Science & Justice 59 (2019) 217–227

Contents lists available at ScienceDirect

Science & Justice

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

Developmental validation of an enhanced mRNA-based multiplex system for body fluid and cell type identification

Patricia Pearl Albani[a,b], Rachel Fleming[a,*]

[a] *Institute of Environmental Science and Research Limited, Auckland, New Zealand*
[b] *School of Chemical Sciences, The University of Auckland, Auckland, New Zealand*

ARTICLE INFO

Keywords:
Body fluid identification
Forensic
mRNA
Multiplex
PCR
Validation

1. Introduction

In forensic genetics, scientific knowledge regarding genes and hereditary characteristics that confer inter- and intra-individual variation is applied to the resolution of legal conflicts. The highest priority is often allocated to answering sub-source level questions, i.e. whether a person of interest is the source of or a contributor to a DNA profile [1]. In some cases, drawing source and activity level inferences from the presence or absence of bodily fluids may provide crucial additional information to an investigation. Conventional methods have been employed for decades to determine the somatic origin of a DNA sample. However, these often possess limited applicability. Newer methods for body fluid identification have focused on nucleic acid analysis, including messenger RNA (mRNA) [1–4], microRNA [5], microbial RNA [6], and DNA methylation profiling [7]. mRNA analysis complements conventional body fluid testing due to its high sensitivity, specificity, and an extensive range of well-characterised markers [8]. Co-extracting RNA and DNA from the same sample allows for combined interpretation of genotyping and body fluid profiling results in simple mixtures (e.g., semen and vaginal epithelia).

The most commonly employed mRNA-based technique for body fluid typing to date is the *reverse transcriptase* polymerase chain reaction (RT-PCR). Laborious, sample-consuming uniplex reactions have been replaced by multiplexed systems that assess multiple markers of interest simultaneously. Such assays typically detect and discriminate between circulatory blood, saliva/buccal mucosa, semen (containing spermatozoa), seminal fluid (azoospermic), menstrual fluid (MF), and vaginal material (VM) [2,4]. Some also target nasal or general mucosa and skin cells [1,9,10].

The body fluid markers to be included in multiplexed RT-PCR assays were commonly selected from literature or database searches [2,3,11–14], or microarray studies [15–17]. Recent advances using whole transcriptome sequencing (RNA-Seq) have enabled new possibilities in the search for novel, possibly unknown differentially expressed mRNAs [18–20]. Previously, we reported the characterisation of new body fluid markers, partly identified from RNA-Seq [21]. Using a combination of these and well-known body fluid markers, a set of three new mRNA-based endpoint RT-PCR multiplex assays for the detection and identification of forensically relevant body fluids was developed, optimised, and validated. This set included a duplex for the detection of saliva/buccal mucosa, a quadruplex for the detection of circulatory blood and menstrual fluid, and a pentaplex for the detection of semen, seminal fluid, and vaginal material. The performance of the optimised system was assessed by conducting developmental validation studies in reference to SWGDAM guidelines [22]. These evaluated the specificity, sensitivity, and precision of each assay, as well as the ability to identify mixtures, aged, and case-type samples.

2. Materials and methods

2.1. Sample collection

Human bodily samples were obtained from healthy volunteers with

---

[*] Corresponding author at: Institute of Environmental Science and Research Ltd, Mt Albert Science Centre, Hampstead Road, Private Bag 92021, Auckland 1025, New Zealand.

E-mail address: rachel.fleming@esr.cri.nz (R. Fleming).

https://doi.org/10.1016/j.scijus.2019.01.001
Received 7 August 2018; Received in revised form 16 December 2018; Accepted 20 January 2019

full informed consent. Samples for specificity testing included circulatory blood, liquid saliva, semen (containing spermatozoa), azoospermic seminal fluid, menstrual fluid, and vaginal material for RNA, as well as blood from a male individual for DNA. Donors were between 24 and 53 years of age and included males and females for circulatory blood and saliva. Blood was placed on sterile Cultiplast® rayon swabs (LP Italiana SPA, Milano, Italy) in aliquots between 5.0 and 0.05 μL. Saliva and semen were deposited on swabs in aliquots of 10.0–0.25 μL, and 2.0–0.25 μL, respectively. Semen donors included two azoospermic individuals (vasectomised). MF and VM were obtained by volunteers themselves using swabs provided for them. Volunteers donating semen, MF, or VM were asked to abstain from sexual intercourse for one week prior to sample collection. A total of 120 human samples was collected for this study.

Mixtures of body fluids were prepared by adding increasing volumes of blood or semen (1 μL, 5 μL, and 10 μL) to 1/3 of a MF swab. Likewise, 1 μL, 5 μL, or 10 μL saliva was added to 1/3 of a VM swab, as well as to 2 μL semen placed on a swab. Finally, 2 μL semen and 10 μL saliva were added to a VM swab. All samples were prepared in duplicate, except for mixtures of MF and semen.

For the sensitivity study, decreasing volumes of circulatory blood (2.5–0.05 μL), saliva (5.0–0.25 μL), semen, and seminal fluid (both 1.0–0.05 μL) were extracted. Decreasing RNA concentrations (RNA stock sample extracted from one whole swab) were reverse transcribed for MF and VM. All samples were reverse transcribed using 10 μL and 1 μL RNA, each prepared in duplicate.

For the species specificity testing, circulatory blood and saliva were collected at a suitable opportunity (e.g. during a necessary health check or surgery) from 24 species, including, among others, primates, monkeys, birds, cat, chicken, dog, guinea pig, otter, rabbit, sheep, and wallaby. Samples were kindly supplied by pet owners, veterinarians, and Auckland Zoo staff. A total of 41 samples (20 circulatory blood and 21 saliva/buccal mucosa) were obtained from 24 different species. DNA fractions were retained from all species.

2.2. DNA/RNA co-extraction and RNA purification

DNA/RNA co-extractions were carried out as described previously [21] using the Promega® DNA IQ™ System (Promega Corporation, Madison, WI, USA), following the manufacturer's instructions. DNA was eluted in 50 μL elution buffer.

Crude RNA lysates were further processed using the ReliaPrep™ RNA Cell Miniprep System (Promega) as published [21]. RNA was eluted in 45 μL nuclease-free water. Purified RNA samples were immediately DNase treated using the TURBO DNA*free*™ Kit (Ambion®). The manufacturer's instructions were followed, adding 4.5 μL 10× TURBO DNase Buffer and 2 μL TURBO™ DNase to each sample.

2.3. Quantification of RNA and DNA samples

The absence of genomic DNA in RNA samples of human origin was verified using the Quantifiler® Human DNA Quantification Kit (Applied Biosystems®) as described in [21]. If residual genomic DNA was detected in an RNA sample, the extract was again DNase treated and requantified. This was repeated (no more than three times) until no human genomic DNA was detectable in both quantification duplicates of the same sample.

The DNA concentration of the human body fluid sample was also determined via use of the Quantifiler® System as described above. Animal DNA was quantified using the Qubit® 2.0 Fluorometer and Qubit® dsDNA High Sensitivity Assay Kit (Molecular Probes® by Life Technologies, Inc.). Reactions were performed according to the manufacturer's instructions using 2 μL of each sample.

2.4. Reverse transcription of RNA samples

DNA-free RNA samples (10 μL or 1 μL) were reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems®) according to the manufacturer's instructions. Each reaction comprised a total volume of 20 μL.

2.5. Selection of body fluid markers

Using a candidate mRNA approach, potential body fluid markers were evaluated for their ability to detect and discriminate between forensically relevant body fluids. These included circulatory blood, saliva/buccal mucosa, semen containing spermatozoa, seminal fluid (azoospermic), MF, and VM. A total of 38 candidate mRNAs were selected using RNA-Seq data analysis from previous work [23], a literature review [2,15,19,24,25], and database searches [26–30].

A comprehensive evaluation process with respect to specificity and sensitivity identified the 11 most informative body fluid markers within this set. These were *haemoglobin delta* (*HBD*) [21] and *solute carrier family 4* (*anion exchanger*), *member 1* (*Diego blood group*) (*SLC4A1*) [21] for circulatory blood, *histatin 3* (*HTN3*) [3] and *follicular dendritic cell-secreted protein* (*FDCSP*) [15] for saliva/buccal mucosa, *transition protein 1* (*TNP1*) [21] and *protamine 1* (*PRM1*) [24] for spermatozoa, *microseminoprotein beta* (*MSMB*) [15] and *kallikrein-related peptidase 2* (*KLK2*) [21] for seminal fluid, *matrix metallopeptidase 10* (*MMP10*) [14] and *stanniocalcin 1* (*STC1*) [21] for menstrual fluid, as well as *cytochrome P450 family 2*, *subfamily B*, *member 7*, *pseudogene* (*CYP2B7P*) [19] for vaginal material. Thus, each body fluid or cell type (except vaginal material) was targeted by two independent mRNAs. Detecting multiple markers per body fluid can strengthen the confidence in profiling results [8,13]. However, for vaginal material, only one suitable mRNA marker emerged from the candidate screening.

The development refrained from including constitutive transcripts (so-called housekeeping genes) in the multiplex system. Research has shown that their use is inadequate for normalisation of PCR results due to high variability among donors and body fluids [31]. This corresponds with our experience in the application of CellTyper, the multiplex previously used in our laboratory [2], which targeted three constitutive genes: either or all three mRNAs may remain undetected when body fluid markers are detected, particularly in saliva samples.

2.6. Primer and multiplex design

Primers for *HBD*, *SLC4A1*, *FDCSP*, *HTN3*, *MMP10*, *STC1*, and *CYP2B7P* were designed to target transcript stable regions (StaRs) [18] using the OligoAnalyzer 3.1 online tool (Integrated DNA Technologies, Inc., Coralville, IA, USA). Sequencing coverage maps were viewed in Geneious v.5.6.7 (Biomatters Ltd., Auckland, New Zealand) and regions of high read coverage were selected for primer design. Primers for *TNP1*, *KLK2*, and *MSMB* were designed using conventional primer design strategy, whereas primers for *PRM1* were adopted from the literature [24]. The specificity of all primers to their intended mRNA target was verified using Primer-BLAST (National Center for Biotechnology Information, U.S. National Library of Medicine, Bethesda, MD, USA).

Primers were compiled into three multiplex assays: 1) a duplex combining *FDCSP* and *HTN3* (multiplex D), 2) a quadruplex including *HBD*, *SLC4A1*, *MMP10*, and *STC1* (multiplex Q), and 3) a pentaplex combining *PRM1*, *TNP1*, *KLK2*, *MSMB*, and *CYP2B7P* (multiplex P) for the identification of saliva, blood, and sexual body fluids. The main advantage of this approach is to maintain individual efficiency of target detection compared to one large multiplex. This is due to reduced competition among primers and avoidance of potentially suboptimal PCR conditions for some markers.

Optimised primer concentrations were as follows: 1) 0.05 μM *FDCSP* and 0.012 μM *HTN3*, 2) 0.05 μM *HBD*, 0.04 μM *SLC4A1*, 0.04 μM

Table 1
Primer sequences and expected amplicon sizes of all markers included in the three multiplex assays.

Target body fluid	mRNA marker	Accession no.	Locus	mRNA length [bp]	Primer concentration [μM]	Primer sequence (5′-3′)	Amplicon size	Ref.
Saliva/buccal mucosa	*FDCSP*	NM_152997.3	4q13	566	0.05	F: HEX-CTCTCAAGACCAGGAACGAGAA R: GGGCAGATTCAGGTATTGGAATAG	170	[15]
Saliva/buccal mucosa	*HTN3*	NM_000200.2	4q13	775	0.012	F: HEX-AAGCATCATTCACATCGAGGCTAT R: ATGCGGTATGACAAATGAGAATACAC	138	[3]
Circulatory blood	*HBD*	NM_000519.3	11p15	785	0.05	F: ACTGCTGTCAATGCCCTGTG R: FAM-ACCTTCTTGCCATGAGCCTT	176	[21]
Circulatory blood	*SLC4A1*	NM_000342.3	17q21	4965	0.04	F: HEX-AACTGGACACTCAGGACCAC R: GGATGTCTGGGTCTTCATATTCCT	102	[21]
Menstrual fluid	*MMP10*	NM_002425.2	11q22	1758	0.04	F: HEX-CCCACTCTACAACTCATTCACAGAG R: GGTTCCTCAGTAGAGGCAGG	108	[14]
Menstrual fluid	*STC1*	NM_003155.2	8p21	3883	0.02	F: FAM-CTGCCCAATCACTTCTCCAACA R: TTTCTCCATCAGGCTGTCTCT	105	[21]
Spermatozoa	*PRM1*	NM_002761.2	16p13	424	0.03	F: HEX-GCCAGGTACAGATGCTGTCGCAG R: GTGTCTTCTACATCTCGGTCTG	150	[24]
Spermatozoa	*TNP1*	NM_003284.3	2q35-36	407	0.04	F: GATGACGCCAATCGCAATTACC R: FAM-CCTTCTGCTGTTCTTGTTGCTG	102	[21]
Seminal fluid	*KLK2*	NM_005551.4	19q13	3019	0.14	F: TTCTCTCCATCGCCTTGTCTG R: HEX-AGTGTGCCCATCCATGACTG	135	[21]
Seminal fluid	*MSMB*	NM_002443.3	10q11	488	0.03	F: CTTTGCCACCTTCGTGACTTTATG R: FAM-ACAGTTGTCAGTCTGCCACT	142	[15]
Vaginal material	*CYP2B7P*	NR_001278.1	19q13 *pseudogene*	2956	0.02	F: CCGTGAGATTCAGAGATTTGCTGAC R: HEX-TGAGAAATACTTCCGTGTCCTTGG	113	[19]

MMP10, and 0.02 μM *STC1*, and 3) 0.03 μM *PRM1*, 0.04 μM *TNP1*, 0.14 μM *KLK2*, 0.03 μM *MSMB*, and 0.02 μM *CYP2B7P*. Primer sequences and expected amplicon sizes are listed in Table 1. Body fluid markers and primers are protected under New Zealand Patent application 735997 and subsequent international applications [32].

2.7. Multiplex endpoint PCR

PCR was performed on a GeneAmp PCR System 9700 in 25 μL reactions using 12.5 μL Qiagen® Multiplex PCR buffer, 2.5 μL primer mix, and 2 μL or 10 μL cDNA. Where 2 μL cDNA was used, the total reaction volume of 25 μL was achieved by the addition of 8 μL nuclease-free water. DNA samples were amplified using an input of approximately 1.5 ng, performing dilutions where necessary. DNA from blood was preferred over saliva due to the potential of co-extracting plant material in animal saliva samples.

Amplification negative controls (ANEG) comprised nuclease-free water in place of cDNA. Amplification positive controls (APOS) were prepared from pooled cDNA from known samples per body fluid (five buccal samples for multiplex D, four menstrual fluid samples for multiplex Q, and four semen and four vaginal material samples for multiplex P) from various individuals. Each sample was individually tested for the presence of all target mRNAs prior to pooling. The resulting APOS samples were diluted in TE buffer to display peak heights of around 10,000 relative fluorescence units (RFU) without over-amplification.

The in-house protocol for RT-PCR [2] was optimised by adjusting the annealing temperature and duration, as well as the final elongation time. To allow for the use of a universal amplification protocol, PCR conditions were selected as those which maximised target signals simultaneously in all three multiplex assays. Final optimised PCR conditions were: initial denaturation at 95 °C for 15 min, followed by 35 cycles of 94 °C for 30 s, 60 °C for 3 min and 72 °C for 1 min, final elongation at 72 °C for 10 min, and cooling down to 4 °C.

2.8. Capillary electrophoresis and data analysis

PCR products were separated on a 3500xL Genetic Analyzer (Applied Biosystems®). Briefly, 9.6 μL Hi-Di™ was mixed with 0.4 μL GeneScan™ 600 LIZ® dye Size Standard v2.0 (Applied Biosystems®) per sample, to which 2 μL of PCR product was added. One amplification positive control and one negative control were injected per every 22 samples analysed. Samples were injected at a voltage of 1.2 kV for 24 s. Results were analysed using GeneMapper® ID-X v.1.5 (Applied Biosystems®) and an analytical threshold (AT) of 50 RFU. This value was initially chosen to capture low-level transcript abundance and was subsequently adjusted.

3. Results

3.1. Species specificity

All primate blood samples (except squirrel monkey) produced signals for both of the circulatory blood markers (Supplementary Table 1). Most signals were observed for *HBD*, particularly in primate and rabbit blood. This was expected, since primate mRNA is very similar to human mRNA (e.g., 98% sequence identity between human and northern white-cheeked gibbon *HBD* [33]). Furthermore, haemoglobins are widely expressed in many bird and mammal species, although some only possess a pseudogene [34]. *STC1* was only observed in the grey-headed flying fox sample. A signal the size of *MMP10* plus 2 bp was detected in cat blood. Amplification products of the same size as *CYP2B7P* were detected in the siamang gibbon and cotton-top tamarin samples. This could be the result of *CYP2B7P* expression in primates, whereas humans only possess a pseudogene. The cotton-top tamarin sample also displayed an off-scale *MSMB* peak.

The majority of animal saliva samples did not indicate the presence of target amplification products. Only the bonnet macaque sample produced *FDCSP*, *SLC4A1*, *MSMB*, and *CYP2B7P* signals (Supplementary Table 1). *FDCSP* was also detected in the squirrel monkey and dog samples. The cotton-top tamarin sample displayed *MSMB* and *CYP2B7P* peaks, which were also observed in blood. These were unlikely to originate from residual DNA, since the amplification of DNA did not give rise to comparable signals. Therefore, *MSMB* or low levels of *CYP2B7P* mRNA may be present in circulatory blood or saliva of some primate species. The remaining signals may have originated from amplification of trace amounts of mRNA due to overloading PCR reactions, since sample volumes were difficult to estimate. Additional amplification products outside expected marker positions were observed in most samples. These possibly resulted from unspecific primer binding and may be avoided by further increasing the annealing temperature [35].

Table 2
Body fluid specificity of the three multiplex assays. Values are given in RFU. Expected results are highlighted.

Sample	*FDCSP*	*HTN3*	*HBD*	*SLC4A1*	*MMP10*	*STC1*	*PRM1*	*TNP1*	*KLK2*	*MSMB*	*CYP2B7P*
1 µL saliva (10 µL RNA)	8761	21446	..	..	..	..	..	..	..	..	..
2.5 µL saliva (1 µL RNA)	1411	11608	..	..	..	..	..	..	..	..	..
0.1 µL circulatory blood (10 µL RNA)	..	..	9847	15279	..	..	..	..	..	..	..
0.5 µL circulatory blood (1 µL RNA)	..	..	4039	9634	..	..	..	..	..	..	..
0.25 µL semen (10 µL RNA)	..	..	..	..	..	..	8160	4491	17917	7654	..
0.5 µL semen (1 µL RNA)	..	..	..	..	..	..	14514	7404	1128	2139	..
1 µL seminal fluid (1 µL RNA)	..	..	..	..	..	..	187	..	6584	13212	..
0.5 µL seminal fluid (1 µL RNA)	..	..	..	..	..	..	..	..	6507	6244	..
Menstrual fluid (1:250, 10 µL RNA)	..	..	11164	24560	16115	4994	..	..	..	..	..
Menstrual fluid (1:150, 10 µL RNA)	..	..	10593	12879	12297	7993	..	..	..	..	692
Vaginal material (1:150, 10 µL RNA)	..	..	..	..	..	..	..	..	..	..	8140
Vaginal material (1:100, 10 µL RNA)	..	..	..	..	..	..	..	..	..	..	8274
DNA	..	..	..	..	60	..	..	54263[1]	..	..	..
APOS	4544	11251	4702	2821	4331	2272	4052	1324	8651	7744	4363
ANEG	..	..	..	..	..	..	..	..	..	..	..

[1] *TNP1* forward primer was placed across an exon/exon boundary.

Animal DNA samples mostly displayed raised baselines and numerous unspecific amplification products of peak heights below 1000 RFU (Supplementary Table 1). Although some peaks were of the same size as expected marker products, this likely occurred by coincidence. The appearance of several unexpected signals in combination with a noisy baseline was a good indicator for the presence of DNA. Signals exceeding 4000 RFU were observed for *TNP1* from bonnet macaque, pygmy marmoset, siamang gibbon, and spider monkey. This may be due to the fact that the *TNP1* primers amplified DNA (see Section 3.2). In addition, *MSMB* was observed in the golden lion tamarin sample.

3.2. Body fluid specificity

Table 2 shows that no cross-reactions from non-target body fluids were observed, except for a *PRM1* signal (187 RFU) in an azoospermic semen sample. However, spermatozoa can sometimes be present in semen following vasectomy [36]. In addition, *CYP2B7P* was undetected in one menstrual fluid sample. Cervical mucus and vaginal discharge contribute little to the total fluid volume lost during menstruation [37], hence corresponding markers may be present below the detection limit.

The human DNA sample produced a peak of 60 RFU for *MMP10* (Table 2). This signal could be attributed to elevated baseline and can be avoided by raising the AT. In addition, *TNP1* was amplified (54,263 RFU). This was likely due to the fact that the *TNP1* forward primer was placed across an exon/exon boundary, with only seven bases aligning to a different exon than the reverse primer. *TNP1* therefore cannot distinguish between mRNA and DNA templates, and a *TNP1* signal is not confirmatory for the presence of semen without additional controls. For example, *reverse transcriptase* negative (RT-) controls can help to verify whether residual genomic DNA may have contributed to a signal. Furthermore, massively parallel sequencing (MPS) could determine amplicon sequences and thus distinguish between templates in the future.

To evaluate the potential for false positives due to excessive sample input, ten samples per body fluid from five donors (10 µL saliva, 5 µL blood, 2 µL semen, and whole MF and VM swabs) were amplified. Target marker signals were typically over-amplified, i.e. in the 70,000–90,000 RFU range (Supplementary Table 2). Exceptions were target signals in menstrual fluid samples from donor R and *CYP2B7P* signals in menstrual fluid samples from all donors, which were considerably lower. This corroborates previous findings of variation in transcript abundance among individuals and samples [8,13].

Low-level cross-reactions were observed for all markers and body fluids, except for *MMP10*, *STC1*, *PRM1*, and *MSMB* in circulatory blood, *HBD*, *SLC4A1*, *PRM1*, and *KLK2* in saliva, and *HTN3* in menstrual fluid (Supplementary Table 2). This confirms previous reports of low transcript abundance in non-target body fluids for all currently known mRNAs [1,10,13,19]. Most signals were below 500 RFU and would likely be absent if a suitable AT were applied and target marker peaks were in the ideal range of 4000–12,000 RFU on a 3500xL instrument. However, cross-reactions exceeding 10,000 RFU were observed for *FDCSP* in two MF samples from two donors, for *MMP10* in two saliva, one semen, and three VM samples, as well as for *MSMB* in one VM sample. This demonstrates relatively higher *FDCSP*, *MMP10*, and *MSMB* transcript abundance in non-target body fluids and consequently lower specificity compared to the remaining mRNAs. Nevertheless, no cross-reactions were observed at ideal sample input (Table 2). It is therefore essential to limit sample input amounts and avoid over-amplification, although this may result in overlooking minor components of body fluid mixtures (see Section 3.6). *HTN3*, *HBD*, *SLC4A1*, and *PRM1* appeared to be the most specific markers. Examples of electropherograms for the three multiplex assays are shown in Fig. 1.

3.3. Sensitivity

The lower limit of detection (LOD) for the three multiplexes was approximately 0.5 µL saliva (multiplex D), 0.05 µL circulatory blood (multiplex Q), 0.05 µL semen containing spermatozoa (multiplex P), and 0.25 µL azoospermic seminal fluid (multiplex P) using 10 µL RNA for cDNA synthesis. These results were similar to other forensic multiplex systems [1,2,10,11,38]. For MF (multiplex Q) and VM (multiplex P), the LOD was difficult to determine since different mRNAs were undetected in less diluted samples and re-appeared in higher diluted samples. Approximately 1/50th of the RNA obtained from a whole swab was limiting, using 1 µL RNA for cDNA synthesis.

3.4. Precision

The precision of the three multiplexes was evaluated by triplicate amplification of the same cDNA samples. Standard deviations (σ) and coefficients of variation (CV), expressed as σ divided by the mean, were calculated from resulting peak heights.

The saliva markers displayed dispersion around the mean of 67% and 39% for *FDCSP*, and 77% and 103% for *HTN3*. This demonstrates a higher level of variability around the mean for *HTN3*, and moderate to low precision for both markers. Variability ranged between 8% and 49% for *HBD*, and between 18% and 36% for *SLC4A1*. Both circulatory blood markers therefore showed higher precision than the saliva markers. Less dispersion appeared to occur in MF samples. *MMP10*, *STC1*, and *CYP2B7P* showed variability between 21 and 24%, 14–16%, and 18–19%, respectively. These values demonstrate moderate to good levels of precision among replicates and samples, particularly for *STC1*.

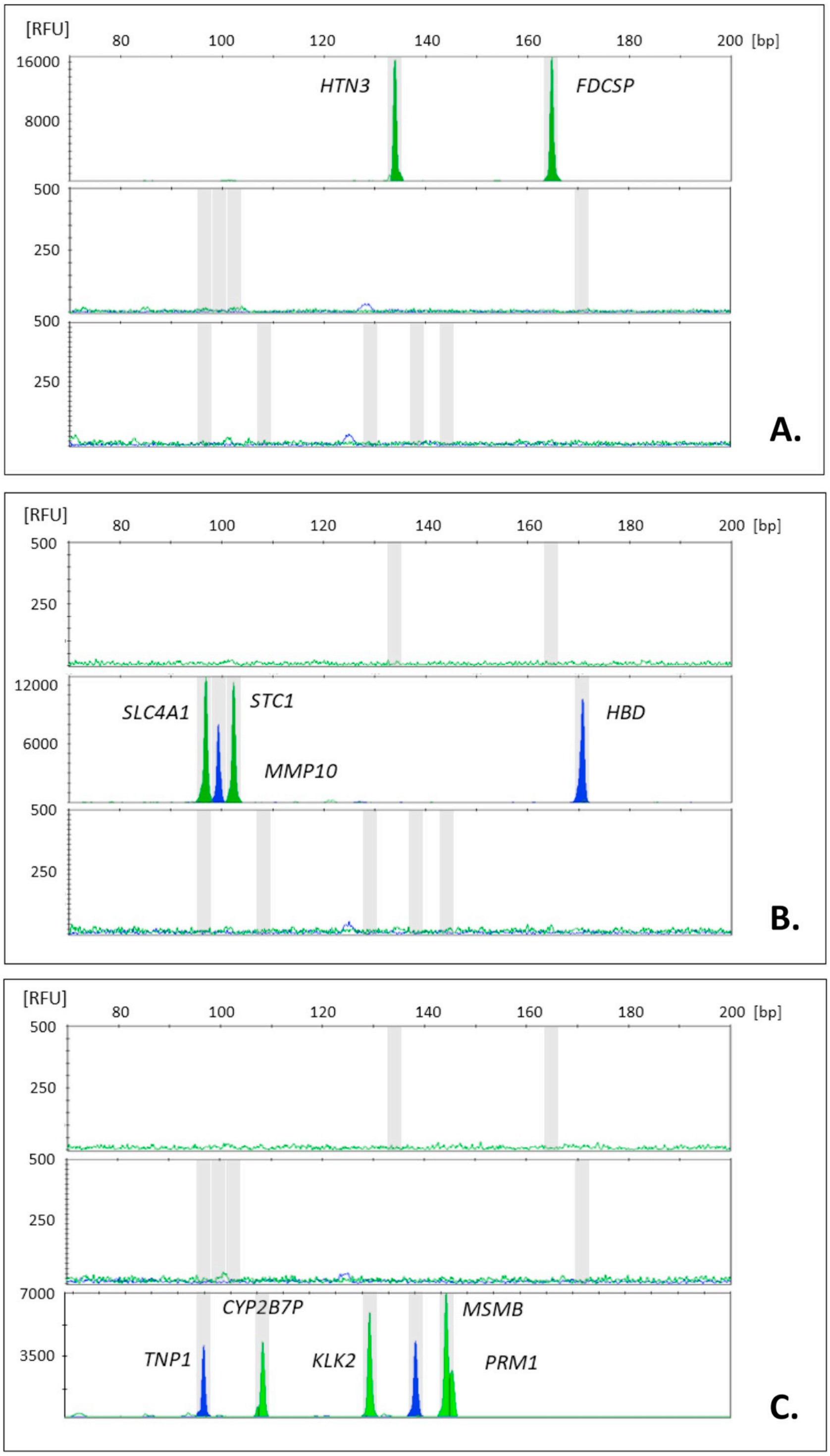

Fig. 1. Electropherograms of A. a buccal sample, B. a menstrual fluid sample, and C. a mixed sample of semen and vaginal material. Each sample was amplified using multiplex D (top), multiplex Q (middle), and multiplex P (bottom).

Variability ranged between 14 and 93% for *PRM1*, 7–53% for *TNP1*, 14–141% for *KLK2*, and 16–51% for *MSMB*. The high dispersion of *KLK2* in one semen sample (141%) was due to failure of amplification in two replicates. *KLK2* was also undetected in one replicate of a second semen sample, whereas all other mRNAs were consistently detected. Although high variability of peak heights is expected for mRNA analysis [39], further research including a greater number of replicates may determine CV values more precisely.

3.5. *The effect of multiplexing*

To investigate the effect that multiplexing has on target detection, 12 samples, i.e. two per body fluid, were amplified for a total of three replicates in both multiplex and uniplex reactions. All samples had previously shown ideal peak heights in multiplex amplifications. As Fig. 2 shows, only *HTN3* exclusively produced higher signals in multiplex compared to uniplex. For most markers and samples, higher average peak heights (APH) were obtained in uniplex reactions. This

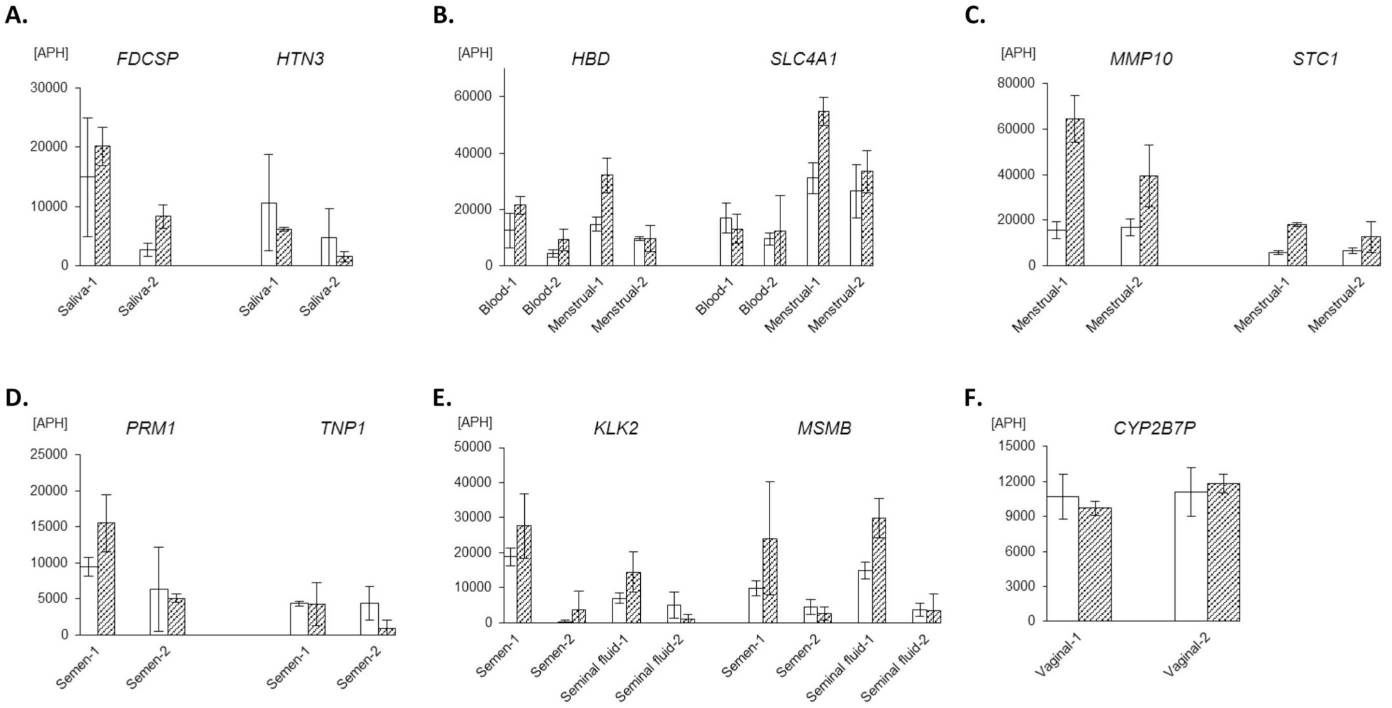

Fig. 2. The effect of multiplexing. APH obtained from three replicates each in multiplex (white bars) and in uniplex reactions (shaded) for A. 0.05 µM *FDCSP* and 0.012 µM *HTN3*, B. 0.05 µM *HBD* and 0.04 µM *SLC4A1*, C. 0.04 µM *MMP10* and 0.02 µM *STC1*, D. 0.03 µM *PRM1* and 0.04 µM *TNP1*, E. 0.14 µM *KLK2* and 0.03 µM *MSMB*, and F. 0.02 µM *CYP2B7P*. Error bars represent one standard deviation.

was expected due to the lack of competition among different primer sets in uniplex amplifications [35]. The strongest negative effect was observed for *MMP10* and *SLC4A1*. APH were up to 4.1- and 1.8-fold lower in multiplex compared to uniplex reactions, respectively. This was likely the result of low heterodimerisation values between primers ($\Delta G \geq -9.76$ kcal/mol). Interestingly however, differences in APH for *SLC4A1* and *HBD* were more pronounced in MF than in circulatory blood.

Whereas no clear tendency towards increased signals in uni- or multiplex was observed for *PRM1*, *TNP1* appeared to perform slightly better in multiplex. This mRNA was consistently detected in multiplex, while two uniplex replicates failed to amplify. *KLK2* and *MSMB* respectively were also undetected in four and two of 12 replicates using uniplex reactions, whereas only three and zero replicates failed in multiplex. The effect for *CYP2B7P* was negligible, although standard deviations were slightly higher in multiplex.

In 60% of 30 marker observations averaged from triplicate amplifications, the target markers exhibited less peak height variance in multiplex than in uniplex (data not shown). *TNP1*, *KLK2*, and *MSMB* exclusively showed higher precision in multiplex. Thus, while multiplexing exerted a negative effect on absolute peak height and therefore target detection, the markers had a tendency towards increased precision and consistent amplification in multiplex. The loss in peak height due to multiplexing was counteracted by the adjustment of primer concentrations, which balanced signals among markers within the same multiplex.

3.6. Resolution of body fluid mixtures

All body fluid mixtures were correctly resolved, except for one sample of 1 µL saliva mixed with 2 µL semen (Table 3). Using the undiluted cDNA sample derived from a 1:8 dilution of the extracted RNA, *FDCSP* and *HTN3* reached 5829 RFU and 3135 RFU, whereas the semen markers ranged between 11,521 RFU for *MSMB* and 40,745 RFU for *KLK2* (data not shown). The circulatory blood and MF markers were undetected in both amplifications. The additional dilution of the cDNA sample to adjust peak heights of the semen markers to the ideal 4000–12,000 RFU range resulted in loss of signal for the saliva markers. This implies that uneven mixtures with an abundant major component and a small minor component may fail to be correctly resolved.

CYP2B7P was not observed in mixtures containing menstrual fluid. This was likely because this mRNA was present below the detection threshold, as observed previously (Section 3.2). *TNP1* was also undetected in two samples containing semen, likely due to amplification failure. Two unexpected signals (*MMP10*, 58 RFU and KLK2, 50 RFU) resulted from elevated baseline. Importantly, greater body fluid volumes did not necessarily produce higher peaks. Although *HBD* signals increased with larger blood volumes in the first set of mixtures with MF, the second set of mixtures did not show this correlation. This probably resulted from differences in template abundance among samples. It is therefore not possible to reliably estimate relative compositions of body fluid mixtures from peak heights alone, although higher template concentrations generally tended to produce higher signals.

3.7. Detection of seminal mRNAs in post-coital vaginal samples

To evaluate the time frame, during which seminal mRNAs could be detected on vaginal swabs collected post intercourse, 24 samples with a time since intercourse (TSI) between one and six days were amplified using multiplex P. The results are shown in Table 4.

All four seminal markers were consistently detected for up to three days post intercourse. The lowest signal from a TSI 3 d sample was 1469 RFU for *PRM1* (sample D19). Swabs collected four days post intercourse also exhibited all four seminal markers, except sample D10, which did not show a *KLK2* signal. This possibly resulted from amplification failure, as *KLK2* was also undetected in the APOS sample, whereas the remaining markers displayed ideal peak heights. Replicate amplification could recover *KLK2* in this sample. Improved primer sequences may help to avoid amplification failure in the future.

The two samples collected after five days (D11 and D26) each displayed *MSMB* and one additional marker. Whereas one sample with a TSI of six days (D12) was undetected, the second sample (D27) showed

Table 3
Resolution of body fluid mixtures. Values are given in RFU. MF was collected on day 2 of the uterine cycle from a naturally cycling donor. Samples were 14 weeks old when further components were added. VM was collected on day 19 of the uterine cycle from a naturally cycling donor. Samples were 11 weeks old when further components were added. For samples containing MF, VM, or semen as component 1, the RNA was diluted 1:75, 1:50, and 1:8, respectively, prior to RT. Further dilution of cDNA samples was carried out for MF-blood, MF-semen (5 µL and 10 µL), and semen-saliva mixtures to adjust peak heights. SA = saliva, SM = semen, ENEG = extraction negative control. Expected results are highlighted in light grey, expected results not detected are highlighted in dark grey.

Component 1	Component 2	Component 3	*FDCSP*	*HTN3*	*HBD*	*SLC4A1*	*MMP10*	*STC1*	*PRM1*	*TNP1*	*KLK2*	*MSMB*	*CYP2B7P*
1/3 MF swab	1 µL BL	-	..	..	1571	3664	7086	9807	..	..	..	..	..
1/3 MF swab	5 µL BL	-	..	..	9753	20413	8472	21721	..	..	..	..	..
1/3 MF swab	10 µL BL	-	..	..	12315	19872	8601	12956	..	..	..	..	..
1/3 MF swab	1 µL BL	-	..	..	6249	16318	19536	25545	..	..	..	..	..
1/3 MF swab	5 µL BL	-	..	..	2295	13001	12188	19191	..	..	..	..	..
1/3 MF swab	10 µL BL	-	..	..	3166	16881	4428	5460	..	..	..	..	..
1/3 MF swab	1 µL SM	-	..	..	6062	6601	30329	29343	14917	12288	14164	12421	..
1/3 MF swab	5 µL SM	-	..	..	311	2045	10582	11989	1025	..	14952	7074	..
1/3 MF swab	10 µL SM	-	..	..	764	6136	15707	33366	10836	4419	37190	19283	..
1/3 VM swab	1 µL SA	-	1621	231	..	..	..	..	..	..	..	..	4282
1/3 VM swab	5 µL SA	-	23238	5822	..	..	..	..	..	..	50	..	17977
1/3 VM swab	10 µL SA	-	5242	5582	..	..	..	..	..	..	..	..	4183
1/3 VM swab	1 µL SA	-	53	986	..	..	..	..	..	..	..	..	3338
1/3 VM swab	5 µL SA	-	50406	6040	..	..	58	..	..	..	..	..	4881
1/3 VM swab	10 µL SA	-	4033	6744	..	..	..	..	..	..	..	..	8114
2 µL SM	1 µL SA	-	..	..	..	..	..	..	16620	7496	7958	4032	..
2 µL SM	5 µL SA	-	8437	3136	..	..	..	..	20254	6410	24811	16595	..
2 µL SM	10 µL SA	-	4399	2852	..	..	..	..	15859	851	4834	3413	..
2 µL SM	1 µL SA	-	2216	467	..	..	..	..	24749	..	18672	6716	..
2 µL SM	5 µL SA	-	20124	1156	..	..	..	..	10871	2604	30234	4758	..
2 µL SM	10 µL SA	-	11780	2374	..	..	..	..	21566	14463	5409	4016	..
VM swab	2 µL SM	10 µL SA	23670	2281	..	..	..	..	10326	7714	34099	5536	11779
VM swab	2 µL SM	10 µL SA	15709	2962	..	..	..	..	21109	18189	34287	18756	12074
ENEG			..	..	..	..	..	..	..	..	..	..	..
APOS			13151	21233	1815	13105	5979	7181	5034	1362	15315	15746	10890
ANEG			..	..	..	..	..	..	..	..	..	..	..

Table 4
Amplification of post-coital vaginal samples using multiplex P. All samples originated from the same two donors.

Sample	TSI [d]	*PRM1* [RFU]	*TNP1* [RFU]	*KLK2* [RFU]	*MSMB* [RFU]	*CYP2B7P* [RFU]
D7	1	53,019	52,613	51,435	53,441	45,548
D8	2	38,110	40,106	31,701	23,211	39,685
D9	3	13,230	17,152	7805	6388	39,692
D10	4	1540	68	..	1199	40,162
D11	5	..	..	2427	1505	38,527
D12	6	..	..	..	..	39,487
D13	1	39,420	40,949	38,386	41,219	40,931
D14	1	19,750	25,374	23,331	14,049	45,938
D15	2	3042	4408	9076	3382	43,595
D16	3	4565	6674	5387	2350	36,805
D17	1	11,109	12,920	14,528	19,963	38,775
D18	2	6976	9695	8050	11,864	37,921
D19	3	1469	3780	8480	3351	40,988
D20	4	778	4640	4178	1725	43,034
D21	1	42,714	42,941	11,261	17,857	42,579
D22	1	44,767	44,720	3923	45,009	44,564
D23	2	36,215	40,950	18,612	13,093	45,131
D24	3	28,344	20,011	7064	17,859	39,485
D25	4	3563	3382	4165	3180	35,026
D26	5	1138	..	..	3257	38,958
D27	6	903	..	..	..	35,199
D28	1	38,872	39,961	40,210	40,676	40,186
D29	2	34,086	41,455	20,364	20,989	41,503
D30	3	24,524	19,618	12,684	21,045	45,680
APOS	–	900	3330	..	4147	3394
ANEG	–	51	..	..	..	51

a *PRM1* peak (903 RFU). Hence, the detection of seminal mRNAs in post-coital samples using the pentaplex is possible for up to six days. These results demonstrate a considerable enhancement of marker detection in post-coital samples compared to previous studies [13], which reported that the detection of seminal mRNAs was limited to samples with a TSI ≤ 1 d. Two signals in the ANEG sample resulted from elevated baseline.

3.8. *Stability studies*

The forensic literature reported the successful mRNA amplification from body fluids up to 56 years after deposition [40]. In this research, the ability to detect and identify aged body fluids, aged RNA, and aged cDNA samples was investigated. Five single-source samples for each of these three categories were selected with regard to storage time and amplified using all three multiplex assays, performing cDNA dilutions where necessary. In addition, an aged cDNA sample obtained from a nosebleed was analysed. The results are shown in Table 5.

All aged circulatory blood samples (17–25 months old) were correctly identified, with no cross-reactions observed. Aged RNA samples (29–35 months old) correctly exhibited all target markers, except for *CYP2B7P*, which was absent from the menstrual fluid sample. Aged cDNA samples (15–30 months old) were also successfully amplified, with no cross-reactions present. In the aged MF cDNA sample, the menstrual fluid marker *STC1* was undetected, however a strong *CYP2B7P* signal provided additional confidence in the vaginal origin of the sample.

The nosebleed sample correctly exhibited signals for *HBD* and *SLC4A1*, whereas *PRM1*, *TNP1*, and *KLK2* were undetected. *FDCSP* and *HTN3* signals were also lacking, although the detection of saliva makers in nasal mucosa has been reported [10]. However, *MMP10*, *STC1*, *CYP2B7P*, and in particular *MSMB* were observed. This may be problematic, since these results falsely indicate the presence of a mixture of MF and semen. One previous study also reported the amplification of *CYP2B7P* from nasal mucosa [10]. An AT of 400 RFU would prevent false positive identification of *STC1* and *CYP2B7P*, but still allow for *MMP10* and *MSMB* to be identified. Caution is therefore warranted in the interpretation of mRNA profiling results in the possible presence of nasal mucosa. Consequently, a *MMP10* signal without detecting *STC1* or *CYP2B7P* was considered not confirmatory for MF (unless the *MMP10* peak height exceeds those of the circulatory blood markers), whereas *MSMB* must be accompanied by a second semen marker to confirm the presence of semen. Further research using MPS

Table 5

Marker detection in aged samples. Peak heights (RFU) were obtained from aged body fluid samples, aged RNA, and aged cDNA, stored at room temperature or frozen for 15 to 35 months. ENEG = extraction negative control. Expected results are highlighted in light grey, expected results not detected are highlighted in dark grey.

Sample Description	Age	Storage	*FDCSP*	*HTN3*	*HBD*	*SLC4A1*	*MMP10*	*STC1*	*PRM1*	*TNP1*	*KLK2*	*MSMB*	*CYP2B7P*
20 µL circulatory blood deposited on a rayon swab	24 months	Dark, room temperature	..	..	3916	19040	..	..	..	..	..	..	..
20 µL circulatory blood deposited on a rayon swab	25 months	Dark, room temperature	..	..	5668	7574	..	..	..	..	..	..	..
20 µL circulatory blood deposited on a rayon swab	24 months	Dark, room temperature	..	..	3936	16586	..	..	..	..	..	..	..
20 µL circulatory blood deposited on a rayon swab	23 months	Dark, room temperature	..	..	546	9533	..	..	..	..	..	..	..
20 µL circulatory blood deposited on a rayon swab	17 months	Dark, room temperature	..	..	696	6537	..	..	..	..	..	..	..
RNA from 50 µL circulatory blood	35 months	Frozen	..	..	17636	26586	..	..	..	..	..	..	..
RNA from a buccal swab	33 months	Frozen	2099	4249	..	..	..	..	..	..	..	..	..
RNA from 5 µL semen containing spermatozoa	29 months	Frozen	..	..	..	..	..	..	11812	4084	4393	498	..
RNA from a menstrual fluid swab	30 months	Frozen	..	..	5774	1867	64475	20765	..	..	..	..	..
RNA from a vaginal swab	33 months	Frozen	..	..	..	..	..	..	..	..	..	..	4645
cDNA from 50 µL circulatory blood	30 months	Frozen	..	..	15378	25925	..	..	..	..	..	..	..
cDNA from a buccal swab	29 months	Frozen	9627	7857	..	..	..	..	..	..	..	..	..
cDNA from 5 µL semen containing spermatozoa	29 months	Frozen	..	..	..	..	..	..	26703	12911	21623	8243	..
cDNA from a menstrual fluid swab	15 months	Frozen	..	..	2201	3372	512	..	..	..	..	..	14560
cDNA from a vaginal swab	29 months	Frozen	..	..	..	..	..	..	..	..	..	..	903
cDNA from a nosebleed (blood and nasal mucosa)	18 months	Frozen	..	..	11294	13047	1049	162	..	..	..	28333	192
ENEG			..	..	..	..	..	..	..	..	..	..	..
APOS			15908	22935	2250	4464	4450	758	1782	1847	..	5725	3258
ANEG			..	..	..	..	..	..	..	..	..	..	..

technologies could determine whether amplicons detected in different body fluids are derived from the same mRNA sequences. Additionally, the multiplex system may be extended to include nasal mucosa markers.

3.9. Case-type samples

Case-type samples were processed in a blind study, in which sample sources were withheld from the researcher. A total of twelve samples (six swabs (samples 1–6) and six tape lifts (samples 7–12)) were analysed. All samples were initially amplified using 10 µL RNA and 10 µL cDNA. Subsequent cDNA dilutions were performed where necessary. Based on the results obtained in the previous sections, an AT of 400 RFU was applied for peak allocation and dilutions were performed when the highest peak exceeded 20,000 RFU. This fourfold increase in peak height acceptable for calling (compared to 100–5000 RFU for CellTyper) corresponded with the roughly fourfold increase in the dynamic scaling range of the 3500-series instruments compared to the 31xx generation. All samples or highest dilutions thereof were also amplified using CellTyper to evaluate and compare multiplex performance. The results are displayed in Table 6. RT- controls were prepared for all samples. None of these displayed any marker peaks (data not shown).

Three samples (3, 8, and 11) exhibited no marker peaks using either multiplex system. Sample 3 was a saliva sample from a chicken, and therefore correctly lacking mRNA results. Sample 8 was obtained from the inside of the crotch of a pair of men's undergarments from an azoospermic male. Hence, the presence of seminal fluid was probable. Sample 11 was a tape lift from a coffee cup and therefore expected to contain saliva. The collected material may have been insufficient to produce a result for these two samples.

Samples 1 (vaginal swab), 2 (skin swab of saliva and blueberry juice), 7 (inside of the crotch of a pair of men's undergarments), and 12 (bloodstain) were undetermined using CellTyper. The new multiplex system confirmed the presence of vaginal material for sample 1. This demonstrates that *lactobacilli* can be unreliable VM markers in some individuals. The detection of *CYP2B7P*, however, enabled determination of the source of this sample. A *TNP1* signal (611 RFU) was obtained for sample 2. This should be interpreted with caution, since it could have originated from residual genomic DNA, although the RT- control was devoid of target signals. Replicate quantification and/or amplification (where sample volumes permit) could verify the possible presence of DNA residues in such samples. For sample 7, the new multiplex assay confirmed the presence of seminal fluid. *TNP1* added strong support for the presence of semen but should be interpreted with some caution due to the risk of amplification from DNA. *MMP10* was not informative, since no corresponding mRNAs were detected. Finally, *HBD* and *SLC4A1* were observed in sample 12 (Fig. 3), which correctly confirmed the presence of circulatory blood. These results demonstrate improved body fluid detection using the new multiplexes compared to CellTyper in three of the four samples.

Sample 4 was identified as VM using the new multiplexes. Although this was a correct result, the assay failed to detect saliva as the second component (Table 6). In contrast, only saliva was confirmed in sample 5. This swab also comprised a mixture of saliva and VM. Saliva had been applied after (sample 5) or before (sample 4) collecting the VM sample. This could indicate that the cell lysis during the extraction process is most likely to remove cellular material from the outermost surface of a swab. Another explanation may be that the body fluid proportions were too uneven to be resolved. CellTyper detected saliva in both samples. This demonstrates higher sensitivity for saliva compared to the new assay. In turn, however, CellTyper failed to identify VM in either sample.

Both assays correctly confirmed the presence of saliva in sample 6. This sample further contained traces of blood, which neither assay detected. The possible presence of saliva was also expected for sample 9 (tape lift from the neck and upper front of a T-shirt). The new multiplexes detected *FDCSP*, *MMP10*, and *MSMB*. These signals were insufficient to confirm the presence of a body fluid. CellTyper detected corresponding marker types (*STATH* and *MMP11*), which also did not confirm a body fluid. It appears that mRNA background levels may be present on some everyday objects, which could be addressed by further research.

The improved multiplex system confirmed the presence of circulatory blood in sample 10. *MMP10* was also observed but was not informative due to the absence of additional mRNAs. This sample was collected from the inside of the crotch of a pair of men's undergarments, with traces of blood applied. CellTyper detected *TGM4*, which indicated the presence of seminal fluid, but failed to detect blood. Overall, the new multiplex assay seemed to be more sensitive for circulatory blood

Table 6
Analysis of case-type samples. Expected results are highlighted in light grey, expected results not detected are highlighted in dark grey.

Sample	Appearance	Expected result[1]	*FDCSP*	*HTN3*	*HBD*	*SLC4A1*	*MMP10*	*STC1*	*PRM1*	*TNP1*	*KLK2*	*MSMB*	*CYP2B7P*	Results from CellTyper amplification[2]
1 (swab)	/	VM (VM swab)	..	..	..	..	..	..	..	..	..	..	12279	NR
2 (swab)	Slightly red	SA (skin swab of saliva with blueberry juice present – possible interference)	..	..	..	..	..	..	..	611	..	..	..	NR
3 (swab)	/	NR (swab from the inside of the mouth of a chicken)	..	..	..	..	..	..	..	..	..	..	..	NR
4 (swab)	/	VM + SA present (VM swab with saliva applied first)	..	..	..	..	..	..	..	..	..	..	12069	SA confirmed (*HTN3*)
5 (swab)	Slightly red	VM + SA present (VM swab with saliva applied afterwards)	18482	879	..	..	..	..	..	..	..	..	..	SA confirmed (*HTN3* + *STATH*)
6 (swab)	Slightly red	SA ± BL (saliva with traces of blood)	15794	1345	..	..	..	..	..	..	..	..	..	SA confirmed (*HTN3* + *STATH*)
7 (tape lift)	/	SP and SF (inside of the crotch of a pair of men's undergarments)	..	..	..	..	431	..	..	1659	1623	9718	..	NR
8 (tape lift)	/	SF (inside of the crotch of a pair of men's undergarments – azoospermic)	..	..	..	..	..	..	..	..	..	..	..	NR
9 (tape lift)	/	SA (neck and front of a t-shirt)	3047	..	..	..	11539	..	..	..	..	972	..	Possible SA (*STATH*) MMP11 inconclusive
10 (tape lift)	/	SP, SF, BL (inside of the crotch of a pair of men's undergarments, traces of blood)	..	..	9225	1475	1710	..	..	..	..	..	..	Possible SF (*TGM4*)
11 (tape lift)	/	SA (top of a coffee mug)	..	..	..	..	..	..	..	..	..	..	..	NR
12 (tape lift)	/	BL (bloodstain)	..	..	3766	10373	..	..	..	..	..	..	..	NR

[1] Expected results were disclosed after completion of mRNA analysis. BL = circulatory blood, SA = saliva, SP = spermatozoa, SF = seminal fluid, VM = vaginal material, NR = no result.

[2] CellTyper amplifications were performed as published [2]. PCR products were separated on a Genetic Analyzer 3130*xl*, with a peak amplitude threshold of 100 RFU.

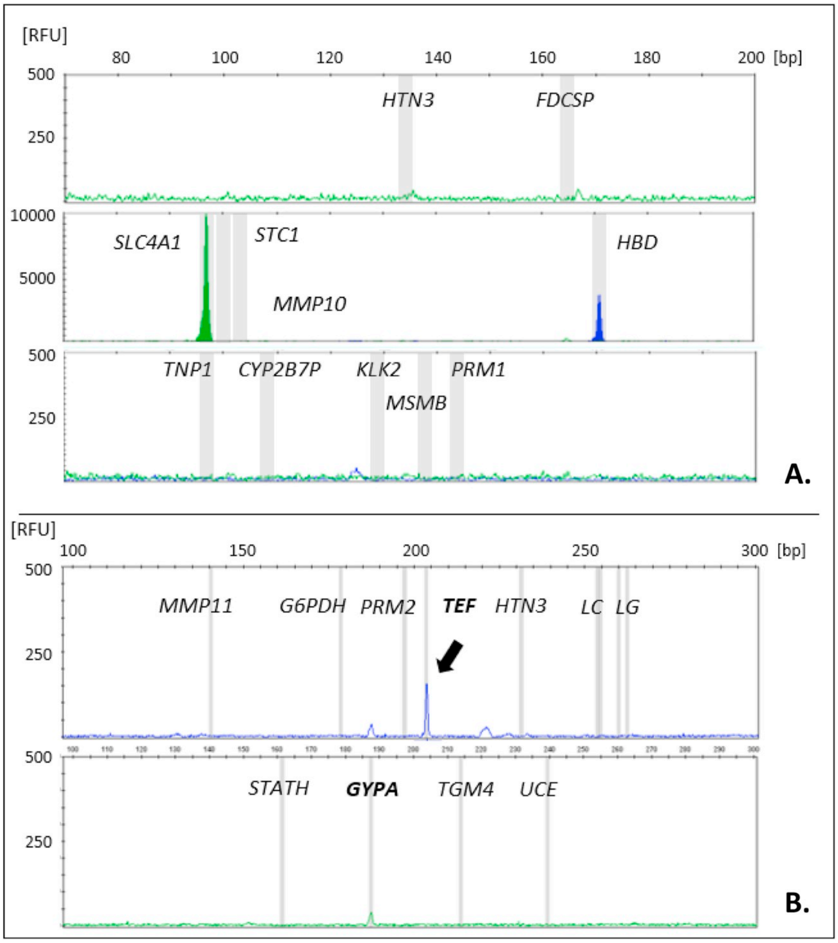

Fig. 3. A. Electropherograms of case-type sample 12 amplified with the three new multiplexes D (top), Q (middle), and P (bottom). *SLC4A1* and *HBD* were detected using multiplex Q, thus confirming the presence of circulatory blood. B. Electropherogram of case-type sample 12 amplified with CellTyper. Only one peak was present (marked by an arrow), which corresponded to the housekeeping gene *TEF*. The circulatory blood marker *GYPA* was undetected. Dye channels were superimposed for multiplexes Q and P and separated for CellTyper. The sample originated from a bloodstain.

and seminal mRNAs, whereas CellTyper was more sensitive for saliva. Further adjustment of primer concentrations may increase the sensitivity of the new multiplex system for saliva.

4. Discussion

The results of this validation study demonstrate successful application of the three endpoint RT-PCR multiplex assays to the identification of low abundance and aged body fluid samples, as well as to the resolution of mixtures and case-type samples. The optimised system showed similar specificity and sensitivity to other forensic multiplex systems [1,2,38], with improved performance on 4/12 case-type samples compared to CellTyper. Samples were amplified with all three assays in the new multiplex system to avoid false positive marker identification resulting from detecting spurious transcript levels in non-target body fluids. Samples had to be diluted and the amplification repeated when the highest peak in either of the three assays exceeded 20,000 RFU. Although this increases the workload within the laboratory, separation of markers into three small assays maintains target sensitivity and efficiency of amplification.

The species specificity study highlighted that some primer sequences were not human-specific. *HBD* was frequently amplified from non-human blood samples, particularly from primates, cat, and rabbit. Large, red stains should therefore be analysed with caution. Cotton-top tamarin, bonnet macaque, and siamang gibbon samples also readily produced false positives for *CYP2B7P* and *MSMB*. Saliva samples gave less false positives, although *FDCSP* was recovered from dog saliva. The occurrence of multiple extra peaks in an electropherogram was a strong indicator of the presence of genomic DNA. The analyst should therefore carefully review the framework of the case and consider whether samples may be giving false positive results. The absence of a DNA profile can additionally indicate the presence of a non-human body fluid. If the presence of animal body fluids is suspected, additional species testing should be carried out.

Across all human body fluids, higher volumes of body fluid, RNA, and cDNA generally produced higher signals. There was no indication of inhibitory effects at increased template amounts, although high-template samples may show increased baseline noise and non-specific peaks that could fall into marker windows. False positives readily occurred in overloaded PCR reactions. These may be caused by low-level gene expression in non-target body fluids or artefact formation resulting from non-specific primer annealing. It was therefore essential to adjust input amounts (i.e. limit peak heights to ≤20,000 RFU on a 3500xL) to establish marker specificity. Replicate amplifications may be useful to identify cross-reactions. RT- controls can provide additional information on whether DNA may have contributed to a signal. An AT of 400 RFU for 3500xL data is recommended to additionally help prevent false positive marker identification using the new multiplex assay.

Throughout this study, high inter-individual and inter-sample variation was observed, although the body fluids detected were consistent among replicates. This was expected due to the multitude of factors that affect gene expression [8] and the inability, at present, to measure the human-specific RNA concentration in a sample [41]. The impact of this variation was further exacerbated by low precision among replicates. Multiplexing increased overall precision but had a detrimental effect on absolute peak height for most markers. Additionally, stochastic effects were prominent in low-template samples. Drop-out was observed for various markers at low RNA concentrations, whereas the same markers re-appeared at even lower RNA concentrations.

Mixtures of vaginal material and semen in samples collected post intercourse were successfully identified for up to six days. It is important to note, however, that mixtures with uneven proportions may not be fully resolved. Whereas the major component was successfully detected in all mixtures analysed, minor component(s) may be undetected because of low abundance, resulting in signals below the detection threshold. This is a general limitation of the technique. In view of the above results, the newly developed multiplex system provides an enhanced method for body fluid and cell type assessment of forensic samples using mRNA profiling.

Acknowledgements

The authors are grateful to all volunteers for their participation in this study. Special thanks to Auckland Zoo for providing animal samples. The authors also acknowledge the advisory contribution of J Patel, as well as critical reading of the manuscript by L Melia and D Foskin. Funding for this work was granted by the Institute of Environmental Science and Research Ltd. (ESR) Core Fund. This research was completed in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Auckland, Auckland, New Zealand.

Compliance with ethical standards

This study was approved by the University of Auckland Human Participants Ethics Committee (UAHPEC), reference number 011733. All samples were collected with full informed consent.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scijus.2019.01.001.

References

[1] 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, For. Sci. Int. Genet. 6 (2012) 565–577.

[2] R.I. Fleming, S. Harbison, The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids, For. Sci. Int. Genet. 4 (2010) 244–256.

[3] J. Juusola, J. Ballantyne, Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification, Forensic Sci. Int. 135 (2003) 85–96.

[4] C. Haas, B. Klesser, C. Maake, W. Bar, A. Kratzer, mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR, For. Sci. Int. Genet. 3 (2009) 80–88.

[5] E.K. Hanson, H. Lubenow, J. Ballantyne, Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs, Anal. Biochem. 387 (2009) 303–314.

[6] R.I. Fleming, S. Harbison, The use of bacteria for the identification of vaginal secretions, For. Sci. Int. Genet. 4 (2010) 311–315.

[7] J.H. An, A. Choi, K.J. Shin, W.I. Yang, H.Y. Lee, DNA methylation-specific multiplex assays for body fluid identification, Int. J. Legal Med. 127 (2013) 35–43.

[8] T. Sijen, Molecular approaches for forensic cell type identification: on mRNA, miRNA, DNA methylation and microbial markers, For. Sci. Int. Genet. 18 (2015) 21–32.

[9] M. van den Berge, A. Carracedo, I. Gomes, E.A. Graham, C. Haas, B. Hjort, P. Hoff-Olsen, O. Maronas, B. Mevag, N. Morling, H. Niederstatter, 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, For. Sci. Int. Genet. 10 (2014) 40–48.

[10] 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, For. Sci. Int. Genet. 20 (2016) 119–129.

[11] J. Juusola, J. Ballantyne, Multiplex mRNA profiling for the identification of body fluids, Forensic Sci. Int. 152 (2005) 1–12.

[12] J. Juusola, J. Ballantyne, mRNA profiling for body fluid identification by multiplex quantitative RT-PCR, J. Forensic Sci. 52 (2007) 1252–1262.

[13] 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. Legal Med. 127 (2013) 707–721.

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

[15] S.M. Park, S.Y. Park, J.H. Kim, T.W. Kang, J.L. Park, K.M. Woo, J.S. Kim, H.C. Lee, S.Y. Kim, S.H. Lee, Genome-wide mRNA profiling and multiplex quantitative RT-PCR for forensic body fluid identification, For. Sci. Int. Genet. 7 (2013) 143–150.

[16] D. Zubakov, E. Hanekamp, M. Kokshoorn, W. van Ijcken, M. Kayser, Stable RNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples, Int. J. Legal Med. 122 (2008) 135–142.

[17] R. Fang, C.F. Manohar, C. Shulse, M. Brevnov, A. Wong, O.V. Petrauskene, P. Brzoska, M.R. Furtado, Real-time PCR assays for the detection of tissue and body fluid specific mRNAs, Int. Congr. Ser. 1288 (2006) 685–687.

[18] M.H. Lin, P.P. Albani, R. Fleming, Degraded RNA transcript stable regions (StaRs) as targets for enhanced forensic RNA body fluid identification, For. Sci. Int. Genet. 20 (2016) 61–70.

[19] E.K. Hanson, J. Ballantyne, Highly specific mRNA biomarkers for the identification of vaginal secretions in sexual assault investigations, Sci. Justice 53 (2013) 14–22.

[20] Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics, Nat. Rev. Genet. 10 (2009) 57–63.

[21] P.P. Albani, R. Fleming, Novel messenger RNAs for body fluid identification, Sci. Justice 58 (2018) 145–152.

[22] Scientific Working Group on DNA Analysis Methods, Validation Guidelines for DNA Analysis Methods, https://www.swgdam.org/publications (Last accessed June 2017 (Updated 2016)).

[23] M.H. Lin, D.F. Jones, R. Fleming, Transcriptomic analysis of degraded forensic body fluids, For. Sci. Int. Genet. 17 (2015) 35–42.

[24] M. Bauer, D. Patzelt, Protamine mRNA as molecular marker for spermatozoa in semen stains, Int. J. Legal Med. 117 (2003) 175–179.

[25] R. Fleming, S. Harbison, The development of an mRNA amplification positive control for the identification of body fluids, Can. Soc. For. Sci. 47 (2014) 37–45.

[26] X. Liu, X. Yu, D.J. Zack, H. Zhu, J. Qian, TiGER: a database for tissue-specific gene expression and regulation, BMC Bioinformatics 9 (2008) 271.

[27] S.J. Xiao, C. Zhang, Q. Zou, Z.L. Ji, TiSGeD: a database for tissue-specific genes, Bioinformatics 26 (2010) 1273–1275.

[28] J. Pan, S. Hu, D. Shi, M. Cai, Y. Li, Q. Zou, Z. Ji, PaGenBase: a pattern gene database for the global and dynamic understanding of gene function, PLoS One 8 (2013) e80747.

[29] C. Wu, C. Orozco, J. Boyer, M. Leglise, J. Goodale, S. Batalov, C.L. Hodge, J. Haase, J. Janes, J.W. Huss, BioGPS: An extensible and customizable portal for querying and organizing gene annotation resources, Genome Biol. 10 (2009) R130.

[30] M. Safran, I. Dalah, J. Alexander, N. Rosen, T. Iny Stein, M. Shmoish, N. Nativ, I. Bahir, T. Doniger, H. Krug, A. Sirota-Madi, T. Olender, Y. Golan, G. Stelzer, A. Harel, D. Lancet, GeneCards Version 3: the human gene integrator, Database (2010) baq020.

[31] A. Whitehead, D.L. Crawford, Variation in tissue-specific gene expression among natural populations, Genome Biol. 6 (2005) R13.

[32] P.P. Albani, J. Patel, R. Fleming, Method for body fluid detection, New Zealand patent application 735997 (2017).

[33] D.A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, E.W. Sayers, GenBank, Nucleic Acids Res. 41 (2013) D36–D42.

[34] R.C. Hardison, Evolution of hemoglobin and its genes, Cold Spring Harb. Perspect Med. 2 (2012) a011627.

[35] O. Henegariu, N.A. Heerema, S.R. Dlouhy, G.H. Vance, P.H. Vogt, Multiplex PCR: critical parameters and step-by-step protocol, BioTechniques 23 (1997) 504–511.

[36] G.E. Lemack, M. Goldstein, Presence of sperm in the pre-vasectomy reversal semen analysis: incidence and implications, J. Urol. 155 (1996) 167–169.

[37] I.S. Fraser, G. McCarron, R. Markham, T. Resta, Blood and total fluid content of menstrual discharge, Obstet. Gynecol. 65 (1985) 194–198.

[38] C. Haas, B. Klesser, A. Kratzer, W. Bar, mRNA profiling for body fluid identification, For. Sci. Int. Genet. Suppl. Ser. 1 (2008) 37–38.

[39] J. Harteveld, A. Lindenbergh, T. Sijen, RNA cell typing and DNA profiling of mixed samples: can cell types and donors be associated? Sci. Justice 53 (2013) 261–269.

[40] H. Nakanishi, M. Hara, S. Takahashi, A. Takada, K. Saito, Evaluation of forensic examination of extremely aged seminal stains, Legal Med. 16 (2014) 303–307.

[41] A. Lindenbergh, P. Maaskant, T. Sijen, Implementation of RNA profiling in forensic casework, For. Sci. Int. Genet. 7 (2013) 159–166.