling in forensic genetics I: Possibilities and limitations, Forensic Sci. Int. 203 (2010) 71–75, https://doi.org/10.1016/j.forsciint.2010.07.006.

[24] S.R. Head, H.K. Komori, S.A. LaMere, T. Whisenant, F. van Nieuwerburgh, D. R. Salomon, P. Ordoukhanian, Library construction for next-generation sequencing: overviews and challenges, 66, 68, passim, Biotechniques 56 (2014) 61–64, https://doi.org/10.2144/000114133.

[25] M. Morey, A. Fernández-Marmiesse, D. Castiñeiras, J.M. Fraga, M.L. Couce, J. A. Cocho, A glimpse into past, present, and future DNA sequencing, Mol. Genet. Metab. 110 (2013) 3–24, https://doi.org/10.1016/j.ymgme.2013.04.024.

[26] M.L. Metzker, Sequencing technologies - the next generation, Nat. Rev. Genet 11 (2010) 31–46, https://doi.org/10.1038/nrg2626.

[27] A.E. Escalante, L. Jardón Barbolla, S. Ramírez-Barahona, L.E. Eguiarte, The study of biodiversity in the era of massive sequencing, Rev. Mex. De. Biodivers. 85 (2014) 1249–1264, https://doi.org/10.7550/rmb.43498.

[28] M.O. Pollard, D. Gurdasani, A.J. Mentzer, T. Porter, M.S. Sandhu, Long reads: their purpose and place, Hum. Mol. Genet. 27 (2018) R234–R241, https://doi.org/10.1093/hmg/ddy177.

[29] T. Sijen, Molecular approaches for forensic cell type identification: On mRNA, miRNA, DNA methylation and microbial markers, Forensic Sci. Int. Genet 18 (2015) 21–32, https://doi.org/10.1016/j.fsigen.2014.11.015.

[30] S. Ingold, G. Dørum, E. Hanson, A. Berti, W. Branicki, P. Brito, P. Elsmore, K. B. Gettings, F. Giangasparo, T.E. Gross, S. Hansen, E.N. Hanssen, M.-L. Kampmann, M. Kayser, F.-X. Laurent, N. Morling, A. Mosquera-Miguel, W. Parson, C. Phillips, M.J. Porto, E. Pośpiech, A.D. Roeder, P.M. Schneider, K. Schulze Johann, C. R. Steffen, D. Syndercombe-Court, M. Trautmann, M. van den Berge, K.J. van der Gaag, J. Vannier, V. Verdoliva, A. Vidaki, C. Xavier, J. Ballantyne, C. Haas, Body fluid identification using a targeted mRNA massively parallel sequencing approach - results of a EUROFORGEN/EDNAP collaborative exercise, Forensic Sci. Int. Genet 34 (2018) 105–115, https://doi.org/10.1016/j.fsigen.2018.01.002.

[31] C. Haas, E. Hanson, M.J. Anjos, R. Banemann, A. Berti, E. Borges, A. Carracedo, M. Carvalho, C. Courts, G. de Cock, M. Dötsch, S. Flynn, I. Gomes, C. Hollard, B. Hjort, P. Hoff-Olsen, K. Hríbiková, A. Lindenbergh, B. Ludes, O. Maroñas, N. McCallum, D. Moore, N. Morling, H. Niederstätter, F. Noel, W. Parson, C. Popielarz, C. Rapone, A.D. Roeder, Y. Ruiz, E. Sauer, P.M. Schneider, T. Sijen, D. S. Court, B. Sviežená, M. Turanská, A. Vidaki, L. Zatkalíková, J. Ballantyne, RNA/DNA co-analysis from human saliva and semen stains–results of a third collaborative EDNAP exercise, Forensic Sci. Int. Genet 7 (2013) 230–239, https://doi.org/10.1016/j.fsigen.2012.10.011.

[32] C. Haas, E. Hanson, R. Banemann, A.M. Bento, A. Berti, Á. Carracedo, C. Courts, G. de Cock, K. Drobnic, R. Fleming, C. Franchi, I. Gomes, G. Hadzic, S.A. Harbison, B. Hjort, C. Hollard, P. Hoff-Olsen, C. Keyser, A. Kondili, O. Maroñas, N. McCallum, P. Miniati, N. Morling, H. Niederstätter, F. Noël, W. Parson, M.J. Porto, A. D. Roeder, E. Sauer, P.M. Schneider, G. Shanthan, T. Sijen, D. Syndercombe Court, M. Turanská, M. van den Berge, M. Vennemann, A. Vidaki, L. Zatkalíková, J. Ballantyne, RNA/DNA co-analysis from human skin and contact traces–results of a sixth collaborative EDNAP exercise, Forensic Sci. Int. Genet 16 (2015) 139–147, https://doi.org/10.1016/j.fsigen.2015.01.002.

[33] C. Haas, E. Hanson, W. Bär, R. Banemann, A.M. Bento, A. Berti, E. Borges, C. Bouakaze, A. Carracedo, M. Carvalho, A. Choma, M. Dötsch, M. Durianciková, P. Hoff-Olsen, C. Hohoff, P. Johansen, P.A. Lindenbergh, B. Loddenkötter, B. Ludes, O. Maroñas, N. Morling, H. Niederstätter, W. Parson, G. Patel, C. Popielarz, E. Salata, P.M. Schneider, T. Sijen, B. Sviezená, L. Zatkalíková, J. Ballantyne, mRNA profiling for the identification of blood–results of a collaborative EDNAP exercise, Forensic Sci. Int. Genet 5 (2011) 21–26, https://doi.org/10.1016/j.fsigen.2010.01.003.

[34] M. Bauer, D. Patzelt, A method for simultaneous RNA and DNA isolation from dried blood and semen stains, Forensic Sci. Int. 136 (2003) 76–78, https://doi.org/10.1016/S0379-0738(03)00219-6.

[35] M. Alvarez, J. Juusola, J. Ballantyne, An mRNA and DNA co-isolation method for forensic casework samples, Anal. Biochem. 335 (2004) 289–298, https://doi.org/10.1016/j.ab.2004.09.002.

[36] G. Dørum, S. Ingold, E. Hanson, J. Ballantyne, L. Snipen, C. Haas, Predicting the origin of stains from next generation sequencing mRNA data, Forensic Sci. Int. Genet 34 (2018) 37–48, https://doi.org/10.1016/j.fsigen.2018.01.001.

[37] A.P. Salzmann, M. Bamberg, C. Courts, G. Dørum, A. Gosch, T. Hadrys, G. Hadzic, M. Neis, P.M. Schneider, T. Sijen, M. den van Berge, P. Wiegand, C. Haas, mRNA profiling of mock casework samples: Results of a FoRNAP collaborative exercise, Forensic Sci. Int. Genet 50 (2021) 102409, https://doi.org/10.1016/j.fsigen.2020.102409.

[38] E. Hanson, G. Dørum, M. Zamborlin, S. Wang, M. Gysi, S. Ingold, R. Lagace, C. Roth, C. Haas, J. Ballantyne, Targeted S5 RNA sequencing assay for the identification and direct association of common body fluids with DNA donors in mixtures, Int J. Leg. Med 137 (2023) 13–32, https://doi.org/10.1007/s00414-022-02908-9.

[39] M. Neis, T. Groß, H. Schneider, P.M. Schneider, C. Courts, Comprehensive body fluid identification and contributor assignment by combining targeted sequencing of mRNA and coding region SNPs, 1872-4973 73 (2024) 103125. ⟨https://doi.org/10.1016/j.fsigen.2024.103125⟩.

[40] C. Forsberg, L. Jansson, R. Ansell, J. Hedman, High-throughput DNA extraction of forensic adhesive tapes, Forensic Sci. Int. Genet 24 (2016) 158–163, https://doi.org/10.1016/j.fsigen.2016.06.004.

[41] T. Magoč, S.L. Salzberg, FLASH: fast length adjustment of short reads to improve genome assemblies, Bioinformatics 27 (2011) 2957–2963, https://doi.org/10.1093/bioinformatics/btr507.

[42] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (2009) 1754–1760, https://doi.org/10.1093/bioinformatics/btp324.

[43] S. Wang, G. Shanthan, M.M. Bouzga, H.M. Thi Dinh, C. Haas, A.E. Fonneløp, Evaluating the performance of five up-to-date DNA/RNA co-extraction methods for forensic application, Forensic Sci. Int. 328 (2021) 110996, https://doi.org/10.1016/j.forsciint.2021.110996.

[44] M. van den Berge, A. Carracedo, I. Gomes, E.A.M. Graham, C. Haas, B. Hjort, P. Hoff-Olsen, O. Maroñas, B. Mevåg, N. Morling, H. Niederstätter, W. Parson, P. M. Schneider, D.S. Court, A. Vidaki, T. Sijen, A collaborative European exercise on mRNA-based body fluid/skin typing and interpretation of DNA and RNA results, Forensic Sci. Int. Genet 10 (2014) 40–48, https://doi.org/10.1016/j.fsigen.2014.01.006.

[45] E.J. Lee, T.D. Schmittgen, Comparison of RNA assay methods used to normalize cDNA for quantitative real-time PCR, Anal. Biochem. 357 (2006) 299–301, https://doi.org/10.1016/j.ab.2006.06.011.
[46] C.A. Lewis, T.R. Layne, S.J. Seashols-Williams, Detection of microRNAs in DNA Extractions for Forensic Biological Source Identification, J. Forensic Sci. 64 (2019) 1823–1830, https://doi.org/10.1111/1556-4029.14070.
[47] N. Gao, Y. Huo, D. Yu, F. Cheng, T. Wang, X. Zhang, L. Zhang, W. Hu, J. Li, P. Yuan, J. Liu, Y. Wang, J. Yan, Evaluation of reverse transcription yield of RNA standards and forensic samples based on droplet digital PCR, Biochem. Biophys. Res. Commun. 711 (2024) 149909, https://doi.org/10.1016/j.bbrc.2024.149909.
[48] S.A. Bustin, Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems, J. Mol. Endocrinol. 29 (2002) 23–39.
[49] W.M. Freeman, S.J. Walker, K.E. Vrana, Quantitative RT-PCR: pitfalls and potential, 124-5, Biotechniques 26 (1999) 112–122, https://doi.org/10.2144/99261rv01.
[50] D.P. Chandler, C.A. Wagnon, H. Bolton Jr, Reverse transcriptase (RT) inhibition of PCR at low concentrations of template and its implications for quantitative RT-PCR, Appl. Environ. Microbiol. 64 (1998) 669–677.
[51] A. Ståhlberg, J. Håkansson, X. Xian, H. Semb, M. Kubista, Properties of the reverse transcription reaction in mRNA quantification, Clin. Chem. 50 (2004) 509–515, https://doi.org/10.1373/clinchem.2003.026161.
[52] A. Ståhlberg, M. Kubista, M. Pfaffl, Comparison of reverse transcriptases in gene expression analysis, Clin. Chem. 50 (2004) 1678–1680, https://doi.org/10.1373/clinchem.2004.035469.
[53] C. Cossu, U. Germann, A. Kratzer, W. Bär, C. Haas, How specific are the vaginal secretion mRNA-markers HBD1 and MUC4? Forensic Sci. Int.: Genet. Suppl. Ser. 2 (2009) 536–537, https://doi.org/10.1016/j.fsigss.2009.08.063.
[54] C. Haas, E. Hanson, A. Kratzer, W. Bär, J. Ballantyne, Selection of highly specific and sensitive mRNA biomarkers for the identification of blood, Forensic Sci. Int. Genet 5 (2011) 449–458, https://doi.org/10.1016/j.fsigen.2010.09.006.
[55] J. Juusola, J. Ballantyne, mRNA profiling for body fluid identification by multiplex quantitative RT-PCR, J. Forensic Sci. 52 (2007) 1252–1262, https://doi.org/10.1111/j.1556-4029.2007.00550.x.
[56] D. Zhao, X. Chen, Z. An, E. Hanson, J. Ballantyne, Cytochrome c oxidase subunit 1-based human RNA quantification to enhance mRNA profiling in forensic biology, J. Forensic Sci. Med 3 (2017) 115, https://doi.org/10.4103/jfsm.jfsm_63_17.
[57] K. Grisedale, A. van Daal, Linear amplification of target prior to PCR for improved low template DNA results, Biotechniques 56 (2014) 145–147, https://doi.org/10.2144/000114148.
[58] E. Reimann, K. Abram, S. Kõks, K. Kingo, A. Fazeli, Identification of an optimal method for extracting RNA from human skin biopsy, using domestic pig as a model system, Sci. Rep. 9 (2019) 20111, https://doi.org/10.1038/s41598-019-56579-5.
[59] M. van den Berge, G. Ozcanhan, S. Zijlstra, A. Lindenbergh, T. Sijen, Prevalence of human cell material: DNA and RNA profiling of public and private objects and after activity scenarios, Forensic Sci. Int. Genet 21 (2016) 81–89, https://doi.org/10.1016/j.fsigen.2015.12.012.
[60] A. Schuster, C. Tang, Y. Xie, N. Ortogero, S. Yuan, W. Yan, SpermBase: A Database for Sperm-Borne RNA Contents, Biol. Reprod. 95 (2016) 99, https://doi.org/10.1095/biolreprod.116.142190.
[61] S. Ingold, G. Dørum, E. Hanson, D. Ballard, A. Berti, K.B. Gettings, F. Giangasparo, M.-L. Kampmann, F.-X. Laurent, N. Morling, W. Parson, C.R. Steffen, A. Ulus, M. van den Berge, K.J. van der Gaag, V. Verdoliva, C. Xavier, J. Ballantyne, C. Haas, Body fluid identification and assignment to donors using a targeted mRNA massively parallel sequencing approach - results of a second EUROFORGEN / EDNAP collaborative exercise, Forensic Sci. Int. Genet 45 (2020) 102208, https://doi.org/10.1016/j.fsigen.2019.102208.
[62] 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 (2014) e100123, https://doi.org/10.1371/journal.pone.0100123.
[63] K. Sakurada, T. Akutsu, K. Watanabe, Y. Fujinami, M. Yoshino, Expression of statherin mRNA and protein in nasal and vaginal secretions, Leg. Med. (Tokyo) 13 (2011) 309–313, https://doi.org/10.1016/j.legalmed.2011.07.002.
[64] M. Bamberg, L. Dierig, G. Kulstein, S.N. Kunz, M. Schmidt, T. Hadrys, P. Wiegand, Development and validation of an mRNA-based multiplex body fluid identification workflow and a rectal mucosa marker pilot study, Forensic Sci. Int. Genet 54 (2021) 102542, https://doi.org/10.1016/j.fsigen.2021.102542.
[65] T. Akutsu, H. Saito, K. Watanabe, K. Toyomane, T. Yamagishi, H. Iwase, Evaluation and simultaneous determination of rectal mucosa markers by multiplex reverse transcription-PCR for biological evidence of sexual assault with anal penetration, Forensic Sci. Int. Genet 59 (2022) 102712, https://doi.org/10.1016/j.fsigen.2022.102712.
[66] H. Johannessen, E. Hanson, P. Gill, C. Haas, E.F. Bergseth, J. Ballantyne, A. E. Fonneløp, Body Fluid Identification in Samples Collected after Intimate and Social Contact: A Comparison of Two mRNA Profiling Methods and the Additional Information Gained by cSNP Genotypes, Genes 14 (2023) 636, https://doi.org/10.3390/genes14030636.
[67] A. Lindenbergh, P. Maaskant, T. Sijen, Implementation of RNA profiling in forensic casework, Forensic Sci. Int.: Genet. 7 (2013) 159–166.
[68] A.D. Roeder, C. Haas, mRNA profiling using a minimum of five mRNA markers per body fluid and a novel scoring method for body fluid identification, Int J. Leg. Med 127 (2013) 707–721, https://doi.org/10.1007/s00414-012-0794-3.
[69] D. Iacob, A. Fürst, T. Hadrys, A machine learning model to predict the origin of forensically relevant body fluids, Forensic Sci. Int.: Genet. Suppl. Ser. 7 (2019) 392–394, https://doi.org/10.1016/j.fsigss.2019.10.025.
[70] Illumina, Calling Sequencing SNPs.
[71] W. Keller, J. Wolf, A. Gerber, Editing of messenger RNA precursors and of tRNAs by adenosine to inosine conversion, 0014-5793 452 (1999) 71–76. ⟨https://doi.org/10.1016/S0014-5793(99)00590-6⟩.
[72] D.B. Stern, M. Goldschmidt-Clermont, M.R. Hanson, Chloroplast RNA metabolism, Annu. Rev. Plant Biol. 61 (2010) 125–155, https://doi.org/10.1146/annurev-arplant-042809-112242.
[73] R.T. Libby, J.A. Gallant, The role of RNA polymerase in transcriptional fidelity, Mol. Microbiol. 5 (1991) 999–1004, https://doi.org/10.1111/j.1365-2958.1991.tb01872.x.
[74] J.F. Sydow, P. Cramer, RNA polymerase fidelity and transcriptional proofreading, Curr. Opin. Struct. Biol. 19 (2009) 732–739, https://doi.org/10.1016/j.sbi.2009.10.009.
[75] P.R. Buckland, Allele-specific gene expression differences in humans, Hum. Mol. Genet 13 (Spec 2) (2004) R255–R260, https://doi.org/10.1093/hmg/ddh227.
[76] M. Li, I.X. Wang, V.G. Cheung, Response to comments on "Widespread RNA and DNA sequence differences in the human transcriptome", Science 335 (2012) 1302.
[77] S. Wang, Z. Wang, R. Tao, F. Song, Y. Hou, Validating the consistency of cSNPs analysis results between DNA and RNA using SNaPshot method, Forensic Sci. Int.: Genet. Suppl. Ser. 7 (2019) 76–78, https://doi.org/10.1016/j.fsigss.2019.09.030.
[78] A.H. Jepsen, M.-L. Kampmann, S.B. Jacobsen, C. Børsting, J.D. Andersen, Identification of individuals from low template blood samples using whole transcriptome shotgun sequencing, Forensic Sci. Int. Genet 72 (2024) 103089, https://doi.org/10.1016/j.fsigen.2024.103089.
[79] A. Gimelbrant, J.N. Hutchinson, B.R. Thompson, A. Chess, Widespread monoallelic expression on human autosomes, Science 318 (2007) 1136–1140, https://doi.org/10.1126/science.1148910.